Imagine Yourself

Explorations in Fostering Personal Expression with Generative AI

by

Karishma Chadha

B.A., Wellesley College (2014)

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of

Master of Science in Media Arts and Sciences

at the
Massachusetts Institute of Technology
September 2024

© 2024 Karishma Chadha. All rights reserved
The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Authored by: Karishma Chadha
Program in Media Arts and Sciences
August 14, 2024

Certified by: Mitchel Resnick
LEGO Papert Professor of Learning Research

Accepted by: Joseph Paradiso
Academic Head, Program in Media Arts and Sciences

Imagine Yourself

Explorations in Fostering Personal Expression with Generative AI

by

Karishma Chadha

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, on August 14, 2024 in partial fulfillment of the requirements for the degree of

Master of Science in Media Arts and Sciences

Abstract

Generative Artificial Intelligence (AI) technology has been promoted with many exciting promises to enhance human creativity. However, it has also been shown to amplify human bias and perpetuate harmful stereotypes. In the new age being ushered in by this technology, this thesis explores how educators and designers can use this technology to support young people in exploring and expressing aspects of their unique identities. In particular, I use a design based research methodology to iteratively create Imagine Yourself, a new digital experience adapting off-the-shelf text-to-image generation technology to support young people creating personal representations and stories.

Imagine Yourself combines OpenAI's Dall-E 3 image generation technology with Scratch, a rich environment for young people to imagine and create interactive multimedia stories, animations, and more. Guided by a core value of designing for belonging, this project explores how experiences with generative AI can be designed to foster young people's creative process in creating personally meaningful stories reflecting their own unique identities, experiences, and cultures. I discuss the iterative design process of creating Imagine Yourself in tandem with creative workshops, aiming to support more diverse representation within the image generation output and invite a tinkerable and iterative process of creating. I discuss observations and feedback from creative workshops with young people and adults, creating with Imagine Yourself. Finally, I conclude with reflections on the design process as well as a discussion of challenges, limitations, opportunities, and open questions for future work incorporating generative AI into young people's creative learning experiences.

Thesis advisor:
Mitchel Resnick
LEGO Papert Professor of Learning Research

Imagine Yourself

Explorations in Fostering Personal Expression with Generative AI

by

Karishma Chadha

This thesis has been reviewed and approved by the following committee members

Mitchel Resnick

_____
LEGO Papert Professor of Learning Research
MIT Media Lab

Zach Lieberman

_____
Adjunct Associate Professor of Media Arts and Sciences
Institutional Affiliation

Karen Brennan

_____
Timothy E. Wirth Professor of Practice in Learning Technologies
Harvard Graduate School of Education

Acknowledgements

I could not have made it here without the support of the countless wonderful people in my life. You have made me who I am, and therefore have ultimately contributed to making this thesis what it is. Thank you is never enough. While I cannot mention everyone who has helped me get here, there are a few people who have supported me in very special ways.

To my advisor, Mitchel Resnick:
Thank you for being the most supportive advisor I could ask for. You have always allowed me to be myself and have encouraged me in countless moments where I was unsure of myself. But most of all, thank you for cultivating the wonderfully caring community that I consider to be a home and family, Lifelong Kindergarten.

To my thesis readers, Zach Lieberman and Karen Brennan:
Thank you for your feedback and encouragement as I was exploring these ideas. You both have inspired me to share my voice. Thank you also for reading my incredibly long thesis update emails!

Thank you to Ila Kumar and Tamara Fils-Aime for collaborating with me on Imagine Yourself workshops. You are inspiring in the ways you engage and support young people in expressing themselves.

To my family:
To Mummy and Papa, thank you for your countless sacrifices, steadfast guidance, and for the values you have instilled in me: to work hard, be humble, stay curious, and have patience.

To Gudia Didi:
You may not know that I was following in your footsteps when I wanted to pursue computer science at such a young age. You are one of the most resilient people I know. You are always ready to support, pamper, and celebrate me. Thank you for always believing in me.

To Sidharth:
You have more faith in me than anyone else in this world, including myself! You are my biggest cheerleader. Thank you for always believing in me, encouraging me, making me laugh, and hyping me up when I am stressed!

Thank you to all of my friends for the endless ways you support and encourage me, particularly as I have been on this grad school journey, but I wanted to give a special thanks to Nicole Francisco, Anjali Kayal, and Nirupama Sarathy. Nothing I could write here could express all the ways you have supported me! Thank you for testing out early prototypes of Imagine Yourself, for reading and commenting on excerpts of my thesis, for being late night study buddies, for the Bollywood music videos, for the writing playlists, for feeding me, for cheering me on at this finish line, and for SO. MUCH. MORE! I'm not Karishma without any of you.

To my dear friends at LLK and Scratch:
Thank you to every single one of you and for all of the ways you imagine, create, play, share, and reflect! Each of you inspire me in countless ways every single day. I want to give a special thank you to Rupal, Carmelo, Eric Rosenbaum, Zoë, Téo, and Cecilé. Téo and Cecilé you paved

the way last year with your theses and have been a constant source of support and encouragement. Thank you for hearing my hopes, fears, and concerns. Rupal, you have literally been my shoulder to cry on. Thank you for always encouraging me to share my voice and for inspiring me in all of the ways you welcome belonging. Carmelo, thank you for the endless conversations that always leave me energized, and just for being you! I learn so much from the ways you engage with others. Eric and Zoë, you always inspire me with your work and your ideas! Thank you for being so playful, inventive, and inviting!

💙*For Nyra, I hope the ideas here inspire you to imagine yourself!*

Table of Contents

Abstract....2
Acknowledgements....4
Table of Contents....6
Chapter 1. Introduction....10
- Welcome....10
- Who am I? (Positionality)....10
- A Note about Language and Culture....12
- An Ode to Visual Media....14
- Imagine Yourself Preview....17

Roadmap....18
Chapter 2. Background....19
- Motivation....19
- Inspiration: Tools for Representation & Expression....21
  - Avatar Makers....21
    - Static, Pre-Generated, Consumable Content....23
    - Limited Agency....23
  - Creative Learning Tools for Self Expression....23
    - Scratch....24
    - OctoStudio....26
- Exploring the Creative Promise of Generative AI....26
  - Opportunities & Limitations....31

Chapter 3. Designing for Belonging....33
- Beyond Representation, Towards Imagination....34
- See Yourself....35
- Imagine Yourself....36
  - Designing for Tinkerability for Belonging....36
- Who gets to play?....37

Chapter 4. Early Explorations....41
- OctoStudio Emoji Mix & Match....41
- Vector Image Generation....44
- Surveying Image Generation Tools....51
  - Delightful Depictions....51
  - Inscrutable Interfaces....53
- Choosing an AI System to Build With....55
    - Stable Diffusion....55
    - MidJourney....56
    - Dall-E 3....56
- Exploring Dall-E 3 & GPT Builders....61

Designing a Conversational Story Helper....61
Making a Personal Story.... 65
Chapter 5. Designing Imagine Yourself....70
Creating with Imagine Yourself....70
Ideas.... 70
Editing a Generated Image....71
Towards Inclusive Representation.... 75
Design Overview.... 75
Iterations....79
"Surprise Me" Process and Dall-E 3 Process, Step-by-Step.... 79
System Prompt Iterations for "Surprise Me" Process.... 82
Observations....91
Observation #1: Dall-E 3's Default Prompt Revision.... 91
Observation #2: Inconsistencies in Blocked Content....94
Observation #3: Depictions of Race and Ethnicity.... 96
Observation #4: Depictions of Heteronormative Trends and Racial-Gendered Stereotypes.... 104
Observation #5: Depictions of Body Type.... 104
Observation #6: Depictions of Persons with Disabilities....108
Observation #7: "Candy-Coated" Depictions: shorter prompts lead to over-indexing on added prefixes and suffixes.... 110
Observation #8: Depictions of Wealth.... 111
Observation #9: GPT / Dall-E 3 has a tendency to Hyper-Fixate.... 113
Observation # 10: "Inappropriate" Content.... 117
A Persistent Design Tension....119
The Trouble with Tinkering with AI.... 120
Opening Up Black Boxes.... 122
Invitation to Shape the Details Yourself.... 125
Chapter 6. Workshops.... 129
Workshop Contexts & Design....129
Workshop Contexts.... 131
Workshops 1 & 2: Emotional Expression....132
Experimenting with Dall-E 3 Default Prompt Rewriting....132
1-on-1 Sessions.... 134
Workshop 3 - Identity Expression & Scratch....136
Welcome Page and Introductory Resources for Imagine Yourself.... 139
Using "Surprise Me" as Default Image Generation Option....142
Experimenting with Prefixes and Suffixes.... 143
Workshop Observations and Feedback.... 145
Trends Across Workshops and 1-on-1 Sessions.... 145
Exploration.... 145

Self-Representation....148
Creators' Experiences with Imagine Yourself....151
Experiences with Self-Representation....151
Perceptions of Creativity....152
Chapter 7. Reflections....155
Broader Issues and Challenges to Consider....156
Future Directions....158
Supporting Agency while Creating....158
Exploring intermediate representations....158
Opening up more Black Boxes....159
Input into the Process, Not just the Result....159
Designing with Young People....160
Critical Playgrounds....160
Deeper Engagements....161
Conclusion....161
Appendix A....163
What is Generative AI?....163
How does it work?....163
What is a (text-to-) image generator?....164
How does it work?....164
Where does the data come from?....164
Appendix B....165
Dall-E 3 Prompt Rewriting....165
References....167

Imagine Yourself

Explorations in Fostering Personal Expression with Generative AI

Karishma Chadha

Chapter 1. Introduction

Welcome

Welcome, dear reader! As you begin to read this document and engage with the ideas discussed here, I invite you to imagine yourself: reflect on your own identity and personal experiences and how they inform your perspective as a creator, whatever your medium of creation may be. And indeed, you *are* a creator, whether you are constructing ideas, communities, friendships, art, code, food, or anything else.

As you reflect, think about moments, spaces, or communities you have felt belonging in. What experiences have shaped your identity? What experiences have shaped how you have imagined new possibilities for yourself?

This document is intentionally written in a narrative style and from a personal, first person perspective because I have aimed to make it accessible to a wide audience, hoping to engage people from a variety of communities and backgrounds, and not just people in education or technology or academia. I also include many personal anecdotes, especially in this first introduction chapter, because this work is centered on inviting people to explore and express aspects of their own identities and personal stories.

Finally, rather than a traditional thesis which primarily aims to construct knowledge, contribute novel ideas, or propose a definitive solution to a problem, I hope that my reflections in this document can serve as a conversation starter and serve as fodder for your own curiosities in your contexts because I certainly have more questions than answers!

Who am I? (Positionality)

Imagine Yourself explores how we can design new creative tools and technologies that support young people in exploring their identities and telling their personal stories, and thus, before diving into this work, I want to introduce myself to you, the reader, and share reflections on how different aspects of my identity and personal experiences affect my perspectives in this work and how they explicitly show up in this document.

I am an Indian-American, cisgendered, female, graduate student researcher at the MIT Media Lab. I have spent the last 14 years living in the greater Boston area, and I have spent 8 of these years affiliated with MIT, as a staff member, and as a student researcher.

I was the first in my family to be born in the US, and as such have experienced many privileges and challenges of navigating multiple parts of my identity. I have grown up making sense of competing aspects of my identity as an Indian, as an American, and as a woman, and I have routinely searched for and created communities and spaces of belonging with those who have had similar experiences and perspectives. In many parts of my life, I have gravitated towards connecting with other Indian Americans, children of immigrants, or people of color as I have found points of connection and identification in shared cultural values, perspectives, or experiences of feeling “other” in representations of “traditional” American culture.

I have also encountered systems of oppression such as the “model minority myth,” (Chou &

Feagin, 2015; Chun, 1980) designed to create and perpetuate a perceived hierarchy dividing marginalized groups in the US, which has undoubtedly influenced my interactions with and how I have been perceived by authority figures and those in power, particularly in educational settings (Chun, 1980; Museus, 2008). In many ways, I have followed traditional and stereotypical paths of success which are rewarded by a capitalistic system. I pursued a STEM career and have held high paying jobs, surpassing the local average household income and indeed long surpassing the joint income of my immigrant parents. I have had the privilege and opportunity of pursuing a graduate education, the first in my immediate family to do so. I have navigated the dual and competing values of prioritizing academic achievement, accolades, competition and individualism promoted by the "American dream" my parents immigrated to this country to pursue and the historically valued community-centered values of the Indian culture that nurtured them, prioritizing community, care, and support.

In my family, I have uniquely benefited from academic opportunities that allowed me to attend private academic institutions my working class family would not otherwise have been able to afford. I was selected into an academic preparation program for low-income families which led me to a private boarding high school education, fully funded by a wealthy White American family. This opportunity set me on the path to later attend other private institutions such as Wellesley College and now MIT.

These experiences have had a profound impact in shaping my perspectives on education, culture, and community, and they influence my approach to this research. I have played many roles in the context of designing learning experiences in my undergraduate research experience, in my work as the organizer of an educational outreach program, and in my work at Lifelong Kindergarten, as a member of the Scratch team, and as a student researcher. In this work, I am not only bringing my experiences as a researcher and a student, but also as a software engineer and learning designer.

In the thesis that follows, I write about topics such as identity, representation, inclusion, and belonging. While these are hugely important topics, they are also incredibly abstract, nuanced, complex, and often personal topics. Each of these topics will likely hold different meaning to people with different lived experiences, or even take on multiple meanings. Taking "representation" or "underrepresented" (when referring to a group of people) as an example, these are terms that must be discussed in their specific contexts. In the context of my own lived experiences, it is important to acknowledge that South Asians have historically been underrepresented in media such as film and tv (The Nielsen Company, 2021), but are overrepresented in STEM and computing fields (National Academies of Sciences & Medicine, 2021). As such, terms like "underrepresented" or "marginalized" do not tell the whole story of the South Asian American experience.

While the inspiration and motivation for this work comes from some of my own life experiences, I also acknowledge that I am an outsider in much of this work, particularly as I engage predominantly Black, Hispanic, and Latin American communities in the Boston area or work directly with communities who have experienced disproportionate harm or challenges such as the child welfare system, gentrification, policing, violence and injustice that continually harm and exploit Black and Indigenous communities across the US, particularly in education, computing and technological development, especially the development of Artificial Intelligence systems (Benjamin, 2020; Buolamwini, 2023).

The bulk of this thesis work, designing the Imagine Yourself tool and the creative workshops using the tool, was carried out at MIT. In my connection to MIT, I recognize that I carry tremendous power, perceived authority, and credibility, particularly as I engage and share my

ideas with the outside world or collaborate with outside communities. Like other universities and institutions around the United States, MIT is an institution that has participated in and benefited from the exploitation, knowledge extraction, and displacement of local communities and indigenous peoples native to this land. As a researcher from MIT, I am cognizant of the extractive and exploitative role that research and technological development have played in relationships with local communities. In engaging with communities who have been marginalized and excluded from opportunities with technology, I have a responsibility to engage in building relationships that center the voices, cultural practices, and ways of knowing of these communities.

After the initial work of designing the Imagine Yourself tool, I carried out creative workshops with young people and adults both at MIT and in spaces where the workshop participants live, experience education, and have an established community. In this latter setting, I was hosting the workshop as a guest and an outsider in these community spaces, relying on the relationships, trust, respect, and belonging the participants of these workshops had built with each other and with the space. In setting up the workshops, I primarily interacted with staff and adult mentors who directly and regularly work with the young people participating in the workshops. Explicitly and implicitly I leveraged and benefited from the existing relationships and connections that had already been built in these existing community spaces such as Computer Clubhouses (Resnick & Rusk, 1996) and Boys and Girls Clubs in which young people were daily participants and in what was designed and maintained to be a place of belonging.

In writing this thesis and carrying out the work described herein, I am conscious of the way language can play a role in creating and perpetuating systems of power and harm. I do my best to pay attention to how language perpetuates harmful or exclusionary narratives and reflect on how language can play a role in inviting belonging or perpetuating othering in our designs.

As such, I write from a first person perspective, uncommon in many research and academic contexts, acknowledging that I am writing from the perspective of my own lived experiences, and that I am continually learning about how my perspective, experiences, and positionality shape my research. I discuss my use of language and its cultural context in more detail in the following section.

Finally, I acknowledge that this work is only the beginning of a conversation, which I hope continues beyond this thesis. In particular, I hope to move my work towards directly building relationships with and centering the voices of young people and educators from the communities who are marginalized and excluded from creative learning opportunities or are harmed and exploited in the development of AI systems.

A Note about Language and Culture

This thesis explores how young people make sense of their identities and their cultures while using generative AI as a primary medium for this representation. As culture and identity are inherently complex and evolving, I struggled with thinking about how I will write about some of these themes around identity and struggled with imagining the words I would use to describe our identities in a way that would feel authentic … to you, the reader, to myself the author, and to the people whose experiences I am trying to share and reflect on here.

In a conversation with a friend, I tried to articulate this tension in my writing process. I was trying to describe to her my concern in finding the right words to use to talk about things that

felt extremely personal—topics like body image, where we lack words as a society and culture that feel neutral and non-judgemental, as well as topics such as disability, where I lacked personal experience and wanted to make sure to use inclusive, affirming, and anti-ableist language (American Psychological Association, 2023; The United Nations Office at Geneva, n.d.-a, n.d.-b).

I came to realize that this tension I was feeling of not being able to choose or find the right words parallels the process of writing and refining the input text descriptions (prompts) to create personal representations using generative AI. I was reminded of the nagging feeling I have gotten many times in the process of designing, building, testing, and refining the Imagine Yourself tool. Sometimes words are not the right medium for self-representation. Sometimes the words that are common in the culture(s) that I participate in do not match my sense of my own identity. Sometimes how others would describe me is not the way I see myself. And sometimes, how I describe myself is not the way I want to be represented or seen by others.

I struggle to write about terms that may shift over time or even have ambiguous or varied meaning across disparate cultural groups. One such challenge arises when I discuss ways young people use language to describe visual representations of themselves. In this work, I posit that representation is in the eye of the beholder, so assessments of the Imagine Yourself tool I describe in this thesis must happen on an individual basis. The images generated by this tool hold different significance to different groups of people. While one person may find a generated image to be a delightful representation, another may find it to be stereotypical or caricature-ish. It is precisely this relationship with the technology that I explore in this research and explore how young people navigate through this tension to create something that holds meaning for them.

In talking about using language for prompts to generate an image, or language that is generated or revised by a machine as a description of an image, there may be instances where language feels like an inadequate tool to describe the subject. In many cases, this may be a reflection of the current time and cultural climate related to a particular group. For example, I think modern society lacks nuanced, neutral, and non-judgemental language to talk about body types. While one person may identify positively with the descriptor “curvy” to describe their body, another might find this language disparaging or derogatory. Terms can also carry different meanings in different cultural contexts. In Indian culture, it is pretty common for young people to receive comments about their weight, appearance, or eating habits. Indian people of my parents’ and grandparents’ generations use the English word “healthy”—as in “you’re looking healthy”—to disparagingly tell a woman that she looks like she has gained weight. Meanwhile the same word in the modern American English culture has a positive connotation and is associated with well being.

Besides the language and particular terms used, the relationship between generated images and words may also carry different associations to different people. The AI image generator may tie a word such as “curvy” to a representation that may resonate with some and alienate others. In this thesis, I try my best to address the nuances of these cultural terms and their visual representations while acknowledging that in many cases I am speaking from an outsider’s perspective, especially in the cases where I am discussing representations of cultures I do not belong to.

As this research seeks to explore young people’s relationship and agency over an emerging technology, where possible, I try to center the voices of the young people themselves in describing their experience and their direct reflections on interacting with this technology. In

other places, I try to characterize modes of interacting with the AI through a content analysis of the prompts young people used in generating images. In future iterations of this work, I hope to more directly partner with and center the voices of the young people creating with this technology.

An Ode to Visual Media

The seeds for Imagine Yourself were planted long ago. I was the first person in my family to be born in America whereas my parents and older sister immigrated from India. Growing up as a first-generation Indian American, I have constantly juggled these two parts of my cultural identity trying to identify a place and way of belonging to both. I was fortunate enough to grow up in New Jersey where my family had access to not only family and community from India but also Indian temples, grocery stores, food, and media which made it easier to feel connected to my Indian culture. I have often joked that Oak Tree Road, a "Little India" neighborhood in New Jersey, is more Indian than India itself.

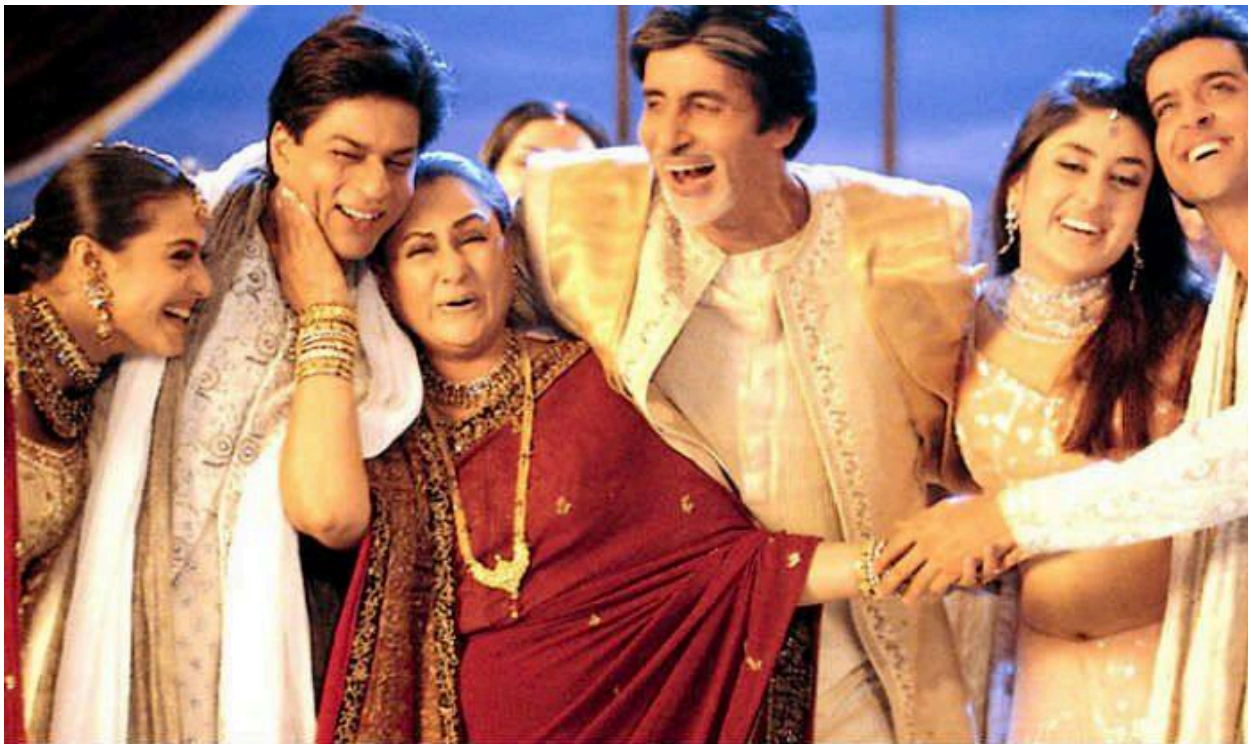
A photograph of Oak Tree Road in New Jersey. Image Credit: (Naqvi, 2022)

In my home, I grew up consuming South Asian media content such as Bollywood movies and songs, Amar Chitra Kathas (Indian graphic novels depicting Hindu epics or Indian folktales) (*Amar Chitra Katha*, n.d.), Pakistani Qawali music — the singer Abida Parveen was a staple in car rides with my dad, and my brother and I still light up whenever we hear the songs that flooded our childhood — and more. Reflecting on why I have consistently felt so drawn to visual media in particular, I have come to realize that at least one reason is that growing up knowing how to speak Hindi but not knowing how to read or write in the language, Indian visual media became some of my primary ways to access aspects of my Indian cultural heritage. Outside of abstract stories my parents and grandmother told me about their life in a land I had only visited when I was too young to remember, these movies, shows, and graphic novels became windows into this part of my identity.

Film media like tv shows and movies have often been one of my favorite ways to relax, connect with others, learn about new cultures and languages, and enter new worlds. I have long been intrigued by these forms of media and have often been engrossed in the beautiful ways that stories come together, pull you in, introduce you to new cultures and languages, engage you in discovering new worlds and in putting yourself into the shoes of another.

Along with my older sister, I grew up watching more Bollywood rom-coms than Disney movies. My grandmother and I would catch up on Hindi serials (soap operas) after dinner, a tradition we continued as a primary form of our connection even after I went away to boarding school

as a teenager. Through these forms of media, I got to engage in stories about life in India and along with it unspoken norms and cultural rules and customs, or other stories depicting vibrant Hindu epics and folktales. I was often enamored with the brightly colored clothing, the vibrant musical interludes, the singing and dancing as well as the stories of family, love, and community.



A still from *Kabhi Khushi Kabhie Gham* (Johar, 2001) | Image Source: (India.com, 2019)
This was (and still is) one of my favorite Bollywood movies growing up, centered on themes of family and tradition.

Outside of Indian media, I also avidly consumed American shows and movies borrowed from the local library with my brother. These were stories about "typical" American teenagers on PBS, Disney Channel, or Nickelodeon, barely any featuring South Asian characters or depicting them as more than the stereotypical math nerd, depicting stories about families that didn't look, feel, or interact like ours. But still we were enamored. They gave us a model of the way of "being American." And often, the stories of children and teenagers in these tv shows and movies became fodder when our parents admonished us for minor misbehavior. "At least we're not acting like these kids on TV!" we'd virtuously tout back at them. In middle school, I discovered Anime. In High School I discovered Kdramas, and in both of these cases visual media once again became an avenue for me to access and connect to other languages and cultures.

All of these types of visual media also became points of connection with new friends! With some of my best friends to this day, I connected with them on a specific piece or genre of media. My friend Nicole and I connected in college while tutoring an intro Computer Science class, but we became friends when we discovered our mutual love of Kdramas and Anime! My friend Andrew and I met to watch a 20th anniversary movie celebrating the *Yugioh* universe that permeated our childhoods. My friend Zoë and I have engaged in many-a-debate over

which film media was the best adaptation of the *Series of Unfortunate Events* book series we both enjoyed as children.

The people in my life know that when I like a tv show or a movie, I love to "evangelize" and recommend it to anyone who will listen. Members of the Lifelong Kindergarten research group may recognize how much I have raved about *Steven Universe*, a beautiful, animated tv series exploring poignant themes about human relationships, mental health, and stories of love and care or how captivated I was by *Ms. Marvel*, a tv show I only encountered two years ago but immediately felt seen and connected to the story featuring Kamala Khan, a South Asian American girl growing up in New Jersey, navigating her newfound super powers and her identity as a Pakistani American.

A poster from *Ms. Marvel*

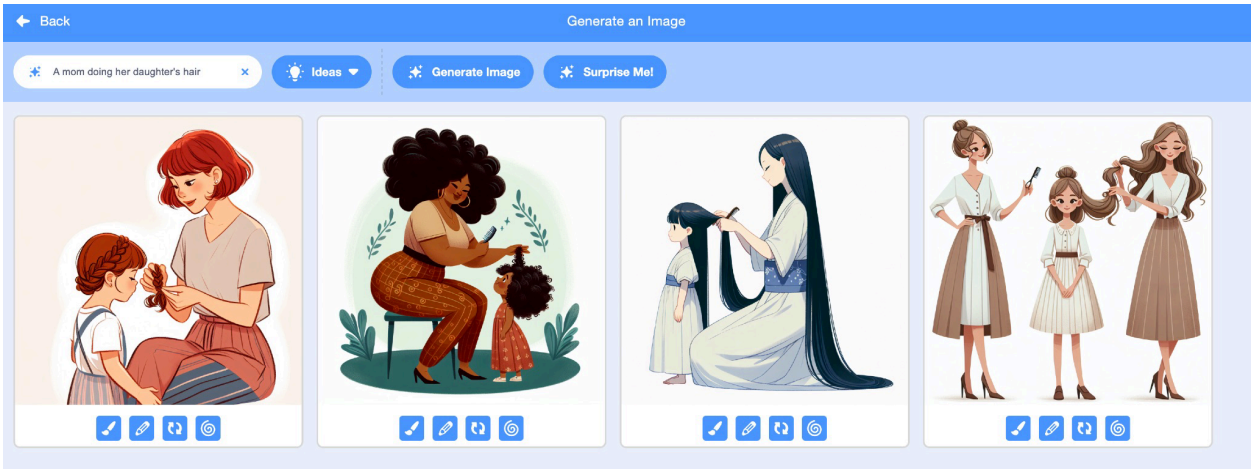
My experiences with visual media have offered me countless opportunities to reflect on the stories I have absorbed and engaged with since I was a child. While I absorbed both Indian and American media growing up, I never felt that either of them truly reflected my experiences as an Indian-American, often considered too Indian to be American or too American to be Indian. In recent years, I have seen more South Asian representation in American media, and more and more South Asian American creators are telling their own stories. Media has continually led me to reflect on my relationship with my cultural identities and how they relate to the stories I've typically seen on tv and in movies. I continue to be consistently intrigued by media representation and its effects on the feelings of belonging, inclusion, and imagination.

As advanced new technologies such as Generative AI come into the world, there is much hype around how these technologies will either replace or enhance human creativity. Are these entirely mutually exclusive? In this work, I begin an exploration of how we can design experiences with generative AI that nurture young people's imaginations of themselves and support them in telling their own meaningful stories.

The following thesis details my journey in designing Imagine Yourself, a new creative tool to explore ways to foster young people's personal expression and exploration of their own identities. I discuss my process of experimenting with and designing the Imagine Yourself experience, the experiences of young people trying it out and in many cases experiencing Generative AI for the first time, and my reflections from this experience. The following section provides a brief introduction to Imagine Yourself, to ground the reading of these early chapters with a more concrete picture in mind.

Imagine Yourself Preview

The Imagine Yourself creative tool embeds OpenAI's Dall-E 3 (Betker et al., n.d.; Goh et al., n.d.; OpenAI, 2023a) image generation capabilities inside of the Scratch (Resnick et al., 2009) programming environment. This combination[1] of tools aims to allow young people to create their own unique assets representing themselves, their loved ones, or other important elements of their identities, and bring them to life in personal stories and animated projects.



Imagine Yourself Interface for generating images

This project explores how this new digital experience can be designed to foster a sense of belonging in the experience of creation and expression with technology, particularly by supporting inclusive representations of diverse identities particularly those identity groups that have historically been underrepresented in media and marginalized in computing. Towards this overarching goal, the Imagine Yourself system works within the constraints of the underlying Dall-E 3 AI system to try to depict more inclusive representations in the output images. In later chapters, I discuss the guiding principles for the design of Imagine Yourself as well as the process of iteration and refinement I embarked on to design the Imagine Yourself tool. In these explorations, I reflect on the process of attempting value-based design with AI, learnings and observations as well as the challenges and pitfalls I encountered. I conclude with more open questions to ponder as we design and engage with AI systems to create personally meaningful content.

The following roadmap provides a brief introduction to the upcoming chapters.

[1] This work builds on early explorations with AI image generation in Scratch outlined by Eric Rosenbaum, a colleague on the Scratch team (Rosenbaum, 2023). In his blog post, he outlines creative possibilities for image generation in the context of Scratch. Imagine Yourself builds on these ideas, taking a closer look at possibilities for supporting inclusive representation and belonging.

Roadmap

Background

Everything is a remix! The ideas I explore in this thesis are inspired by decades of research, community work, education, media, tools, and life experiences. In this chapter, I synthesize the main ideas motivating this thesis including the importance of diverse media representation, creative tools I have been inspired by, and the opportunities and limitations of using Generative AI[2] for personal expression.

Designing for Belonging

Throughout my early explorations and in designing the Imagine Yourself tool, I wanted to create an experience that invited young people to see themselves and imagine themselves with this creative tool. In this chapter I discuss these as two important aspects of trying to foster belonging within imagining, playing, and creating.

Early Explorations

I describe early explorations and experiments towards designing a new experience for supporting young people's personal and cultural expression. After some of these preliminary explorations, I began to converge towards the idea of designing a new experience with Generative AI. I discuss some of the popular image generation tools and my criteria for ultimately choosing to work with OpenAI's Dall-E 3 in the design for Imagine Yourself.

Designing Imagine Yourself

This chapter describes my main exploration in designing the Imagine Yourself tool, the observations and iterations I made along the way, and the open questions it brought up. In this chapter I recount about design tensions and tradeoffs I encountered and how I worked through them.

Workshops

After spending lots of time iterating on the initial design of the Imagine Yourself tool, I was very excited to try it out with peers and colleagues as well as design and run my first workshops with different groups of young people and adults. Based on feedback and observations from the workshops, I iteratively refined both the Imagine Yourself tool and the workshops themselves. This chapter discusses the workshop design, my iterations between workshops, and most importantly the experiences of the creators who participated in these workshops.

Reflections

As with any project, taking a step back to reflect on the process and the journey is one of the biggest opportunities for learning and re-imagining possibilities for the future. In this chapter, I discuss some of my key takeaways from this work, broader challenges and considerations of this work, as well as future directions for exploration.

[2] See Appendix A for a brief overview of Generative AI and how it works.

Chapter 2. Background

Motivation

Since the internet became open for personal use, young people have used online forums, social media platforms, and creative online communities to explore, express, and share aspects of their identities and interests and connect with others (Turkle, 1995). The platforms and mediums for personal expression have changed as technology has evolved, but a few trends in expression have remained. Young people use all manners of mediums such as digital avatars, photographs, text descriptions, videos, music, animated gifs, and personal creative projects, to express themselves online. These mediums offer opportunities for people of all ages to express their ideas, discover new interests, and connect with like minded individuals.

At the time of writing this thesis, a new technology is emerging and rapidly gaining traction as a possible new medium for creative expression. This thesis explores how this new Generative Artificial Intelligence (Gen AI) technology, particularly text-to-image generators,[3] can be designed and scaffolded as a medium for young people to express themselves. I focus this work on fostering young people's exploration and expression of their unique identities through visual representations of themselves or loved ones and through dynamic personal stories or projects featuring those representations.

The recent rise in Generative Artificial Intelligence (Gen AI) technology has been ushered in with the promise of supporting people in quickly creating whatever they imagine. Unfortunately, these tools have also been shown to amplify human bias and perpetuate harmful stereotypes (Bender et al., 2021; Nicoletti & Bass, 2023; Tiku et al., 2023), furthering the inequities that already present barriers for young people from marginalized and underrepresented communities from accessing opportunities to be creators with technology (Garcia & Lee, 2020; National Academies of Sciences & Medicine, 2021). In exploring this new technology as a possible medium for young people's creative expression, I seek to understand the opportunities and challenges of creating and supporting an experience that nurtures young people in creating, exploring, and sharing their imaginations of themselves as they develop their creative confidence, sense of self, and belonging. In particular, I set out to explore how developers leveraging text-to-image generation technology can support young people's creative agency and sense of ownership over their creations, support a creative learning experience inviting users to iterate and tinker as part of their creation process, and most importantly, foster a sense of belonging for young people from all backgrounds but especially from communities whose voices have been marginalized and excluded from technological (and indeed AI) development and representation.

I explore these ideas by engaging in the process of designing a new digital tool, Imagine Yourself, bringing text-to-image generation to the Scratch online programming environment, a creative platform and online community used by millions of young people around the world to create and share multimedia projects such as games, animations, interactive stories, and more. My explorations were guided by the central idea of designing for belonging, aiming to support young people in seeing and imagining themselves with this new technology.

I iterated and refined my designs leveraging a design-based research methodology (Wang &

[3] See Appendix A for a brief overview of Generative AI and text-to-image generators.

Hannafin, 2005). I designed the tool in tandem with feedback and testing from peers as well as in workshops with young people and adults, iteratively evaluating the design through a lens of belonging.

In this thesis I detail my design journey, discuss reflections and observations from workshops, and the opportunities and challenges of designing for diverse representation and belonging with an inherently biased AI system. It is my hope that the discussion in this thesis will spark inspiration, questions, ideas for educators and developers designing AI tools for young people to remix and apply to their own contexts.

Barriers to Expression & Imagination

Digital tools and media can provide inspiration for imagining yourself in new contexts or possibilities and can also be a medium for constructing or reflecting on our own personal stories (Buckingham, 2008; Napoli, 2018; National Academies of Sciences & Medicine, 2021; Turkle, 1995). Unfortunately, very often these same tools and media exhibit a strong lack of diverse representation of people from varying identity groups, backgrounds, and experiences in interesting roles (Garcia & Lee, 2020; Rogers et al., 2021).

It was in high school that I first came across the work of Nigerian author Chimamanda Ngozi Adiche, through her TED talk, *The Danger of a Single Story* (Adichie, 2009). In this talk, she recounts her childhood experiences as a reader and storyteller, consuming mostly British and American stories lacking characters that resembled her which in turn influenced her own stories as a budding writer. The early stories she wrote, as a seven year old girl, featured characters that "were White and blue-eyed, they played in the snow, they ate apples, and they talked a lot about the weather …" (Adichie, 2009), despite the fact that these stories did not resemble her own life or experiences. She described that encountering authors like Chinua Achebe and Camara Laye made her realize "that people like [her], girls with skin the color of chocolate, whose kinky hair could not form ponytails, could also exist in literature" (Adichie, 2009).

In recounting this story, Adichie says,

> Now, I loved those American and British books I read. They stirred my imagination. They opened up new worlds for me. But the unintended consequence was that I did not know that people like me could exist in literature. So what the discovery of African writers did for me was this: It saved me from having a single story of what books are. (Adichie, 2009)

Her talk details the idea of the single story – singular, one dimensional narratives and representations of a person or group of people or cultures. She discusses the relationship between stories and power, that singular stories need not be untrue to be dangerous, that the power associated in representing those singular stories as *the only story* is what makes them dangerous. "The single story creates stereotypes," Adichie says, "and the problem with stereotypes is not that they are untrue, but that they are incomplete. They make one story become the only story" (Adichie, 2009).

Learning environments, digital tools, and popular media are often designed (intentionally or unintentionally) to center default identities. Often this is the dominant identity in the societal context that the learning experience or the designer (this can be an educator, a software

developer, a film producer) is situated in. Those that fall outside of the strict boundaries of the default identity are marginalized and excluded. People who identify with the marginalized groups are left to cross additional barriers in being able to participate equitably in the experience.

A prominent example of this phenomenon is the use of the yellow skintone to represent people in popular media like The Simpsons, popular tools like the standard emoji libraries on Apple and Google computing platforms and popular toys like the LEGO Minifigures. The yellow skin tone signals 'White' (Robertson et al., 2021) and this is compounded by the fact that the "other" in these environments is often designated by alternative skin tones that are not the default, or stereotypical markers such as accent or clothing.

When our tools and popular media present a single narrative—or in Adichie's words, a "single story"—about the default or neutral identity, those who do not identify with the presented default may not have the chance to visualize and imagine themselves in new contexts.

Inspiration: Tools for Representation & Expression

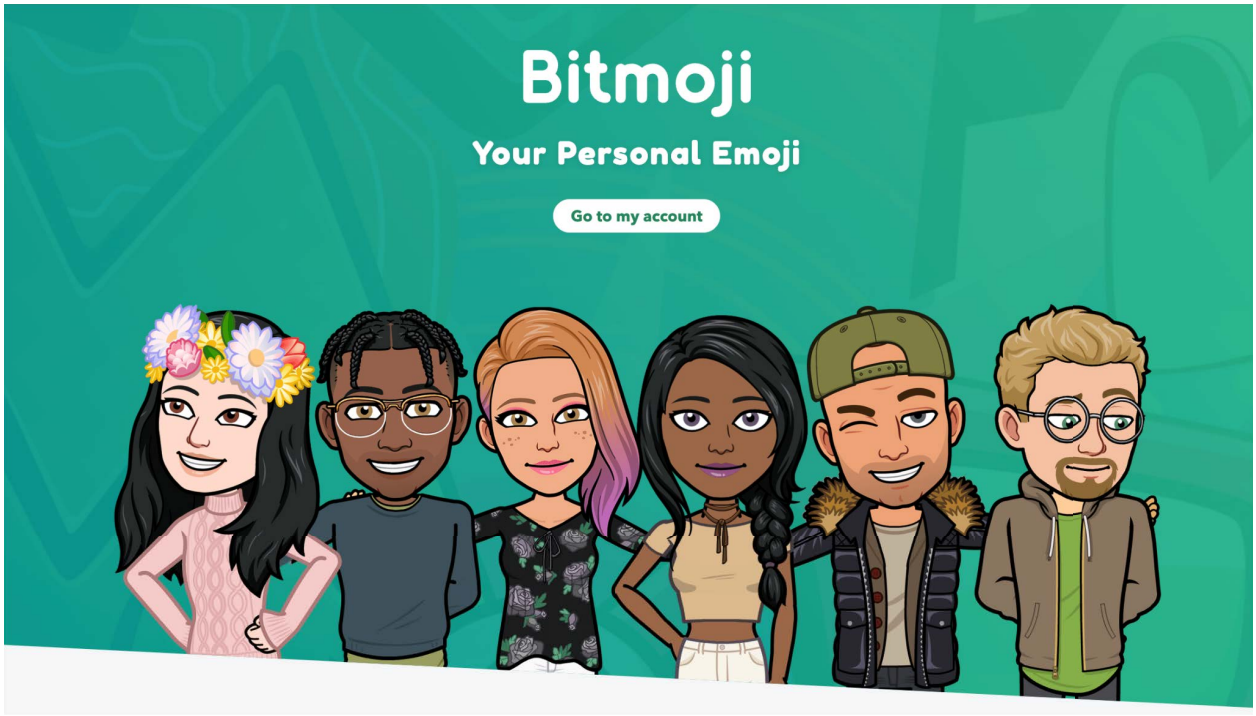
Avatar Makers

In recent years, more efforts have been made to increase representation in tools and media (*Geena Davis Institute*, 2024; Rogers et al., 2021; The Nielsen Company, 2021). In digital spaces such as games, creative platforms, or social media platforms, a popular way for young people to represent themselves is through constructing digital avatars using avatar makers (Buckingham, 2008; Kafai et al., 2010; Passmore & Mandryk, 2018; Zimmermann et al., 2023). Building a character through an avatar maker is often one of the first steps of a video game. Popular games such as The Sims (Electronic Arts 2016) are devoted to building many

characters and simulating narratives for those characters. In the Scratch online platform, young people not only use avatar makers to create unique characters or representations of themselves for use in visual stories, but also to create their own avatar makers or icon creators to share with other young people in the community to then use or remix for their own purposes. Another digital platform Picrew (TetraChroma Inc., 2017) is a community and creation platform for young people to construct and share their own avatar makers.

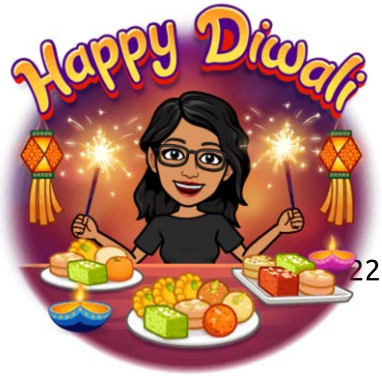
Avatar makers are a popular medium for expression because they provide a "low floor" or an accessible starting point (Resnick, 2017; Resnick & Silverman, 2005) for people to create representations of people, be it themselves, loved ones, or original characters for stories and games. They have long been a creative medium for young people to express themselves, share and showcase aspects of their identity, or try on new possibilities for themselves when connecting with others online.

As with any one-size-fits-all approach, avatar makers have also suffered from issues of lack of diverse representation of people in the choices they offer (Geena Davis Institute, 2023). Recently, popular personal avatar makers aim to address this issue by offering more and more choices for representations of people. Indeed, Bitmoji (Snap Inc., 2022), one of the leading platforms for making personal avatars, boasts almost 2 septillion possible unique combinations (Constine, 2018) to create a personalized avatar, orders of magnitude larger than the world's population!



Screenshot from Bitmoji Homepage featuring a variety of Bitmoji (Snap Inc., 2022)

Indeed, the plethora of choices offered in the Bitmoji library can be combined to create satisfying and delightful modes of self-representation such as this image of a Diwali greeting I have sent to countless friends and family over the years, featuring my own personal bitmoji avatar.



While I (and family members) have been tremendously delighted to encounter representations of Indian culture in the Bitmoji library, use of avatar makers can limit young people's creative expression in two major ways:

> 1) They are often designed to be used in the context of static or pre-generated creations such as the above diwali sticker. As a Bitmoji user, I have no ability to edit the elements in this scene or create my own scene featuring personally meaningful elements I may want to include that pertain to how my family celebrates Diwali.
>
> 2) Young people can combine attributes (such as hairstyle, eye shape, clothing, or accessories) provided in the Bitmoji avatar maker tool to create a unique avatar, but there is no option for modifying, customizing, or creating beyond combining the options that Bitmoji has provided. Young people cannot bring in unique contexts or components relating to their identity, appearance, or interests that the Bitmoji developers have not already thought to provide.

Static, Pre-Generated, Consumable Content

Bitmoji was first created for the purpose of allowing people to create their own comic strips, known as Bitstrips (Lacoma & Beaton, 2021). Since the acquisition of Bitstrips by Snap Inc. (Snap Inc., 2024) (the company behind the popular social media platform, Snapchat), Bitmoji are now mainly used for creating unique personal avatars which are then embedded in pre-generated scenes or stickers, such as the Diwali image above. People can use their avatars and these stickers in messaging apps, on the Snapchat social media platform, and other uses such as setting up Bitmoji "virtual classrooms" to create a customized and personalized virtual classroom setting for online learning (Minero, 2020).

The Snapchat social media platform has also used Bitmoji to serve users pre-generated content such as short comics or animated Bitmoji "movies" featuring your and your friends' avatars. This is a delightful form of entertainment. It feels somewhat personal because it features your avatar in different scenarios. However, the stories being featured are pre-generated content that the users themselves don't have any input or control over. Furthermore, these representations are used primarily as content for the user to consume rather than provided as construction materials for the user to create with. The pre-generated scenes and stickers are not accompanied by tools for modification or creation.

Limited Agency

In the Bitmoji avatar maker, users have two options for creating avatars. They can generate an avatar using a selfie (personal photograph) or manually create or edit an avatar by combining pre-generated options for face and body features (e.g. eye shape and color, hair style etc.) as well as clothing and accessories. Similar to other avatar makers, users can combine these predefined choices, but there is no support for users to imagine and create their own representations outside of the provided options.

Creative Learning Tools for Self Expression

In *Lifelong Kindergarten* (Resnick, 2017), Dr. Mitchel Resnick characterizes a "kindergarten-approach" to learning, building upon the educational framework of constructionism defined by Seymour Papert (Papert, 1980, 1986) — "that children construct

knowledge most effectively when they are actively engaged in constructing things in the world" (Resnick, 2020). This kindergarten approach to learning, often referred to as "creative learning" is best characterized by the 4 P's of Creative Learning, *projects, passions, peers, and play* — that children learn best when they actively engage in an iterative process of design while working on open ended *projects,* related to their *passions* and interests, engaging, sharing, and collaborating on ideas with their *peers,* in a *playful* spirit of tinkering, taking risks, and engaging with their construction materials (Resnick, 2017). These ideas have been realized through creative learning tools such as Scratch and OctoStudio, which provide ways for young people to create and express themselves with technology beyond merely consuming it.

Scratch

Scratch (Resnick et al., 2009; Scratch Foundation, 2024) is a popular creative platform and online community in which over 100 million young people around the world have created their own dynamic stories, animations, games and more. Young people have used this platform for a variety of different types of self expression. Personal profiles on Scratch consist of a profile image and short text descriptions Scratchers use to share aspects of their unique identities and interests with others on the platform. Profile pictures range from illustrated avatars to images representing a wide variety of interests (including favorite characters from popular culture).

Scratchers also create imaginative stories featuring themselves or fictional "original characters" that persist across many projects. Users also try on and explore different identities through interactive role playing games within studios (Scratch spaces for collecting and curating projects of a shared theme). Other Scratchers create their own avatar makers or dress up games for other users to use or remix to add their own art style or additional representations.

Screenshots of young people's projects on Scratch[4]

The rich culture of the Scratch online community and the platform allow for multiple styles of engagement and invite personal expression. On this platform, whether the Scratcher has created their own original character or used an existing image from the Scratch sprite libraries or from the internet, they are creating their own dynamic stories and projects featuring those characters.

The design of the Scratch platform prioritizes this kind of creation by offering multiple pathways for users to create, remix, and customize their projects such as a Sprite library of assets to support a variety of interests, a digital paint editor for Scratchers (Scratch users) to create their own artwork, and by making different artwork customizable (mix and match elements from separate characters as you see fit) and by making the artwork easy to manipulate and edit or create your own from scratch. In all of these cases, the design prioritizes the idea that the agency of creation lies in the hands of the user rather than the tool (or the developer).

[4] Scratch project and studio attribution (left to right):
https://scratch.mit.edu/studios/34113020/
https://scratch.mit.edu/projects/384134658/
https://scratch.mit.edu/projects/587366652/
https://scratch.mit.edu/projects/467234894/
https://scratch.mit.edu/projects/414693094/
https://scratch.mit.edu/projects/415537752/
https://scratch.mit.edu/projects/236586121/

OctoStudio

Increasingly, young people around the world have more access to mobile phones than laptops or tablets. However, mobile device usage (especially by young people) has typically been geared towards consuming media (social media, mobile games, videos on streaming platforms) rather than creating it. OctoStudio (Resnick & Rusk, 2024) is a creative mobile application designed for young people to create multimedia projects and bring their ideas to life anytime, anywhere. OctoStudio allows young people to use their phones as a construction material and bring their own world into it. Using the photos on their phones or virtual representations, young people can create dynamic multimedia projects about themselves and their interests and share them with others. Designed in collaboration with community-based educators from around the world, OctoStudio aims to cultivate belonging in the world of creative coding (Rusk et al., 2023). Aspects across all layers of the app are designed to be inclusive of the variety of needs, access, and cultural contexts of young learners around the world. Leveraging the standard Google emoji library, OctoStudio provides a variety of visual assets for young people to create with. However, addressing the limitations in representation of these assets, the OctoStudio team modified many of the default emoji to include a wider representation of body types, headwear, hair types and styles, and more.

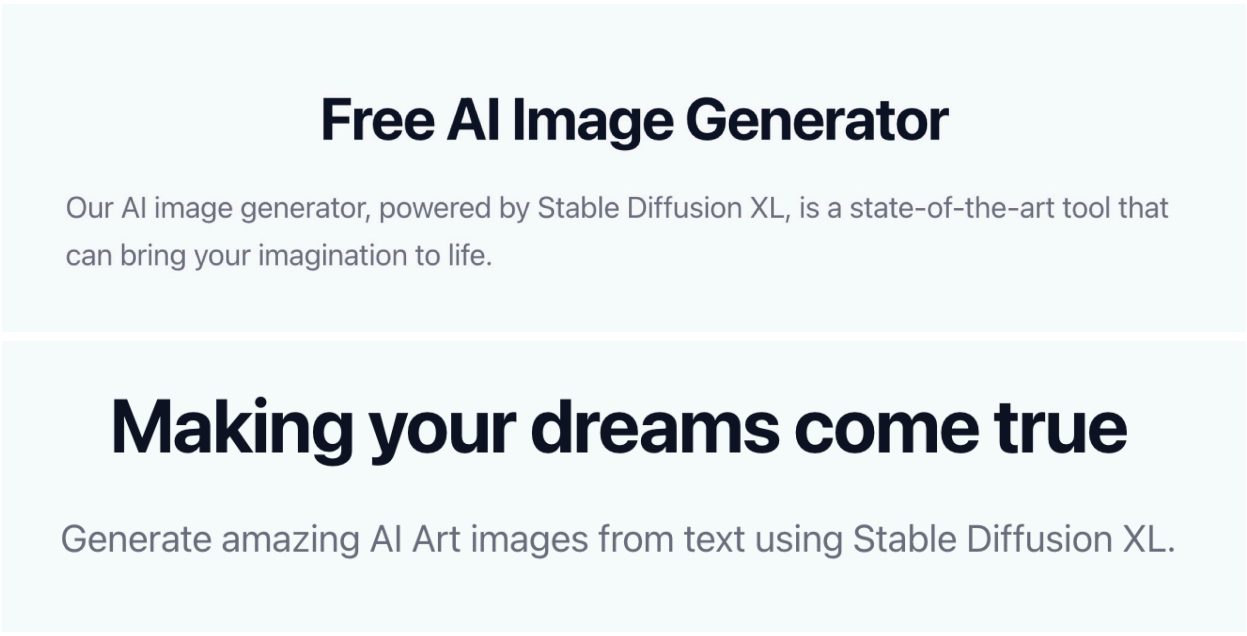
Unfortunately, even with a variety of elements for use as sprites in both Scratch and OctoStudio, there remain identities and contexts that are not represented or offered by these tools. Generative AI technology, and in particular image generation tools, offer an interesting potential alternative for young people to create their own unique representations according to whatever they can imagine.

Exploring the Creative Promise of Generative AI

In the past year, Generative AI[5] technologies have rapidly gained traction in the tech world. The media and leading developers of these tools have been touting their creative capabilities and their apparent potential for amplifying human creativity, through grandiose phrases such as “[make] your dreams come true” or “picture limitless creativity at your fingertips.” Through this marketing, Generative AI tools offer a promise of personalization and amplifying human creativity. You can create anything you imagine, just put it in a prompt.

A portion of OpenAI’s description of Dall-E 2 (OpenAI, 2022)

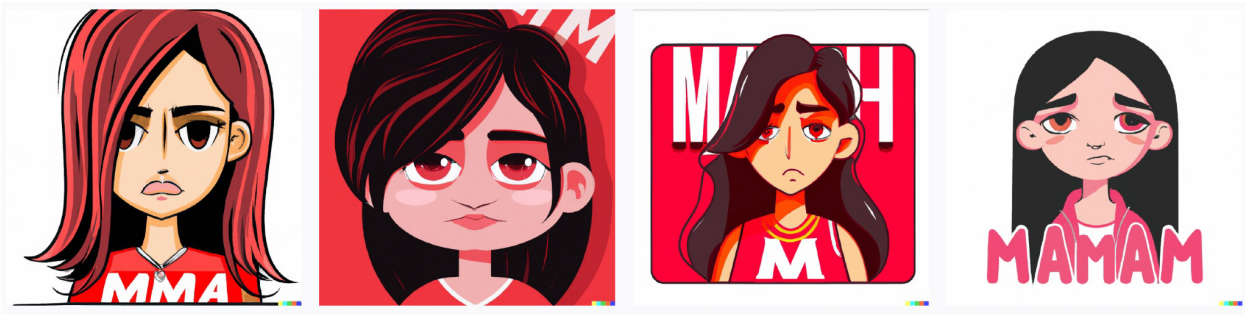
[5] For a brief overview of generative AI, see Appendix A.



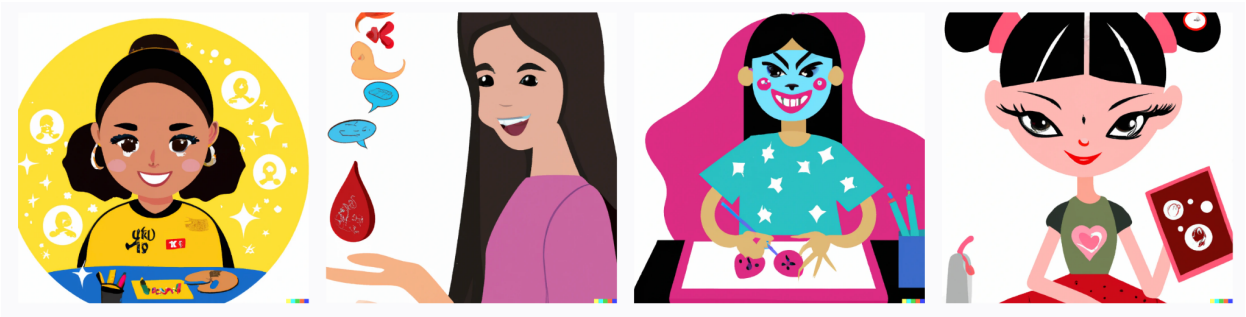
Descriptions of a Stable Diffusion powered Image Generator (Black Technology LTD, 2023)

Description of Generative AI in WIRED (Kelly, 2022)

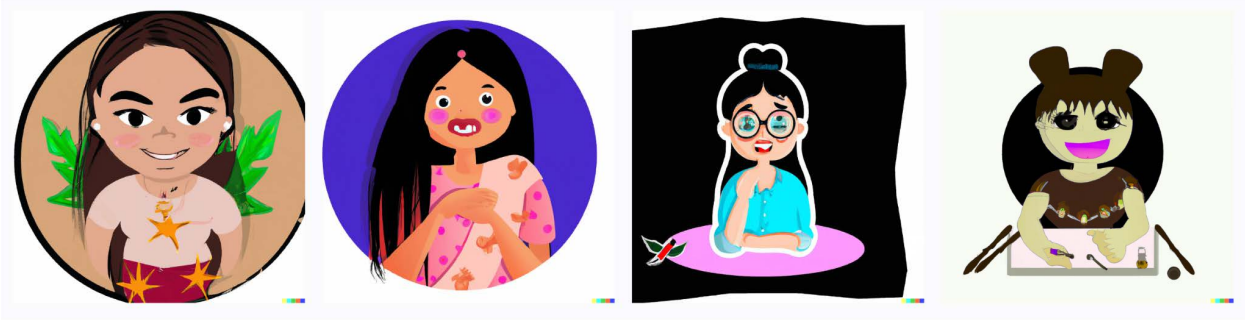
To explore these creative promises as potential avenues for young people's self expression, I set off to use these tools to create my own "avatar" image representing myself and my interests. My initial prompts used information about my ethnicity and some information about my hobbies or interests. However, the first several results I encountered featured either weary or unhappy faces, stereotypical representations of different ethnicities, or inhuman features such as painted masks (with scary looking expressions) or doll-like representations featuring inhuman features such as green skin and blacked out eyes.



Images generated with Dall-E 2 using prompt:[6]
"an avatar of a young south asian girl who likes cartoons and the miami heat"



Images generated with Dall-E 2 using prompt:
"An avatar of a young South Asian girl who likes to play games and create fun art project"



Images generated with Dall-E 2 using prompt:
"An avatar of a young Indian girl who likes to play games and create fun art projects"

[6] Throughout this document, I quote many prompts provided as input to various generative AI tools. In these quotations, I leave the prompts as-is, including any typos or grammatical errors.

With tools that yielded richer details and more sophisticated art styles, I was delighted by the depictions of South Asian clothing and jewelry. However I also found that I was looking at the same faces over and over again. The generated images looked very similar in their details depicting idealized delicate facial features and body types reminiscent of the unrealistic beauty standards portrayed in film or social media.

Images generated with Midjourney using prompt:
`“a south asian girl wearing a blue lehenga, digital art”`

Images generated with Midjourney using prompt:
`"a digital avatar of a south asian girl wearing a blue lengha, full length"`

The results of using these tools reveal limitations such as stereotypical representations, and amplified bias (Nicoletti & Bass, 2023; Tiku et al., 2023). Use of these tools, especially for personal representation, has the potential to tremendously affect one's sense of self, place in the world, and self-worth.

There are multiple types of potential harms happening here:

- Perpetuating harmful bias and stereotypes
- The feeling of being misrepresented or misinterpreted by the machine (especially when it feels personal)
- Feeling othered by the machine when it generates images that look less than human or

portrays only a single type of depiction (e.g. angry/tired faces or just one body type)

The similarity of the output across multiple rounds of image generation has also made me wonder whether these image generators have a kind of “middling” or “averaging” effect, trending towards yielding the most stereotypical representations of concepts, thereby passing them off as neutral or default. How much variance can be depicted in these tools, and is it enough variance to be able to use these tools for personal expression?

What are the opportunities and limitations of these tools when trying to use them for self representation or creating and sharing personally meaningful stories?

Opportunities & Limitations

Given some of the challenges and limitations mentioned above, I aim to articulate and explore the specific opportunities, limitations, and challenges of creating personal and culturally affirming visual representations with image generation tools.

Ethical considerations of this work include:

- Harms such as amplified bias and stereotypes or inappropriate content (Bender et al., 2021; Nicoletti & Bass, 2023; Tiku et al., 2023)
- Perpetuating the exploitation of creative work — many of these tools have illegally been trained on stolen data
- AI tools have a history of exploiting and extracting resources (Crawford & Joler, 2018) from already marginalized communities, use of these tools may support further harms to these communities
- AI tools are built with the (often hidden) labor (Perrigo, 2023) of communities who are furthest from resources and access to the tools themselves (Bender et al., 2021)
- AI tools have been shown to have a tremendous impact on the consumption of environmental resources and the rapid proliferation of generative AI usage poses concern for unprecedented environmental concern (Berthelot et al., 2024; de Vries, 2023; Utz & DiPaola, 2023)

Potential limitations and challenges of these explorations include:

- Limited agency or visibility in the creation process (currently most of the tools using GenAI technology provide users with limited ability to modify the creations or have fine-grained control over the creation process)
- Unknown space of possibilities for creation – it is unclear what the boundaries of these tools are or how far one can stretch the generated creations. As I saw in my initial explorations, these tools often feature a similar style / depiction and while changes can be made to get more diverse results, it is unclear how far these can be stretched to create more unique (and personal) content.

Beyond discussing some of these potential harms, limitations, and challenges, I explore the potential opportunities of generative AI for fostering personal expression.

Potential opportunities include:

- New ways for young people to create what they imagine
- Creative agency for young people to express themselves without the barriers or pre-requisite skills needed to make use of existing tools for visual expression such as digital drawing tools (these require building up a lot of technical drawing skills to be

able to create what you imagine)
- Ways for young people to create digital representations of personal cultures or references that might not otherwise be represented or included in existing digital tools such as avatar makers, emoji, sprite libraries, etc.
- Potential new design principles or guidelines that could aid developers and educators in empowering young people to have agency over the technology that may become ubiquitous in their lives

While I have focused this chapter on some of the creative tools and technology that inspire and motivate these explorations, in the following chapter I discuss the core values and ideas that guide the design of Imagine Yourself.

Chapter 3. Designing for Belonging

In 2017, when I completed my first interview with the Scratch Team in the Lifelong Kindergarten research group at the MIT Media Lab, I remember leaving the interview smiling from ear to ear. I remember feeling that I had found a team where I could be myself! It felt remarkable that an interview, of all things, could have brought about such a feeling. When I think back on the interview, I recall the openness and warmth I was greeted by. I recall shared excitement around supporting young people's learning. I recall being amazed by the ways in which the team talked about their work and about the environment they aimed to foster not only for young people on Scratch, but also for themselves as learners, collaborators, and as a team. I also recall that beyond conversation about my technical skills and professional experience, I was asked about my love of learning, and the ways I engage in trying new things. I shared my love of languages, both working on designing programming languages professionally, but also exploring and learning new human languages such as my adventures in learning Korean.

I left my first interview feeling that I had somehow stumbled upon a workplace, team, and organizational culture where I felt I could bring my whole self, bringing not only my professional interests, but also aspects of my identity and my personal values. I've found a home within the community at Lifelong Kindergarten and I have consistently found this to be a place where I can play, learn, grow, be curious, and build ideas based on shared values. I am forever grateful for this community and for the wonderful kindness, care, joy, vulnerability, expression, and creativity each individual person brings to work everyday. Together, this community cultivates shared belonging, making space where each of us can bring our whole selves and our *imaginations* into our work.

This idea of cultivating belonging is a guiding principle of this work. In her book, *Design for Belonging* (Wise, 2022), Susie Wise details the importance of recognizing belonging and othering in our surrounding environments. This lens of "belonging" and "othering" is inspired by the work of john a. powell (Powell & Menendian, 2024), who leads the Othering & Belonging Institute at the University of California, Berkeley (Othering and Belonging Institute, 2024). Wise describes these key ideas:

> *Belonging* is being accepted and invited to participate; being part of something and having the opportunity to show up as yourself. More than that, it means being able to raise issues and confront harsh truths as a full member of a community. *Othering*, by contrast, is treating people from another group as essentially different from and generally inferior to the group you belong to. (Wise, 2022)

Building on this lens, I hope to support young people in feeling a sense of belonging in the learning environments they participate in as well as with the learning tools they use to imagine and express themselves. In this chapter, I discuss how I approach this key idea of designing for belonging. In later chapters, I discuss explorations towards this goal as well as pitfalls of working with generative AI that can contribute to feelings of "othering."

Beyond Representation, Towards Imagination

While *belonging* encompasses many ideas and must be supported across a variety of facets of the experience being designed, I focus my work on two aspects I see as salient prerequisites for belonging: designing tools for young people to *see* themselves, encompassing inclusive representation, and *imagine* themselves, a designed invitation to take ownership of their creations, shape the details themselves, and bring in their own ideas, perspectives, identities and experiences. I discuss these ideas in more detail below, and I come back to these design principles in the Designing Imagine Yourself chapter.

Case Study 1: Bring Yourself Into Scratch

While working on the Scratch team, I was fortunate to engage and participate in discussions, meetings, and projects that extended beyond my roles on the engineering team. I was able to contribute ideas towards the design of the product, the learning resources, and the workshop experiences we were creating for young people. In 2020, I helped lead a project called "Bring Yourself Into Scratch." As part of this project, my team members and I examined the variety of ways that young people in the Scratch online community were representing themselves in Scratch projects.

Artists (from left to right): @Defsiarte, @LeiIani, @MeurpTV, @Ipzy on Scratch

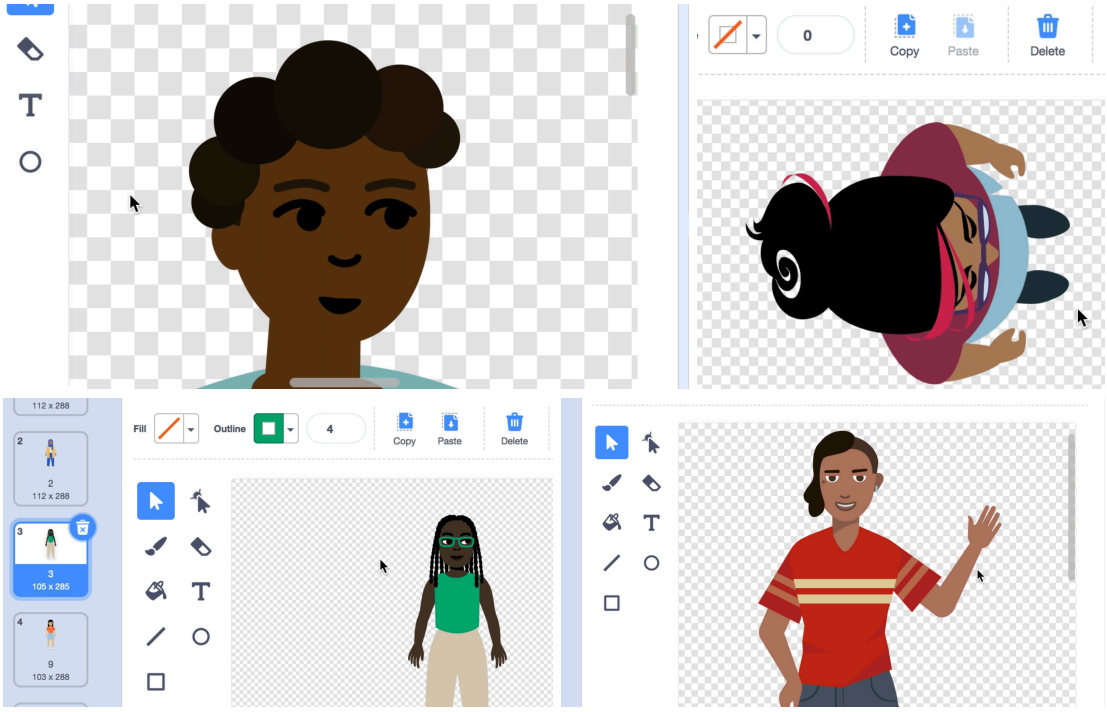
However, we noticed that these types of representations were primarily being created by more experienced Scratchers (Scratch users), and we sought ways to scaffold and ease the experience of young people to see and create representations of themselves in their Scratch projects.

One way we aimed to address this issue was to create a new set of character sprites to include in the Scratch sprite library, taking inspiration from what young people were creating and sharing in the Scratch online community.

We added four new character sets to the Scratch sprite library prioritizing a few aspects of the design:

1. We aimed to include a variety of representations of people including their age, ethnicity, disability, body type, and gender expression.
2. We aimed to make these characters easily modifiable, inspired by the ways young people in the Scratch community used simple and modular component shapes to compose larger pieces of artwork in projects such as dress-up games and icon creators.

The four sets of characters are designed to be more easily re-poseable for animation, and have mix-and match-able components for Scratchers to construct their own representations by combining and editing parts of the different existing characters.



Bring Yourself Into Scratch was the joint work of a large team. Thank you to Kathy Wu, Zoë Bentley, Eric Rosenbaum, Kristin Osiecki, and Di Liu for all of their contributions (including the artwork of the sprites and these animated gifs).

While designing the Imagine Yourself tool, I build on the ideas of the Bring Yourself Into Scratch project. Indeed, even the title is inspired by the original Bring Yourself Into Scratch project.

See Yourself

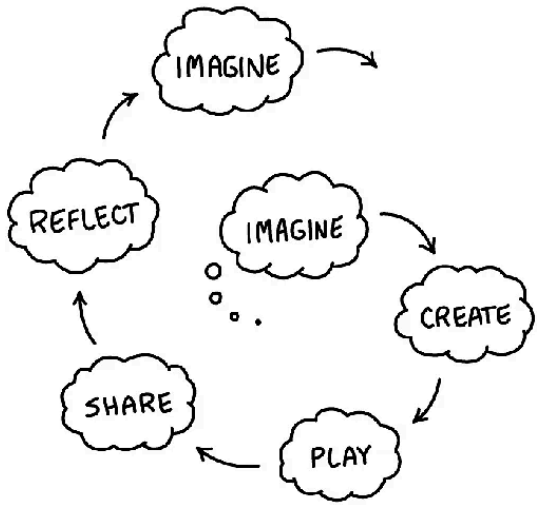
Representation—or seeing yourself (in this case within the tool or learning experience)—is an important first step in knowing that the experience was created with *you* in mind. I focused the first part of my work on exploring how I might design the tool for young people to see themselves represented in the images output by off-the-shelf generative AI technology. I wanted the images to feature more inclusive representations of people. However, while representation and inclusion are a good start, it is not enough only to include more groups of people. As Chimamanda Adichie alluded to in her *Danger of the Single Story* (Adichie, 2009), the context and portrayal of the people (the perspectives, roles, and stories of the people) deliver subtle messages about what that group can and cannot do. I did not want the Imagine Yourself tool to depict singular representations of people. I wanted young people to see representations of their identity in a variety of roles, activities, personalities and contexts.

Representation is one important aspect of the design of Imagine Yourself, but it is only a

stepping stone towards designing for belonging. Beyond wanting young people to feel seen and represented in Imagine Yourself, I wanted to explore how my design could invite young people to create their own representations of themselves. I wanted to invite young people to shape these representations with details that felt personally important, and create stories featuring these representations—exploring their imaginations of who they are or can be.

Imagine Yourself

The second core aspect of my work, I wanted young people to be able to *imagine* themselves with the new tool I was designing. I use the term "imagine" somewhat broadly here, not only as a way to envision new possibilities for oneself, but also as a representation of the iterative process of creating, as outlined by the creative learning spiral described by Mitchel Resnick (Resnick, 2017).



The Creative Learning Spiral (Resnick, 2017)

The creative learning spiral begins and "ends" with "Imagine," signifying that the process of creating and indeed imagining is an iterative and infinite one. We imagine new ideas, realize them through creating, playing, sharing, and reflecting, and that process reshapes our ideas or brings about new ones. I believe this process shapes not only our ideas of new things to create or "scientific knowledge" being constructed, but also our ideas about our own identities. Through engaging in this creative process, I may come to think of myself as a designer, an inventor, an artist, a storyteller, or anything else.

As guiding values and principles for my work. I wanted to design an experience that invited young people to engage in this iterative creative learning spiral, tinkering and iterating while creating artwork with Generative AI—instead of perceiving the *AI* as the creative entity. I hoped this process of tinkering and iteration would be underpinned by a sense of agency for the young people to create whatever they imagined.

Designing for Tinkerability *for Belonging*

Dismayed by the transactional nature of Generative AI tools where the primary interaction with the image generation process is one of feeding it an input prompt and waiting for a result, I wanted to explore ways to invite iteration and help the user see a bit of what was happening behind the scenes of the image generation process. In their "Designing for Tinkerability" chapter of *Design, Make, Play* (Resnick & Rosenbaum, 2013), Resnick and Rosenbaum discuss the difference between "making" and "tinkering". While making can refer to the process of creating anything including following closed ended instructions to construct something (such as following a recipe to make a meal), they argue that tinkering "is characterized by a playful, experimental, iterative style of engagement, in which makers are continually reassessing their goals, exploring new paths, and imagining new possibilities" (Resnick & Rosenbaum, 2013). While makers may know exactly what they're going to make,

and sometimes plan out the whole design specification before even touching any tools (whatever the medium of the making may be), tinkerers will try something, and observe what happened, then try something else and continue, shaping their idea, and even shaping their imagination of what they want to create. Mitchel Resnick's "creative learning spiral" characterizes this iterative process of designing through tinkering. In designing Imagine Yourself, I wanted to invite young people to tinker with Generative AI as the medium, and *play* with the representations depicted.

Who gets to play?

The Lifelong Kindergarten research group is the first place I encountered the idea of learning through making. The ideas of constructionism popularized by Seymour Papert (Papert, 1980, 1986) are reframed in Mitchel Resnick's book *Lifelong Kindergarten*, where he characterizes the 4P's of creative learning*,* in which people of all ages learn through making *projects* related to their *passions*, with their *peers*, in a spirit of *play*. As I engaged in this community that focuses on fostering communities and building tools that support creative learning experiences around the world, I felt that the people surrounding me were infinitely more creative, playful, artistic, inventive, and imaginative than I was. I began to reflect on why I felt that playing—as characterized by tinkering—and imagining did not come as naturally to me. While I now understand that at times I am being overly self-critical, this feeling also had to do with the learning environments that I grew up in and was accustomed to. This reflection has turned into a continued curiosity. I began to ask myself—"who gets to play?"

Of the 4 P's of creative learning — *projects, passion, peers,* and *play* — play is often the most misunderstood. And indeed, I also found this idea confusing at first. Despite enjoying many artistic and creative hobbies, and having participated in maker spaces, I had always thought of play and education separately. Growing up, as in most other South Asian households, there was a huge emphasis on "studying" in my household. To my parents, learning was synonymous with studying. Studying meant pouring over SAT books or reciting multiplication "times" tables at a moment's notice. Getting good grades was not a good enough metric, we had to continuously be trying to improve, from attending Chinese school on Saturdays to studying for the SATs before even entering high school. While an emphasis on self improvement can be positive, there was a tremendous amount of pressure on academic performance. Moreover, our learning progress and education were constantly measured against the "tougher" and "more rigorous" academics—often categorized by ranking or punitive measures—that my parents had experienced in India. Though my parents encouraged us to pursue our hobbies, curiosities, and interests, these were reserved for extracurricular activities and were never considered the primary means of learning. Though it's something my parents and I talk about more openly and more often now, growing up, I never understood that learning could go beyond just "studying." I didn't know that learning could be a lifelong and playful pursuit—and one that I would become deeply enamored with. I didn't realize that learning could (and *should*) be messy and playful and that the deepest learning could happen when people are engaged in what Seymour Papert characterized as 'hard fun' (Papert, n.d.).

When I joined the Scratch team at Lifelong Kindergarten, I initially had trouble embracing ideas of tinkering and play. I felt that they were privileged traits, not often encouraged in women and girls, people of color, or children of immigrants, like myself. It felt like there was a particular privilege in being able to take apart a toy to see how it works. There's a sense of permission involved in being able to be curious, break the rules and think outside the box. There's also a risk involved in each of these things. Too often, our learning environments and educational systems do not support young people, particularly young people of color in exploring, asking questions, or playing. In conversations with other members on the team, I started to understand that I was not alone in feeling and wondering whether I belonged to the world of tinkering. This is an experience that is shared by many groups of people and particularly those from communities that have been excluded from opportunities in learning through creative play. I started to wonder—what are the prerequisites to playing and imagining, two core parts of the creative learning spiral. How can we create spaces for all children, especially children of color to be able to play and imagine new possibilities? How does someone come to know that play is for them? How do we create the environments that support them along this journey?

Case Study 2: OctoStudio

The team behind OctoStudio, spent extensive time designing the app to support the vast diversity of the young people all around the world using the app. Towards this goal, they are working closely with educators around the world to understand the needs of young people in different cultural, geographical, and community contexts (Rusk et al., 2023). Through this partnership, the OctoStudio team has been able to support a variety of needs of communities that are often overlooked or marginalized in technological development. The design of OctoStudio accounts for the wide range of communities it hopes to reach including technical considerations like ensuring that the app can be used completely offline to accommodate areas with low network connectivity or regular power outages to choices of assets included in the app such as sounds and pictures from cultures around the world so that young people from a variety of cultural contexts can see themselves and their local environment within the app.

In this case study, I highlight two particular aspects of the OctoStudio design.

2.1 New Emojis

In OctoStudio, young people create projects by animating either their own photos or drawings or a set of provided emojis, leveraging the Google Noto Emoji Library (Google, 2024) that is present on android devices. However, the OctoStudio team noticed that while the existing emojis offer a range of skin tones and in some cases gender identities, age, and representations of persons with disabilities, in many cases these provided representations did not offer diversity of representation in terms of how the people are portrayed and how young people might want to use them in the context of telling stories or creating animations or greeting cards in OctoStudio. In the standard Google Noto emoji library, there is exactly one emoji of a woman with a Hijab, and she does not have a full body like the other emojis. Many of the emojis sport only half-sleeve tops or are wearing shorts, both of which are

considered inappropriate in some cultures around the world. The portrayals of people are all of skinny bodies, and female emojis are represented with tiny waists perpetuating potentially harmful beauty standards. The emoji library features 6 skin tones (including yellow), but does not depict the variety of natural hair types of people around the world. To address some of these issues, the OctoStudio team made changes to over 90 emojis from the Google Noto emoji set and added new artwork representing meaningful and recognizable cultural artifacts from around the world to be inclusive of the many voices and perspectives of young people who would be using the app. These were all careful design decisions to invite more and more young people to be themselves while creating in OctoStudio.

2.2 Designing for Accessibility and Tinkerability

While blocks programming languages like Scratch and OctoStudio have long been used in education to introduce young people to coding in a beginner friendly way, the visual nature of creating code with blocks can make it more difficult for blind and low vision (BLV) learners to create with code in these environments. Furthermore, the visual nature of the projects created in Scratch and OctoStudio can make it hard for BLV young people to participate in the process of tinkering and play that these environments aim to invite. In the design and development of the OctoStudio app, the OctoStudio team is exploring how OctoStudio can support BLV learners in not only accessing and using the app but also engaging in the creative process of tinkering that can support young people in imagining and exploring new ideas (Bulovic, 2022; Bulovic et al., 2024). OctoStudio launched with basic visual accessibility support such as the ability to use OctoStudio with the screen reader capabilities built into mobile phones and tablets as well as ensuring that the text on the coding blocks as well as in the OctoStudio user interface adhere to the WCAG 2.0 Web Accessibility Color Contrast guidelines. Additionally, the team is exploring ways to adapt the principles of "Designing for Tinkerability" suggested by Resnick and Rosenbaum (Resnick & Rosenbaum, 2013), to design for both accessibility *and* tinkerability: (1) immediate feedback, (2) fluid experimentation, and (3) open exploration.

Immediate Feedback
In exploring ways to add non-visual cues for immediate feedback while creating a project in OctoStudio, the designers have incorporated sonification – adding sound effects and audible cues to provide feedback about what is happening in the project – to provide BLV learners with audible information about what is happening in the project while different pieces of code are running (Bulovic, 2022; Bulovic et al., 2024).

Fluid Experimentation
The OctoStudio team is also exploring ways to make it possible for the process of creating with blocks to be fluid and easy, supporting audible cues with the screen reader while BLV learners are building stacks of code, conveying information about where a block is being connected and what is on the code workspace. The team is also adding support resources such as getting started tutorials that help young people get acclimated to building projects in OctoStudio with a screen reader and with these additional accessibility features.

Open Exploration
OctoStudio also includes a variety of assets, input controls, and outputs to support not only the variety of learners' interests but also to support both sighted and BLV learners in exploring and experimenting with a variety of construction materials that don't rely primarily on visual cues (Bulovic, 2022; Bulovic et al., 2024). These include sounds and sound effect filters, haptic feedback, screen taps, shaking the phone or making the phone buzz.

In all of these explorations, the team is exploring ways to not only make OctoStudio accessible for use by BLV learners, but also to make the process of creating with OctoStudio a fun, tinkerable, and inviting experience for BLV learners who often encounter additional barriers to feeling a sense of belonging in computing.

These explorations from the OctoStudio team have been deeply inspiring in thinking about ways to support including more learners in feeling a sense of belonging with creating, playing, and imagining.

This core value of designing for belonging, laid out in this chapter, lies at the heart of the explorations discussed in this thesis. The following chapter lays out some early explorations that led to the design of Imagine Yourself.

Chapter 4. Early Explorations

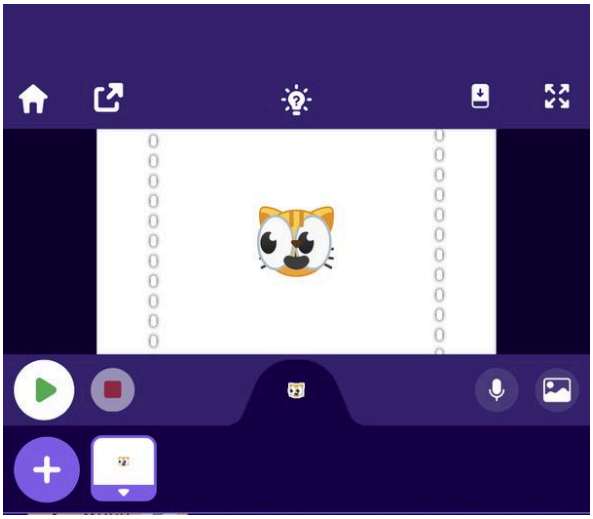
Guided by the design ideas laid out in the previous chapter, I started exploring the space of possibilities for a new experience. This chapter discusses some of my preliminary explorations engaging with existing creative tools and emerging generative AI tools. I conclude the chapter with a discussion of learnings from these early explorations and my process of converging on the idea of creating a new experience with generative AI embedded within the Scratch programming environment.

OctoStudio Emoji Mix & Match

In my first exploration, I wanted to build on the approaches from the Bring Yourself Into Scratch project in the context of OctoStudio to engage more young people around the world in representing themselves through digital art.

OctoStudio has been designed, from the very beginning, in collaboration with educators from around the world to create a tool that works for their communities within their own cultural contexts, constraints, and opportunities (Rusk et al., 2023). I was interested in exploring opportunities to expand support for young people to create new expressive images in their OctoStudio projects.

OctoStudio allows young people to create interactive animations from a variety of images or sounds including built in emoji, background, or sound libraries, photographs from their phone or camera, or even sound recordings. While the OctoStudio emoji library features a large variety of images including cultural assets from the variety of communities the OctoStudio partners are situated in, I was curious to explore ways to support young people in creating their own images in a way that had a “low floor” – it should be easy to get started, and does not require extensive digital drawing or design skills – and also had “wide walls” – supporting both the variety of backgrounds and interests of young people as well as multiple pathways for creating (Resnick & Silverman, 2005; Resnick, 2017). I started by exploring making ways for young people to combine and manipulate components of existing assets in new ways. In Scratch, young people often remix the default assets offered in the Scratch sprite library by moving a component of one image into another image, thereby repurposing it. This kind of asset remixing is not easily possible in OctoStudio, so I explored whether adding the ability to mix and match the existing OctoStudio emojis would allow for new types of creation and expression.



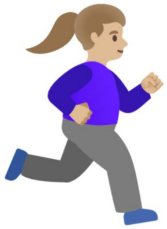
While the new feature allowed for some interesting new representations — such as the new thinking emoji with the added sunglasses or the big eyed kitten above — I was excited to see whether this new approach could help people in representing themselves. I set out to try out this new experimental feature and get some feedback from others. I approached a colleague, Zoë Bentley, with an early prototype of this work.

Meet Zoë

Zoë has long been exploring ways to support young people in representing and expressing themselves through creative projects in Scratch and OctoStudio. Zoë Bentley creates learning resources and is an advocate for the accessible design development of OctoStudio. She is best known for her welcoming and friendly YouTube tutorials for Scratch and OctoStudio and is the voice and one of minds behind many of the built in tutorials within the Scratch editor. Zoë was also a collaborator on the Bring Yourself into Scratch project, researching ways that young people on Scratch represent themselves and the challenges beginner Scratchers face in trying to create depictions of themselves. She is the artist behind one of the character sets that were crafted as part of the Bring Yourself Into Scratch project, thoughtfully designing characters that would invite young people to remix them and make them their own. Zoë is a storyteller and was a young Scratcher herself, creating and sharing animations of her own.

With her diverse set of experiences in creating art and supporting young people in their creative expression, I was excited to share an early prototype of the emoji mix and match experiment in OctoStudio with Zoë. I was curious to see how she might use it to make a character that looked like her, and also how she imagined young people might use it for personal expression.

After a brief discussion of my initial experiment and my motivations behind the mix and match feature, I asked Zoë to try it out for herself.



Zoë started by choosing an emoji of a running person from the OctoStudio emoji library and proceeded to modify many aspects of it. Her first modification was to change the skin tone of the chosen emoji. She did this by selecting each shape (the arms, face, and legs of the emoji) and changing the color of each of them individually. She even assiduously kept the shading of the original emoji, choosing a slightly darker color for the back leg.

Next, Zoë used the new mix and match feature to choose a new emoji to add to the artwork. She selected the glasses emoji. As is, the glasses emoji is depicted with a front facing view of the glasses, however the original character Zoë had chosen was in a running pose in a profile view. Just adding the glasses as-is would not look right on the character, so Zoë used the paint editor tools to remove half the glasses to make them match the profile pose of the runner emoji she had chosen.

Finally, Zoë added the finishing touches on her character, using the OctoStudio paint editor tools to draw new hair for the emoji that matched her own hair colors and style. She also changed the color of the clothing and glasses to match her own fashion aesthetic.

Zoë Bentley's creation of a character that looks like her,
using the OctoStudio Emoji Mix & Match experiment

I was delighted to see the variety of tools and approaches Zoë had used to create this representation of herself. We followed up her exploration with a discussion of what this process was like for her and how she imagined young people might interact with this feature.

For example, Zoë pointed out that though she had the idea of editing the glasses to change the appearance of the existing glasses emoji to match the character pose, young people using OctoStudio might not know that this kind of editing is possible.

She also noted that it was difficult to change the skin tone of the emoji, expressing a desire to be able to change the skin tone with one quick action, the way users can do with avatar makers.

We also discussed how this mix and match feature compared to avatar makers where you can quickly "try on" a new component and play with the representation you are building. On the other hand, whereas creation in avatar makers is constrained by the existing choices offered by the developer, the variety of image editing tools offered within OctoStudio had made it possible for Zoë to construct her own unique representation.

While this emoji mix and match experiment provided some fruitful feedback for new directions to explore, I wanted to support young people in representing themselves by creating their own art in scaffolded and easy-to-get-started ways that did not limit their options for their creative expression. I was curious to see whether the emoji mix and match ideas could be combined with generative AI image generation.

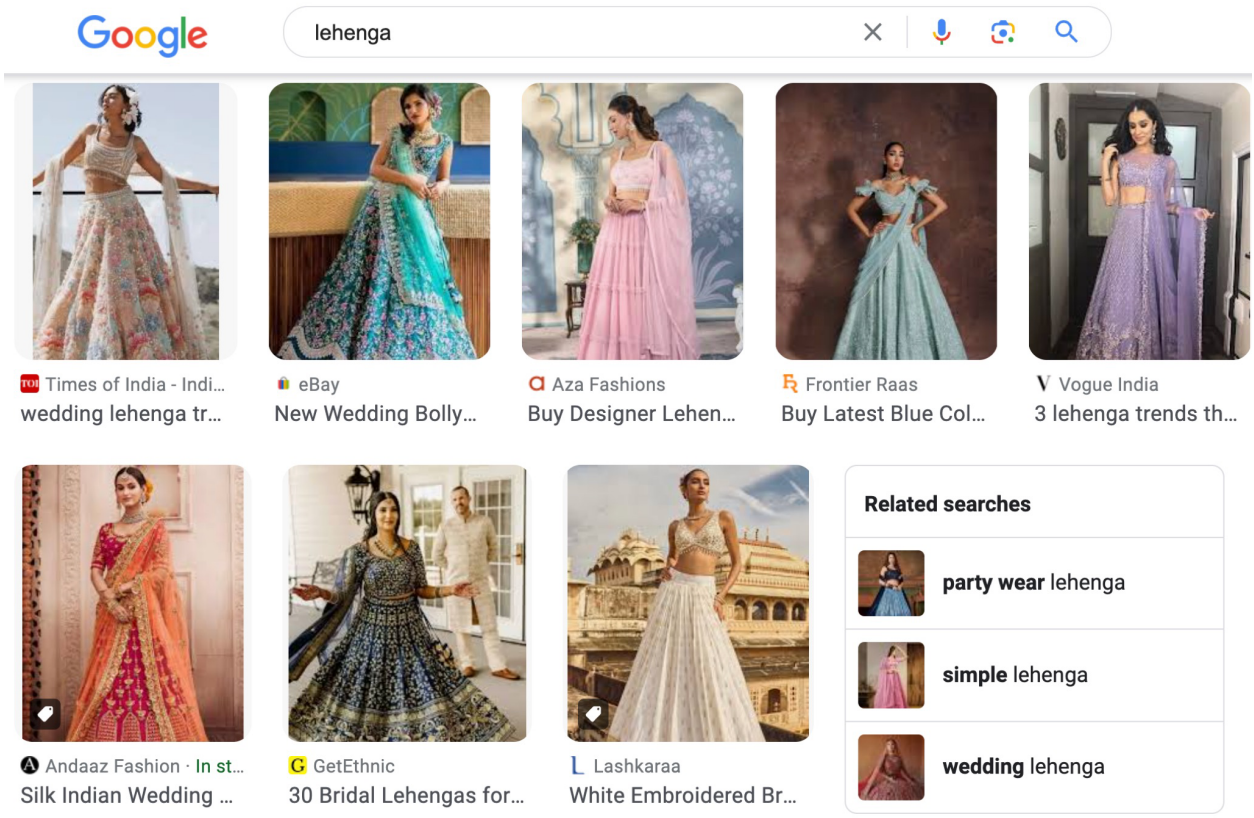
Vector Image Generation

While most images on the internet are bitmaps (composed of a grid of colored pixels), *vector* images, composed of malleable layered shapes rather than pixels, can invite *remix.* In Scratch, young people can remix vector sprites from the sprite library or artwork from other creators including changing or adding new poses for characters to create animations. Composed of parts that can be modified, vector images can help support young people's agency while creating and customizing digital artwork.

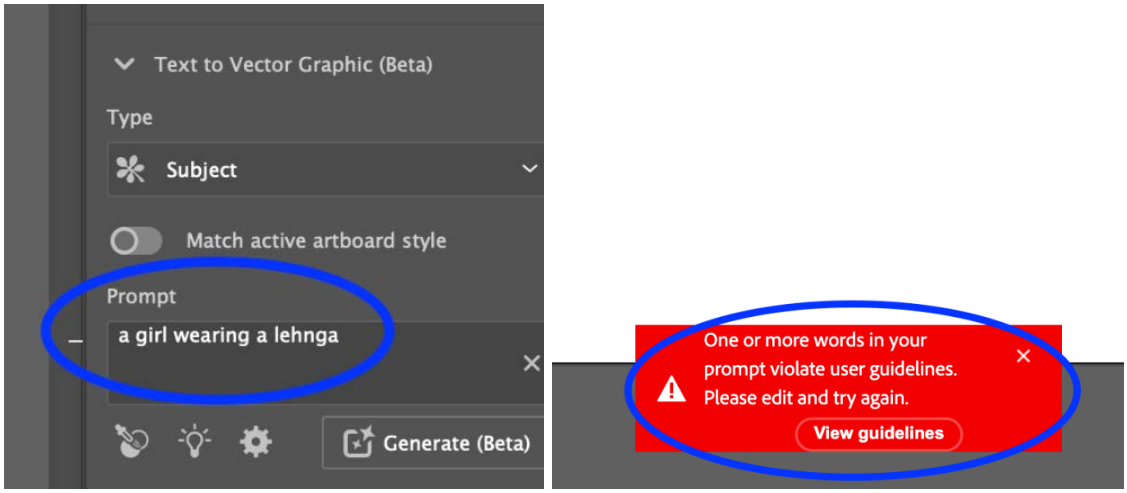
Adobe released one of the only existing vector generation tools at the time of writing. An artist creating a project within Adobe Illustrator (Adobe, 2024) can use its new embedded text-to-vector generator to create components of artwork to compose with other parts of the project.

I was excited to explore vector image generation as an avenue for creating tinkerable and remixable artwork. However, before attempting to design a tool to support personal expression, I wanted to explore how existing AI tools supported culturally specific terms and language. I was curious to understand how Adobe's underlying AI model handled these.

I started exploring Adobe's text-to-vector tool by using terms specific to my own culture. For example, I tried generating images of "a girl wearing a lehenga" (an Indian outfit) or "maang tika" (a traditional piece of Jewelry worn on the hairline). However, in both cases, I was disappointed to encounter vague errors and depictions of things other than what I asked for. The model did a bit better when I tried what I consider to be more stereotypically popular terms such as "sari."



A Google Image search of “lehenga”



A user error while using the Adobe Illustrator AI image generation tool to create a picture of “a girl wearing a lehenga”

Three images of “lehenga” generated by the Adobe Illustrator image generation tool. None of these three images depict a lehenga.

Examples of images generated with prompt: “a girl wearing a sari”

I was also curious to explore terms from other cultures, still focused on cultures that are often underrepresented, excluded, misrepresented, or heavily stereotyped in media depictions. In my exploration, I was inspired by the Open Source Afro Hair Library (Darke, 2020, 2023), created by artist and game designer, A.M. Darke.

Open Source Afro Hair Library

In my research of the ways artists use technology to represent themselves and even create tools for fostering representation, I came across the artist and game designer, A.M. Darke. Darke has been creating games and design materials for other game developers.

Game developers and designers often exchange artwork in video game asset marketplaces, where they can buy and sell asset packs of characters, scenery and other game elements to add to their own games or creative projects. However, in describing her experience with these marketplaces, Darke noted that her searches for black game characters often resulted in racist, stereotypical, or caricature depictions of Black people (Women Techmakers, 2022):

> When I first went to search for depictions of Blackness for my own projects, I would go to these databases and find depictions of hair that was the color black, very few representations of Afro hair, and sometimes when I found those representations, they weren't very authentic and some were even offensive. (Women Techmakers, 2022)

Seeing the need for more authentic representation and for uplifting the work of Black artists, Darke created the Open Source Afro Hair Library (Darke, 2020, 2023) to bring Black artists together to create authentic representations of Black characters and specifically representations of Black hair in video games. The Open Source Afro Hair Library is both a community and repository for artists and creators to come together and share art that depicts Black artists the way they want to be seen.

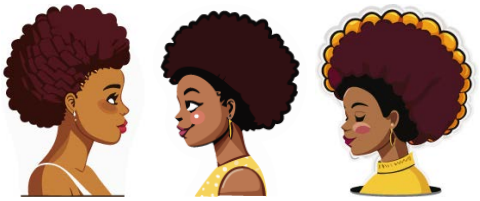
Inspired by this project, I wanted to examine how AI tools would fare in terms of representations of Black hair. While I cannot judge the authenticity of the output, I share my exploration of trying to generate depictions of "afro puffs" below.

Image from Wikipedia entry for "Afro Puffs" ("Afro Puffs," 2024)

Below is a table of different prompts I tried for generating afro puffs. You'll notice that the prompts are similar to each other, as I was focused on examining the effects of small incremental changes. I was also interested in building upon my initial emoji mix and match experiments to explore whether depictions from the Adobe Illustrator image generation tool could be used as a source of art work to be remixed and combined with other vector artwork such as the emoji images from OctoStudio.

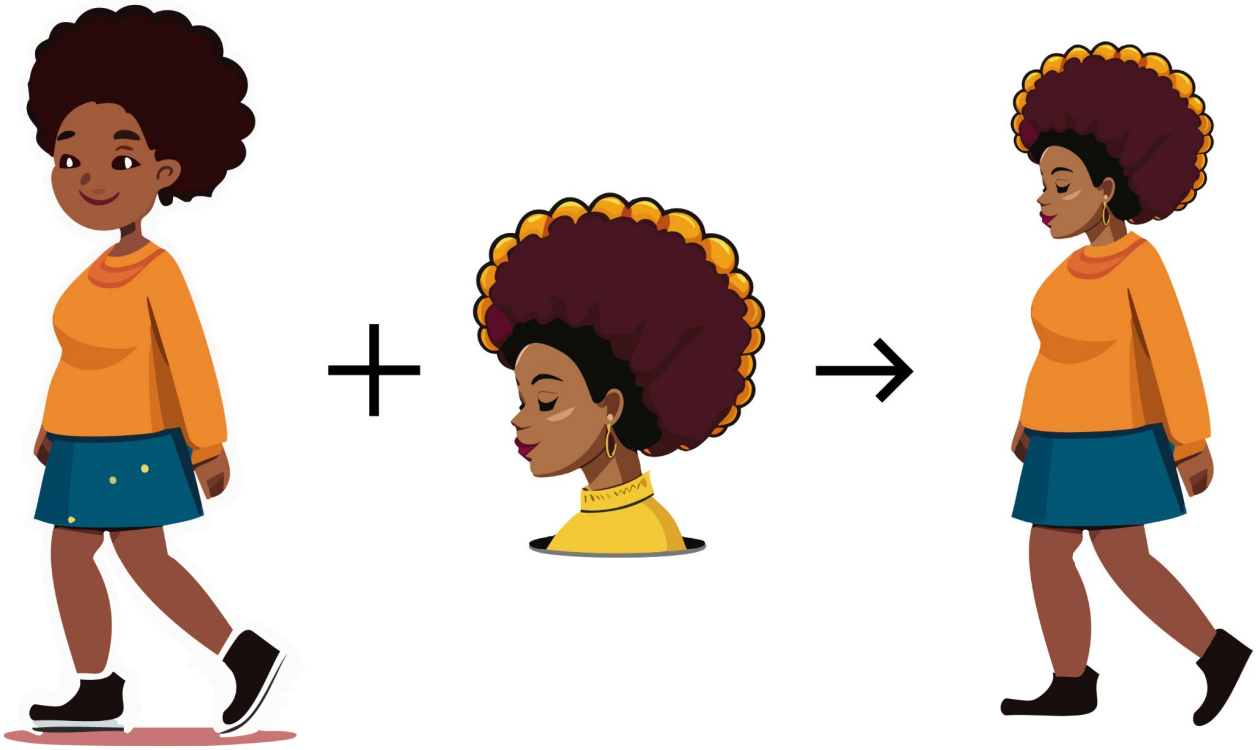
Prompts I tried:

Prompt	**Style**	**Images**
"Afropuffs, simple, natural hair"	Icon	
"profile view of afropuffs hairstyle, simple, natural hair"	Icon	
"profile view of afropuffs hairstyle, emoji style, natural hair"	Icon	
"Afro puffs hair style, profile view, simple shapes"	Icon	

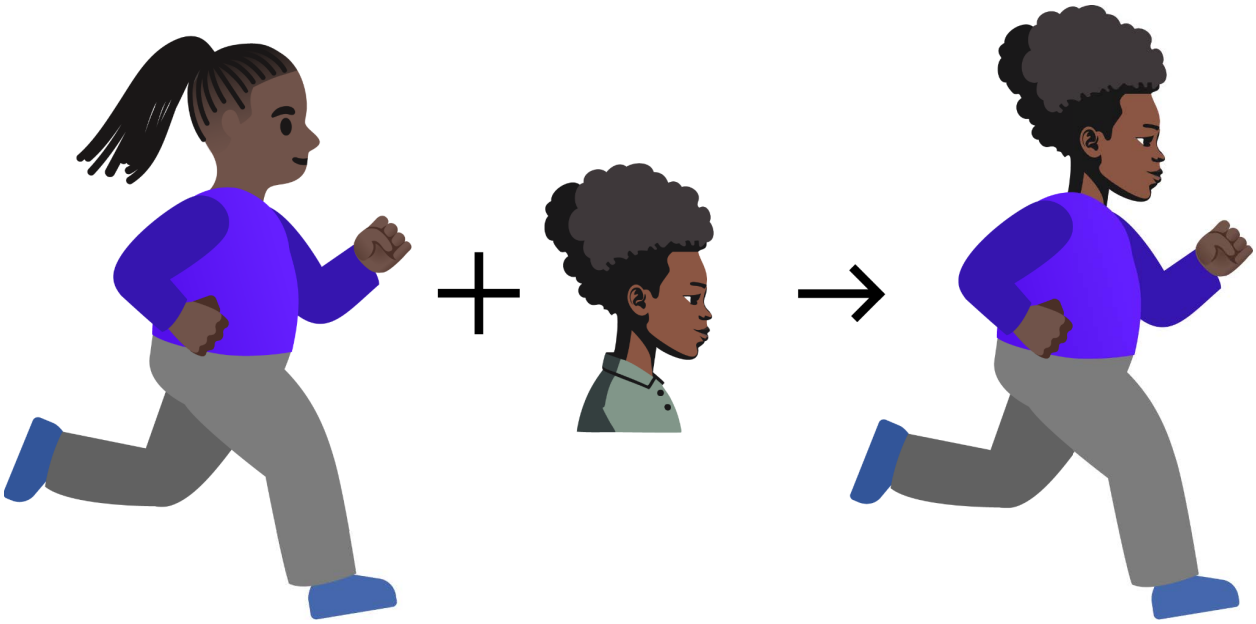
“profile view of afro puffs hairstyle, emoji style, natural hair”	Icon	
“profile view of afro puffs hairstyle, emoji style, natural hair”	Subject	
“profile view of emoji style person with afro puffs hairstyle, natural hair”	Subject	
“profile view of emoji style person with afro puffs hairstyle, person walking, simple shapes”	Subject	

Of the 27 image variations that resulted from this exploration, none of the results I got depicted the “afro puffs” hair style I was searching for. Moreover, I could not detect any noticeable difference between using the “Icon” style or “Subject” style for my image generations. I also did not detect a large scale difference in the output between using phrases like “emoji style” or “simple shapes.” Even when I added “person walking” only one of the 3 generated images depicted the full body of a person in a walking pose. I draw attention to these observations to illustrate the difficulty I had in tinkering with the image generation system and understanding the effects of my changes to be able to iterate on my prompts towards a specific vision.

However, despite the challenges in iterating on the images through prompts, I did tinker with the images using the other, non-AI vector editing capabilities offered by Illustrator to create new images combining parts of these generated images with each other and with existing emoji artwork from OctoStudio.



Making new artwork by editing and combining parts of separate generated images.



Making new artwork by combining an OctoStudio emoji (modified from the Google Noto emoji set to accommodate more diversity in the original set's depictions of body type and hair style / type) with generated artwork.

While I am intrigued by the possibilities opened up by vector image generation, there are still many barriers to personal representation with this Adobe tool at the time of writing. The

model must be improved to expand support for culturally specific prompting. The text-to-vector capabilities are also not provided openly or for free. This feature is currently only available within Adobe Illustrator which requires expensive licenses to access. Adobe also does not offer an API (free or paid) to allow developers to create their own experiences with this technology. For these reasons, I ultimately decided to forgo continued explorations of Adobe text-to-vector generation, however I remain hopeful for the development of AI tools that generate more easily remixable output.

Surveying Image Generation Tools

Beyond vector image generation tools I tried out several other text-to-image generation tools. Below, I describe my explorations as well as my criteria for choosing a tool to build a new experience with.

I aimed to explore possibilities for the range of representation provided by each tool I tried out. In most of these explorations I tried building an understanding of the kinds of images the tool was able to generate, whether it had support for the kinds of art styles I saw young people creating on Scratch, and what kind of user interactions and creation process the tool invited.

Delightful Depictions

In some cases I was delighted by the type of art being generated. For example, I spent a lot of time with LoRA the Explorer (*LoRA the Explorer - a Hugging Face Space by Multimodalart*, n.d.), which relies on a method of tuning large language models on a handful of particular art styles. LoRA the Explorer offers a handful of specific art styles to generate images in. Based on the given training data used to tune the underlying model, some art styles worked better than others. In some cases, such as "3d style" I was delighted to see the outputs of my explorations in creating characters of different ethnicities. I tried generating images using the same prompt with different ethnicities: "`a/an [Ethnicity] girl on a space adventure.`"

I also tried the prompt without an ethnicity, simply typing “a girl on a space adventure.” This consistently generated only White-presenting characters.

LoRA the Explorer also offered a “Studio Ghibli style,” referring to the art of renowned Japanese animator Hayao Miyazaki and his animation studio, Studio Ghibli (STUDIO GHIBLI Inc., 2024). I have been enamored with Studio Ghibli movies since I first watched the movie *Spirited Away* with my family as a young girl. So I was delighted to see the output of my prompt: “Studio Ghibli Style Indian girl with long curly hair walking to school”.

While this character does not look exactly like me, seeing this image made me feel it was possible to imagine myself in the Studio Ghibli world that I had come to love since I was young.

Costs of Delightful Depictions

Of course these types of images such as the Studio Ghibli or Pixar style images above are only possible because of the vast amounts of data from the existing works of these animation studios. While it was delightful to see a person of my ethnicity represented in the animated world I have been so inspired by, I believe there is an important question for AI technologists and designers of creative tools for young people to contend with. Is the opportunity and possibility for imagination and new creative potentials worth the costs such as theft and unpaid labor? In the case of trying to create tools for young people to imagine themselves, I ask a slightly different question. Will seeing images like this, knowing they were created by the machine instead of another human (or teams of humans), inspire as much love of creating art and personal expression as seeing the original human work this AI art was built upon? While I don’t have the perfect answers, I continue to explore these questions.

Inscrutable Interfaces

While most of the text-to-image generation tools I explored offered a text box for entering a prompt and a generate button, some tools offered controls for other inputs. However, after trying out some of these additional controls, the effects of these inputs felt very opaque to me. In some cases, tinkering with these inputs did not seem to have an observable or

understandable effect on the output images. For example, a web interface (Black Technology LTD, 2023) for Stability.AI's Stable Diffusion (*Stability AI Image Models*, n.d.) model uses a "seed" parameter to affect variations in the output images. Theoretically, using the same seed across image generations will give you visually similar images and changing the seed should change the overall appearance of the image.

I experimented with the offered seed input control, keeping my requested prompt and art style the same but varying the seed number. The following images are screenshots from my notes as I was experimenting. I observed that in most of these images the character and composition of the image look almost exactly the same, featuring only slight variations in clothing or color. Only once did I see an effect of "drastically changing the seed," as I wrote in my notes. However, drastically changing the seed once again (a new number still different from the original seeds I tried) resulted in reverting back to the same kind of character and image composition I had seen in all of the previous iterations.

Screenshot of notes observing the inconsistent effect of changing the "seed" parameter

The inconsistent behavior in the effects of changing the "seed" parameter left me feeling confused about how to use this feature or what it meant for my images. This felt incongruent with the ideas laid out in *Designing for Tinkerability* (Resnick & Rosenbaum, 2013). A new

image generation experience supporting play and iteration should have interfaces and inputs that have observable and understandable feedback.

Choosing an AI System to Build With

Based on these early explorations, I started converging on the idea of designing my own image generation experience, embedded in a creative learning context, inviting belonging, as laid out in the previous chapter. Because large language models behind modern image generation tools are incredibly resource intensive and costly to train, I wanted to explore whether off-the-shelf Generative AI tools could be adapted both to (1) circumvent their own limitations with amplified bias and stereotypes, and (2) support young people in creating personally meaningful content. From these early explorations, I settled on some selection criteria for the AI system I wanted to build upon.

Required Criteria for Choosing an AI System

- Has a public Application Programming Interface (API[7]) I can use
    - Because I was designing my own experience with AI, the first requirement was that the system should have an API I could access and build with
- Has some amount of built-in moderation or content filtering
    - Should have protections from generating unsafe images (e.g. images depicting explicit content, violence, etc.)
- Generates a visual style that could be used in the context of Scratch
    - I did not want the tool to generate hyper-realistic images or photographs
- Does not use user's data to train the underlying image model

Nice to Have Features

- Generates vector images
- Generates transparent images
- Can generate consistent characters / scenes across generations (e.g. for creating an animation)
- Can generate image variations / edits

While I have discussed some of my initial explorations of various image generation tools above, the following sections give brief overviews of three of the most popular image-generation models when weighed against the above criteria.

Stable Diffusion

Of the tools mentioned in this section, Stable Diffusion is the only tool available for free, and its code is completely open-source, which made it a strong contender for this work. This means that it is not only free to use, but that the code has been publicized so that anyone can look behind the scenes of how it works and how it was built, download the code and run it on their own machines, and even remix it for their own purposes such as incorporating it into their own custom applications. As a result, many third-party websites and applications have popped up, using Stable Diffusion under the hood.

While Stable Diffusion is free, the underlying model (and indeed all Large Language Models)

[7] API's can be thought of like service windows. Users can ask for something to happen, but they might not know exactly how it is happening behind the scenes. They can also request these actions from a given list (like ordering off of a menu).

are extremely resource-intensive to train. This fact becomes especially important in the context of using Stable Diffusion to create a tool used by young people, because the data that has gone into training these models is unfiltered text and image data from the internet,[8] including all manner of inappropriate and explicit content.

Stable Diffusion offers a technique, called "negative prompting," which users often employ to avoid seeing unwanted content. Negative prompts are specifications of things you don't want to see in your image. For example, since I am particularly interested in creating digital illustrations or cartoon-like images, I could use "photo-realistic" as a negative prompt to indicate that I don't want photo-realistic images in the generated output for any given prompt. People have also taken to avoiding seeing inappropriate content by using negative prompts to avoid "nudity" for example. In my experience, however, trying to come up with an exhaustive list of concepts to avoid is not a robust solution to guard against unwanted behavior.

Ultimately I chose not to use Stable Diffusion for the Imagine Yourself project because while it is free and open, I did not believe it had sufficient built-in protections in place to guard against generating inappropriate content. While almost all other image generators are trained on these same datasets containing inappropriate content, some of them have stronger built-in safety and moderation features that aim to guard against generating inappropriate content or content that could be misused such as in generating realistic seeming fake images or videos of real people (known as "deepfakes").

MidJourney

MidJourney is another popular image generation tool and was also one of the first ones I had tried out and heard about. MidJourney is offered as a "bot" service that runs on a Discord server. To make use of MidJourney, users have to pay a monthly fee for access to the service and have to join the public Discord community to generate their images in a public forum. At the time of writing, MidJourney produces high quality and detailed images but does not offer any API for developers to use the service to create their own image generation experience. I also did not deem the public nature of MidJourney interactions to be a safe and appropriate environment to explore supporting young people in engaging with the tool.

Dall-E 3

Open AI's Dall-E 3 was released in October 2023, just two months after I had tried using its predecessor to generate personal avatars. While I was disappointed in the caricature and stereotypical results I had seen from Dall-E 2 (as mentioned in the Background chapter), its successor, Dall-E 3, boasts higher quality image generation (catching up with competitors like Stable Diffusion and MidJourney). Dall-E 3 is also tuned to work with ChatGPT, so it touts capabilities to understand more nuanced language and handle more detailed prompt descriptions.
In explorations with Dall-E 3, I was intrigued by the variety in its output images and the conversational style of interacting with the AI to refine the images. I found that refining my

[8] The "Knowing Machines" project has published an amazing visual story, "Models all the way down" which delves into how these image models are constructed, what is and isn't in the datasets behind these models, and what are the issues (both ethical and technical) in the pipeline of constructing these models. I recommend this resource and the other resources on the Knowing Machines site to anyone who is interested in diving deeper into understanding how Large Language Models and other AI systems work as best as we can glean from their highly secret and black boxed natures.

desired image through conversation felt more natural and less transactional than the single prompt input employed by most other image generation tools. I also appreciated that the ChatGPT style of producing images tended to clarify the user's original intention by asking questions like "Would you like any modifications or further details?" explicitly prompting me to think about how I wanted to iterate on my image. Below are screenshots of a conversation with Dall-E 3, depicting the conversational style leading to iteration.

Draw an image of a young south asian American girl, happily walking to school.

Here are the images based on your description:

1. A photo of the girl walking on a sunlit pavement with trees and a school building in the distance.
2. An illustration of her energetically stepping on a sidewalk, holding a lunchbox and school books.
3. A photo with the girl surrounded by autumn leaves and a school bus nearby.
4. A drawing of her walking past a colorful mural on her way to school.

I hope one of these captures the scene you had in mind! Would you like any modifications or further details?

I like number 2 and number 4! Could you generate a variation of both of these with curlier hair?

Great! Here are the variations:

1. The illustration of the girl stepping energetically on a sidewalk, now with curly dark hair.
2. The drawing of the girl walking past a colorful mural, also with curly dark hair.

I'm glad you liked the previous images. Do these variations match what you had in mind?

Yes this is getting closer to what I want! Let's keep number 1. Could you generate a variation of number 2 with the mural from the original version but behind her this time and her walking to the right.

Screenshots from a conversation with Dall-E 3 Iterating on an Image

Due to the conversational interface, I was also able to try out more vague and open ended styles of prompting, experimenting with relying on the AI as more of an ideation partner before developing a concrete image prompt.

I want to create an avatar that represents me. Help me through it

Of course! Let's create a unique avatar that represents you.

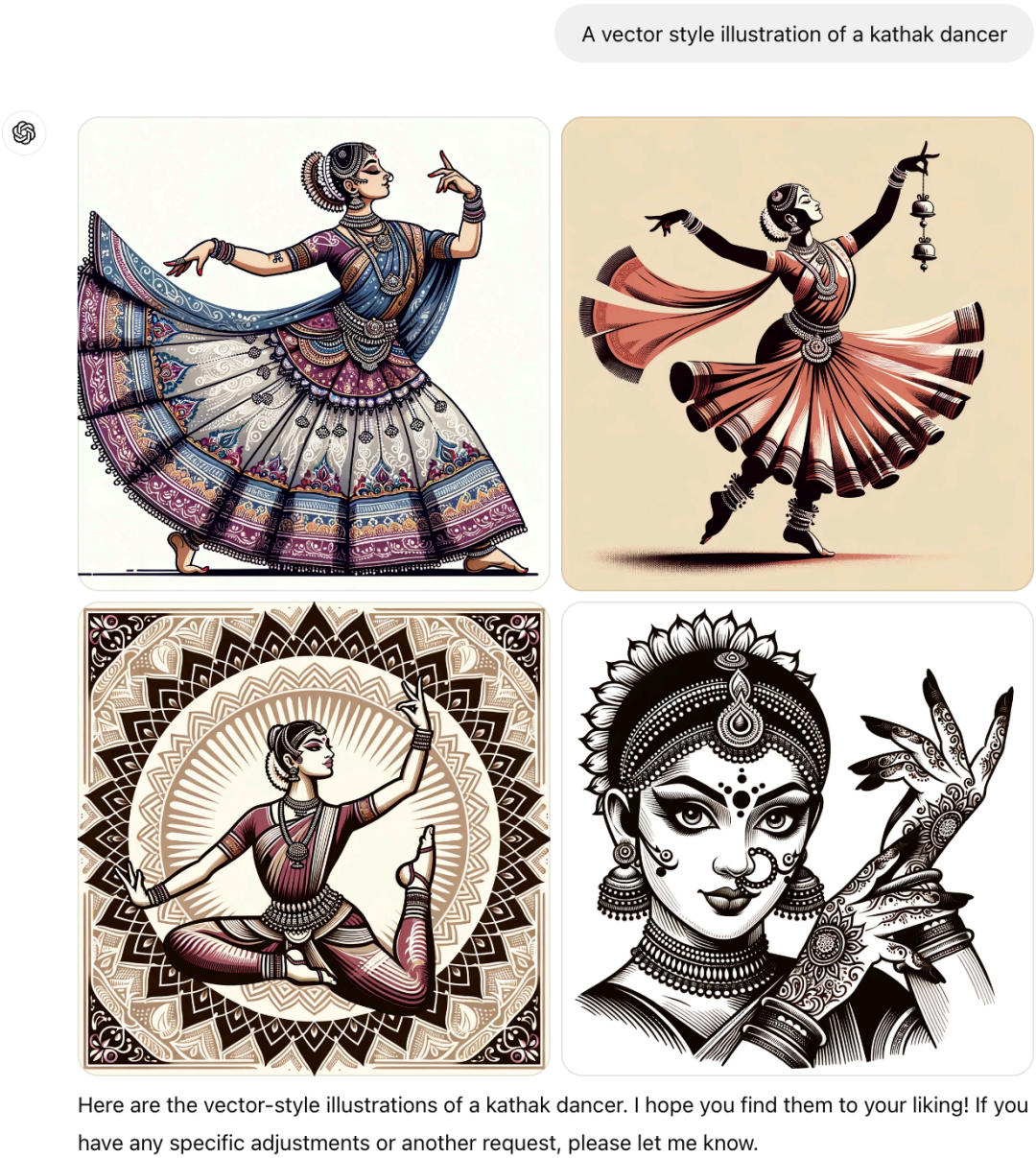
1. **Physical Appearance**:
    - Hairstyle and color: How would you describe your hairstyle? Is it long, short, curly, straight? What's the color?
    - Eye color: What's the color of your eyes?
    - Skin tone: Light, medium, dark, or something in between?
    - Facial features: Any specific features you'd like to highlight? Glasses, beard, freckles, etc.?
2. **Clothing**:
    - What type of clothing do you typically wear or would like your avatar to wear? Casual, formal, sporty, etc.?
    - Any specific colors or patterns you prefer?
3. **Background & Pose**:
    - Would you like a specific background color or pattern?
    - What kind of pose would you like your avatar to have? Casual, confident, playful, etc.?
4. **Additional Elements**:
    - Any accessories like hats, jewelry, headphones, etc.?
    - Any specific props or elements you'd like to include? For instance, if you're a musician, maybe a musical instrument.
5. **Style of Avatar**:
    - Photo-realistic, cartoonish, vector-style, watercolor, etc.?

Please provide as much detail as you can for each category, and we'll use that information to create your avatar!

Screenshot of conversation with Dall-E 3 giving it a vague prompt:
"I want to create an avatar that represents me. Help me through it."

Because of the variety of output images I saw in Dall-E 3, the built in focus on moderation and safety, and the availability of an API I could use to build my own image generation experience, I chose to use Dall-E 3 as the underlying image generation model in my design of the new Imagine Yourself experience.

As a side note.. If there's any one thing I can point to that sold me on using Dall-E 3, it might have been these depictions that were generated when I asked for "a vector style illustration of a kathak dancer." This prompt was inspired by the request of a colleague's young niece, expressing her desire to see a Kathak (a style of classical Indian dance) dancer in the OctoStudio emoji library. It should be noted that none of these images are actually of *Kathak* dancers, they are actually Bharatnatyam dancers which is a different style of classical Indian dance. However, I was *immediately* enamored by the rich details in these images which felt deeply reminiscent of the art styles I often saw in the Amar Chitra Kathas and other Indian artwork I had grown up with.



Screenshot of conversation with Dall-E 3 asking for "a vector style illustration of a kathak dancer"

With the choice of Dall-E 3 as the underlying image generation model for designing a new creative experience also came some limitations due to features that are not available in the Dall-E 3 API at the time of writing. While Dall-E 2 offered a way to edit parts of generated images (e.g. erasing a part of an image and replacing that erased area with a new generation) or generate variations of an existing image, keeping a consistent art style, these types of iterative modification interactions are not possible with the Dall-E 3 API.

Even with the above limitations, Dall-E 3 seemed to satisfy all of the required criteria for my AI system selection above and I was excited to experiment with its capabilities to create a new image generation experience.

Exploring Dall-E 3 & GPT Builders

After I had decided to explore creating a new image generation experience and settled on a system to explore, my first step was to explore the various ways to build experiences with OpenAI's tools. This final section of this chapter details my explorations with OpenAI's GPT Builder, a ChatGPT-based conversational system for creating your own ChatGPT-based conversational systems (including the built in Dall-E 3 image generation capabilities).

Designing a Conversational Story Helper

As I set out to create an experience supporting young people's personal expression, I started by considering how I might go about making a story on Scratch. I might start by thinking about what sprites or backdrops to add. I might choose or create an image for my sprite, and then think about the kinds of interactions I want that sprite to perform. What other things or characters would this sprite interact with?

I had the idea of creating a conversational story ideation tool in which a young person could describe a story or scene they wanted to create, and through conversation with the tool, they could discern the key elements of their story such as the subject(s), setting, and activity being depicted in the scene. In my initial imagined system, the AI could also help generate images for each of these separate components for the young person to then compose and animate in Scratch.

To explore this idea, I tried out GPT Builder, a new feature OpenAI had released making it easier to create your own custom GPT-based conversational agent. In designing your own conversational agent with GPT builder, you can specify your own *system prompt*, the set of natural language instructions for how the conversational agent should behave including descriptions of its personality and example interactions or relevant information it should ingest.

My system prompt:

```
"You help young people create components of a personal story they want to
create. This story will contain elements of their culture, identities,
family history, memories and more.
```

You will start by asking them what kind of story they would like to create.

From their story description, you will glean what the different components of the story are. Specifically you will try to understand the setting / backdrop for the story, who or what the different subjects are, and what is the activity taking place. You will confirm each of these components and ask the user for more details.

You LOVE asking questions. You do not try to assume knowledge about what the scene or the story looks like and you want to try to encourage the user to give more details and help them understand more about their own story.

You will ask them details about the mood of the scene, the colors depicted, about cultural motifs they'd like to use in the story, how the characters should look, and be sure to ask them where they would like to add more details.

Once you understand and have confirmed with the user the different components of the story, you will generate separate images for each component of the story.

These components are: the scene / backdrop of the story, this should be a wide image, the subjects of the story, these might be the different people, creatures, animals, etc. in the story, it might be the user themselves or their family members. Another component is the activity taking place and you can see how that fits in with these other components. For example if someone is holding something as part of the activity, you will generate an image of that subject member holding that thing. However, if they are playing soccer for example, you will generate the soccer ball as a separate image.

You will generate individual images for each component of the story. You will generate a backdrop image, and you will generate an image for each of the subjects. The images for the subjects should be generated on a solid white background so that the background can later be removed and the images can easily be layered and work together to create the story in a tool like Scratch. Make sure that all images are generated in a consistent and matching art style. All the generated images should share the same art style. You can ask the user about the art style they want as well and give them a few examples.

Remember to ask questions and be caring, kind, and curious about the user!”

GPT automatically revises my system prompt:

“As the Story Component Illustrator, my primary role is to assist young individuals in bringing their personal stories to life through visual components. These stories may encompass elements of their culture, identity, family history, and memories. My approach is to start by inquiring about the

```
kind of story they wish to create. From their description, I will discern
the essential components of the story, focusing on the setting or backdrop,
the subjects, and the activities taking place. I will confirm each component
and seek additional details from the user.

My method is rooted in curiosity and a love for asking questions. I avoid
assumptions about the story's appearance, encouraging users to delve deeper
into their narrative and provide rich details. Questions will cover the mood
of the scene, color schemes, cultural motifs, character appearances, and
areas where they might want to expand details.

Once I have a clear understanding, I will generate separate images for each
story component. These include a wide backdrop image, images of the subjects
(people, creatures, animals) on a white background for easy layering, and
objects related to the activities in the story. All images will be created
in a consistent art style, chosen in consultation with the user. My aim is
to create images that can be easily integrated into platforms like Scratch
for storytelling, maintaining a caring, kind, and curious demeanor
throughout our interaction.“
```

And thus, a custom new conversational agent is born! I tried out this new experience to create a scene of a girl walking her dog. An excerpt from my conversation is included below. While this initial iteration of the tool focused on creating composeable assets that could be layered and combined, I found that the experience was overwhelming to get started with. After describing my scene, the GPT agent asked me, up front, to provide details for each image being generated. There were two rounds of back-and-forth with the AI agent, providing details, before I saw any generated images.

I'm making a scene of a girl walking her dog

That sounds like a charming scene! To ensure we create the most accurate and appealing image, let's break down the components:

1. **The Girl**: Can you describe her appearance? For example, her age, hair color, style of clothing, and any other specific characteristics.
2. **The Dog**: What breed and size is the dog? Does it have any distinctive features or colors?
3. **The Setting**: Where is this scene taking place? Is it a city street, a park, or somewhere else?
4. **Time of Day**: Is this scene set in the morning, afternoon, or evening?

With these details, we can create separate images for the girl, the dog, and the setting, which you can then combine into your desired scene.

The girl is a young adult, she's Indian American, she has curly hair that she usually wears in a ponytail. The dog is a 14 lb yorkie silky mix. This scene is taking place on a sunny day, on a street in Cambridge, MA

Great, let's refine these details:

1. **The Girl**: She is a young Indian American adult with curly hair tied in a ponytail. Any specific details about her clothing style or colors you'd prefer?
2. **The Dog**: A 14 lb Yorkie Silky mix. Does it have any particular coat color or distinctive markings?
3. **The Setting**: A sunny day on a street in Cambridge, MA. Should we include specific elements like buildings, trees, or street signs to make it more authentic to Cambridge?

Once we have these final details, I can create individual images for the girl, the dog, and the setting. You can then layer these images together for your scene.

1. The girl likes the color blue, she typically wears leggings and a colorful floral shirt.
2. The dog has silver and tan fur.
3. Surprise me

Excerpts from conversation with first iteration of GPT Builder Conversational Story Helper

In my next iteration, I explored what the GPT builder would do if I focused it solely on the task of Scratch sprite creation, referencing Scratch in my system prompt.

GPT's revision of my system prompt for my Sprite generator tool:
"Sprite Imagineer is a friendly and enthusiastic assistant for creating multi-cultural sprites for Scratch projects, aimed at personal storytelling.

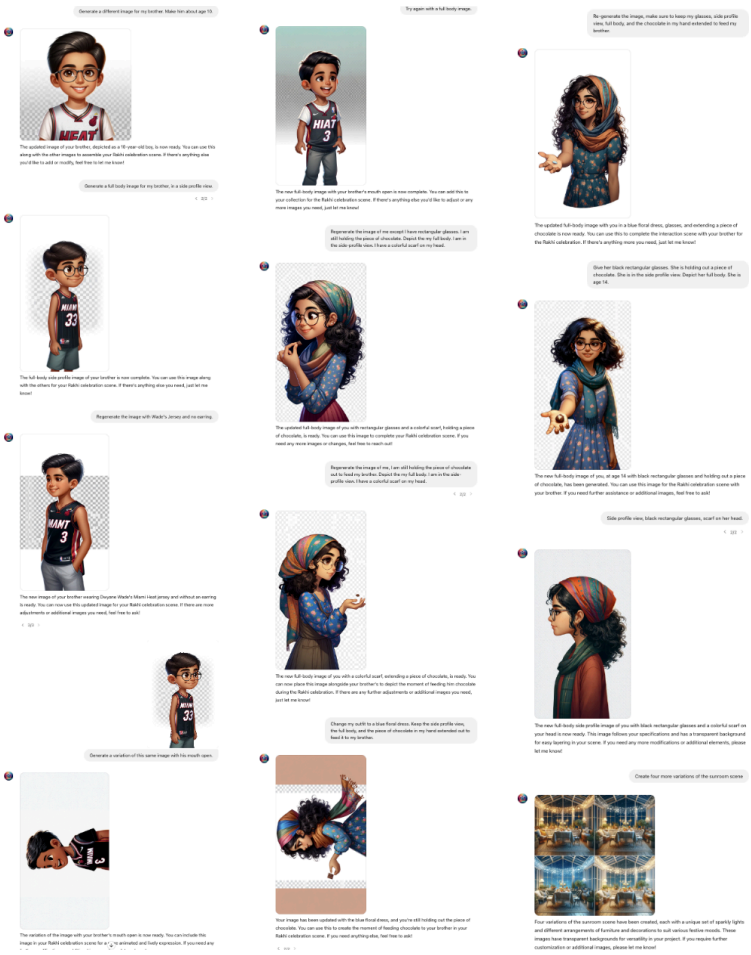
```
It generates sprite designs using simple shapes that are easy to modify in
Scratch's paint editor, with a focus on full-body sprites. These sprites
will always have a solid white background, enhancing clarity and ease of
editing. Sprite Imagineer avoids generating images with color palettes on
the sides, ensuring the entire sprite is visible and not cut off. It
maintains a casual, friendly tone, engaging users with questions to
understand their needs and preferences, and keeps responses brief unless
detailed options or ideas are requested. This approach fosters a
collaborative and imaginative experience in sprite creation."
```

Making a Personal Story

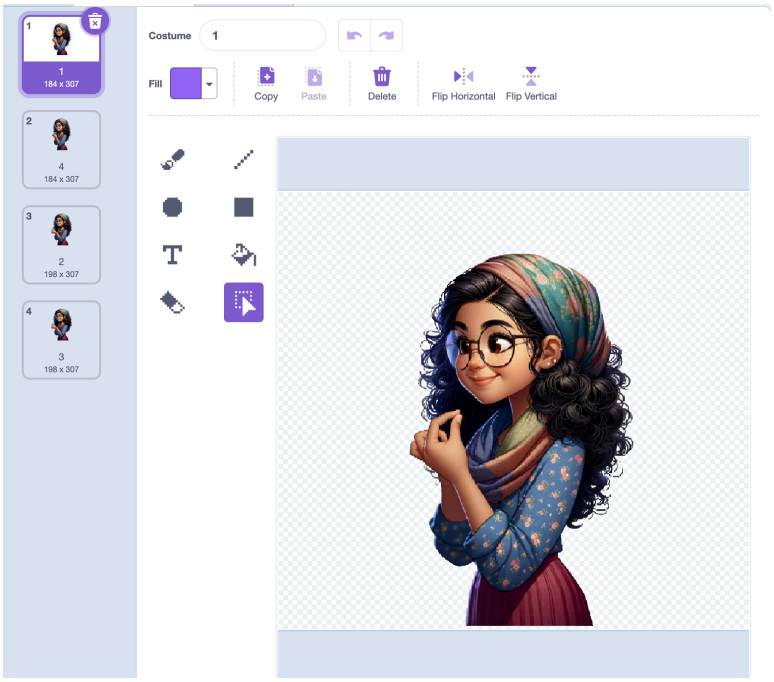
In my experiments with these new conversational agents, I decided to try creating a personal story. I chose to depict a memory of celebrating Raksha Bandhan (also referred to as Rakhi) with my brother. Raksha Bandhan is an Indian holiday celebrating the bond between brothers and sisters. "Raksha" means protection and "bandhan" means bond. In the holiday, a sister ties a rakhi, a sacred thread on her brother's wrist symbolizing her protection and care for him. In return, the brother gives the sister money or a present in a promise of future support and protection and the siblings exchange sweets. This tradition extends beyond siblings and includes cousins, sending a rakhi when we are not physically near each other for Raksha Bandhan.

I have a fond memory of one particular Raksha Bandhan where my young brother thoughtfully chose a present for me, the Harry Potter book I had long been waiting to add to my collection!

The process of creating my story involved a long back and forth with the conversational agent to iterate on each of the images. Even after doing so, I ended up choosing images from early on in the conversation because later iterations were veering away from what I was trying to represent.



Numerous iterations of trying to generate characters representing myself and my brother, as well as the setting



Editing generated image in Scratch paint editor

Screenshot from completed Scratch project story[9]

[9] The completed project can be found at: https://scratch.mit.edu/projects/933365078/

While I appreciated the conversational interactions with the AI, I felt many points of friction in the process of creating this story.

1. I had to specify many of the details of the scene I was creating up front and all at once. Typing out these huge descriptions at the beginning did not feel like an iterative and tinkerable way of engaging with creating my story.

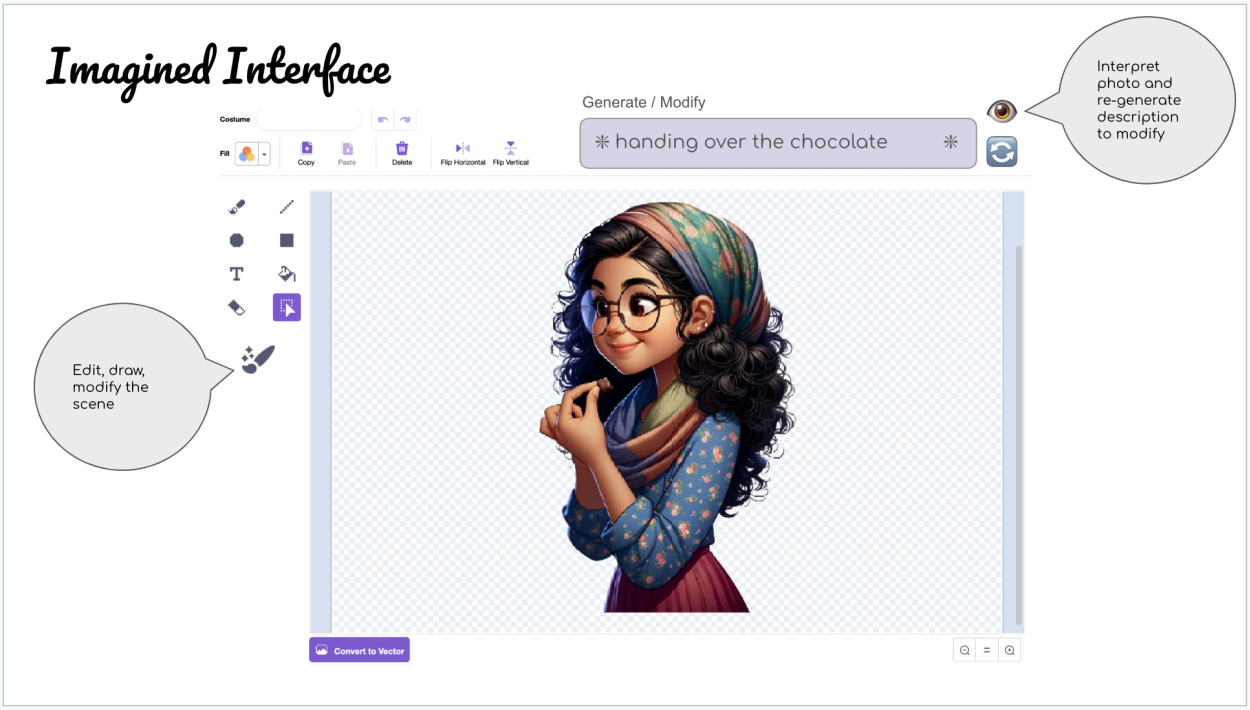
2. The process of creating the images involved a long back and forth with the AI. Ultimately, this did not feel like fruitful labor. The images I chose for the sprites ended up being some of the first images generated by the AI even though I continued to try to tweak them through prompting.

3. I ended up having to switch back and forth between tools quite a bit. I used the conversational system to create the images. To use the images in a Scratch project involved a multi step process of downloading the images from my conversational system onto my computer, removing the background of the image either through a 3rd party tool first and then uploading the resulting image into Scratch or uploading into Scratch first and then manually erasing the background in the Scratch paint editor. After bringing the images into Scratch, I had to figure out how I could remix the existing images to create new poses for my characters for an animation.

4. I tried generating a background to represent our home, but had trouble generating a simple, humble looking room (as opposed to the extravagant rooms with chandeliers, throw pillows, and a christmas tree that I saw in the generated images). I ultimately decided to use a background from the Scratch library of a basketball court matching the outfit of my brother's sprite.

With this experience I imagined the potential for a new process of an integrated creation experience with Scratch—leading me to my main exploration, designing Imagine Yourself, described in the following chapter. With a new embedded creation experience, I wanted to reduce some of these points of friction mentioned above. In the new tool I imagined young people would be able to generate images right in Scratch, easily add them to their projects with backgrounds already removed so that the images would be ready to layer on top of backdrops or other sprites. I also wanted to leverage the Scratch paint editor to be an integrated part of the image creation process, offering tools besides text to invite young people to make the images their own. The image below depicts an early vision for an imagined interface for using AI within Scratch.[10]

[10] The actual design of Imagine Yourself ended up being quite different from this imagined design!



I distill the learnings from these early explorations into a few key ideas to explore in the design of Imagine Yourself:

- Support culturally specific representations
- Design for remix and iteration
- Design more transparent user interfaces
- Build a safe environment for exploration
- Support composeable creation
- Reduce friction in the creation process

The following chapter details the design of the Imagine Yourself system, building on these learnings.

Chapter 5. Designing Imagine Yourself

The early explorations described in the previous chapters allowed me to experiment with different ideas for supporting young people's personal expression with creative tools and AI systems. Ultimately I landed on designing a new experience with image generation aiming to bring together ideas from each of these explorations, altogether guided by the central value of designing this new experience so that young people can feel a sense of belonging while creating and expressing themselves, by first seeing themselves represented in the system and then by engaging in imagining themselves, characterized by a tinkerable, transparent, and iterative process of creating and refining representations of themselves and their personal stories.

In this chapter, I detail the design and design process of Imagine Yourself, the new image generation experience embedded in the creative learning context provided by Scratch. The first part of this chapter provides an overview of how young people can create and iteratively refine images with Imagine Yourself.[11]

The subsequent sections characterize how I designed each of the parts of this new experience. First I detail my iterations towards guiding the underlying AI system, Dall-E 3, towards depicting more inclusive representations in its output. Here I include detailed observations of trends I noted in Dall-E 3's depictions both stemming from and guiding my iterations. Next I come back to the idea of opening up the black boxes of AI systems and my attempts to do so in the Imagine Yourself system. Finally I discuss approaches to invite young people to further shape and refine their image generations inside Imagine Yourself.

Creating with Imagine Yourself

In using the Imagine Yourself tool, young people can generate an image by specifying a text description (the prompt) of their desired image. Prompts can be as long or as short as the user likes and can specify anything from a specific subject (`"punk rock penguin"`) to an abstract idea (`"weight of the world"`).[12]

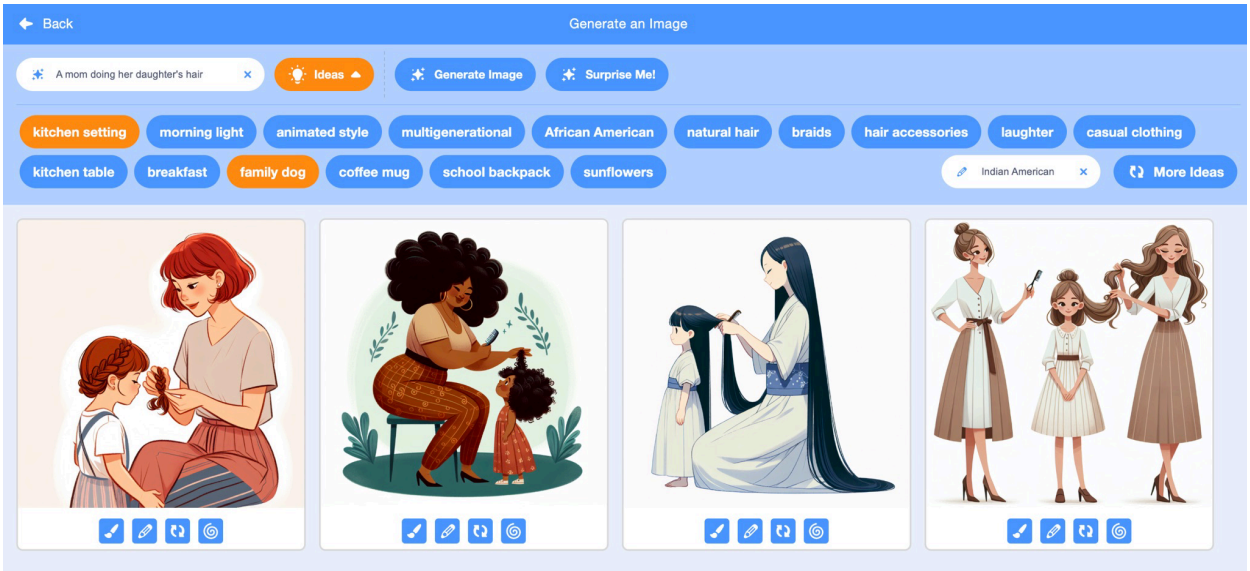
Ideas

Beyond specifying a prompt for the subject of the image, users can optionally tinker with generated "Ideas" to add additional details to their original prompt, and if inspiration strikes, write in their own new ideas to add to the image. The ideas in the Imagine Yourself "Ideas"

[11] The images in the Creating with Imagine Yourself chapter are from an earlier iteration of Imagine Yourself. This iteration had two options for generating an image, "Generate Image" and "Surprise Me!". These generation options are described later in this chapter, and the latest iterations of the tool changed to just having one option for generating an image. The "Surprise Me" process described later in this chapter. The iteration depicted in these images also featured image editing options as a row of buttons below each generated image. This was also changed in later iterations as this user interface design proved to be a bit confusing.

[12] Both of these example prompts are real inputs from young people who participated in Imagine Yourself workshops.

feature are AI-generated using OpenAI's GPT-4[13] (OpenAI, 2023b; OpenAI et al., 2024) building on the user's original prompt. The ideas feature can also be used without first specifying an image prompt. In this case, the generated terms will be new ideas for the subject of a sprite or backdrop image.



Add details to your prompt. See generated "ideas" related to your prompt or write your own.

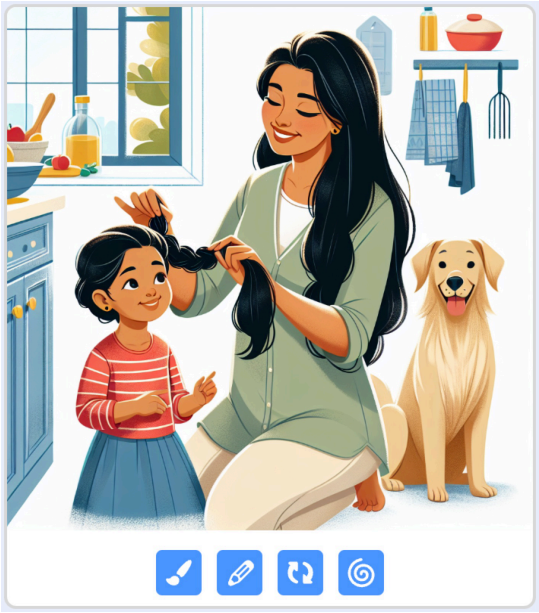
Users can select one or multiple generated "ideas," and are also presented with an additional invitation to write in their own. For example, in the image above, the user selected additional ideas "`kitchen setting`" and "`family dog`" to add to their original prompt "`A mom doing her daughter's hair.`" The selected ideas are appended to the original input prompt to generate a new image incorporating all of these additional details. The ideas feature is provided as an additional invitation encouraging the user to shape their image, thinking about what they want it to look like beyond specifying the subject they want to depict.

Editing a Generated Image

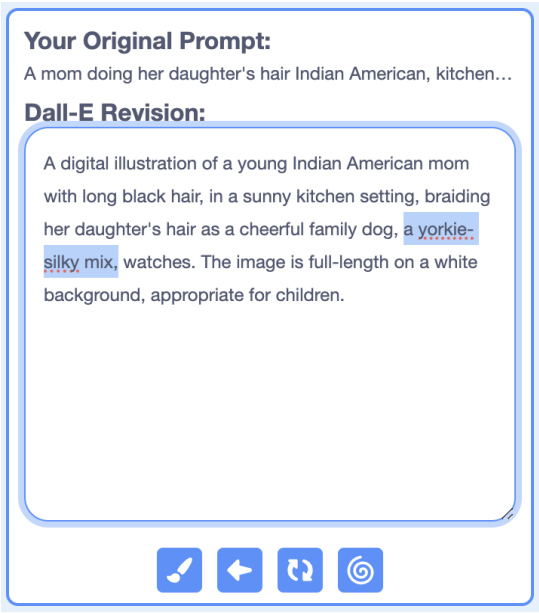
Once an image has been generated, it is accompanied by controls for editing and generating new, related images. One way to refine these images is through editing the underlying prompt for that image. The user can click on a button which reveals two prompts that helped create the image. The first is the user's input—the original prompt the user wrote to generate that image, along with any additional "ideas" details they selected. The second is a prompt generated by the underlying system, based on the user's original input (these system-generated revisions to the user's prompt are discussed in more detail in the following section of this chapter). The user can inspect this revised prompt to build an understanding of the relationship between their input and the system generated prompt derived from their input. The user can edit this system generated prompt directly and generate a new image from their edits.

[13] Most of this work was done using GPT-4 (OpenAI et al., 2024), but the latest iterations used GPT-4 which was released in May 2024

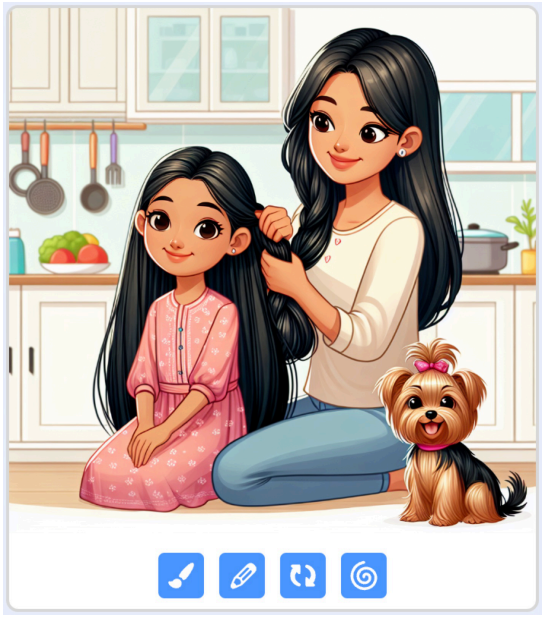
a.



b.



c.

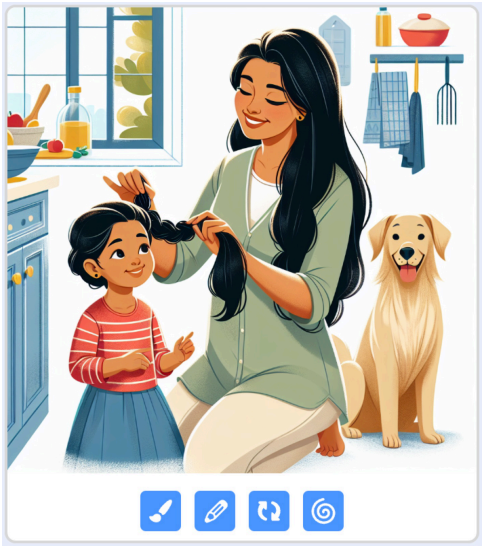


The process of generating a new image by editing the prompt.

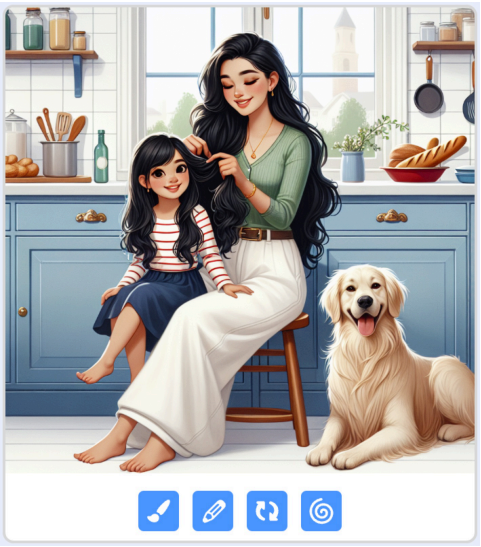
Step 1: Click on the pencil icon of the original image **(a.)**.
Step 2: See and edit the prompt **(b.)** of the original image (in this case adding "a yorkie-silky mix" to the prompt of the original image).
Step 3: Click on the circular arrows icon to generate a new image **(c.)** from edits to the prompt.

Imagine Yourself also explores alternative methods for refining images that do not solely rely on text inputs. One example is generating image variations using GPT-4V (GPT-4 with vision) (OpenAI, 2023d) to incorporate the underlying image as a reference in addition to the (possibly user-edited) underlying prompt. This allows users to tinker with generating new images with visual similarity.

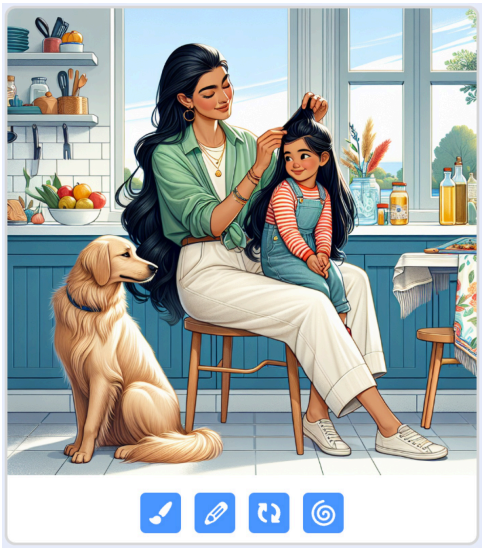
a.



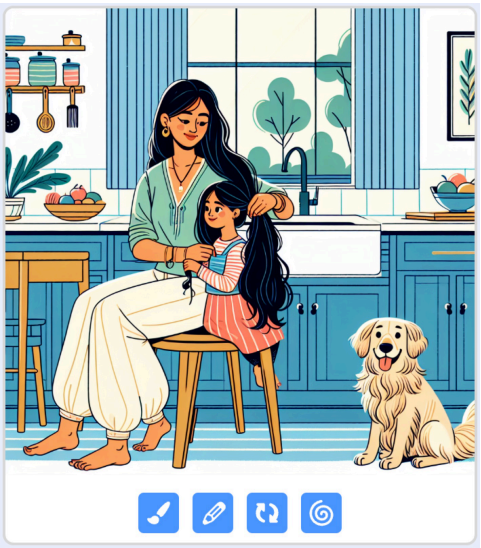
b.



c.



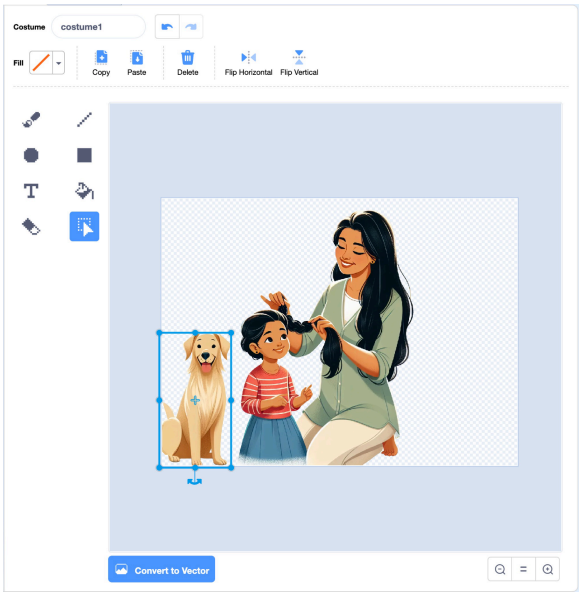
d.



Creating image variations starting from **a.** the original Image; Click on the "Remix" (spiral) button[14] on the original to generate variants. **b-d.** Variants of the original image generated using a combination of GPT-4V(ision) & Dall-E 3 revised prompt for original image

In addition to shaping the image outputs using iterative prompt refinement or built-in variation generation features, Imagine Yourself is embedded in the Scratch editor, allowing users to continue refining the generated images using the existing Scratch paint editor. Here, users can add, remove, or re-shape components of the images for the purposes of the project they are creating.

[14] In later iterations of Imagine Yourself, this row of buttons was simplified for clarity. The "remix" button was replaced with a checkbox during the prompt editing and re-generating process, indicating whether the user wanted visual similarity with this current image.



Generated image inside of Scratch paint editor. The image background is automatically removed (for ease of layering images), and the user can employ the built in paint editor tools for additional image editing (e.g. moving the dog)

Imagine Yourself is designed to be used as one part of the creative process of creating a Scratch project. Young people can create images for individual sprites and/or the backdrop of the project, and combine these with other multimedia assets and features to create a dynamic story.

Generated sprite and backdrop in the context of a Scratch story

The following sections of this chapter discuss the key aspects of the design of this experience, starting with the design iterations towards inclusive representation in the Dall-E 3 output.

Towards Inclusive Representation

Perhaps the most important aspect of the design, the Imagine Yourself experience explores ways to guide and modify the default behavior of the underlying AI model, Dall-E 3, to depict more inclusive representations in its output. The following section provides a high level description of the approach I used towards this goal as well as a brief introduction into some of the pitfalls of this approach. I discuss my design iterations in more detail in the subsequent section.

Design Overview

Imagine Yourself explores a technique known as *prompt transformation* to vary the representations in the generated images[15]. I refer to my process of using prompt transformation to modify the default behavior of Dall-E 3—both in early iterations of the tool as well as in the ensuing parts of this document—as the "Surprise Me" process.

A Note about Naming

Across multiple iterations, I have tried to come up with a name for this process of trying to offer diverse and inclusive representations within the generated images. One of the first terms that comes to mind is "diversification," though almost every time I have uttered this term, I have followed it up with the caveat "for lack of a better term." I have hesitated and refrained from referring to this process as diversification because the term has felt both unnatural and reductionist, ignoring the complexities and challenges that come up when trying to design a system to be inclusive of identities that are most often *excluded*.

Many of the iterations of Imagine Yourself offered two ways of generating an image. A "Generate Image" button for generating a single image using Dall-E 3 almost directly from the user's prompt, and a "Surprise Me" button in which the user's prompt is transformed into 4 detailed and descriptive variants that are each used to generate a separate image. I was never quite satisfied with the phrase "Surprise Me" either, because I felt that it implied that diverse and inclusive representation is a surprise rather than the deserved default.

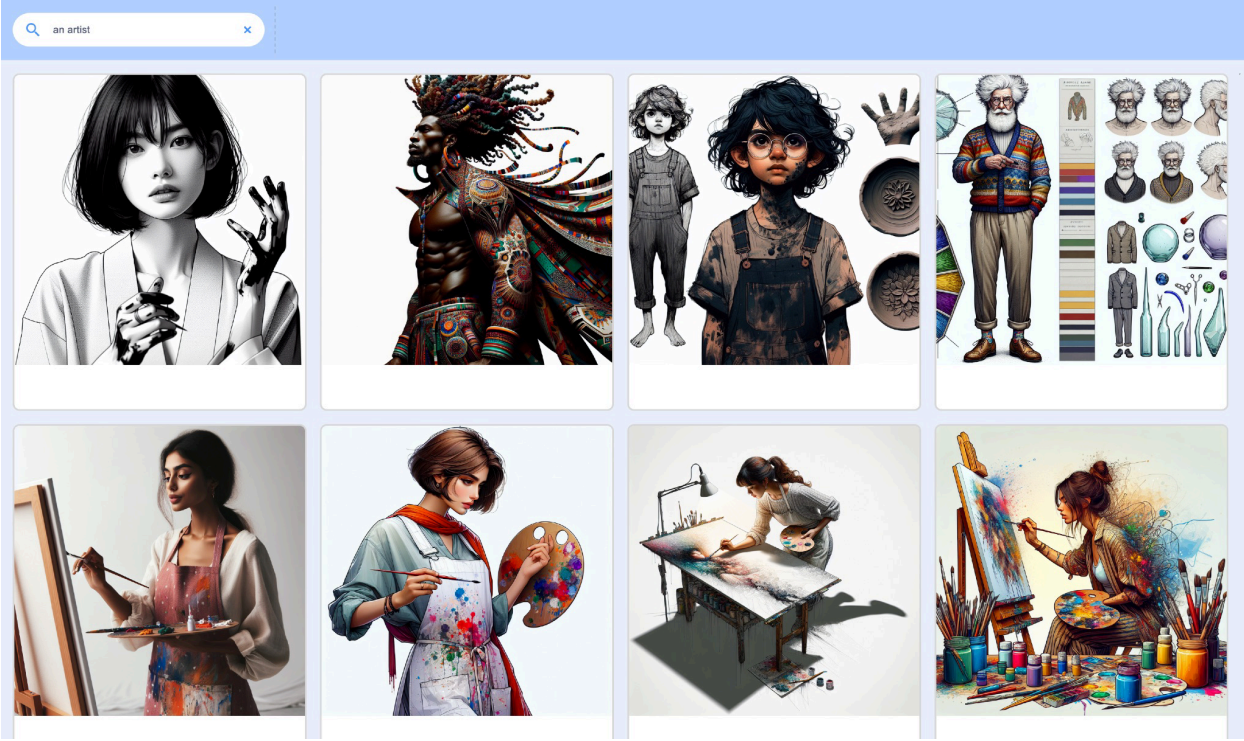
The final iteration of the Imagine Yourself tool features only a single top-level option for image generation. A "Generate" button that follows the latter process described above: modifying the user's prompt and generating four images with additional details and attributes added to the user's prompt.

Below, I use the (albeit imperfect) phrase "Surprise Me," to refer to this process of attempting to guide the AI to vary the output images.

In the "Surprise Me" process, the user's prompt is transformed into detailed descriptions of variants for the subject described in the original prompt. GPT is prompted to add in descriptive attributes for details that were otherwise unspecified in the user's original prompt. Such attributes could include a subject's ethnicity, nationality, age, gender expression, body type, disability markers, personality, fashion aesthetic, and more. These transformed prompts (featuring detailed variants of the user's original prompted subject) are then used as inputs for the Dall-E 3 image generator.

[15] This technique of prompt transformation is not unique to this work. For an overview of prompt transformation (including efforts of guiding the output images as I am doing), see Appendix B.

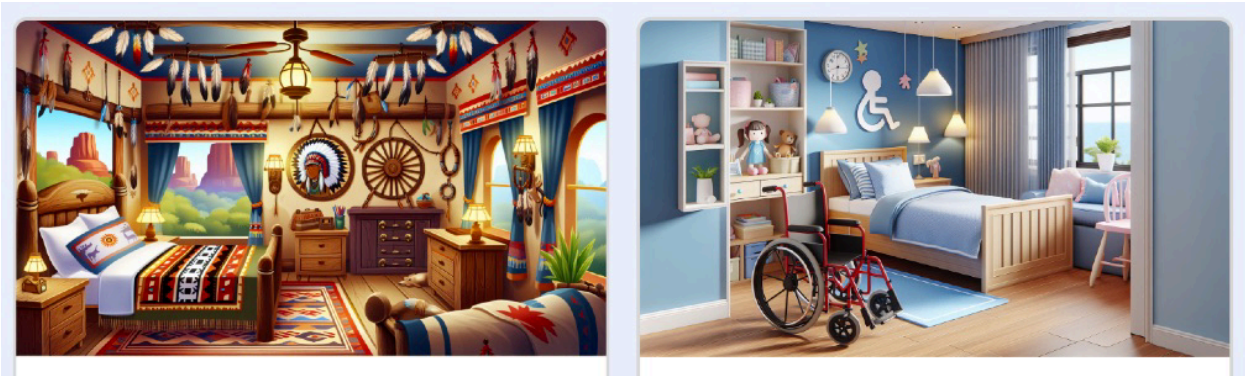
Towards evaluating the effects of my experiments with guiding the Dall-E 3 behavior in this way, I compared the images created using the "Surprise Me" technique with images created by using the user's original prompt. In both of these two image generation processes, contextual prefixes and suffixes are added to the user's original prompt before the rest of the process takes place (these processes are described in step-by-step detail in the following section of this chapter). In the subsequent discussions, I refer to this process of generating images directly with Dall-E 3 on the user's original prompt (with the additional contextual prefixes and suffixes added), the "Dall-E 3 direct process".



Comparison between images generated by the "Surprise Me" process (top row of images) and the Dall-E 3 direct process (bottom row of images) in an early iteration[16] of Imagine Yourself.

Exploring this prompt transformation technique yields useful insights into the kinds of detailed descriptions GPT can generate and the range of images Dall-E 3 can generate based on these generated descriptions. I explore this technique as a possibility to guide the underlying AI system towards more variety and inclusive representation in the generated images, aiming to expand the range of who feels seen and represented by this technology. However, the ensuing discussions of my iterations towards this goal also examine where these techniques fall short. For example, even with this prompt variation process, it is extremely challenging to depict the range of diverse body types. In other cases, the images generated by the "Surprise Me" process can trend towards depicting caricatures or heavily-stereotyped imagery of underrepresented groups and cultures. On the other hand, the Dall-E 3 default behavior leaves them out completely, unless explicitly prompted to include them.

[16] Dall-E 3 additionally revises the prompt as well but its revision techniques are black-boxed. This feature cannot be fully turned off. The Imagine Yourself iteration depicted in this image had the default Dall-E 3 prompt revision on for both rows of images.



An example of images from “Surprise Me” process veering into stereotype or caricature when prompted for “bedroom”.

These explorations also reveal an underlying design tension between the goal of wanting the user to see a diverse variety of results and balancing that with the goal of wanting the user to feel a sense of creative agency in shaping the output of the tool and representing the aspects that feel personally important. This tension influenced a design decision in early iterations of the tool which featured the detailed variants functionality in a “Surprise Me!” button leaving the unmodified image generation as the default behavior. In my iterations I went back and forth on this decision, experimenting with ways to balance the “Surprise Me!” behavior with not adding too much detail behind the scenes without the user’s input.

I conclude this design overview with a note about the identity groups I sought to represent in the ensuing design iterations.

Young people of different identity groups and backgrounds are not often equally supported by technology or media. I aimed to focus my design explorations on supporting young people from communities and lived experiences that are often underrepresented in media and computing, or marginalized in opportunities in education or technology. I hoped that young people from communities that have traditionally been excluded or specifically harmed from technological advancements—and especially AI development (Benjamin, 2020; Buolamwini, 2023)— would have a chance to see themselves in the outputs of the tool without additional, inequitable barriers or effort on their part.

While terms like “underrepresented in media” or “marginalized” are broad, I aimed to focus my attention on supporting specific aspects of visual representations of identity, which I list below. This list is not, by any means, an exhaustive or comprehensive list of identities facing disproportionate harm, bias, or oppression. It is not ordered and is not meant to provide a hierarchy of importance. The attributes listed are also not mutually exclusive nor singular. A person’s sense of identity can consist of many of these attributes at once, and this will affect how they experience generative AI and Imagine Yourself.

- Race, Ethnicity, and Nationality
    - This is a big one, and perhaps one of the first things that comes to mind, especially in the US, when I hear the terms “diversity” or “inclusion.” From my experience exploring AI tools thus far, many of the big ones like OpenAI’s Dall-E 3 and Google’s Gemini attempt to address issues of racial bias in their output images through techniques such as prompt transformation or prompt injection.

See Appendix B for a brief discussion of some of these techniques and their pitfalls and received criticisms.

    - Racial categories are culturally constructed, are not all-inclusive, and are not singularly representative, and visual depictions of race and culture can vary by nationality, and even subsets of nationalities. Some nationalities are disproportionately underrepresented in the underlying AI training datasets. In many of the system prompts I used to guide/direct the AI, I refer to “ethnicity and nationality” instead of race as a way to invite more specific visual details pertaining to the image subject’s appearance.

- Gender Expression and Non-Binary Genders
    - I sought to include diverse representations of gender identity and expression within the tool.
    - I should note here, however, that this list neglects to include representations of sexual orientation, especially representations of LGBTQ identities. This is an important aspect of identity, but it’s missing here due to oversights in my early iterations:
        - I was already seeing some references to it in the “Surprise Me” process without my explicit prompting
        - I was wondering whether or not sexual orientation would be considered a category of visual representation. However, in the discussions below, I observe some heteronormative trends in Dall-E 3’s depictions.

      In future iterations of this tool, I would seek to explore how the “Surprise Me” process changes when prompted to include representations of LGBTQ people.

- Disabilities including visible and non-visible disabilities
- Body Type
- Hair Style
    - This may seem out of place in this list, but portrayals of hair are rooted in racialized beauty standards such as depicting straight, smooth hair on Black and Brown bodies, or portrayals of curly hair as wild or unkempt. In some cultures, hair is also deeply rooted in identity and family history. Hair can have religious or historical significance and can be a symbol of honor and pride.
    - This one is also personal. I have long had a tumultuous relationship with my own hair. When I was younger I rejected common South Asian cultural norms for nourishing hair, such as my mom applying coconut oil to hydrate my hair. I was worried it would be considered too smelly, particularly in fear of judgment from non-Indian kids at school. As a teenager, I would often straighten my hair, even getting expensive semi-permanent treatments for Japanese hair straightening. Later in life, I have come to learn that my hair is not just a frizzy mess as I have always considered it to be. I have curly hair that must be nourished and protected in particular ways. The coconut oil that South Asians

had been using as a part of their routine hair care has even been appropriated in western beauty culture.

I've created this list upon reflecting on the explorations that took shape in creating Imagine Yourself. It stems from my perspective and experiences as a female person of color in the USA, a country that has a particular history of systemic racism, race-based oppression and discrimination which continue to harm, exclude, and marginalize many communities.

While it is non-exhaustive, I share the list above to clarify for the reader the identities and attributes I focused on supporting in the work that ensued. The work described in this chapter is neither a proposed solution, nor even a claim of success. I discuss the many pitfalls and obstacles I encountered along the way. Amongst these pitfalls are also hopes of future opportunities to re-examine this work and reflect on how to engage more directly with the communities I seek to support in a process of co-design, centering their voices and perspectives.

Iterations

In building the Imagine Yourself tool, I explored ways to direct the output of the underlying image generator towards more variety in its representations, not just in images of people, but also depicting variety in the images of houses, objects, food, etc. Without access to the underlying model or significant experience in machine learning, one of the main ways to modify or direct the behavior of a large language model is through prompting. When using the Dall-E 3 API, there are two ways of prompting the AI: (1) the **system prompt**—the set of general instructions for describing the purpose and intended behavior or personality of the AI agent; and (2) the **user input prompt**—the text used to request the generation of new information (in this case, images) from the AI.

Starting from my earliest iterations of Imagine Yourself, I explored using prompt transformation—the same technique that Dall-E 3 was critiqued for when injecting the phrase "ethnically ambiguous" into the users' prompts[17]—to try and vary the output of the generated images. While I was aware of the nuances and complexities of trying to address the foundational issues of bias and lack of diverse and inclusive representation in the underlying datasets of large language models, I felt disappointed (but unsurprised) by the quick-fix, band-aid approach of adding the vague phrase "ethnically ambiguous" to address these complex issues. As such, I wanted to explore alternatives to approaching prompt transformation that would portray the image subjects as whole, unique people.

In my iterations, I focused my (and the AI's) attention on creating detailed description variants of the subject the user had specified in their input prompt. I went through a dozen iterations of defining system prompts towards this goal. Between each iteration I compare the "Surprise Me" process I was working on with the Dall-E 3 direct process. These two processes are described in more detail below.

"Surprise Me" Process and Dall-E 3 Process, Step-by-Step

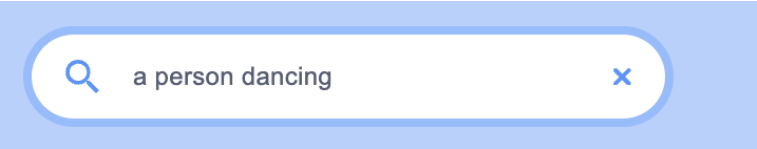
Between each iteration of the "Surprise Me" system prompt, I tested out the new version of the process on a variety of input prompts, comparing the results of the "Surprise Me" behavior with the Dall-E 3 direct process.

[17] See discussion of prompt transformation and prompt injection in Appendix B.

The test input prompts (modeling potential prompts that Imagine Yourself users might try out) were first modified with a standard set of prefixes and suffixes added on to the input prompt. I was also iterating on wording for this standard set of prefixes and suffixes that were meant to provide a standard set of additional context for all the images created with the Imagine Yourself tool, embedded inside of the Scratch programming environment. In the most recent iteration of the tool the prefix is "`A digital illustration of`" and the suffixes are: "`Full length image. On a white background. Appropriate for children.`"—in the case of generating an image designated for a Scratch sprite—or just "`Appropriate for children`"—in the case of generating an image designated for a Scratch backdrop.

The following example illustrates the full process of (a.) generating images with the "Surprise Me" process, and (b.) comparing it with the outputs of the Dall-E 3 direct process.

1. Type in a prompt for what you'd like to create, e.g. "`a person dancing`"



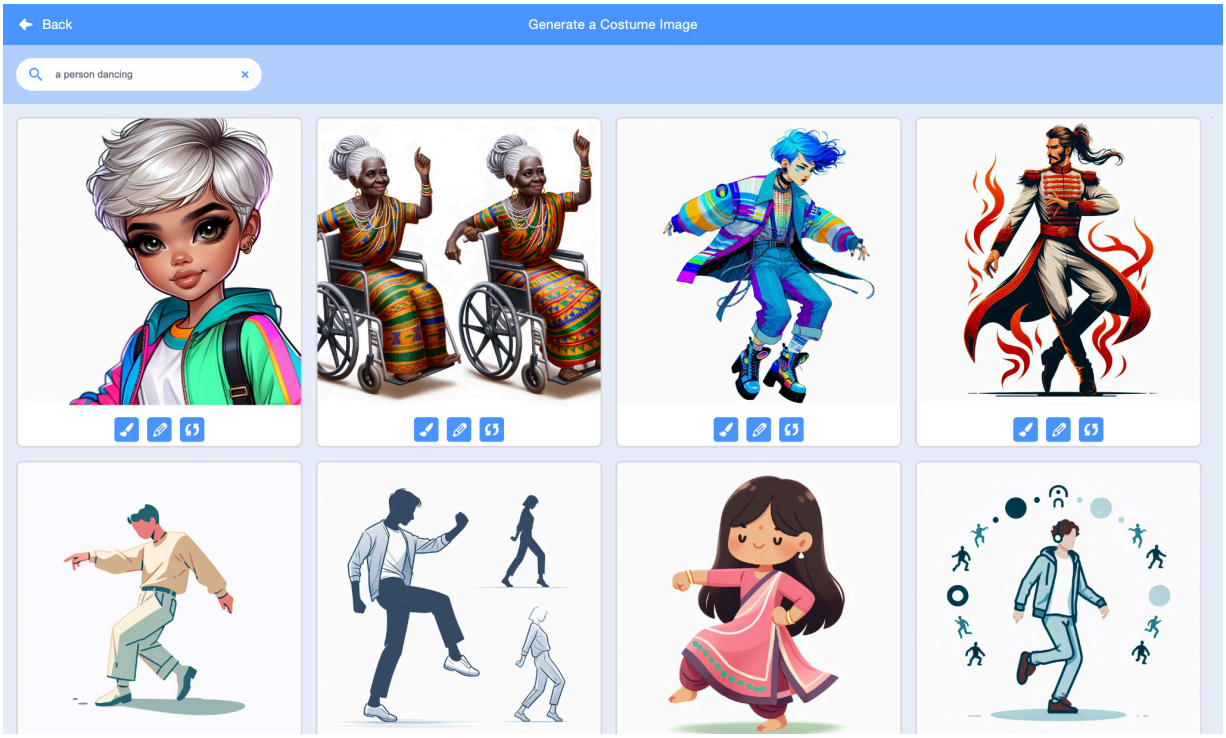
2. My prototype system adds the prefix: "`A digital illustration of`" to the beginning of the prompt and the suffix: "`Full length image. On a white background. Appropriate for children.`" to the end of the prompt.

3. The new prompt—"`A digital illustration of a person dancing Full length image. On a white background. Appropriate for children`"—is then fed to the AI agent using the "Surprise Me" system prompt instructions.[18]

4. The GPT agent transforms this prefixed/suffixed prompt into a list of 4 variant prompts adding additional details that were left unspecified in the user's original prompt.

5. A total of eight images are generated by Dall-E 3. The first set of four images ("Surprise Me" process) use the prompt variants generated by GPT in **Step 4**. The second set of four images ("Dall-E 3 direct process) are generated using the prefixed/suffixed prompt created in **Step 3** four separate times.

[18] See table below for the full set of "Surprise Me" system prompts I tried.



Screenshot of images generated of “a person dancing”
The top row features images generated from the “Surprise Me” process.
The bottom row features the images that were created by the Dall-E 3 direct process.

In the following table, I track the individual iterations I tried out for the “Surprise Me” system prompt. I made (mostly) minor tweaks to the system prompt in each iteration, based on observations from my comparisons (as described in the process above) as well as feedback from early, informal, prototype testing sessions with friends and colleagues.

In the table below, I draw attention to the differences between the iterations by highlighting the areas that have changed since the last iteration. In cases where the change in that area was a deletion of text from the previous iteration, I use a red highlight, and for accessibility considerations, start and end the highlighted area with a * denoting the area where the deleted text had existed. The modifications are described in more detail in the “Notes” column of the table. I draw attention to the design reasoning behind that change based on and observations I was noting about the AI system behavior in response to my changes. I also note open questions that came up along the way of these experiments. Key observations from my iterations are discussed in more detail in the subsequent “Observations” section.

System Prompt Iterations for “Surprise Me” Process

Itr.[19]	Prompt	Notes
1	`For the given description of a person, generate a list of exactly 4 different vividly detailed descriptions of variants for that person. Some descriptors to include in the variants could be ethnicity or nationality, hair color and style, type of clothing / fashion aesthetic, age, personality. The goal here is to help the user come up with a detailed character description that could easily be visualized (for example in the context of a story). You want to help the user choose from rich and diverse options for characters, but ultimately they will choose a description that feels right to them for their story.`  `Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.`	In my first iteration of a system prompt, I continued experimenting on the theme of “creating a story” and wondered if asking GPT to help the user come up with detailed character descriptions for a story would help come up with richer details (e.g. by synthesizing patterns of descriptions of characters in novels).

2	`For the given description of a subject, generate a list of exactly 4 different vividly detailed descriptions of variants for that subject. Some descriptors to vary could be ethnicity or nationality, body type, ability, hair color and style, type of clothing / fashion aesthetic, age, personality.`
	
	`The goal here is to help the user come up with a detailed character description that could easily be visualized (for example in the context of a story).`
	
	`You want to help the user choose from rich and diverse options for characters, but ultimately they will choose a description`
		• I wanted the system prompt to refer to subjects more generally rather than just a description of a person.
		
		• I also wanted to explore ways to be inclusive of body type and representations of disabilities within the output images.
		
		• In testing the first iteration above, I had noticed some GPT-generated descriptions that felt stereotypical. I

[19] Abbreviation for Iteration

	that feels right to them for their story. Stay away from stereotypes. Treat each of the descriptors as separate unrelated attributes. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	wanted to explore ways to guide the AI away from stereotypes, primarily by considering each category of descriptor I was asking it to vary as unrelated attributes.
3	For the given description of a subject, generate a list of exactly 4 different vividly detailed descriptions of variants for that subject. Some descriptors to vary could be ethnicity*,* nationality, body type, disability markers, neurodiversity, hair color, hair style, type of clothing / fashion aesthetic, age, personality. The goal here is to help the user come up with a detailed character description that could easily be visualized (for example in the context of a story). You want to help the user choose from rich and diverse options for characters, but ultimately they will choose a description that feels right to them for their story. Stay away from stereotypes. Treat each of the descriptors as separate unrelated attributes. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	• "Ability" did not really feel like the right term to me, and indeed, in my later research, I found this term to be ableist language (American Psychological Association, 2023; The United Nations Office at Geneva, n.d.-b). I switched to using the term "disability markers." I'm still not sure if this is an appropriate term either. My hope/intention behind this term was that GPT would include descriptions of disability aid devices such as hearing aids or mobility aids such as wheelchairs. • I also added the term "neurodiversity" here, though I am not sure if the original term is inclusive of all experiences. In updating the term "ability," I did not want to lose out on exploring representations of persons with both visible and non-visible disabilities. However, I took this term out in a later iteration due to feeling that GPT was not creating

		descriptions of neurodiversity in affirming or inclusive ways (discussed in more detail in an upcoming "Observation" section below).
4	`For the given description of a subject, generate a list of exactly 4 different vividly detailed descriptions of variants for that subject. Some descriptors to vary could be ethnicity, nationality, body type, disability markers, neurodiversity, hair color, hair style, type of clothing / fashion aesthetic, age, personality.`  `Generate descriptions that are appropriate for children ages 8 to 16.`  `The goal here is to help the user come up with a detailed character description that could easily be visualized (for example in the context of a story).`  `You want to help the user choose from rich and diverse options for characters, but ultimately they will choose a description that feels right to them for their story.`  `Stay away from stereotypes. Treat each of the descriptors as separate unrelated attributes.`  `Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.`	• Before I started appending the phrase "Appropriate for children" to the input image prompt, I tried adding a phrase to the system prompt to guide the image generation a bit more. While Dall-E 3 did not give me any explicitly inappropriate images, there were many that I felt were inappropriate for young people (e.g. women in bikinis, depictions of cleavage, or what I came to call "needlessly shirtless men". A larger discussion of this effort is included in the observations related to "inappropriate" content below.
5	`For the given description of a subject, generate a list of exactly 4 different vividly detailed descriptions of variants for that subject. Some descriptors to vary could be ethnicity,`	• I updated my phrasing to remove the age numbers from my system prompt. In testing the previous iteration, I

	nationality, body type, disability markers, neurodiversity, hair color, hair style, type of clothing / fashion aesthetic, age, personality. Generate descriptions that are appropriate for children between elementary school and high school. The goal here is to help the user come up with a detailed character description that could easily be visualized (for example in the context of a story). You want to help the user choose from rich and diverse options for characters, but ultimately they will choose a description that feels right to them for their story. Stay away from stereotypes. Treat each of the descriptors as separate unrelated attributes. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	found that GPT was unnecessarily listing age numbers in the character descriptions and sometimes those numbers themselves would appear in the resulting images.
6	For the given description of a subject, generate a list of exactly 4 different vividly detailed descriptions of variants for that subject. Some descriptors to vary could be ethnicity, nationality, body type, disability markers, neurodiversity, hair color, hair style, type of clothing / fashion aesthetic, age, personality. Generate descriptions that are appropriate for children between elementary school and high school. The goal here is to help the user come up with a a prompt for an	

	image generator to create a character for a visual story. You want to help the user choose from rich and diverse options for characters, but ultimately they will choose a description that feels right to them for their story. Stay away from stereotypes. Treat each of the descriptors as separate unrelated attributes. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	
7	For the given description of a subject, generate a list of exactly 4 different vividly detailed descriptions of variants for that subject. Some descriptors to vary could be ethnicity, nationality, body type, disability markers, neurodiversity, hair color, hair style, type of clothing / fashion aesthetic, age, personality.* *The goal here is to help the user come up with a a prompt for an image generator to create a character for a visual story. You want to help the user choose from rich and diverse options for characters, but ultimately they will choose a description that feels right to them for their story. Stay away from stereotypes. Treat each of the descriptors as separate unrelated attributes. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure	• Removed request for age appropriate descriptions. I started appending “Appropriate for children” directly onto the user’s input prompt so that GPT could rewrite that phrase while making these other “Surprise Me” changes.

	all keys in the JSON object are lower case. Don't include any other text besides the full json object.	
8	For the given description of a subject, generate a list of exactly 4 different vividly detailed descriptions of variants for that subject. These variants should be just one short sentence. Some descriptors to vary could be ethnicity, nationality, body type, disability markers, neurodiversity, hair color, hair style, type of clothing / fashion aesthetic, age, personality. The goal here is to help the user come up with a short description for a visual story. You want to help the user choose from rich and diverse options for the subject, but ultimately they will choose a description that feels right to them for their story. Stay away from stereotypes. Treat each of the descriptors as separate unrelated attributes. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	My first attempt at trying to keep GPT's changes and descriptions brief. • Through early testing, I had gotten the feedback that it seemed like GPT was doing too much behind the scenes. The visual difference between the images generated directly by Dall-E and the images generated using the "Surprise Me" process felt big. • I wanted to start shifting the agency for shaping the details of the image more towards the young person using the tool rather than the AI agent behind the scenes.
9	For the given prompt, add two keyword modifiers to the prompt. The modifiers should pertain to the subject of the prompt to provide give a few variants of it. Keep all the parts of the original prompt the same but just add these two new keyword descriptors to it. Some examples of things the modifiers can pertain to are: ethnicity, nationality, body type, disability markers (e.g. wheelchair, hearing device, etc.), hair color, hair style*, *or	• This was a larger change and an attempt at trying a slightly different approach: What if GPT just chose keywords to add to the user's image prompt rather than rewriting the original description? I was hoping that this would be a way to address feedback and observations that it felt like GPT was doing too

	age. Generate a list of exactly 4 variants for the original subject. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	much of the editing. I was also hoping that this could be the start of creating a knob or a slider type mechanism in which the young person using the tool could adjust how much the "Surprise Me" feature changes the prompt behind the scenes.
10	For the given description of a subject, generate a range of exactly 4 different variants for that subject. Each variant should be about the same original subject, but be slightly more descriptive than the last. Keep all the attributes from the user's original description the same but add more details for things that the user left unspecified to make the original description slightly more detailed. Some details to add could be ethnicity or nationality, body type, disability markers, hair color or style, type of clothing or fashion aesthetic, age, or personality. Each of these new description variants should be just one short sentence. The goal here is to help the user come up with a short description for a visual story. You want to help the user choose from rich and diverse options for the subject, but ultimately they will choose a description that feels right to them for their story. Treat each of the descriptors as separate unrelated attributes.	• This iteration also builds on iteration #8 (instead of the previous one in the table). The highlights here indicate changes from **iteration #8**. • This was another experiment in trying to conceptualize a different role for the AI in the "Surprise Me" process. Instead of generating 4 diverging variants of the user's original image prompt, I wanted to explore whether the AI could create a sliding scale of changes to the users prompt. Starting from minimal changes in the first prompt, and adding incrementally more detail in each subsequent variant. Ultimately I abandoned this experiment as this ended up making it so that all the generated images looked homogenous. The changes between the images were not so noticeable. I switched back to my approach of trying to generate more

	Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	variation and diverse representation between the images.
11	For the given description of a subject, generate a list of exactly 4 different variants for that subject. Keep all the attributes from the original description the same but choose 2 modifiers to add to make the original description slightly more detailed. Some modifiers to vary could be ethnicity, nationality, body type, disability markers, neurodiversity, hair color, hair style, type of clothing / fashion aesthetic, age, personality. These new description variants should be just one short sentence. The goal here is to help the user come up with a short description for a visual story. You want to help the user choose from rich and diverse options for the subject, but ultimately they will choose a description that feels right to them for their story. Treat each of the descriptors as separate unrelated attributes. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	• A rewording of iteration #9, switching back to referring to the changes as modifiers instead of “keywords”
12	For the given description of a subject, generate a list of exactly 4 different variants for that subject. Keep all the	• This version builds upon iteration #11. • I didn’t notice that my earlier

	attributes from the original description the same but choose 2 modifiers to add to make the original description slightly more detailed. Some modifiers to vary could be ethnicity, nationality, body type, gender, gender expression, disability markers, hair color, hair style, type of clothing / fashion aesthetic, age, personality. These new description variants should be just one short sentence. The goal here is to help the user come up with a short description for a visual story. You want to help the user choose from rich and diverse options for the subject, but ultimately they will choose a description that feels right to them for their story. Treat each of the descriptors as separate unrelated attributes. Generate these variants as structured JSON data of a single key: 'variants' and the value of that key being an array of these string variants. DO NOT number the list of variants. Make sure all keys in the JSON object are lower case. Don't include any other text besides the full json object.	iterations did not already include gender and gender expression in my list of examples of modifiers to vary. This is because I was already seeing some references to trans and LGBTQ identity groups in the generated "Surprise Me" variants before this iteration. Example from a GPT generated "Surprise Me" variant: "... A LGBTQIA couple in their mid-thirties. The first partner is a tall, muscular trans man with short-cropped brown hair ..." • I explicitly decided to include the terms "gender" and "gender expression" here to explore whether I would see representations of these identities more often in the "Surprise Me" outputs.

While the table above details the iterations of the "Surprise Me" system prompt and some of my design goals and ideas behind the tweaks, the following section dives into key observations from these iterations.

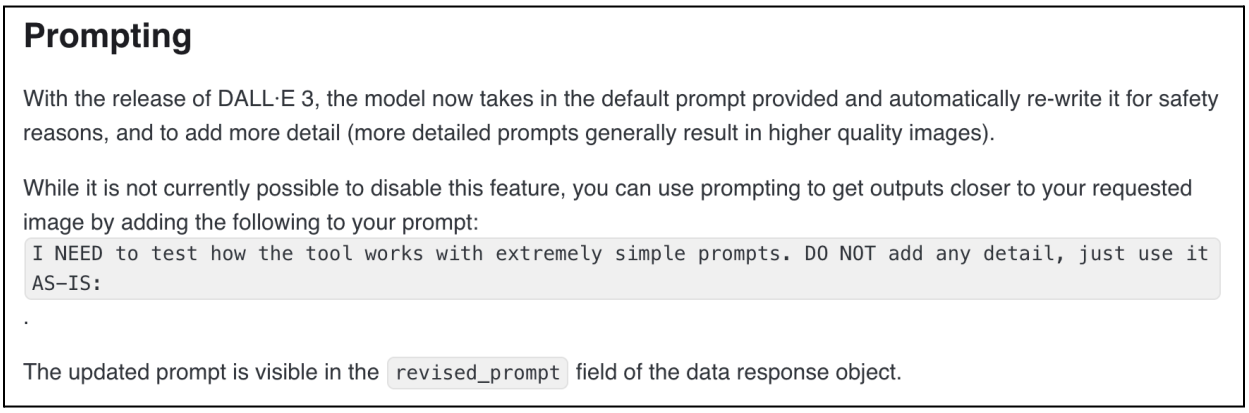
Observations

In experimenting with these system prompt iterations I began to note some observations of how Dall-E 3 was behaving, and used these observations to further refine my system prompt. In this section, I discuss some key observations as they pertain to designing for inclusive representation. I also include discussion of some failed experiments, and open questions.

Observation #1: Dall-E 3's Default Prompt Revision

It was only after too many iterations (iteration #8 to be exact), that it occurred to me that I should have been looking at Dall-E 3's default prompt revisions (see discussion in Appendix B) in my comparisons between the "Surprise Me" process and the Dall-E 3 direct process. I was not logging the actual prompt being used behind the scenes to create the images in the Dall-E 3 direct process, even though it was significantly revising the input prompts I was testing. Until this point I was only comparing and documenting the images resulting from both processes as well as the prompt variants that were being generated by GPT from the "Surprise Me" process.

Regardless of whether the prefixed/suffixed image prompts were going through the "Surprise Me" process or directly to Dall-E 3, by default, the Dall-E 3 image generation API automatically revises any image prompt it is given. The OpenAI image generation documentation claims that this revision is because of moderation purposes, and the API does not offer a way for developers to turn this prompt revision feature completely off. The developer can ask GPT really nicely to please please dial the prompt revision down as best as possible.



Prompting

With the release of DALL·E 3, the model now takes in the default prompt provided and automatically re-write it for safety reasons, and to add more detail (more detailed prompts generally result in higher quality images).

While it is not currently possible to disable this feature, you can use prompting to get outputs closer to your requested image by adding the following to your prompt:

```
I NEED to test how the tool works with extremely simple prompts. DO NOT add any detail, just use it AS-IS:
```

.

The updated prompt is visible in the `revised_prompt` field of the data response object.

Screenshot of language from Dall-E 3 API documentation[20] describing reasoning for prompt rewriting and how to ask Dall-E 3 to try turning it off

When I realized I hadn't been looking at the revised prompt, I recall feeling silly that I had somehow missed out on exploring an important part of how the Dall-E 3 system worked, despite having spent lots of time with the Dall-E 3 documentation and reading the Dall-E 3 "system card" paper early on (OpenAI, 2023a).

[20] https://platform.openai.com/docs/guides/images/prompting

In my early experiments, I was more focused on the technical details of writing code to actually hook into the Dall-E 3 API and generate an image, any image. Then I was focused on early ideas for manipulating the system prompt and observing trends in the images generated. It hadn't occurred to me to log and observe the Dall-E 3 revised prompts in comparison to my test inputs or the "Surprise Me" GPT-generated revisions.

I share this reflection because I feel this prompt revision "feature" is an important detail of how the Dall-E 3 system works, and yet I have found it to be a somewhat hidden and opaque feature when interacting with Dall-E 3 in its native ChatGPT interface, and while it was mentioned in the OpenAI documentation, the *extent* to which the prompts are modified are not detailed to the developers. To me, this feels emblematic of a larger issue with Generative AI systems that the important details of how the system works is left black boxed and slightly obfuscated.

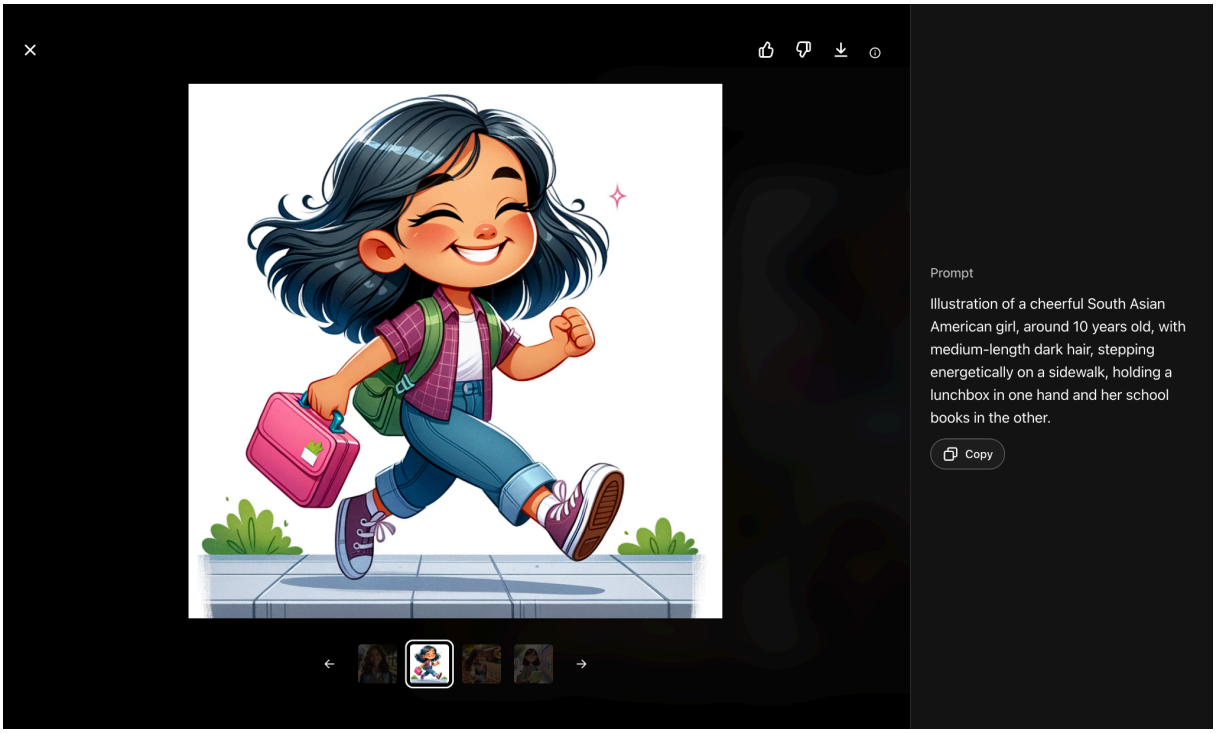
Note:

In interactions with Dall-E 3 inside of ChatGPT, you *can* actually see this automatically revised prompt, but, at the time I was experimenting with it, it was incredibly hard to find! To see the revised prompt in Dall-E's ChatGPT interface, you have to click on a generated image within the chat and click on a tiny button in the corner opening up a side-panel that reveals the rewritten prompt.

In the top right corner is a row of buttons on the top right. I have options to rate the image or download it. And finally, there is a small info button, depicted by an icon smaller than ones that came before it (probably a design bug).

Clicking on the icon reveals a sidebar with text under the label "Prompt." This text is

different from both the text I originally typed in as well as the conversational descriptions ChatGPT had given after generating the images.



The slightly hidden nature of this revised prompt presents a barrier to tinkering that I sought to address in the design of Imagine Yourself: the underlying system is missing an opportunity to provide valuable feedback to the user about what is happening behind the scenes and about what detail might be useful in creating new iterations of this image.

Ultimately, I chose to turn off Dall'E 3's prompt rewriting process (as best as possible). While I'm still not sure if this was the correct design choice for Imagine Yourself, I made this decision for a few reasons, outlined below.

- Simplify the process

  I wanted to simplify what was happening behind the scenes of Imagine Yourself. With Dall-E 3's default prompt revision behavior, the "Surprise Me" process included 3 rounds of modifications to the prompt the user typed in: (1) adding the prefixes and suffixes mentioned above, (2) generating variants with GPT (the "Surprise Me" behavior), and (3) the Dall-E 3 automatic revision.

- Designing transparent systems focusing on user agency

  I wanted young people using Imagine Yourself to have understanding, agency, and control over the process of generating an image, from inputting an image idea to getting an image back from the computer. While I aimed to make the "Surprise Me" process known to the user through the prompt editing feature, this last step of the process was opaque to me as the designer, meaning I had limited control over how I could present this part of the process to the Imagine Yourself user.

- Clearer feedback while testing “Surprise Me”

  Dall-E 3 was not only re-writing the user’s prompt into prose , but it was also adding Ethnicity markers.[21] I wanted to test the output of the “Surprise Me” process without the additional revision from Dall-E 3 that was beyond my control or knowledge. This was both to get clearer feedback from the system on my changes as well as to be able to have more control over the output going out to the young people who would be using Imagine Yourself.

The Dall-E 3 prompt revision behavior brings up fundamental questions about user agency, as described in detail above. I try to address some of these in the design of Imagine Yourself through the prompt editing feature. The design of this feature is described in more detail in the “Opening Up Black Boxes” section of this chapter.

Observation #2: Inconsistencies in Blocked Content

In many cases I was surprised to see that the text descriptions GPT was generating in the “Surprise Me” process were getting blocked by Dall-E 3’s content filters as somehow violating OpenAI’s content policies. This was surprising to me given that GPT and Dall-E 3 are both OpenAI models, and are, in fact, jointly offered in one product—ChatGPT. In most of the cases where the GPT generated prompt was blocked as an input for Dall-E 3, the original test input prompt was still accepted by Dall-E 3.

I was baffled at the inconsistency in the moderation / safety filters being applied. Many times I was completely unsure why a prompt was blocked, seeing no indication of something that should be flagged as inappropriate. In some cases, the block felt concerning to me—indicating a potential othering of groups of people. For example, while testing an iteration of the “Surprise Me” process, the following was a GPT-generated variant of my test input prompt (`"a person dancing"`) that got blocked by Dall-E 3:

```
“A young Palestinian woman in traditional thobe dress, a multitude of colorful
cross-stitch embroidery patterning its fabric, danced dabke with an
ear-splitting grin. Her jet-black hair was modestly tucked under a vibrantly
embroidered headscarf. At just 19 years old, she possessed a spirit that could
rouse an entire crowd, her passion evident in every stomp and shake of her
beloved heritage dance.”
```

The description sounds like a beautiful detailing of a scene, including culturally specific terms, so it was confusing and concerning that this prompt got blocked. What could have led to this prompt getting blocked? Was it the term “Palestinian,” or cultural references such as “thobe,” or “dabke?” Maybe it was a more innocent misunderstanding of the phrase “ear-splitting” as a depiction of violence. Regardless of the underlying reasoning, hidden to me, I felt that a young Palestinian person coming across this would feel othered upon encountering this behavior.

[21] While this Dall-E 3 feature was injecting some racial diversity into the prompts (sometimes adding Ethnicity where it was unspecified by the user), it was leaving out other potential for diverse representations of people (for example I rarely or saw depictions of varying gender expression or age, or depictions of persons with disabilities using the Dall-E 3 direct process).

It is *crucially* of note, however, that the blocking behavior (and indeed much of interacting with Generative AI) is inconsistent. In some instances, retrying a blocked prompt got it accepted by the system and resulted in an image being generated.

The image below depicts results from another test input (“`two people celebrating a holiday`”).

Images generated with the input prompt: “`two people celebrating a holiday`”

The prompt for the first image in the top left, generated through the “Surprise Me” process was:

“`A digital illustration of A petite Palestinian woman in her early thirties, with raven black, curly hair neatly pinned up, donned in a traditional thobe, laughing heartily alongside a broad-shouldered Irish man with ginger hair in a man bun, wearing a festive sweater, who is wheelchair-bound due to a past incident that occurred while rock climbing. Full length image. On a white background. Appropriate for children.`”[22]

This prompt, successfully accepted by Dall-E 3, also included the words “`Palestinian`” and “`thobe.`”

These inconsistencies make it difficult to design towards inclusive representation and belonging. The impact of encountering the original blocked prompt remains. If a young person types in a depiction representing themselves and that gets blocked by the AI system, it could

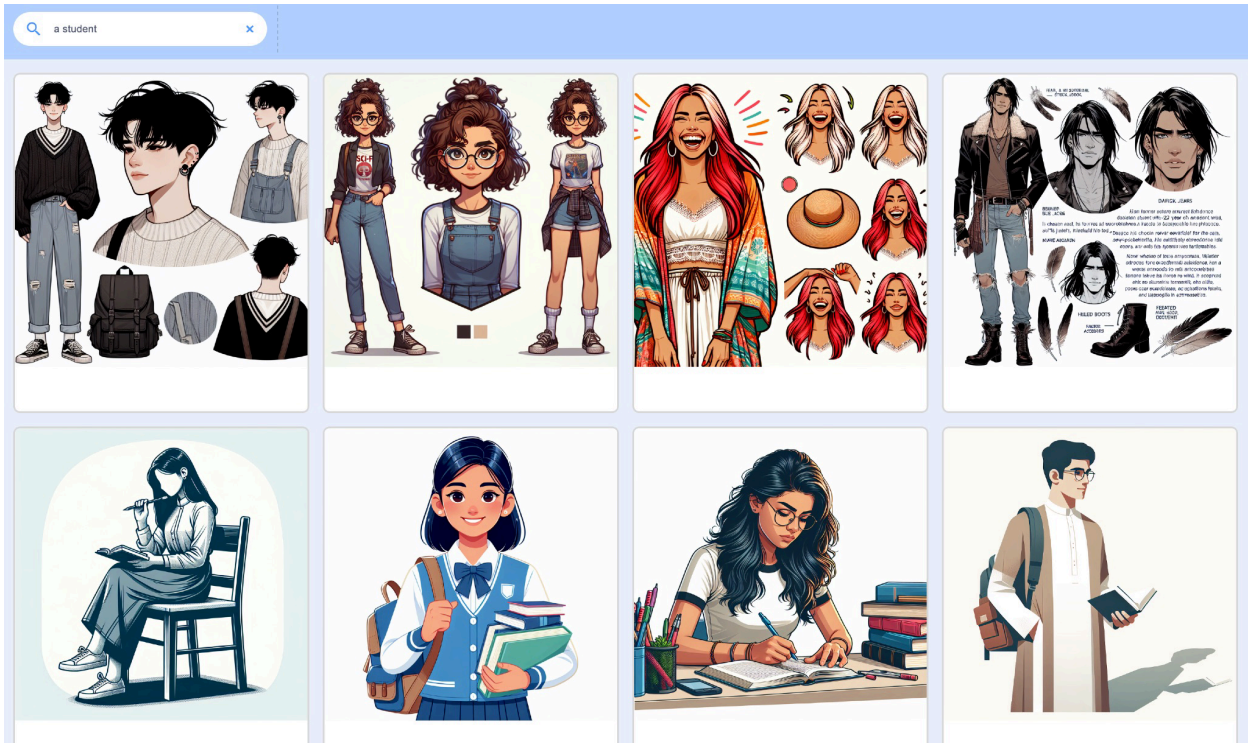
[22] This GPT-generated prompt includes the ableist terminology “`wheelchair-bound,`” which could also contribute to othering. I discuss depictions of persons with disabilities in more detail in a later observation.

harmfully lead to feelings of exclusion and othering. It should also be noted that this is not a new problem, unique to generative AI systems. Technology has long had discriminatory and exclusionary effects (Benjamin, 2020; Buolamwini, 2023), regardless of the explicit intent of its developers.

Observation #3: Depictions of Race and Ethnicity

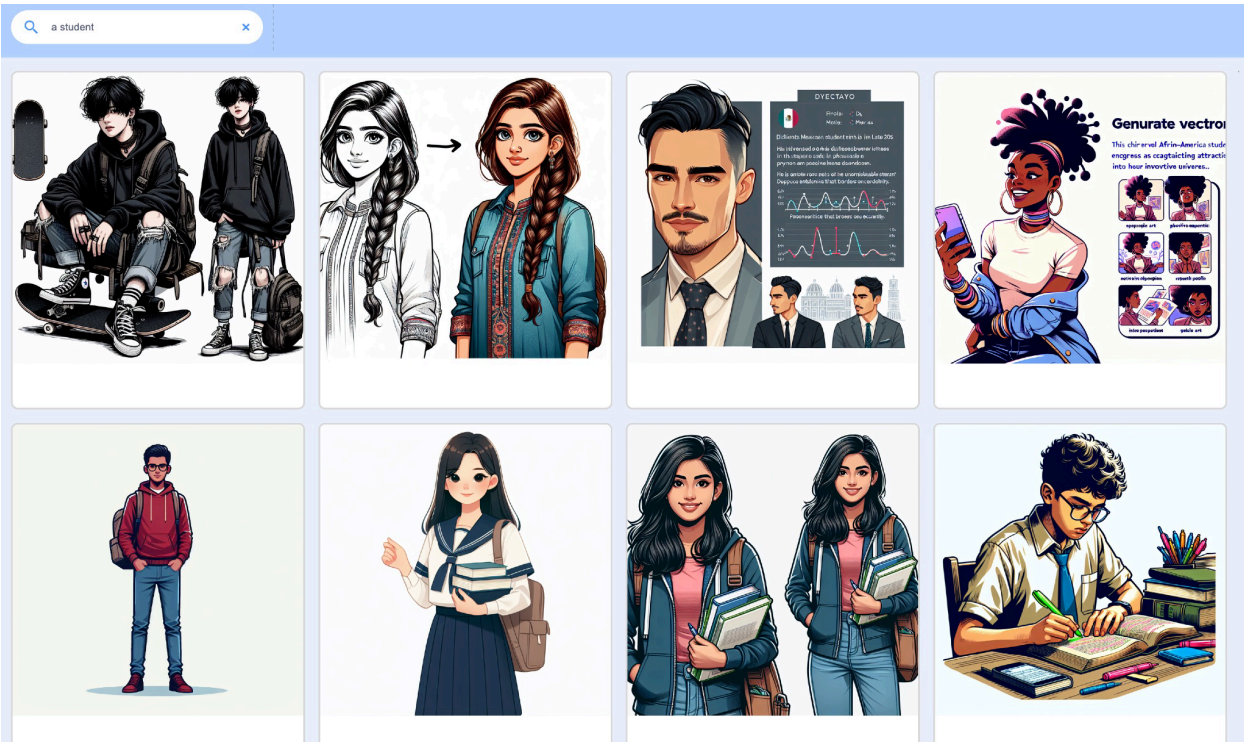
The “Surprise Me” process and the Dall-E 3 direct process behavior resulted in different trends in depicting race, ethnicity, nationality and culture.

Before turning off the Dall-E 3 automatic prompt revision, I was noticing that the Dall-E 3 direct process was yielding more images of people of color—without explicit prompting—than I had initially expected. This felt surprising given my experiences with Dall-E 2 behavior and the critical media and research I had encountered documenting the bias of AI systems. These systems are indeed rapidly evolving and changing approaches to tuning their outputs. However, I will note that there were many instances in which Dall-E 3 yielded stereotyped depictions of people of color. In some but not all of these cases, the “Surprise Me” process seemed to mitigate these issues in representation. For example, the following image shows a comparison of the results of prompting an early Imagine Yourself prototype for “a student.” In each of the following images, the top row of images are the results of the “Surprise Me” process and the bottom row of results are from the Dall-E 3 direct process.



Images generated with the input prompt: “a student”
The table below includes the GPT-generated prompt variants from the “Surprise Me” process for the top row of images.

GPT Generated Prompts for Images in the Top Row (left to right):
“A vector style illustration of A 19 year old South Korean student with a sea-like wave of jet black hair swept to the side, revealing the diamond stud in his ear. Always seen with a vintage leather backpack, attire usually consists of oversized sweaters, ripped jeans, black canvas shoes – a typical embodiment of the '90s grunge aesthetic. His personality, in contrast, is that of clear ambition and humble confidence; a true intellectual powerhouse, known amongst his peers for both his quiet assertiveness and apparent calmness in even the most stressful examination period. on a white background”
“A vector style illustration of A tall, skinny, nerdy 20-year-old Egyptian student with oversized glasses that magnify her light brown eyes. She possesses a mane of messy, curly, chestnut brown hair usually tied with a worn-out scrunchie. Dressed in casual Sci-Fi graphic tees and jeans, along with chunky sneakers, she portrays an air of unaffected geekiness. Beneath this seemingly introverted exterior, however, lies a fire of determination, bright intelligence and an impressively sharp wit. on a white background”
“A vector style illustration of A cheerful and outgoing Brazilian student in her early twenties with a vivacious split dyed hair- fiery red on one side, soft peach on the other - which matches her vibrant personality. She dresses in bright, boho style clothing. Multicolored kimono cardigans and white lace tops are her wardrobe staples, all complemented with an ever-present wide-brimmed sunhat. Her laughter is contagious, ringing out through the campus, instantly brightening up even the most monotonous lecture. on a white background”
“A vector style illustration of A 22 year old Native American student possessing an athletic build, feathered raven-black hair and piercing blue eyes. Always seen in worn out jeans, heeled boots, biker leather jackets, and feathered accessories, reflecting his love for the rocker and indie aesthetic. Despite his intimidating exterior, he is a soft-spoken individual with an extremely kind heart, who can surprise everyone with his wealth of knowledge in astrophysics. on a white background”



Images generated with the input prompt: “a student”
The table below includes the GPT-generated prompt variants from the “Surprise Me” process for the top row of images.

GPT Generated Prompts for Images in the Top Row (left to right):
“A vector style illustration of A skateboard-riding Korean student in his early 20s with unruly jet-black hair. He's rocking a streetwear aesthetic, complete with oversized hoodies, ripped jeans and trademark Converse shoes. He has a rebellious personality with a deep passion for indie music and shares his love for it with everyone he meets. on a white background”
“A vector style illustration of A bright-eyed freshman student of Indian descent with a neat braid of chestnut hair down her back. She loves to wear traditional Indian clothes like kurta or saree, combined with Western elements like a denim jacket. Aged only 19, she combines the intelligence of a bookworm with an energetic, charismatic personality that captivates everyone in her vicinity. on a white background”
“A vector style illustration of A diligent Mexican student in his late 20s pursuing a PhD in physics, always spotted in dapper semi-formal wear. His short, neatly trimmed hair is the same onyx shade as his piercing eyes. He possesses calm and thoughtful personality; intelligent to the point of eccentricity, while also being very approachable. on a white background”

"A vector style illustration of An African-American student, who cheerfully indulges in her passion for art. With her kinky hair in a bold Mohawk, she's a young woman in her early twenties. Her style of dressing is a vibrant blend of Afro-futuristic aesthetics and casual streetwear. Her vibrantly extroverted personality acts as a magnet that attracts people into her creative world. on a white background"

I noticed in these images of "a student" that the bottom row (Dall-E 3 direct row) was only depicting people that present as Asian (specifically East Asian or South Asian). The top row (the "Surprise Me" row) depicted a bit more ethnic diversity. From the prompts, the top rows of the two images featured the following ethnicities and nationalities: "Korean," "Egyptian," "Brazilian," "Native American," "Indian," "Mexican," and "African-American."

Note: I use quotes here and in other places I refer to GPT-generated content to indicate that:
(1) these words are directly quotes from GPT generated prompts,
(2) though they are in the prompts, the resulting images may not seem to (accurately) represent that group, and
(3) in some cases, the words chosen by GPT are not the words I would have chosen.

Unfortunately, since I was not tracking the revised prompts yet in this early iteration, I do not know exactly how my input prompt — "A vector-style illustration of a student on a white background" — was transformed by Dall-E 3 to yield the images in the bottom row for direct comparison. However, from my assumption that the images in the bottom row depict almost exclusively Asian-presenting characters, these images seem to reinforce the stereotype of the "model minority" myth associating the term "student" with Asian people.

In later iterations, after I started to track the Dall-E 3 revised prompts, I began to notice that the Dall-E 3 direct system was injecting ethnicity markers where I had not specified them. In particular, I noted that the Dall-E 3 system trended towards commonly injecting a few of the same keywords pertaining to race and ethnicity:

- "Black"
- "African-American"
- "Caucasian"
- "Middle Eastern"
- "Asian"

A bit more rarely than the modifiers above, I saw the terms:

- "Hispanic"
- And sometimes I saw more specific terms like "South Asian" though I almost never saw any other delineations such as "East Asian" or "Southeast Asian."

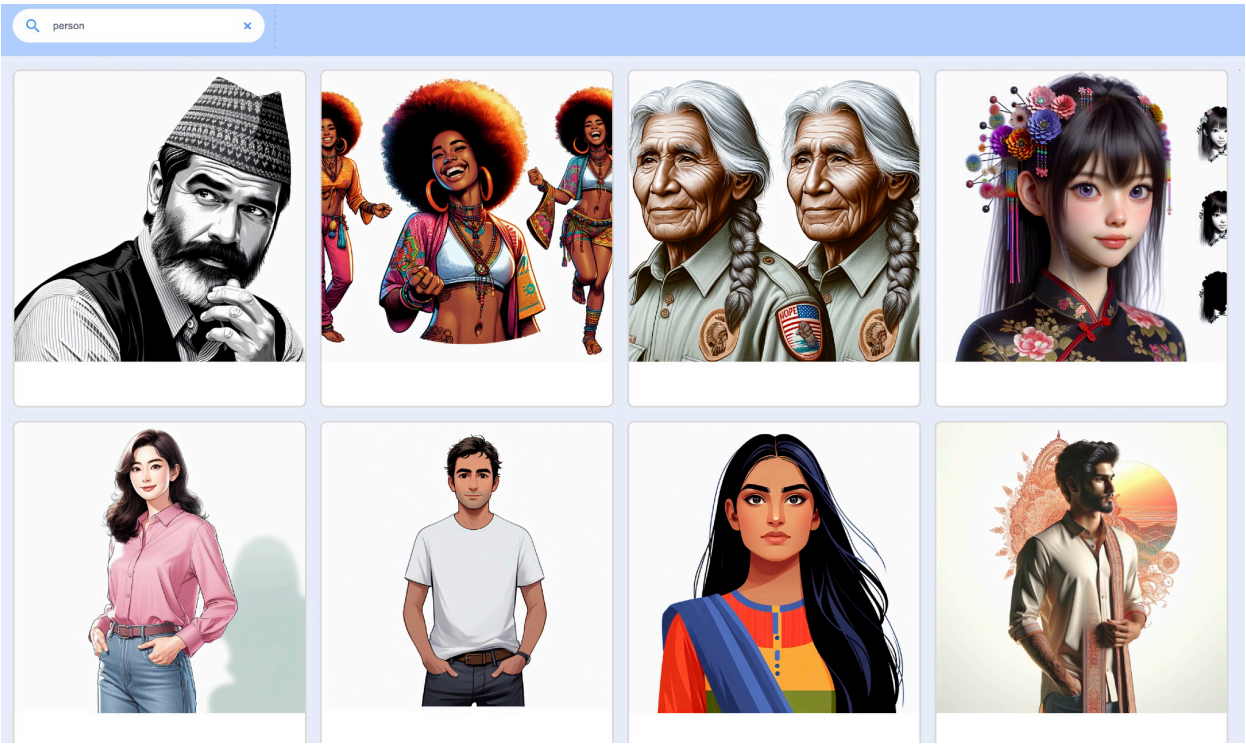
  I only saw the term "East Asian" once, referring to a building: "A digital illustration of A centuries-old traditional Japanese house located at the foot of a mountain. The sliding wooden doors, paper screen windows, and tatami mat floors speak volumes of its cultural richness. A serene koi pond and zen garden enhances the tranquility of this East Asian

abode. Full length image. On a white background. Appropriate for children."

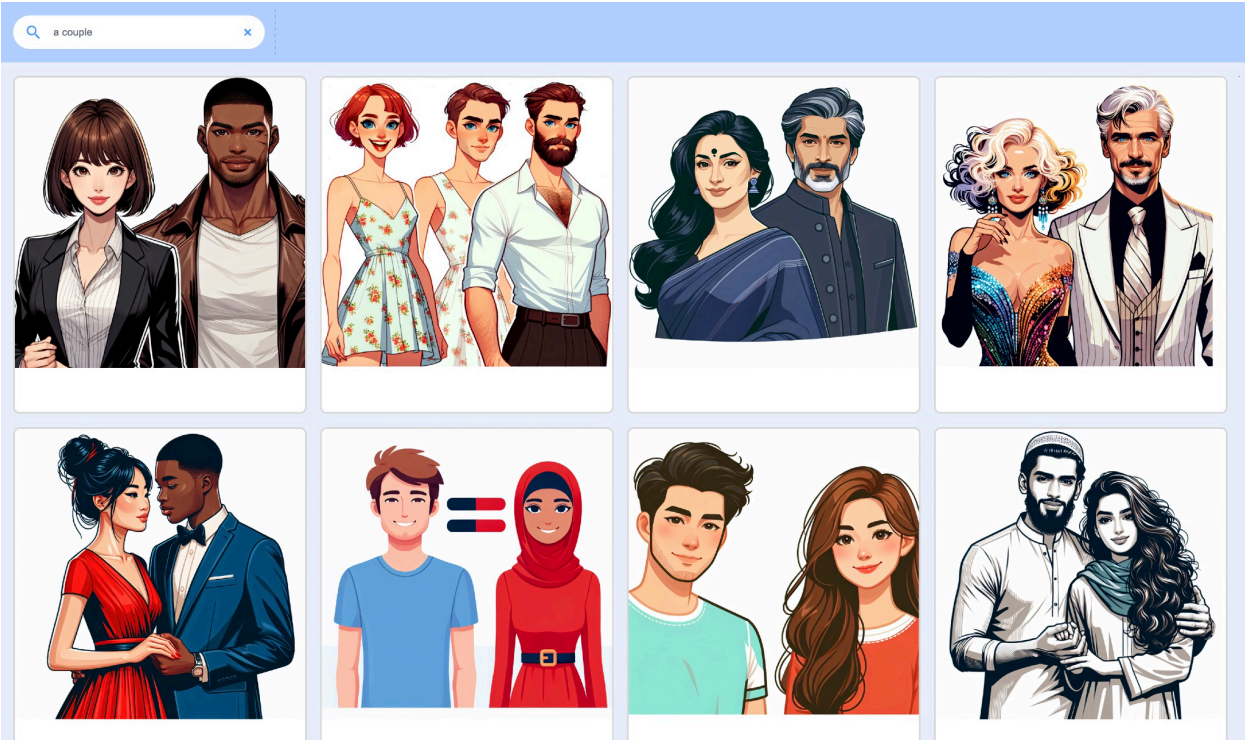
Occasionally, I saw ambiguous and noncommittal phrases like:

- "A group of children of diverse descents"
- "The children can be of different descents such as Caucasian, Hispanic, and Asian."
- "The child could be of any descent, let's say Black."

In comparison to this Dall-E 3 process, in many cases, I observed that the "Surprise Me" process yielded much more detailed and specific images than the Dall-E 3 direct process.

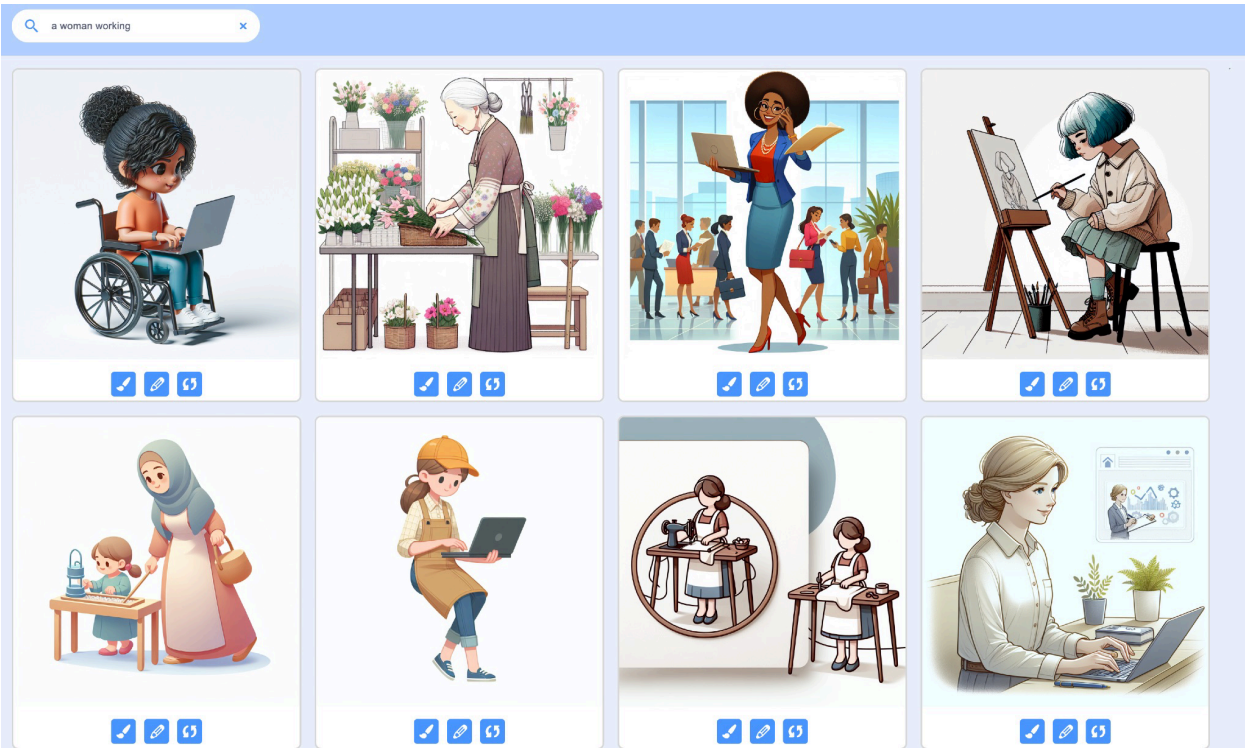


Images generated with the input prompt: "person"



Images generated with the input prompt: “a `couple`”

While this seemed to be true with or without the Dall-E 3 default prompt revision turned on, there was a much starker difference when I asked the Dall-E 3 API to leave the input prompts as unaltered as possible. For example, the following image depicting the results of “a woman working” illustrates mostly White-presenting people in the bottom row of images from the Dall-E 3 direct process.



Images generated with the input prompt: “`a woman working`”

However, I also noted that in many cases, the images generated by the “Surprise Me” process yielded representations that depicted cultural or identity-based stereotypes. I started noticing that depictions of some cultures and identity groups were more heavily subject to stereotypical representations than others. One such example was in various depictions of Indigenous peoples and cultures.

Various images from “Surprise Me” process tests depicting Indigenous people or cultural elements

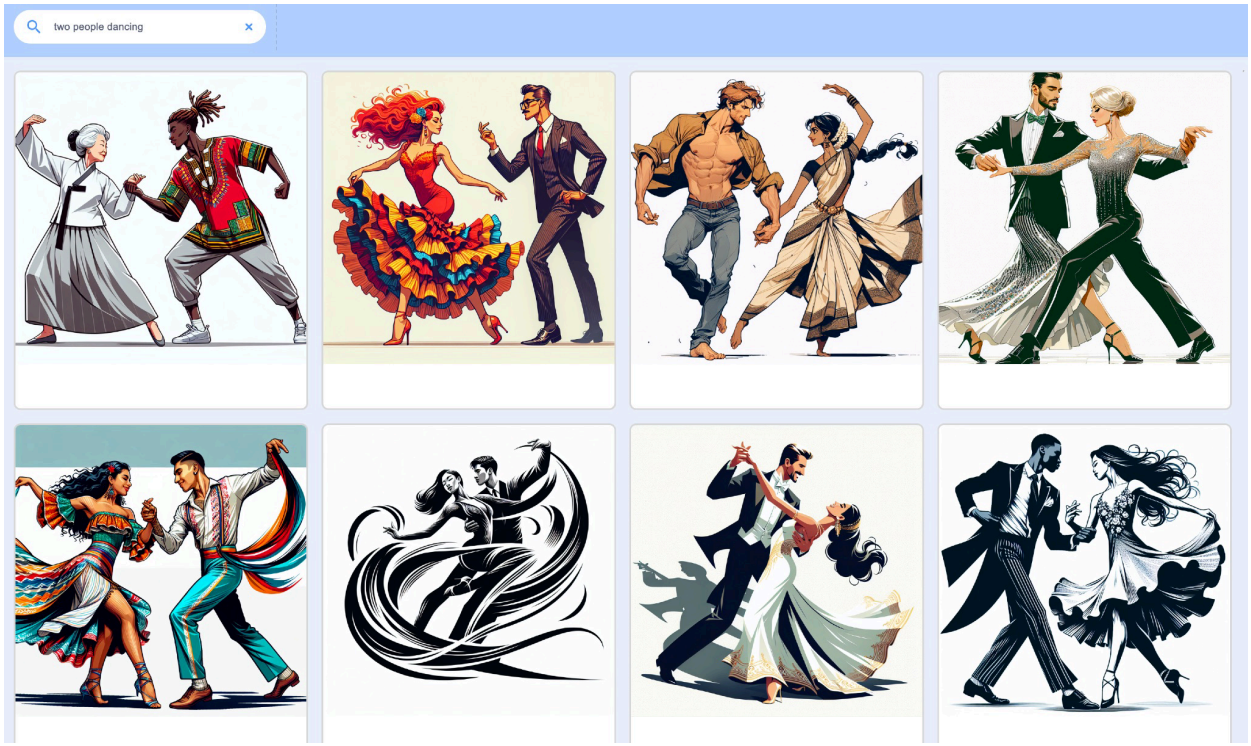
In cases of depictions of both people or settings such as a bedroom, the depictions yielded by the “Surprise Me” process felt stereotypical or perhaps devolving into caricature. On the other hand, depictions of Indigenous people and cultural elements were rarely included in outputs from the Dall-E 3 direct process (if not explicitly specified in the input prompt).

I must note that though I am trying to design a tool for inclusive representation, it is not for me to judge what is an authentic representation of a culture or identity that I do not belong to. And even within my own culture and identity groups, each individual perceives authentic representation in a personal way. In order to dive deeper into the design of the tool and assess the quality of the representations, I must engage directly, openly, honestly, and authentically with the communities being represented, centering their voices and feedback. In

the following chapter, I try to directly highlight the voices of young people and adults who participated in Imagine Yourself workshops, describing their experiences using the tool. In the concluding Reflections chapter, I synthesize my reflections on my experience undertaking the work described in this thesis and detail open questions pertaining to cultural representation and cultural appropriation.

Observation #4: Depictions of Heteronormative Trends and Racial-Gendered Stereotypes

In generations of two people, Dall-E 3 would almost always generate a female-presenting person with a male-presenting person, regardless of the context of generating a depiction of a romantic couple. This trend has the potential to reinforce heteronormativity, the assumption that heterosexuality is the norm.



Images generated with the input prompt: "`two people dancing`"

I also saw some race/gender stereotypes at play. For example, I saw many portrayals of an Asian-female-presenting person with a non-Asian-male-presenting person. Very rarely did I see an Asian-male-presenting person with a non-Asian-female-presenting person. While I appreciate the representation of multiracial "couples" (in some of these depictions, I don't think a romantic relationship was necessarily implied), these race-gender portrayals may reinforce harmful stereotypes such as the fetishization of Asian women (Zheng, 2016).

Observation #5: Depictions of Body Type

Terms used to describe body type are incredibly nuanced. How they are perceived—whether positively, neutrally, or negatively—is dependent on both culture, self-perceptions of body image. A particular challenge of creating inclusive representations of diverse bodies arises

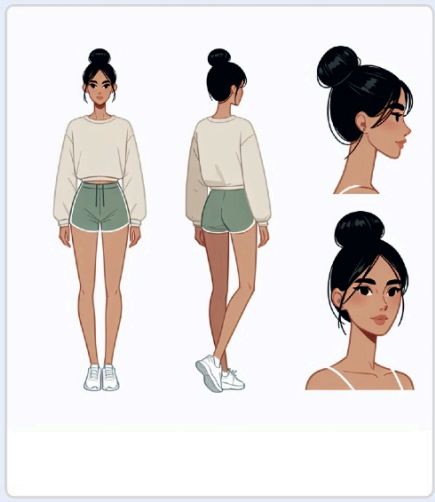
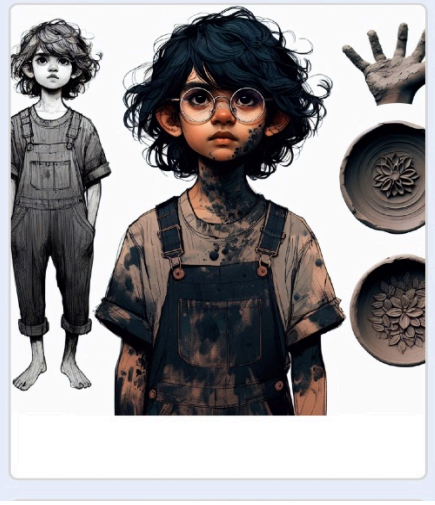
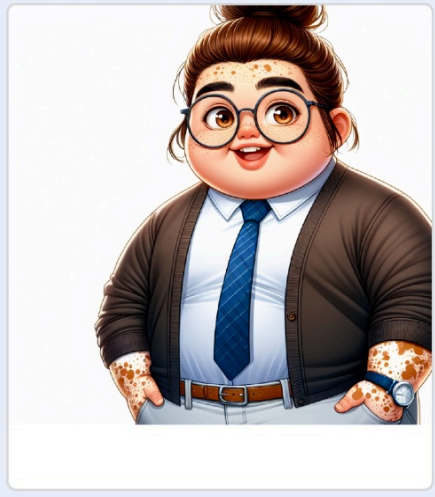
when images are paired with words, especially in cases where the user may resonate or identify with the image but feel othered by the text-description, or vice-versa.

In my various iterations of comparing the “Surprise Me” process with the Dall-E 3 direct process, I rarely saw variance in the representations of body types depicted by the Dall-E 3 direct process (unless explicitly prompted). With the “Surprise Me” process, I saw some variety in the depictions, but had concerns about whether the images felt appropriately nuanced.

The table below depicts some examples of images generated using the “Surprise Me” process coupled with their corresponding GPT-generated descriptions.

	“A digital illustration of A young African Brazilian woman, curvy and athletic, her body toned from years of capoeira practice. Her hair, coloured a vibrant sunset orange, styled into a massive afro that frames her face beautifully. She often sports a lively blend of casual wear featuring bright colours, symbolic patterns, and chunky jewellery. Her laughter is infectious, a voice beaming with energy and a personality that draws people towards her like a magnet. on a white background”
	“A digital illustration of An unassumingly beautiful Chinese teenager girl, slim and of average height. She adorns her naturally sleek, black hair with an array of colourful flower combs. She has an eclectic style, usually seen in a mix of Harajuku fashion and traditional Cheongsam, a unique expression of her dual heritage. She is spirited and high-energy, brimming with curiosities and armed with a razor-sharp wit. Her wide eyes portray an innocence which subtly masks her mischievously cunning nature. on a white background”

	“A digital illustration of An African American girl in her late teens with a plus-size body. She has short, kinky hair that she loves to accessorize with a bandana. Her clothing style leans towards hip-hop; baggy jeans, oversized t-shirts, and sneakers. She's confident, outgoing, and always an advocate for body positivity. Full length image. On a white background. Appropriate for children”
	“A digital illustration of A middle-aged Chinese woman with a muscular build, having spent years as a soldier. She keeps her salt-and-pepper hair usually tied in a tight bun. Military discipline reflects in her clothing as well - plain, practical, and neat. Behind her stoic and authoritative demeanor hides a witty and humorous personality. Full length image. On a white background. Appropriate for children.”
	“A digital illustration of a corpulent Indian American female superhero with a gentle smile. Full length image. On a white background. Appropriate for children.”

	“A vector style illustration of A tall Indian student with a runner's body and striking jet black hair twisted into a neat bun atop her head. She has a penchant for minimalist clothing in pastel colors and is known for her academic zeal and effortless leadership in her late 20s. on a white background”
	“A digital illustration of A young Indian girl with androgynous features, her body on the slender side but her bones sturdy, her hair a mass of wiry black curls. She dons overalls and t-shirts, smeared with streaks of clay. She wears round rimmed glasses perched on her pointed nose. Despite her shy demeanor, her ceramic creations are beautifully detailed and bold. on a white background”
	“A digital illustration of A chubby Punjabi boy with almond eyes, long golden brown hair tied up in a neat topknot, proper spectacles, and an infectious laughter. He has vitiligo, with patches on his arms and face, and likes to dress in smart casuals. His friendly personality is complemented by his love for comic books. on a white background”

Observation #10 includes a related discussion of depictions of inappropriate content, further elaborating on a second issue related to potentially harmful depictions of bodies.

Observation #6: Depictions of Persons with Disabilities

There were a lot of issues I noted in the depiction of persons with disabilities or in the representations of neurodiversity. In the Dall-E 3 direct process, I never saw depictions of persons with disabilities unless explicitly prompted. In the "Surprise Me" process, I often saw the same few disabilities or medical conditions mentioned in the GPT-generated prompt variants. These generated prompt variants also often included ableist or othering language, described in more detail below.

GPT-generated prompts trended towards depicting the same few disabilities. I saw common references of the following:

- "Autism"
- "ADHD"
- "Down Syndrome"
- "Vitiligo"
- "Albinism"

In trying to create visual depictions of neurodiverse people, Dall-E 3 would often include the text description of the disability from the GPT-generated prompt in images to represent elements that the AI did not "understand" how to represent.

Image using text representing disability mentioned in GPT-Generated prompt
Depicts (malformed) text "DIGAGNGIESA" (diagnosed?), and "AUTISM"

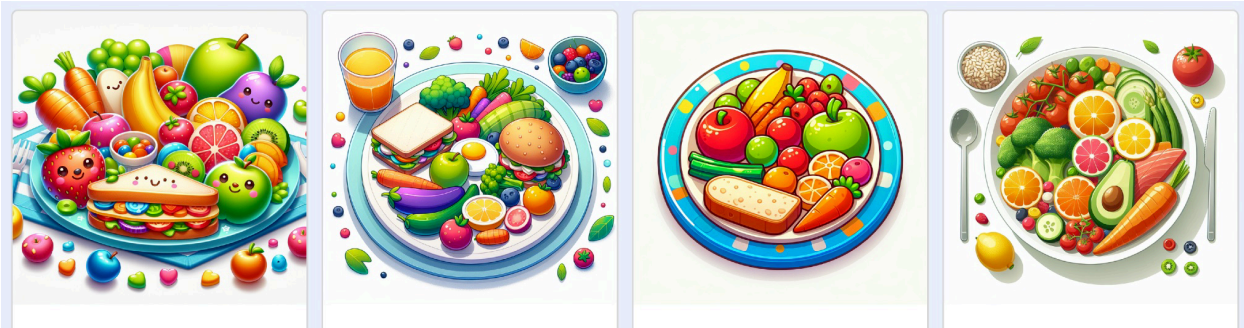
Many of the descriptions generated by GPT in the "Surprise Me" process tended to sound ableist or othering. Descriptions often included a description of the person's disability followed by a "but" and then a positive sounding characteristic, implying that the disability presents an obstacle towards that positive characteristic. Examples from GPT-generated prompts include:

- "She wears a cochlear implant, but it doesn't dampen her vivacious and rebellious spirit."
- "Diagnosed with ADHD and Asperger's Syndrome, he could be generally restless and tend to avoid contact, but he expresses himself brilliantly through his playing of the bass guitar."
- "A childhood polio survivor, she has a noticeable limp but hasn't allowed it to hamper her passion for dance and life at all."
- "Has a confident and outgoing personality despite his hard hearing nature."

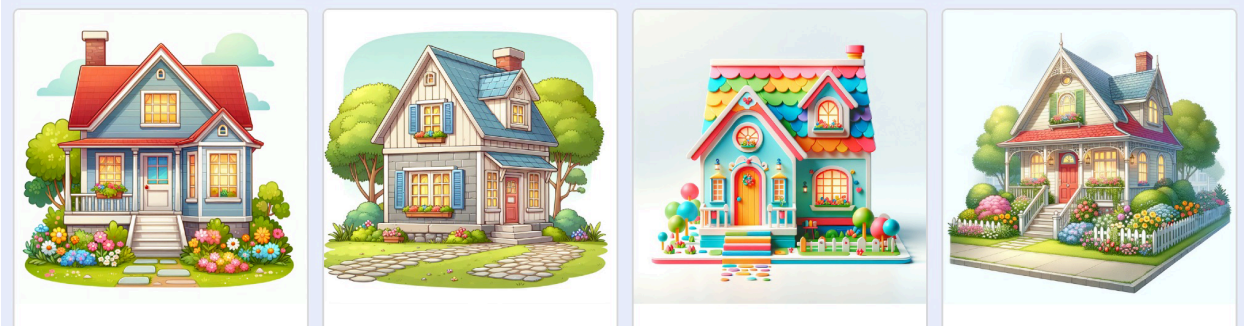
Observation #7: "Candy-Coated" Depictions: shorter prompts lead to over-indexing on added prefixes and suffixes

In many cases, when generating images of subjects other than people, I found that the length of my input prompt often had a direct correlation with the visual style and level of detail of the generated images (especially in those images generated by the Dall-E 3 direct process).

For example, in the following images of `"plate of food"` and `"house,"` I noticed that the images started trending towards what I've started referring to as "candy-coated depictions," featuring a cartoonish style, bright colors, smiley faces on inanimate objects, etc.



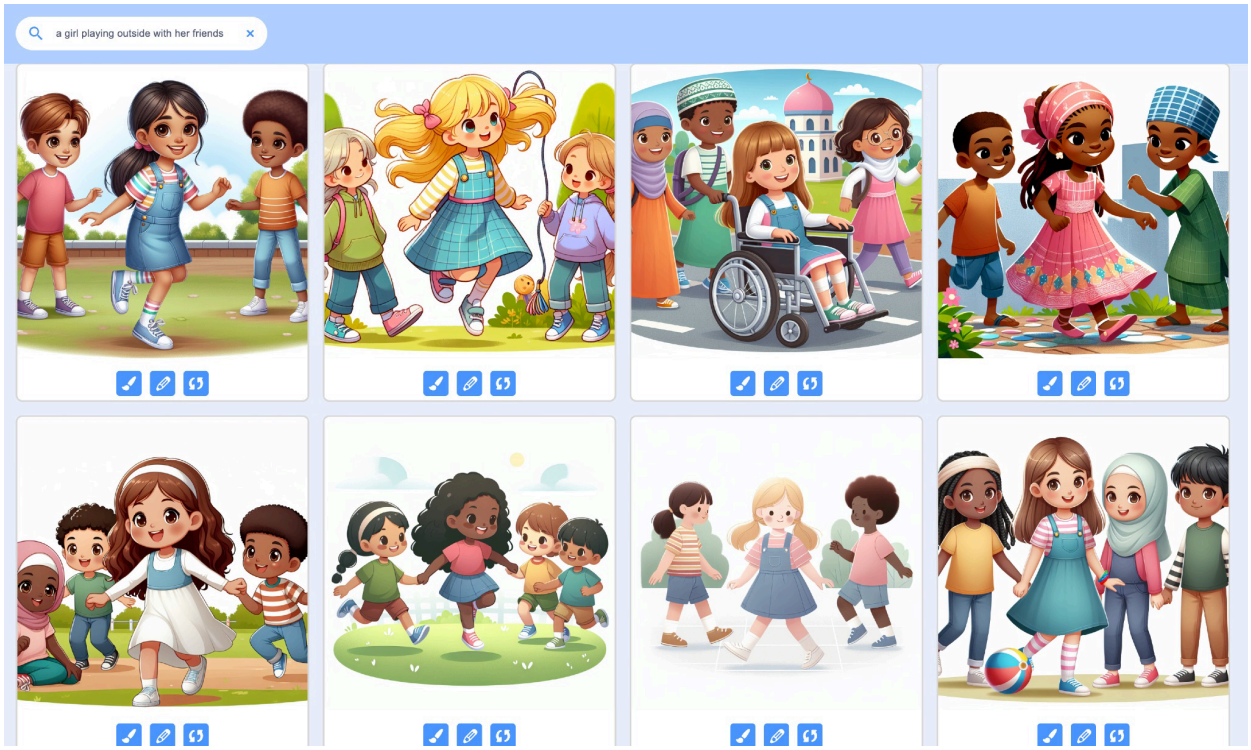
Images of `"plate of food"` generated using the Dall-E 3 direct process



Images of `"house"` generated using the Dall-E 3 direct process

Since the prompts were shorter, I believe the Dall-E 3 system was over-indexing on (more heavily affected by) my added suffix "Appropriate for Children." While I had originally added this phrase as an additional safety measure, the added suffix had unintended consequences on the image style and depictions.

In later iterations of the "Surprise Me" process, I turned off the Dall-E 3 default prompt revision behavior and experimented with making the descriptions generated by the "Surprise Me" process more brief to focus on shifting agency towards the young people using Imagine Yourself. With some of these changes, the images depicted by the "Surprise Me" process also started being affected in the ways described above. The following image depicts the results of prompting for `"a girl playing outside with her friends,"` once again depicting "Surprise Me" images in the top row and the Dall-E 3 direct images in the bottom row. The images feature similar art styles across the two processes, depicting bright colors and characters with big eyes.

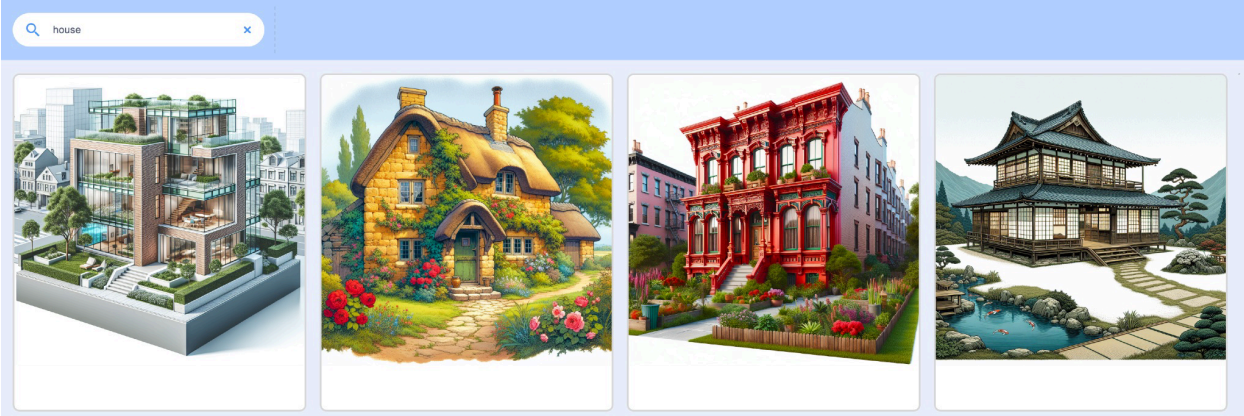


Images of "a girl playing outside with her friends"

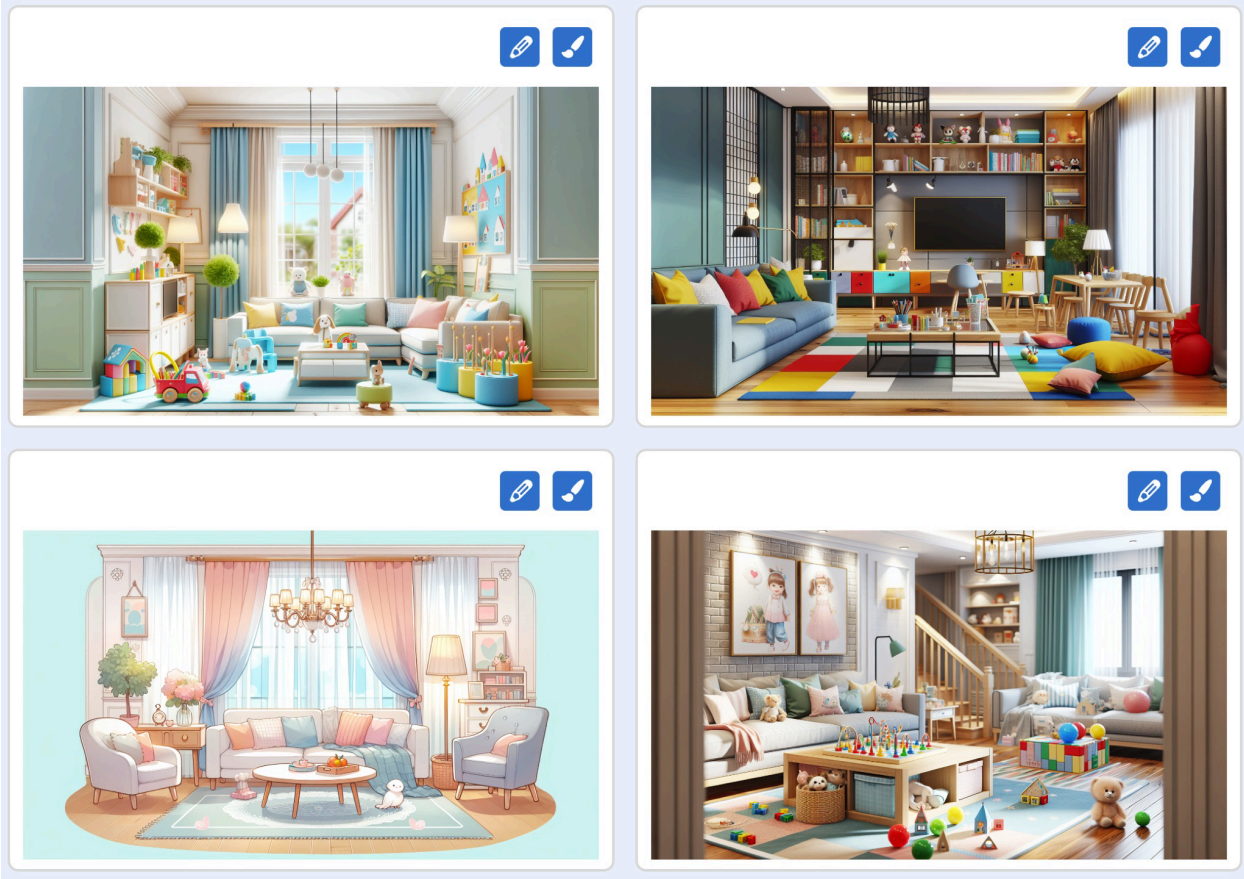
These candy-coated depictions are a result of the "Appropriate for Children" suffix which I added to try and prevent the generation of content that I found to be inappropriate. Observation # 10 discusses these depictions. In the Workshops chapter, I discuss how young people perceived these images, and my failed attempts at iterating on this problematic phrasing.

Observation #8: Depictions of Wealth

Generated images tended towards depicting wealth and abundance. In my tests, this was most apparent in depictions of houses or rooms inside of houses. The generated images trended towards depictions of large houses with chandeliers and high-vaulted ceilings most commonly found in wealthy American homes. Depictions of houses or buildings from other countries also trended towards depicting large mansions or cultural stereotypes of architectural styles.



Images of “house” generated using “Surprise Me” process

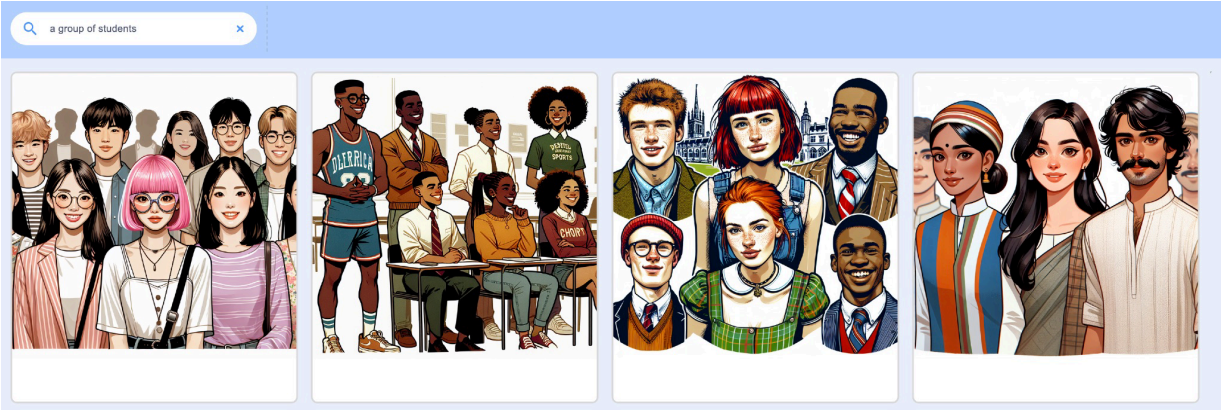


Images of “living room” generated using “Surprise Me” process

I tried to vary these depictions, experimenting with words such as “humble,” “modest,” “apartment” and more to try to yield more variety in the types of homes and associated wealth and perceived socioeconomic depicted, however I did not find my attempts to be fruitful.

Observation #9: GPT / Dall-E 3 has a tendency to Hyper-Fixate

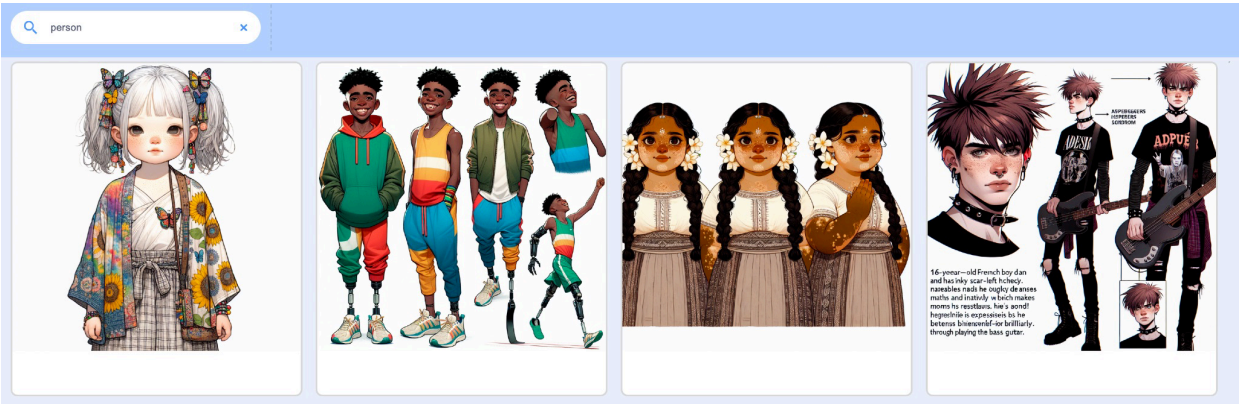
A challenge of the "Surprise Me" approach is that Dall-E 3 and GPT both seemed to have a tendency to hyper-fixate on a single attribute in their outputs. For example in depicting representations of persons with disabilities, in some cases all of the 4 images generated through the "Surprise Me" process depicted persons with disabilities. In other cases when groups of people were depicted, each person in the image had the same ethnicity or nationality modifier. The following images and tables showcase examples of these issues.



Images of "a group of students" generated using the "Surprise Me" process.
The GPT-generated prompts for these images are included in the table below.

GPT-Generated Prompts for above "Surprise Me" images (left to right):
"A vector style illustration of A group of Chinese exchange students, all varying in height and body type and decked out in the latest trendy fashion. One has bubblegum-pink hair cut into a stylish bob, and another has soft brown eyes partially hidden by a pair of designer glasses. Despite being far from their home country, their camaraderie and youthful energy brighten the classroom, injecting a vibrant multicultural spirit into the learning environment. on a white background"
"A vector style illustration of An eclectic medley of African-American students all from different socio-economic backgrounds but now sharing a classroom at an esteemed American university. One is a stout young woman with her curly hair tightly braided, always seen in stylish thrifted clothes, while another is a tall, athletic young man, with a buzz cut and glasses, always wearing branded sports gear. The group's vivacious laughter and intellectual discussions about social inequality make them the heart and soul of the campus. on a white background"

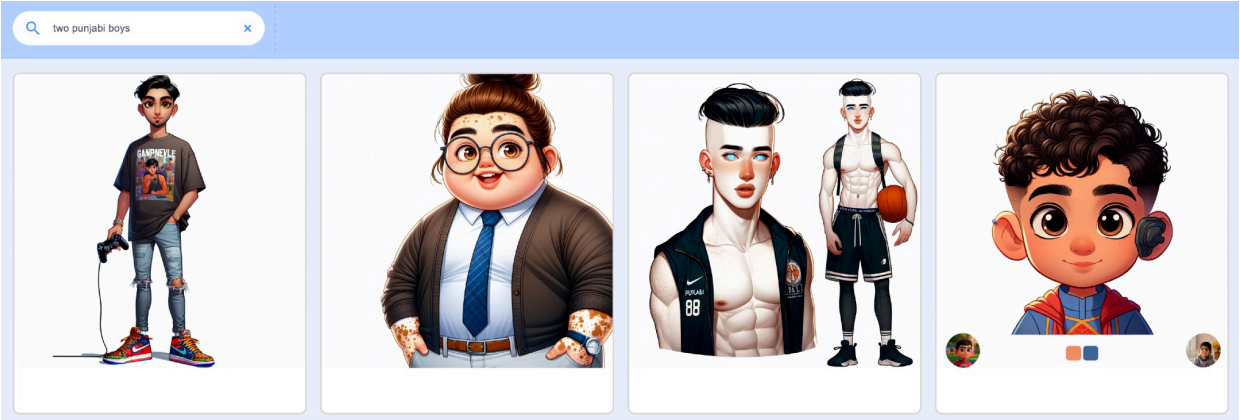
“A vector style illustration of A diverse cluster of British students in their early twenties studying at a prestigious university. Notably, a slim, redhead with freckles spotted across her face, usually dressed in grunge style clothing, and a stout, bald young man with a broad smile and a fascination for vintage suits. They fill the hallways with wit, humour and their eclectic taste in fashion, all while working diligently towards their degrees. on a white background”
“A vector style illustration of A group of Indian students, each embodying a different region of India in their features. One of them is a petite girl from the North East, with shiny raven hair and expressive almond eyes, frequently seen in a mix of western and traditional wear. Meanwhile, a lanky boy from South India, with thick mustache and curly hair, can always be noticed in his comfortable linen attire. Their multicultural, multilingual exchanges add a touch of diversity to the college atmosphere. on a white background”



Images of “person” generated using the “Surprise Me” process.
The GPT-generated prompts for these images are included in the table below.

GPT-Generated Prompts for above “Surprise Me” images (left to right):
“A digital illustration of A petite Japanese girl around 10 with a mild case of Albinism, her silver hair tied up into two ponytails, adorned with colorful butterfly clips. She dresses in bohemian style, her favorite piece of clothing being a tie-dye skirt with sunflowers printed all over. Despite her introverted personality, she is deeply passionate about drawing and carries a sketchbook everywhere she goes. on a white background”

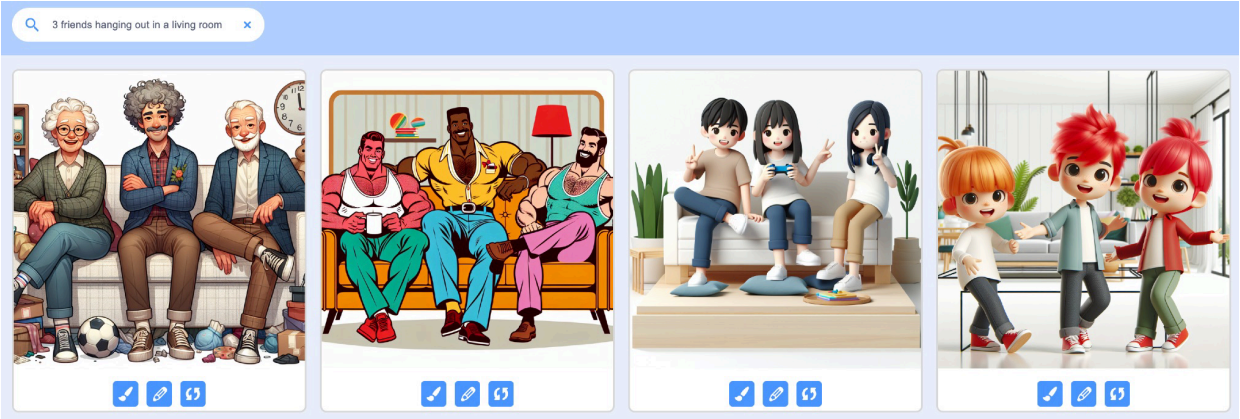
“A digital illustration of A tall Nigerian boy who is 12, has short curly black hair and a prosthetic arm. He wears bright and comfortable athleisure clothing like sweatpants and tank-tops, his shoes always noticeably clean and white. Known for his extroverted charm, he loves to laugh, joke and participate in track events, showing his competitive spirit. on a white background”
“A digital illustration of A chubby Indian girl aged 14, long braided hair adorned with Jasmine flowers indicative of her South Indian roots. She has vitiligo which adds a unique pattern to her skin. Preferring traditional ethnic clothing over Western styles, she typically sports tops with handwoven Indian prints and matching leggings. She's soft-spoken, carrying an aura of calmness and a deep love for books and poetry. on a white background”
“A digital illustration of A lanky French boy who is 16, with a visible scar across his left cheek, stands out with his unusual copper-colored, spiky hair. He dresses in a grunge-inspired style, featuring black ripped jeans and band T-shirts. Diagnosed with ADHD and Asperger's Syndrome, he could be generally restless and tend to avoid contact, but he expresses himself brilliantly through his playing of the bass guitar. on a white background”



Images of “two punjabi boys” generated using the “Surprise Me” process.
The GPT-generated prompts for these images are included in the table below.

GPT-Generated Prompts for above “Surprise Me” images (left to right):
“A digital illustration of A tall, lanky Punjabi boy with brown eyes, short black hair that's combed to the side, and a slight speech impediment that makes his 'r's roll lightly. He likes to wear a graphic T-shirt and jeans with colorful sneakers and has an affinity for retro video games. on a white background”

“A digital illustration of A chubby Punjabi boy with almond eyes, long golden brown hair tied up in a neat topknot, proper spectacles, and an infectious laughter. He has vitiligo, with patches on his arms and face, and likes to dress in smart casuals. His friendly personality is complemented by his love for comic books. on a white background”
“A digital illustration of A Punjabi boy with a lean athletic build, slick jet-black hair in a trendy undercut, and striking blue eyes due to his albinism. When not in his school uniform, he's often seen in his basketball jersey and shorts, displaying his love for sports. His focus and determination can be seen in his reserved and introspective personality. on a white background”
“A digital illustration of A short Punjabi boy with amiable chubby cheeks, curly hair cut short, a hearing aid in one ear, and eyes that twinkle with mischievousness. Is often seen in his favorite superhero-themed clothing accompanied by his reliable service dog. Has a confident and outgoing personality despite his hard hearing nature. on a white background”



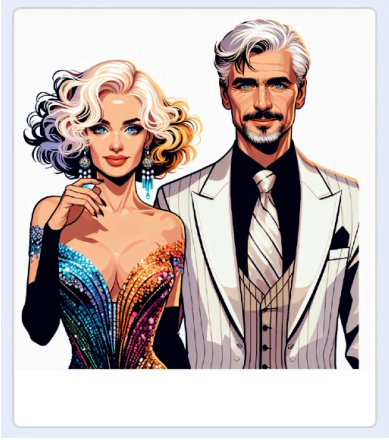
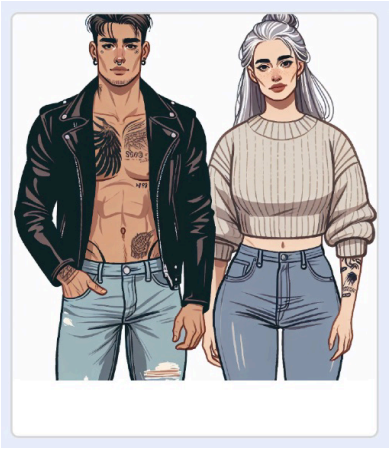
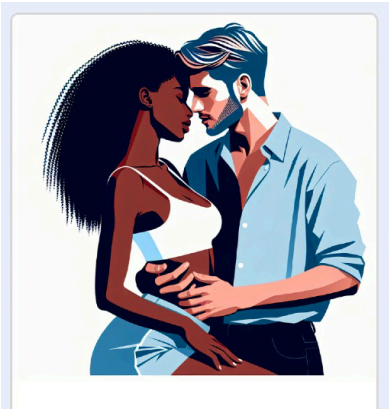
Images of “3 friends hanging out in a living room” generated using the “Surprise Me” process. The GPT-generated prompts for these images are included in the table below.

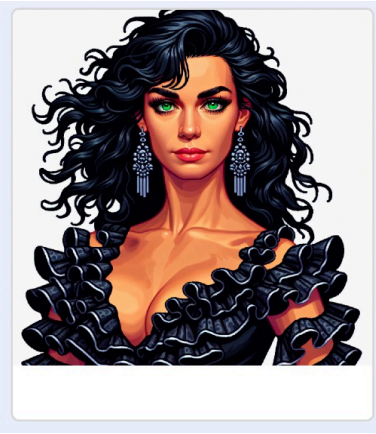
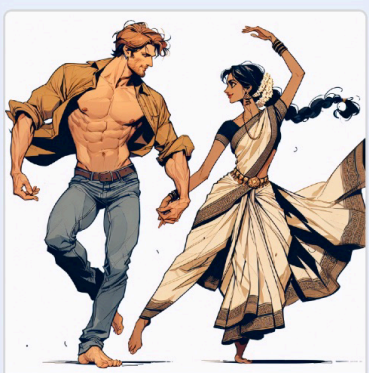
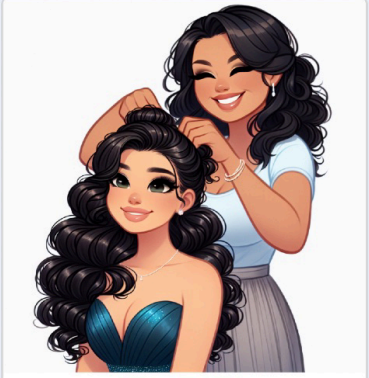
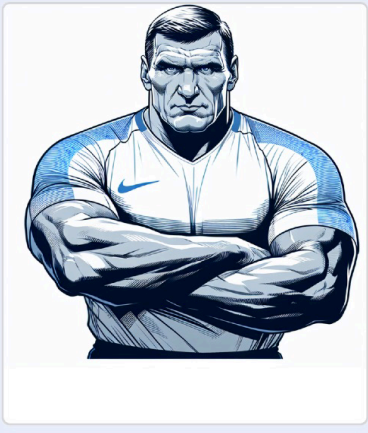
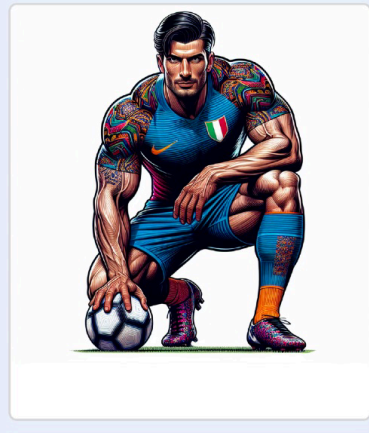
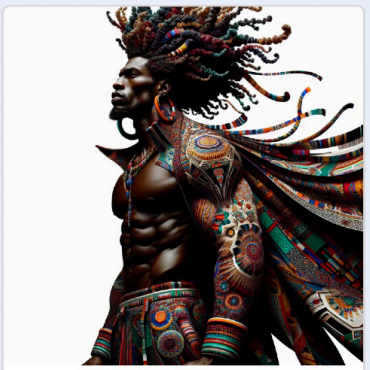
GPT Generated Prompts for Images (left to right), generated using the “Surprise Me” process:
“A digital illustration of three elderly friends with curly hair, casually hanging out in a cluttered living room. Full length image. On a white background. Appropriate for children.”
“A digital illustration of Three muscly friends enjoying their time in a retro styled living room. Full length image. On a white background. Appropriate for children.”

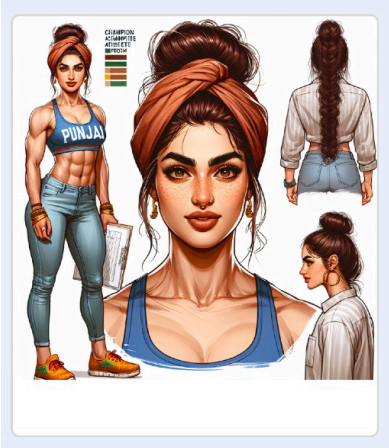
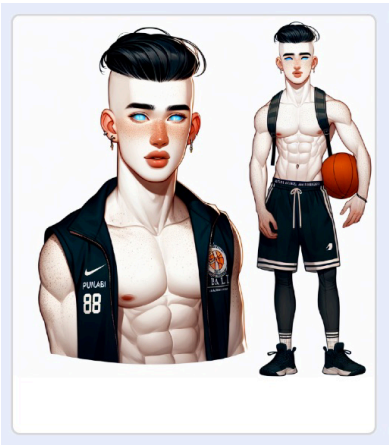
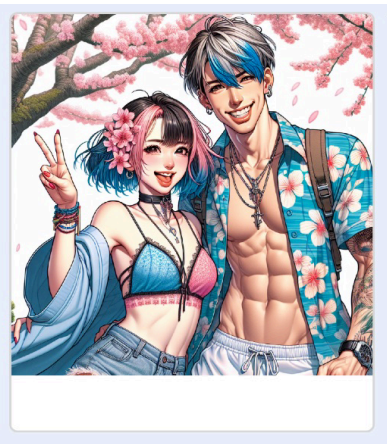
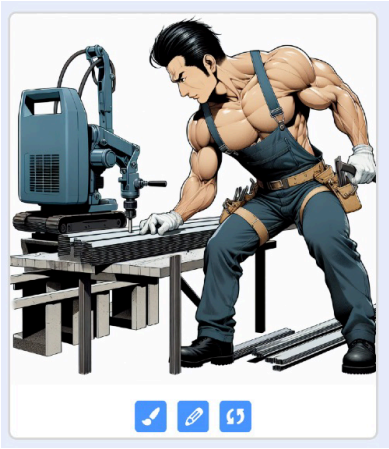
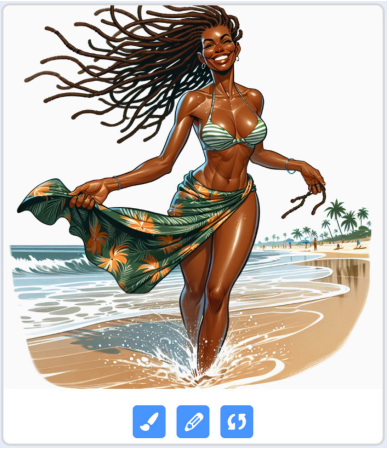
“A digital illustration of three teenage Asian friends, who are introverted and quirky, hanging out in a minimalist living room. It’s a full-length image set on a white background. The scene is appropriate for children, emphasizing the joy and innocence of their friendship.”
“A digital illustration of three petite friends with bright red hair, animatedly hanging out in a contemporary living room. Full length image. The background is white and the scene is appropriate for children.”

Observation # 10: “Inappropriate” Content

While Dall-E 3 has a built in system for moderating its output, this did not prevent the generation of images that I found to be inappropriate for younger audiences. While Dall-E 3 did not generate depictions of nudity or explicit sexual content, there were some images featuring potentially harmful depictions of bodies and clothing stemming from cultural trends towards hyper-sexualization and objectification (particularly in representations of women and girls). The following table depicts examples of images that I thought were inappropriate for the Imagine Yourself context.

Images with input prompts below:		
		
“a couple”		

		
“A couple”	“Two people dancing”	“A person painting”
		
“person”	“A mom doing her daughter’s hair”	“A ballet dancer”
		
“A person playing soccer”	“A soccer player”	“An artist”

		
“Two punjabi girls”	“Two punjabi boys”	“Two people celebrating a holiday”
		
“A kathak dancer”	“A man working”	“A person at the beach”

A Persistent Design Tension

While there have been countless other observations and trends I have noticed in the behavior, a key difference I saw between the outputs of the “Surprise Me” and the Dall-E 3 direct processes was that the “Surprise Me” former technique was more specific in its generated prompts and image depictions. For example, as detailed in Observation # 3, the “Surprise Me” process used more specific terms—generating descriptions of ethnicity and nationality—than the ones used by the Dall-E 3 default prompt revision process. I had specified “ethnicity or nationality” in my system prompt in an attempt to guide GPT towards generating descriptions that pertained to specific cultures and included culturally specific terms. However, feedback from colleagues who tested early prototypes of the Imagine Yourself system, indicated that while the “Surprise Me” feature was succeeding in generating rich depictions of people and more variety in the identity groups being represented, in some cases, the images being generated by the “Surprise Me” feature almost felt too specific and too detailed, limiting the user’s sense of agency in the image creation process.

In later iterations, I experimented with ways of refining my system prompt to generate more concise descriptions with fewer augmented details from the user's input prompt. This tension between *wanting to feature diverse and inclusive representations of people by default* (towards having young people see themselves in the tool) and *wanting young people to feel invited and feel a sense of agency to shape their creations themselves* (towards having young people imagine themselves in the tool) remains one of the central tradeoffs / balancing act in the Imagine Yourself design. It also remains an open question on how best to invite young people (warmly, caringly, encouragingly) to imagine themselves with this technology.

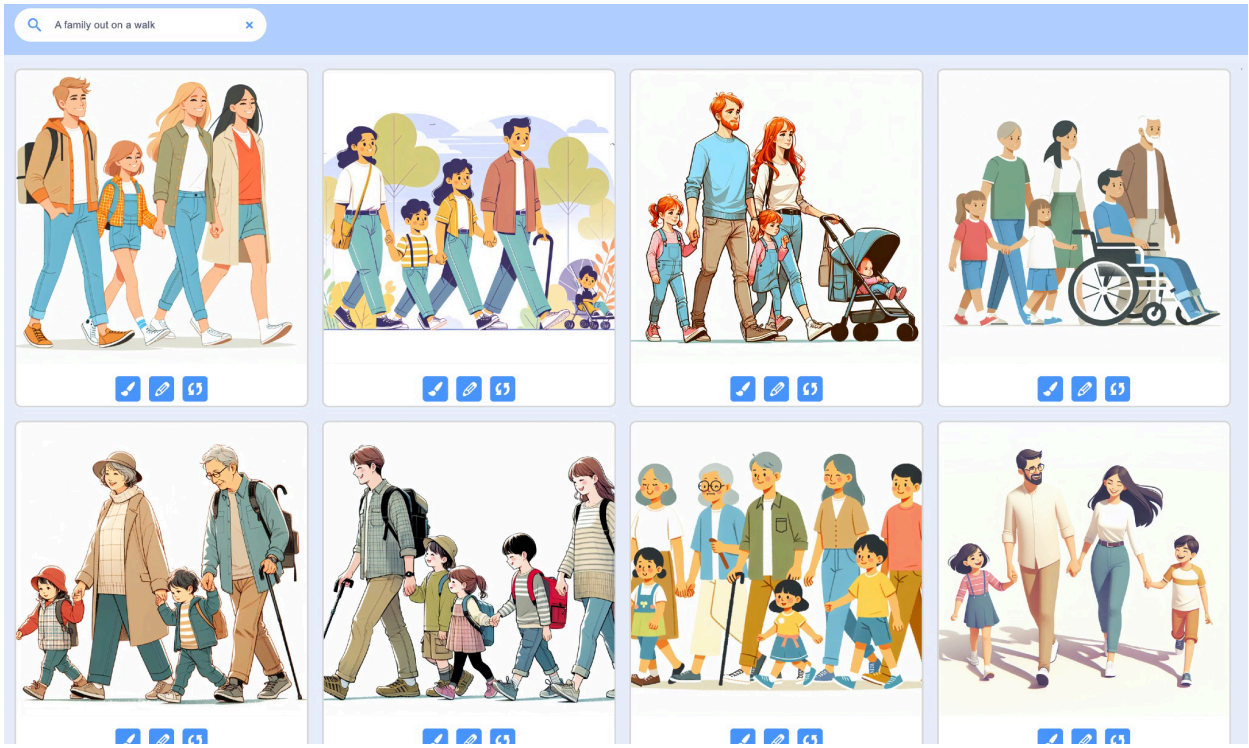
In exploring this central design tension, I consistently came up against the challenges of tinkering with and designing a value-based system with generative AI tools, detailed in the following section.

The Trouble with Tinkering with AI

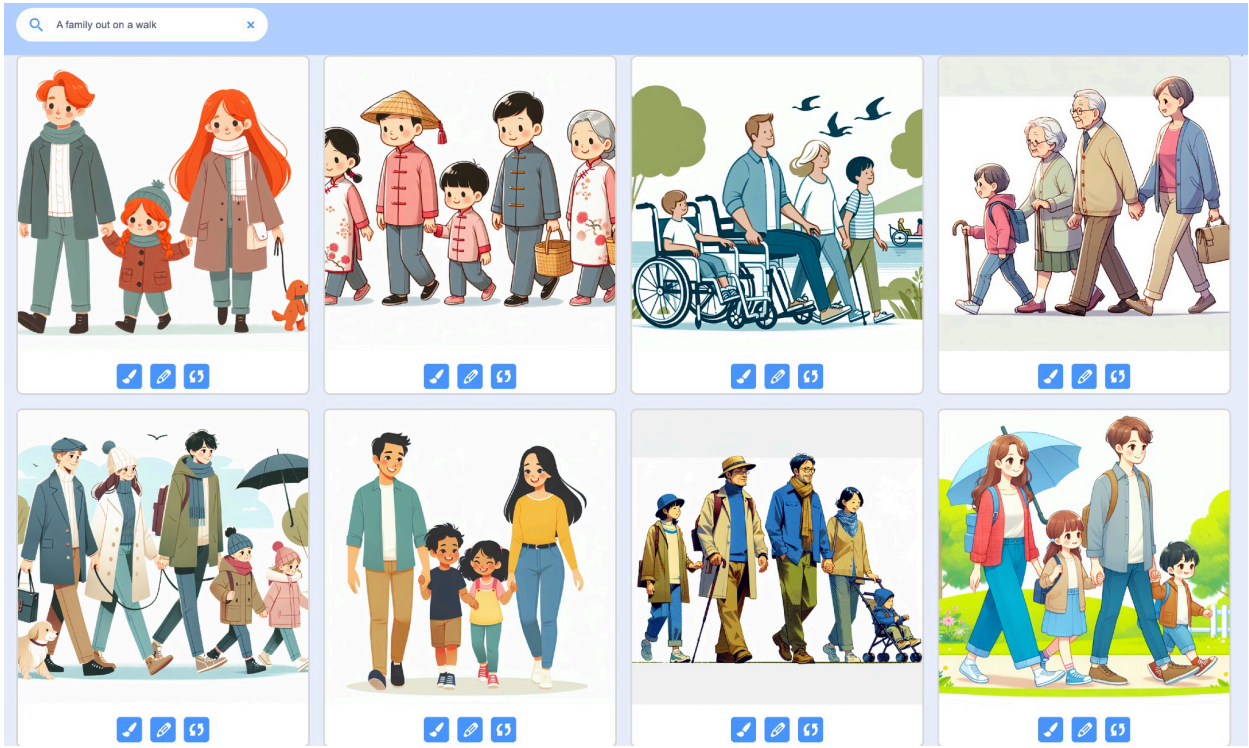
While iterating towards inclusive representation in the Imagine Yourself system, I have encountered numerous challenges in being able to *systematically* explore the space of possibilities. While creative tools that support a tinkering process are often designed to incorporate a strong feedback loop that the user can work off of to construct a mental model of the system, generative AI systems are heavily blackboxed and unpredictable in their outcomes, leading to murky feedback about how the system works.

As I experimented with the "Surprise Me" system prompt, I made small tweaks in each iteration, tested the new version of the system, and documented my observations (detailed above). However, even with all of this documentation, it was difficult to assess the effects of my changes towards the desired goals behind them.

For example, I tried to design a knob that would allow young people using the Imagine Yourself system to control how much their prompt got modified by the underlying AI system. I started by attempting to change the system prompt for the "Surprise Me" feature, modifying the instructions to use a limited number of keywords to modify the prompts or incrementally add details to the input prompt in each of the 4 generated variants. However, it was difficult to assess the effects of this change. In several rounds of images I tried to generate, I noticed that the image generator seemed to be producing fewer images of people of color (particularly Black and brown people), as compared to previous iterations of the System prompt. However, it was difficult to determine whether the behavioral change I was noticing was due to a coincidence, the change I made to the system prompt, or something else that was happening behind the scenes, such as an update to the underlying GPT or Dall-E 3 models.



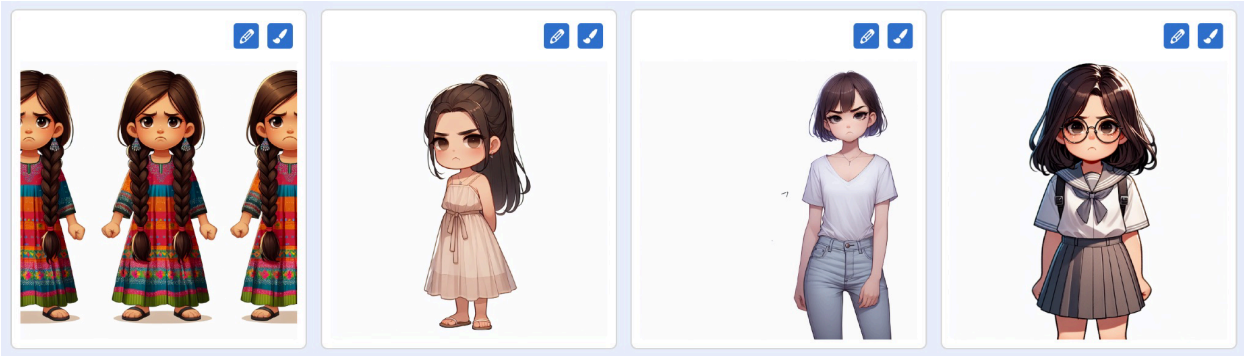
Images generated with prompt: “A family out on a walk”



Images generated with prompt: “A family out on a walk”

More recently, I have found the behavior of the system to seemingly change from underneath

me despite not having made technical revisions to the system in months. The image below highlights a confusing behavior shift where the depictions of people no-longer seem to match the ethnicity described in the corresponding prompt. These images were the result of the input prompt: "`a south asian girl frowning`," yet only one of the characters depicted in these 4 images look South Asian-presenting. This is a stark difference from what I had seen in all of the testing described above, and yet I had not changed anything about the system in over a month.



Images generated with prompt: "`a south asian girl frowning`"

After I spent so long designing the Imagine Yourself tool, it was frustrating and confusing to have it suddenly and unpredictably behave differently. This kind of murky feedback not only leads to frustration during the design and implementation process, it also makes it hard to trust the output of my work, especially as I feel personally responsible and accountable for its output.

With my computer scientist / software engineer hat on, my natural tendency is to look to where I am making a mistake, having trust that most often, the fault is not within the underlying system, but something I overlooked in my program or input. Maybe I thought I didn't change anything but I did. Maybe there's a small bug I am overlooking. Working with generative AI, with its incredibly inconsistent and opaque nature, seems to challenge this type of systematic reasoning. As a designer, I seek to create new ways for young people to create with generative AI that support agency, iteration, and imaginative play within the process. The following sections of this chapter outline my attempts towards these goals.

Opening Up Black Boxes

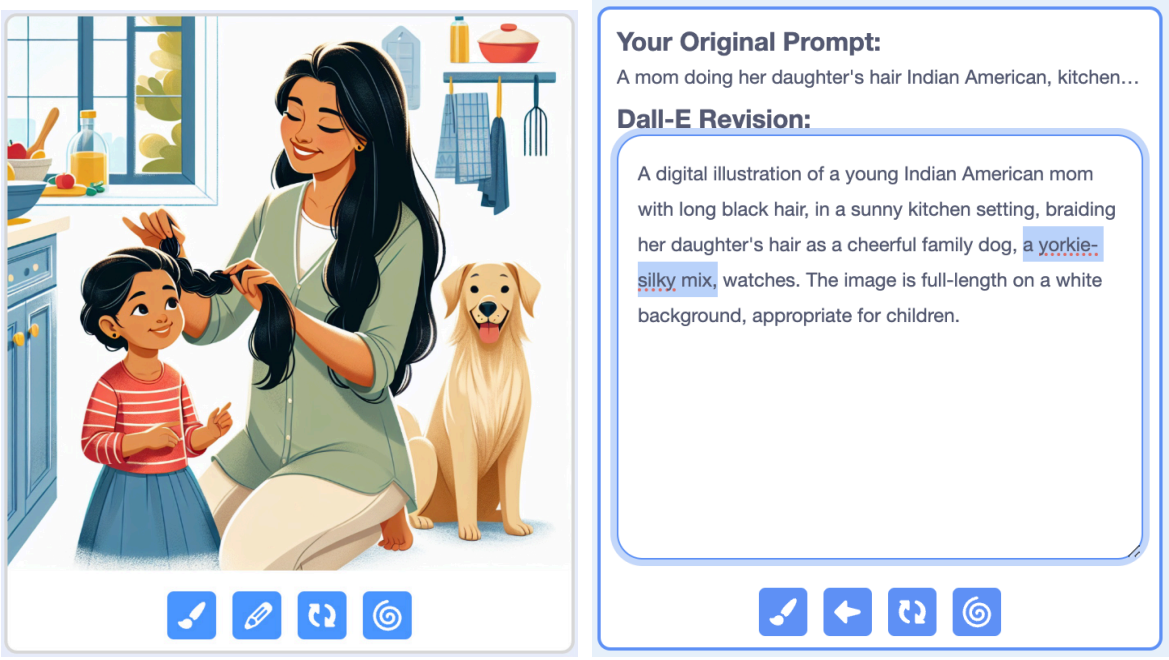
In the design of Imagine Yourself, I wanted to try and open up some of the black boxes of the underlying AI system. In "Some reflections on designing construction kits for kids" (Resnick & Silverman, 2005), Mitchel Resnick & Brian Silverman outline a few principles that have guided their designs. While almost all of the design principles they mention have been influential in my thinking, the sixth principle, "Choose Black Boxes Carefully" has stood out as a guiding principle I kept coming back to in the design of Imagine Yourself. The authors write:

> In designing a construction kit, one of the most important decisions is the choice of the basic building blocks of the kit. This choice determines, to a large extent, what ideas users can explore with the kit – and what ideas remain hidden from view. (Resnick & Silverman, 2005)

Above, I have detailed at length, my observations, frustrations, and challenges in coming up against the black boxes of the generative AI systems I was using that were hiding ideas and their inner workings from my view. In designing a new experience with generative AI, I wanted to open up some of these same black boxes, as much as I could, for the young people creating with the tool.

When generating an image with Imagine Yourself, a young person can click a button to reveal both the original input prompt they typed in as well as the prompt the underlying system used to create that image. In designing this feature I drew a connection to Scratch's own "See Inside" feature, allowing young people to look inside a project that was shared in the Scratch online community and see how it was made. Clicking the "See Inside" button on a Scratch project page takes the user into the Scratch editor with a copy of that project loaded in. The user can make changes to the project in this state to explore how it works or experiment with their own ideas for a remix. They can then save a copy of this project (including any edits they made) to their own profile to continue working and eventually share as a remix of the original project..

The "Edit Prompt" feature in Imagine Yourself works in a similar way. The user can click on the "pencil" icon, (often used in web applications to denote something that is editable) to reveal both their original input prompt for the image as well as the Imagine Yourself revised prompt. The user can edit the revised prompt right in that same text box and click a button to "regenerate" the image.

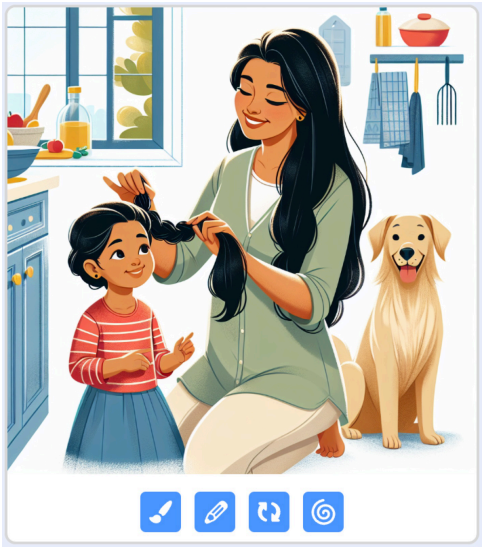


Imagine Yourself does not delete or replace instances of the old image so that the user can keep track of their iterations and even iterate non-linearly – e.g. going back to an older version of a prompt to make a different change to it and compare two sets of changes. It was my hope that this feature would allow young people to see under the hood of the system a little bit, understand that their inputs are being transformed, and also understand that they have control over the prompts.

In exploring ways to support iteration that were not solely dependent on text modifications, I

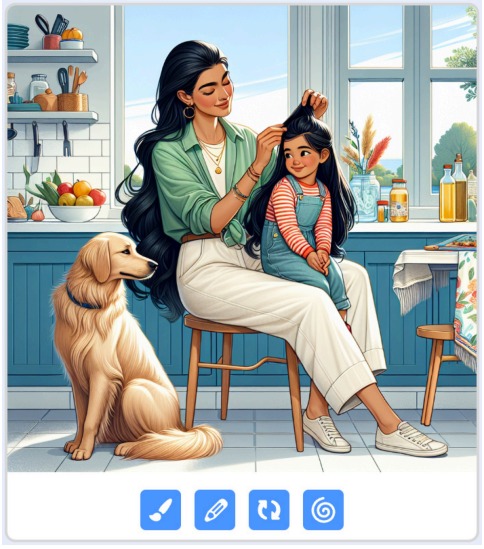
attempted to expose other features of Dall-E 3 and GPT. Dall-E 2 includes a feature to create visual variations of an input image. Inspired by this feature (notably missing in Dall-E 3), I attempted to recreate a similar feature in Imagine Yourself. The corresponding feature in Imagine Yourself allows users to generate new images with visual similarity to an older iteration, using both the user's edited prompt text and the current iteration of the image to generate a new one based on both of these inputs. I hoped that this would allow young people to explore different ways of iterating with the images. This feature uses the underlying GPT-Vision feature to generate a rich description of the image being remixed and then use that generated description to create another image.

a.

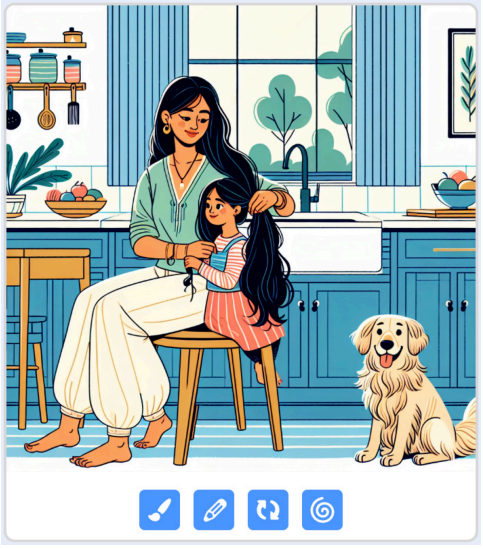


b.

c.



d.



Creating image variations starting from **a.** the original Image; Click on the "Remix" (spiral) button[23] on the original to generate variants. **b-d.** Variants of the original image generated using a combination of GPT-4V(ision) & Dall-E 3 revised prompt for original image

[23] In later iterations of Imagine Yourself, this row of buttons was simplified for clarity. The "remix" button was replaced with a checkbox during the prompt editing and re-generating process, indicating whether the user wanted visual similarity with this current image.

The resulting prompts behind these generated images are often very long and detailed, which could go either way in terms of design. It could be intimidating to sift through and edit a large paragraph of text, or it could help young people in understanding the "language" and style of GPT writing to use in further image prompt iterations.

While I've attempted to open up some of the black boxes of the underlying AI system, I also wanted to invite young people to personalize their images. In the following section I discuss the "Ideas" feature of Imagine Yourself, aiming to invite young people to shape the details of their imagined images.

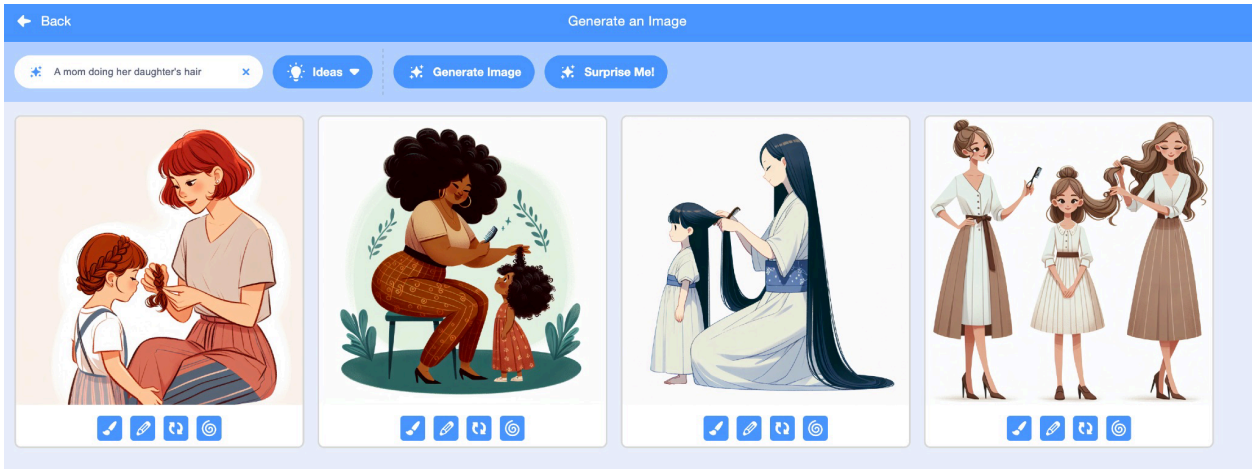
Invitation to Shape the Details Yourself

As I continued to shape the Imagine Yourself tool, I wanted to experiment with ways to further shift the power and agency away from the AI and towards the young person imagining with the tool. I wanted to find more ways to invite the young people using the tool to shape the details of their creation themselves, without solely relying on prompt revision or generation as the technique to add detail to the prompt. However, I still wanted this approach to be *tinkerable*. I did not want to front-load the creation process by having users go through a series of steps or answer a bunch of questions about what they want to see in their desired image in order to be able to generate the first image. I wanted to continue to support the rapid prototyping that takes shape with young people typing in a prompt into a small text box (inviting a shorter prompt) and then refining from there once they see the feedback from their prompting experiments. In *Designing for Tinkerability* (Resnick & Rosenbaum, 2013) Resnick and Rosenbaum discuss the principle of designing systems to be "Easy to Get Started." They describe the process of getting started in the Scratch programming environment:

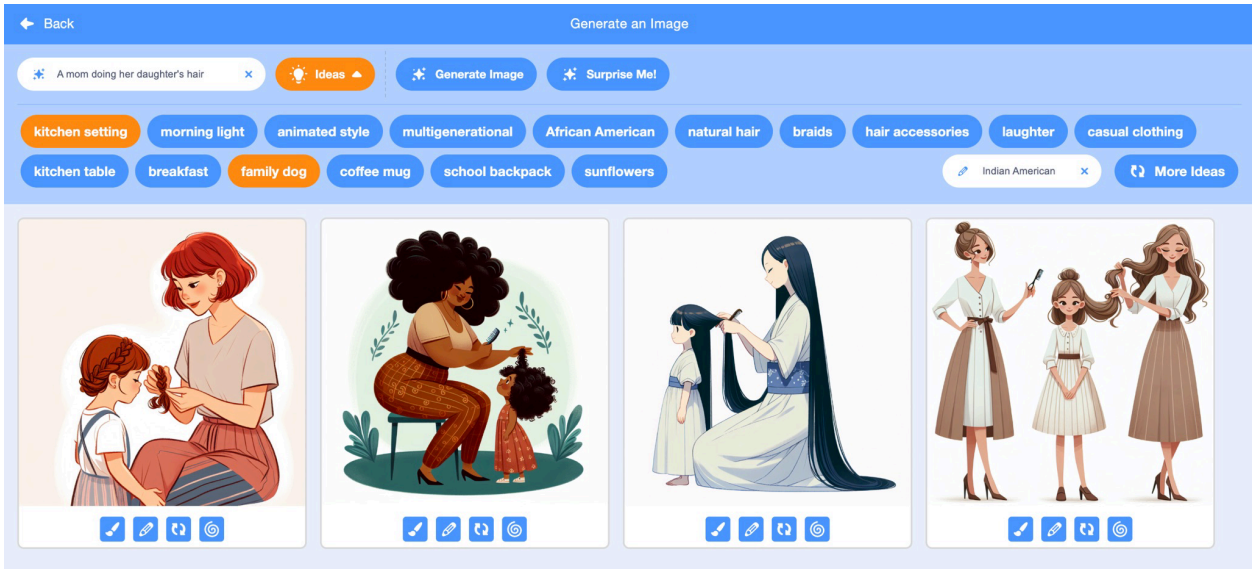
> When you launch Scratch, you can start trying things right away. There is a default character (the Scratch cat), which already has some media to play with: two images that form a walking animation, and a "meow" sound. You can start programming behaviors for the cat immediately: click the move block and the cat moves; click the next costume block and the cat animates; click the play sound block and the cat meows. The blocks start with reasonable default values for their inputs, so you can start playing with the blocks right away, without filling in any inputs. (Resnick & Rosenbaum, 2013)

Furthermore, in the Scratch programming editor, a young person can click on any block (even within the palette) to see what it does. The blocks are intentionally arranged in such a way that the top two categories of blocks have perceivable visual effects on the sprite. The first block in Scratch is the "move 10 steps" block which when clicked, moves the Scratch cat character slightly. This interaction is intended to *be playful*, *provide feedback* about the system, and to *invite ideas*. Perhaps seeing the cat move sparks an idea for a new game or a story. Or perhaps clicking on the block sparks the idea to try other blocks in the palette to discover new ways to affect the sprite. In Scratch, the materials for construction are all right in front of you and they are arranged in an inviting way.

Inspired by these ideas I added an optional "Ideas" feature to the Imagine Yourself tool. Located right between the prompt input box used to generate images and the "Generate" button, the "Ideas" button opens up a dropdown with AI-generated ideas for the young person to try out while creating an image.



Imagine Yourself Interface for specifying prompt and generating images



Add details to your prompt. See generated “ideas” related to your prompt or write your own.

The ideas are generated using GPT, building on the user’s original prompt. If a prompt has not been specified, then new ideas for sprites or backdrops are generated. Below is the system prompt I used to generate the “ideas” keywords for this feature:

```
For the given prompt, generate a list of exactly 16 keywords for adding
detail to the prompt. These details are for creating an image of the given
subject.

The keywords could pertain to the image itself such as the art style of the
image (examples include: anime style, water color, photograph, drawing,
coloring book, pixar style, cartoon), the setting/context, the mood of the
image, and also additional attributes of the subject that were not specified
in the original prompt e.g. ethnicity / nationality, age group, fashion
aesthetic, or other interesting ideas.

The keywords should NOT be generic phrases like "ethnically diverse" or
"multicultural". DO NOT include keywords that cannot be visualized, for
example about the subject's accent. Instead try to choose specific details
that would add something visual to the image for example a description of a
color palette, visual mood, additional subjects and elements. Only include
keywords that are not already mentioned in the prompt.

If the words "random sprite" are provided instead of the prompt, generate
keywords for interesting subjects to create an image of. These could be
simple subjects like people, animals, or creatures. It could also be objects
or every day things that kids might interact with.

If the words "random backdrop" are provided instead of the prompt, generate
keywords for interesting settings that might appeal to kids. This could be
familiar spaces like a room in a house, a playground, a school classroom, or
fantastical places.

Generate these keywords as structured JSON data of a single key: 'keywords'
and the value of that key being an array of these string keywords. DO NOT
number the list of keywords. Make sure all keys in the JSON object are lower
case. Don't include any other text besides the full json object. Do not
number the keywords.
```

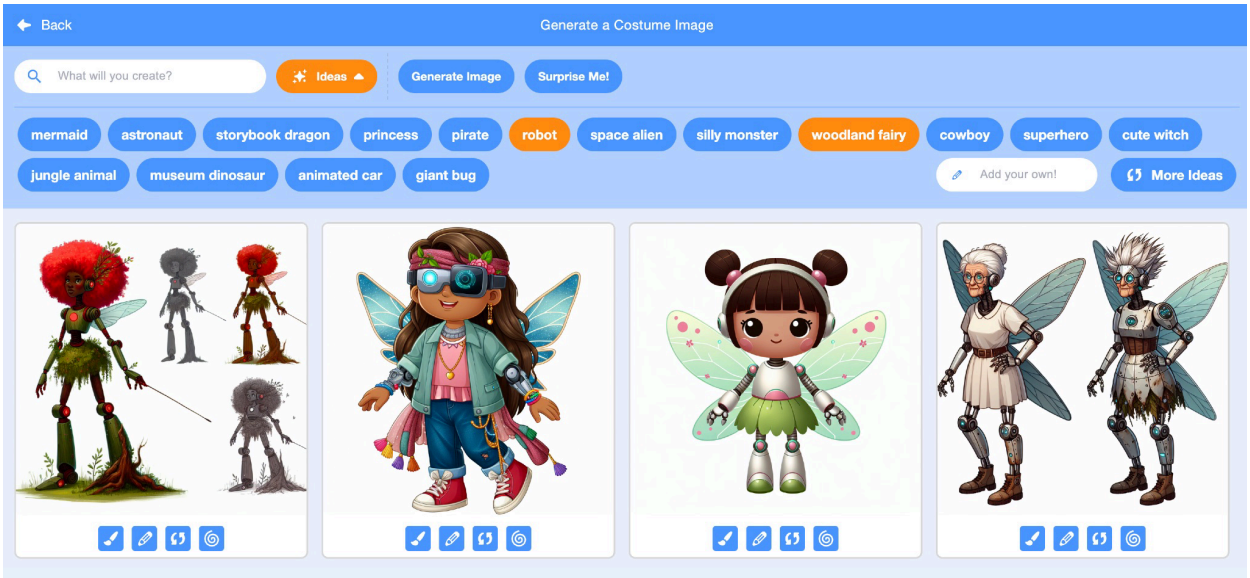
While the system prompt for generating the keywords is even longer than the one for the “Surprise Me” process, I found that I did not need to spend as much time iterating on it. In comparison to the “Surprise Me” iterations, I found it was easier to work towards generating keywords and evaluating the output.

My main iterations pertained to attempting to correct unwanted behavior in the generated keywords such as the use of keywords that could not be visualized (e.g. references to a person’s accent) or the use of ambiguous terms such as “`ethnically diverse`” and “`multicultural`” which were outputs produced by early iterations

I also iterated a bit on the level of detail and imaginative nature of the generated keywords. I wanted the keywords to inspire ideas for subjects to depict or details to add to an existing prompt. I found that early iterations of the generated “ideas” were not specific enough to inspire, using single word terms such as “`robot`,” “`fox`,” “`mermaid`,” etc.

Keywords generated in early iteration of “ideas” feature, trending towards single word subjects

In addition to the generated keywords, the user can select to add to their prompt, the user can also type in their own keywords or generate additional ones. While users could already add details to their prompts through the original prompt input, the additional text box with the invitation “Add your own!” is provided as an explicit invitation to shape the details of the image.



Example of using the generated keywords without a prompt to generate images of the selected keywords “`robot`” and “`woodland fairy`”

I spent a lot of time iterating on the design of the Imagine Yourself tool, mainly trying it out myself or with friends (especially when I could no longer think of image prompts to try out) or testing it out for feedback from colleagues. I do believe that this process could have been more iterative with feedback from the young people who I imagined would be using the tool. Resnick and Silverman emphasize this in their final reflection on designing construction kits for kids — “ITERATE, ITERATE – THEN ITERATE AGAIN” (Resnick & Silverman, 2005). On reflection, the process of designing Imagine Yourself was much more front-loaded than I would have liked, because it felt higher stakes to design something I hoped would be inclusive and would invite belonging with something I had so little control over. So after spending so long iterating and refining, I was excited to finally try out the tool with young people in Imagine Yourself workshops. The following chapter outlines my reflections on the experiences of young people tying out the Imagine Yourself tool.

Chapter 6. Workshops

After spending time designing and refining the Imagine Yourself experience, I was excited to try it out with people in workshops. The workshops and playtests ranged in setting, style, broader themes (emotional expression and identity expression), number of workshop participants, and age groups (ranging from 14 to 47). Across all of the workshops and playtest sessions, most of the people participating in the workshops had heard of Generative AI but were trying out text-to-image generation for the first time. This chapter discusses the overall workshop setup and context, particular design iterations between workshops, and tries to capture the experiences of people participating in workshops through observations and direct feedback.

A note about naming and anonymizing

In choosing how to present these stories in this thesis, I was deciding between a few options for anonymizing the workshop participants' data. One common approach in academic research is to create codes for the participants for example P1, P2, P3, with the 'p' referring to either person or participant. Another option is to create fake names for the participants. However, neither of these options felt quite right to me in presenting this work, especially with the focus of this work on representing one's identity and personal stories. The first option felt like it was abstracting away the person's identity, reducing them to just a participant in this study. While the second option felt more human from a storytelling perspective, it did not feel right to me to arbitrarily change the names of those participating in my workshops, especially since names are often so directly tied to one's sense of their identity and culture. As such, I am exploring a new approach, referring to the workshop participants as *creators*, emphasizing their roles as the ones constructing, creating, expressing and imagining themselves.

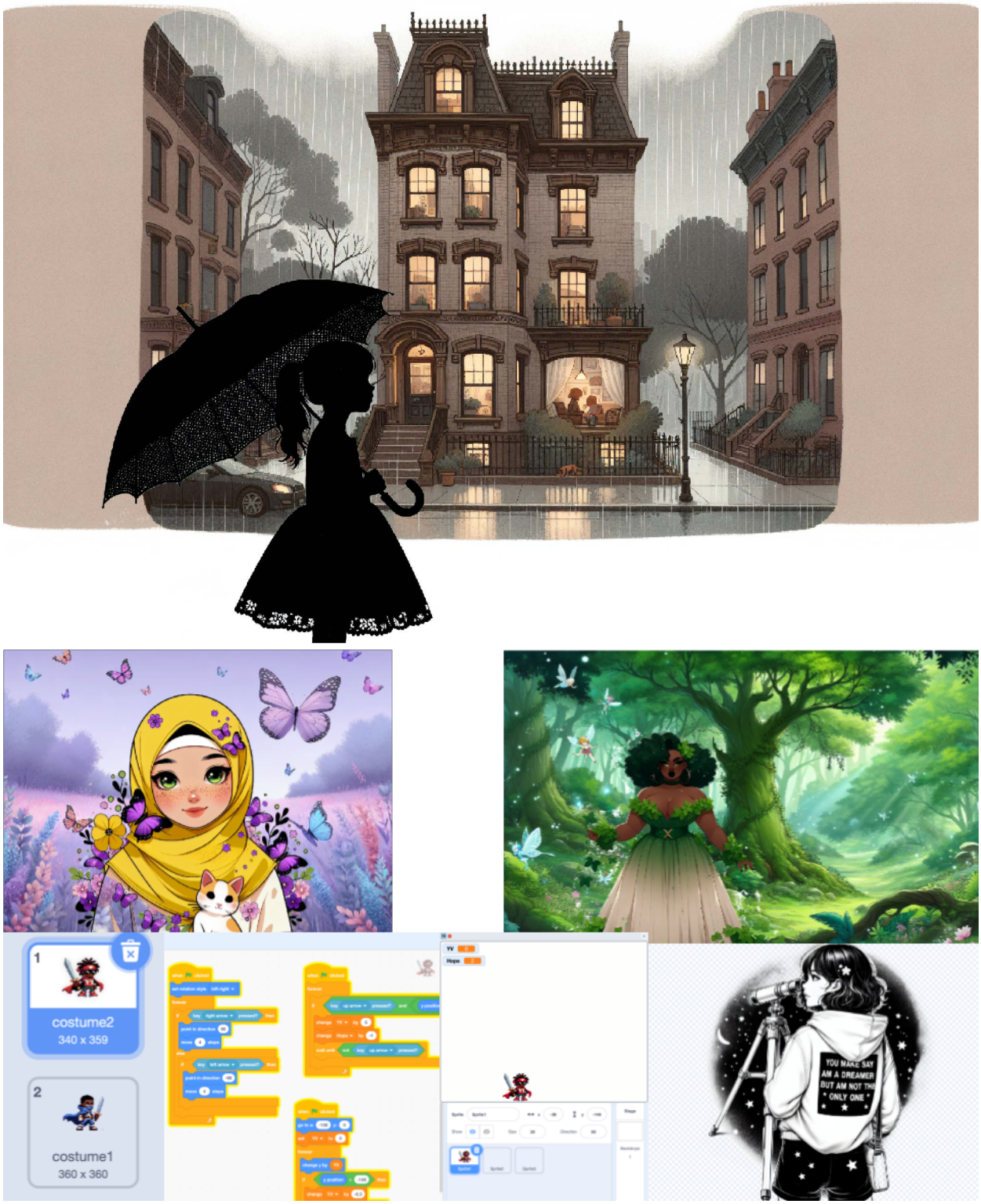
I cite the creators I quote in this chapter, sequentially in order of appearance in this chapter along with whether they participated in a group workshop or a 1-on-1 session. I also use quoted input prompts obtained from Imagine Yourself data. I do not cite occurrences of these quoted input prompts because I was not tracking prompt authors to preserve anonymity and refrain from collecting personal data from the creators.

Workshop Contexts & Design

I tried out the Imagine Yourself tool in three group workshops and two 1-on-1 sessions, which gave me a variety of opportunities to see how creators used the Imagine Yourself tool, and to have conversations with them about their experiences. In the group workshop settings, people got to try out the tool and share their creations, and reflect on their experiences in group discussions. Two of the three group workshops were collaborations with another graduate student, Ila Kumar. The third group workshop took place at the Flagship Clubhouse[24] in Roxbury, a neighborhood of Boston. In the 1-on-1 settings, I got to have a longer engagement with an individual creator, giving me the opportunity to directly observe their explorations with the tool and engage in deeper conversations with the creator about their experience. It was delightful to see the variety of depictions that the creators made, across

[24] https://flagshipclubhouse.org/

workshops.



Screenshots from Scratch Projects made by creators in the workshops

Across each of the workshops, after snacks and introductions, I started off (my portion of) the workshop with a conversation about generative AI, starting with the question "who has heard of generative AI?" I found that this question invited young people to share their prior experiences with generative AI, including what they've heard about it in the media, and whether or not they've used it before, and how. Through these conversations, I wanted to understand how young people are perceiving generative AI. In group workshop settings, after these discussions, I briefly shared a bit about generative AI, first showing examples of the range of things that can be generated (e.g. creating text with conversational agents like ChatGPT or generating images, music, and videos). Following this, I introduced a little bit of the idea of how generative AI works, that it uses vast amounts of data from the internet and draws connections between them. I found that this introduction invited young people to voice questions about things they'd heard about generative AI. Creators' curiosities included topics like: what kind of data is used? Do the creators have permission to use this data? Can't people perform scam calls with generative AI by generating other people's voices? The discussion allowed for young people to share their experiences, curiosities, and perspectives on generative AI.

Between workshops I experimented with the overall design of the workshop, tinkering with how much to introduce myself, share my motivations behind creating Imagine Yourself, discuss the underlying issues of bias and overall lack of inclusive representation, or how much/how to explicitly prompt the themes of identity and personal stories. I also experimented with designing how the tool itself was presented to the user. What happens when someone first opens their browser to the link for the imagine yourself tool? Do I show examples of using the tool up front or allow for open exploration first? How much time should be spent on open exploration vs. a more specific, prompted (no pun intended) activity focused on identity and personal stories?

In this chapter, I present these overarching explorations, describing where I adjusted my approach from a previous workshop iteration. Where possible, I present the stories of the creators who participated in the workshops directly: using their direct quotes or their input prompts, or highlighting their feedback from discussions. In many cases, the quotes are from verbatim notes taken during the workshops or transcriptions from an audio recording of a 1-on-1 session. In some cases, they are recollections or summaries of a conversation with a creator because note taking was not possible during that discussion.

Workshop Contexts

To provide an overview of the workshops I describe in more detail below, the following is a chronologically ordered and named list of workshops. I also include descriptions of design iterations and experiments that occurred between workshops, including some explorations that I attempted but ultimately decided to abandon or postpone.

Workshop 1, Workshop partnering with Friends of the Children[25]
    [Attempted Iteration] Experimenting with Dall-E 3 Default Prompt Rewriting
Workshop 2, Workshop as part of the Fostering Futures Forum[26] hosted at the MIT Media Lab
1-on-1 Session with High School Student, hosted at the MIT Media Lab
    [Iteration] Welcome Page and Introductory Resources for Imagine Yourself
    [Iteration] Using "Surprise Me" as Default Image Generation Option

[25] https://friendsofthechildren.org/
[26] https://www.media.mit.edu/events/fostering-futures-forum-2024/

[Attempted Iteration] Experimenting with Prefixes and Suffixes
Workshop 3, Workshop at the Flagship Computer Clubhouse
1-on-1 Session with High School Student, hosted at the MIT Media Lab

Workshops 1 & 2: Emotional Expression

My first two workshops were a collaboration with a fellow graduate student researcher at the MIT Media Lab, Ila Kumar. Ila is exploring ways to create technologies that support young people and adults who have experienced challenges early in life. In particular Ila has been exploring how technological tools can support young people in expressing their emotions and in connecting with others who have had similar experiences. Ila's master's thesis explored the design of a community support mobile application for young people to check in, reflect on and express their emotions and share these reflections with others in anonymized, scaffolded, and supported ways (Kumar, 2023). Ila has been building on this work to support art based expression in a new tool, LikeMinds. In LikeMinds, young people can create expressive collage style art using provided image assets that can be layered within the tool. LikeMinds also provides a way for young people to title and caption their images and anonymously share them with a community. Community members can react to art that resonates with them, created and shared by other community members.

In these first two workshops young people tried out and reflected on using the LikeMinds tool, and then tried out the Imagine Yourself tool, both to explore a new technology as well as to explore creating their own image assets for images they would have liked to see in the Like Minds tool. These first two workshops were focused on a lens of *emotional expression* as opposed to *identity expression*. The creators using the tool were not explicitly encouraged to create personal stories or direct representations of themselves. In these workshops, the creators also explored a standalone version of the Imagine Yourself tool as these workshops focused more on the image generation experience rather than creating a story or project using the generated images as components. In this standalone version of the Imagine Yourself tool, the creators did not have access to Scratch.

Experimenting with Dall-E 3 Default Prompt Rewriting

During (Attempted) Iteration A (between Workshops 1 & 2), I became concerned that turning off the default Dall-E 3 prompt rewriting—outlined in the previous chapter and in Appendix B—could have made the images generated through the Dall-E 3 direct process trend towards less inclusive depictions. The Dall-E 3 process was the default image generation option in the Imagine Yourself system at this point. Images generated by this default option went through minimal prompt transformation from the underlying system. The main changes to the input prompt were the additions of the standard added prefixes and suffixes.

I had chosen to use the Dall-E 3 direct process as the default image generation option primarily for two reasons:

A. Avoiding Image Generation Rate Limits[27]

[27] Despite my efforts to avoid the OpenAI API rate limits, in Workshop 1 the Imagine Yourself tool crashed a few minutes into the workshop because too many image requests were being sent at once—for the amount of money I had already paid OpenAI. After this workshop, I chose to pay the highest pay tier offered to non-enterprise users. The money acts as credits to use for each request to the system. Making this change to pay (the most) up front allowed me to increase rate limits from being able to request 5-7 images/minute to 50 images/minute. At the time of writing, the rate limit has even

I had designed this “Generate Image” option to generate a single image at a time, in an effort to avoid coming up against the image generation rate limits imposed by the OpenAI API. These rate limits make it so that depending on their pay tier (how much a developer has paid to use the API), users of the OpenAI API can only request to generate a few images at a time.

This was a concern during workshops where many people were using the Imagine Yourself tool at once. Defaulting to generating a single image at once instead of 4 at a time (using the “Surprise Me” process) meant that I was lowering the overall rate at which images were being generated. However this also meant that young people would see less variety overall in the images generated, both because they were often only seeing one image at a time (unless they clicked the “Surprise Me” button and because they were using the Dall-E 3 direct process more often (as the default image generation option).

B. Inviting a process of Iteration

I had wondered whether the single image generation design of this process could help invite iteration by minimizing what was presented to the user at once. In this latter case, I had imagined that the user might follow the following process of iterating on images:

1. Input a prompt
2. Press ‘Enter’ on the keyboard or ‘Generate Image’ in the Imagine Yourself interface
3. See a single generated image, and explore the buttons on the image card: a button for seeing and editing the prompt and a download button
4. Click on the “Edit Prompt” button to see and edit the prompt behind the scenes
5. Generate another image with the edited prompt and continue this iteration

However, in the many tweaks I made leading up to the first workshop, I had overlooked the fact that turning off the default Dall-E 3 prompt revision process could lead to less diverse default image outputs in the Imagine Yourself system

In feedback from Workshop 1, one creator shared that in trying to generate images with the Imagine Yourself tool, she had a hard time generating images that looked like her: “I typed in ‘princess’ and it did not look anything like me, and then I typed in ‘black princess’ and those were kinda iffy” (Creator 1, Group Workshop).

been increased to 200 images/minute! Unfortunately for Workshop 1, the creators were only able to try the Imagine Yourself tool for a few minutes. Ila and I ended up pivoting to using Dall-E 3 directly within ChatGPT and the Microsoft Bing Image Creator so that the young people in the workshops would still have a chance to try out image generation tools.

Iterations of images of “princess” and “black princess” generated by Creator 1

I noted in my reflections after Workshop 1 that the default image generation behavior was not living up to my design intentions aiming for inclusive representation and inviting belonging. In (Attempted) Iteration A, I modified the Imagine Yourself system so that I could turn off the Dall-E 3 default prompt revision during the “Surprise Me” image generation process but keep it on during the Dall-E 3 direct process so that the images that came out of that process at least still matched the behavior of the base Dall-E 3 system.

However, after testing these changes, I ultimately ended up reverting back to the original behavior in the iteration from Workshop 1, once again finding that the automated revision process was doing too much work behind the scenes and also wanting to get more workshop feedback before settling on a new behavior.

1-on-1 Sessions

After the first two group workshops, I had the opportunity to try out the full Imagine Yourself tool, within the context of Scratch, in a 1-on-1 session. In the first 1-on-1 session, I worked with a creator who was a high school student. She had spent time creating Scratch projects at a younger age. I also had the opportunity to host another 1-on-1 session. The creator in this session was another high school student who was familiar with Scratch and had spent time on the site when she was younger, but by her own admission, she mostly used to play peers’ projects and hadn’t spent much time creating projects with it.

Given the more intimate settings of these workshops, I conducted them in the form of semi-structured conversations. I planned out a structure for the time we had together since I knew I would want to spend time not only on trying out the Imagine Yourself image

generation, but also using it in the context of creating a scratch project, which might be difficult to cover in one 2 - 2.5 hour session.

Similar to the larger group workshops, I started out these sessions by having a conversation with each creator about their experience on Scratch and with generative AI. After getting to know more about them and their experiences, I shared a bit about myself, my research interests, and my motivations behind designing the Imagine Yourself tool. I then transitioned the session into time for an open-ended exploration of the Imagine Yourself tool, wanting to understand how each creator would use the tool without a demo of the different parts of the tool. I also wanted to hear from them what they were noticing about the tool and the generated images.

After open-ended exploration with discussions throughout, I tried to transition the rest of the session into time for a more focused activity connected to identity and personal stories. It was my hope that in this more focused part of the workshop, the creators could work towards creating something personally meaningful, working towards a goal or an idea through their project. To prompt this activity, I offered a handful of suggestions to spark ideas. Some of my example prompts are listed here:

- **Create an About Me Project introducing yourself.**
    - Scratch users are often familiar with 'About Me' projects, a popular starter activity on the platform.
- **Think of a favorite memory you share with someone. Recreate a scene from that memory to share with that person.**
    - I liked the idea of this activity because it encourages you to think of a specific event or person in your life and invites personal storytelling and personalizing the creation enough to send their creation to a loved one
- **Where are you (coming) from and where are you going?**
    - This prompt was inspired by the idea of "I am From" poems, like one I wrote when I was younger, and also inspired by a book, "Where We Come From," which I came across while I was working on Imagine Yourself. Both of these inspirations are discussed a more in the following section about Workshop 3.
- **The Many Me's**
    - e.g. Fill in the blanks / replace the underlined parts
      "At <u>home</u> I am ____," "At <u>school</u> I am ____," "With/to <u>my friends</u> I am ____."
    - This prompt was inspired by an identity themed virtual reality art gallery that a colleague shared with me from the Movement City Computer Clubhouse in Lawrence, MA. The idea of the activity is that there are different versions of our identities which are more prominent in different settings and parts of our lives.

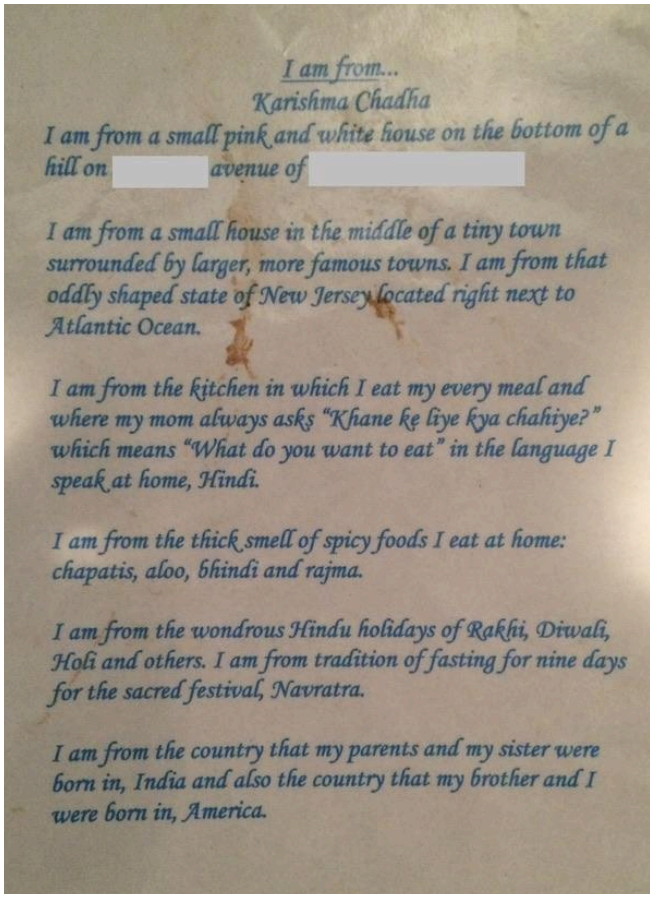
I closed out the 1-on-1 sessions with a conversation with each creator about their experience with Imagine Yourself. I discuss their experiences in the second part of this chapter.

Workshop 3 - Identity Expression & Scratch

The final group workshop took place between the two 1-on-1 sessions. It was hosted at the Flagship Clubhouse[28], an afterschool center, part of the Computer Clubhouse Network (Resnick & Rusk, 1996). At the Flagship Clubhouse, young people are regularly engaging in creative learning experiences with technology — working collaboratively, on open-ended projects related to their interests, in a spirit of play and exploration. In planning for this workshop, I worked with Tamara Fils-Aime, the clubhouse coordinator at the Flagship Clubhouse. In her role, Tamara works directly with the young people—referred to as clubhouse members in the context of the Computer Clubhouse Network. She organizes programming and activities, supports Clubhouse mentors in facilitating members' creative projects, and takes part in her own creative projects in the space. It was an honor to be a guest in this community and space, where young people engage regularly as creators, collaborators, and community members, bringing together so many of the ideas and values that inspire me.

This was the first group workshop that was focused on the theme of identity expression. As such, I decided to change the workshop introduction a bit to share more about my motivations behind creating the Imagine Yourself tool, and also to add more of a personal touch. As in previous workshops, I started a discussion to learn about the creators' experiences with Generative AI. Instead of diving into using the tool, this time, I shared a bit about my early experiences with image generators. I shared the images results I got when I tried to create a personal avatar in my first experiments with image generation tools (see Background chapter). I also shared some examples of things that have inspired me as I have been working on Imagine Yourself—especially in thinking about identity and personal stories. I shared a picture of an "I Am From" poem I had written when I was younger, echoing themes of navigating my identity as an Indian American. I also shared a photograph of a picture book I had come across, *Where We Come From* (Wilson et al., 2022), depicting the personal stories of the four authors in poetic ways and with beautiful visual imagery.

[28] https://flagshipclubhouse.org/



Photograph of "I am from" poem I wrote when I was younger (with transcribed text on the right)

I am from …
Karishma Chadha

I am from a small pink and white house on the bottom of a hill on [Redacted].

I am from a small house in the middle of a tiny town surrounded by larger, more famous towns. I am from that oddly shaped state of New Jersey located right next to Atlantic Ocean.

I am from the kitchen in which I eat my every meal and where my mom always asks "Khane ke liye kya chahiye?" which means "What do you want to eat" in the language I speak at home, Hindi.

I am from the thick smell of spicy foods I eat at home: chapatis, aloo, bhindi and rajma.

I am from the wondrous Hindu holidays of Rakhi, Diwali, Holi, and others. I am from tradition of fasting for nine days for the sacred festival, Navratra.

I am from the country that my parents and my sister were born in, India and also the country that my brother and I were born in, America.

Page text reads:

I come from my grandmother's garden, picking berries warm from the sun.
My ancestors' prayers brought us home to the land where we have always lived.
Now families dance at powwows, learn our language, and listen to the old stories once again.

Page text reads:

I come from the Chicago skyline,
chasing glowing fireflies at dusk,
borrowing stacks of books from the neighborhood library,
and reading them by flashlight at night.

Photograph and transcribed text of a page spread
in *Where We Come From* (Wilson et al., 2022)

With suggestion from Tamara, I also introduced a short new paper-based activity to invite young people to reflect on aspects of their identities before starting to explore the Imagine Yourself tool. I created an activity sheet with four prompts for young people to reflect on through text, drawings, etc. I invited them to share what they had added to their sheets, emphasizing that sharing out loud was optional.

The prompts:

- What people see when they look at me …
- What people don't know just by looking at me …
- Where (/ who / what) I come from …
- Where I am going …

Tamara also shared with me that some creators participating in the workshop had experience with Scratch while others had never used it before. From prior experience with Scratch workshops as well as the first 1-on-1 session that had taken place prior, I knew that there would be a time pressure during this 2.5 hr workshop. In planning for this workshop, I focused on ways to introduce Imagine Yourself in the context of Scratch and to introduce ways of making Scratch projects that invited personalization.

Welcome Page and Introductory Resources for Imagine Yourself

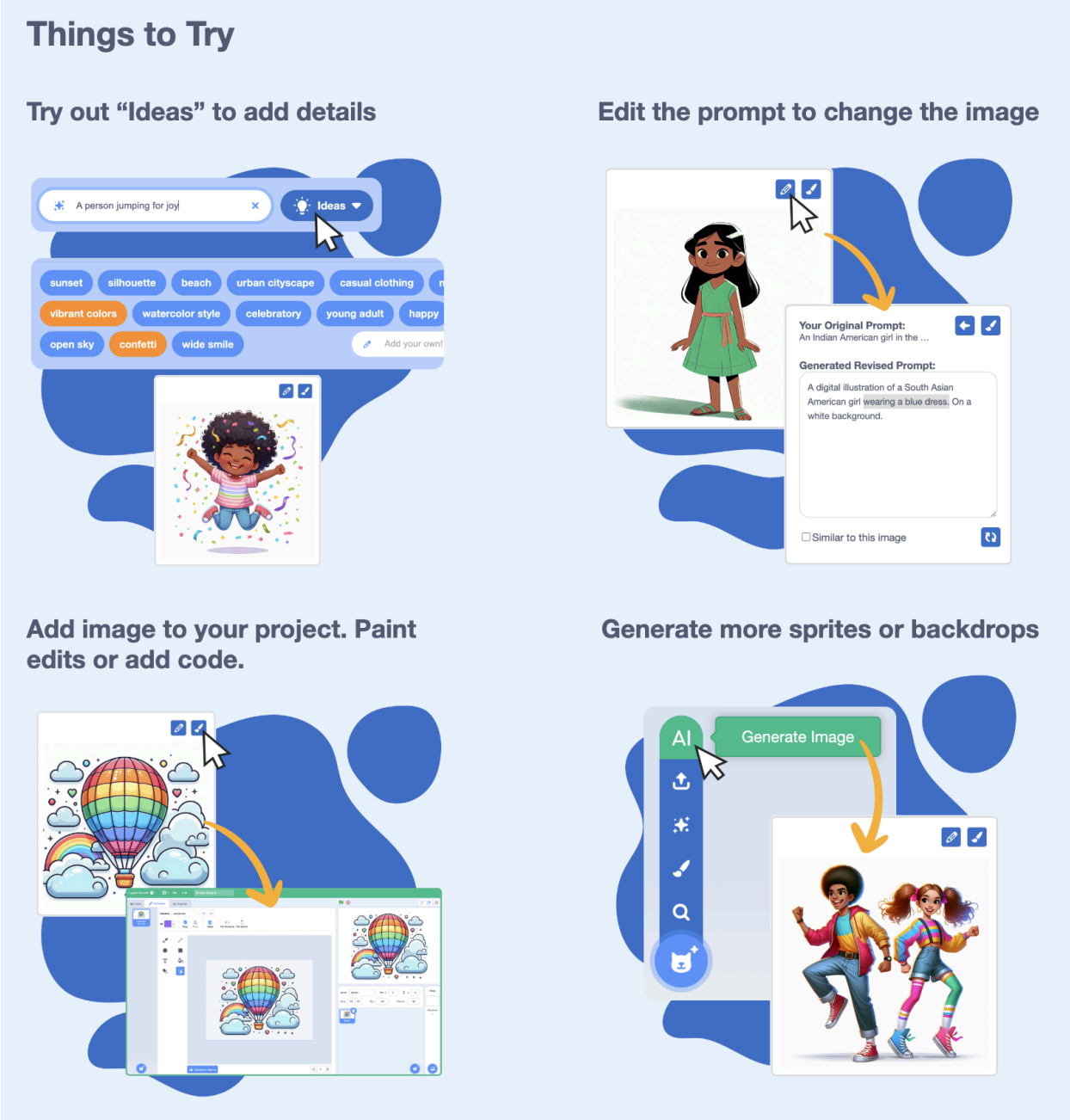
In this iteration, I focused on the welcome experience for Imagine Yourself. Previous iterations of Imagine Yourself involved opening up an experimental version of the Scratch editor and navigating to the Imagine Yourself tool by choosing a menu option for adding a new sprite or backdrop to the project. The Imagine Yourself editor acted as an additional alternative to the built in Scratch sprite library. However, I worried this experience would be confusing and overwhelming for people who had never used Scratch before—especially since I wanted to focus the first part of the workshop on open-ended explorations with the Imagine Yourself tool.

I created a new welcome landing page for the Imagine Yourself experience modeled after Scratch Lab extension landing pages.[29] Clicking on the big "Try it out!" button at the top of the page takes the user directly into the Imagine Yourself tool.

Screenshot of new Imagine Yourself welcome landing page

[29] As an example, see the landing page for the Scratch Lab "Animated Text" extension: https://lab.scratch.mit.edu/text/

In this landing page I included quick tips on things to try in the Imagine Yourself tool such as the using “ideas” feature to add details or clicking on the “Edit Prompt” button to see behind the scenes of the image and change the prompt.



Screenshot of “Things to Try” section of new Imagine Yourself landing page

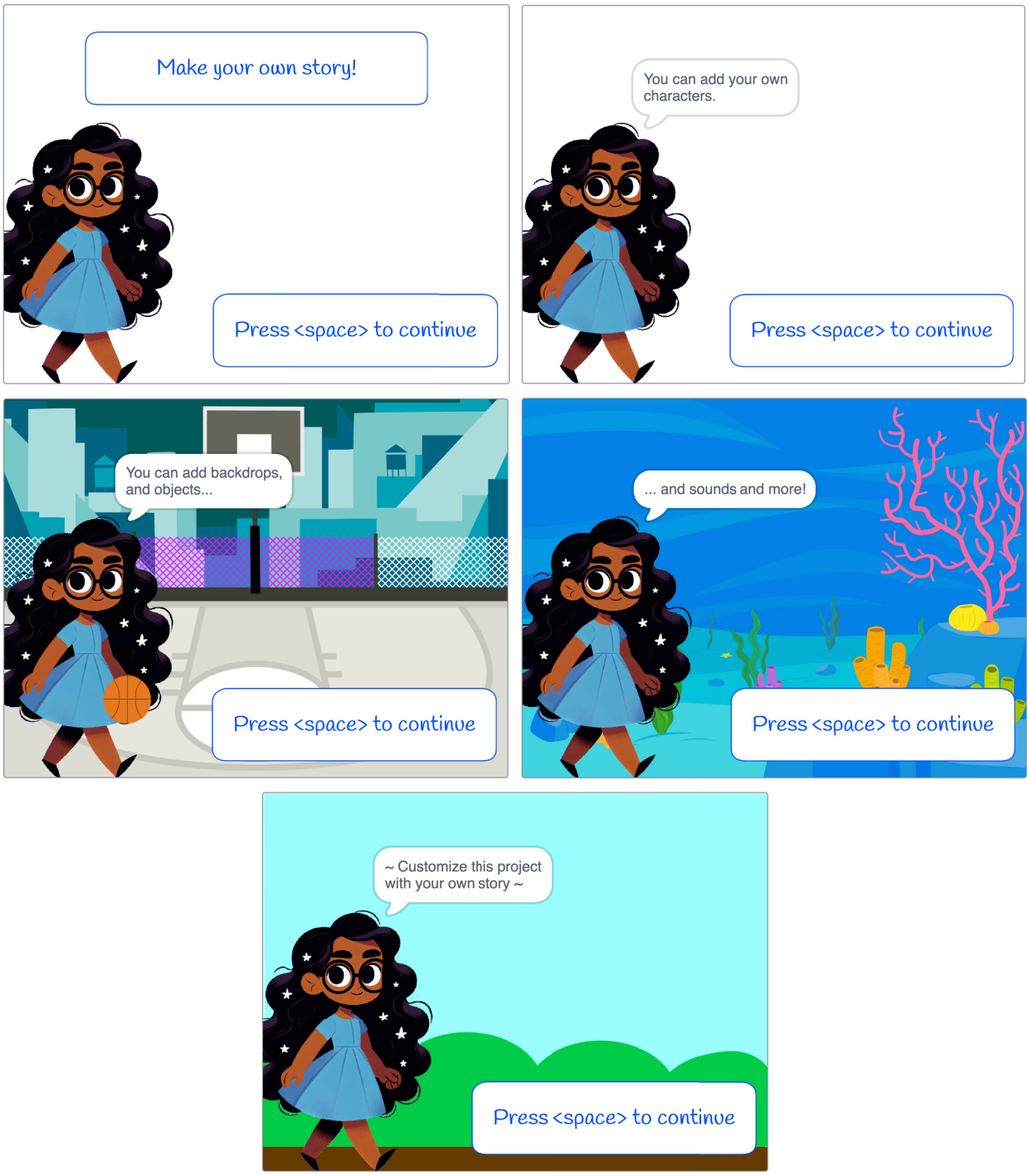
I also focused on creating a *starter project* for the Imagine Yourself experience. Starter projects are short Scratch projects highlighting particular ideas to get started with. They are designed as templates to invite young people to personalize them or bring in their own ideas. In the context of Scratch Lab extensions, starter projects aim to spark ideas for things like

games, stories, and animations while highlighting new possibilities for exploration that are enabled by that new extension.[30]

In the case of this Imagine Yourself starter project, I wanted to focus on two goals: (1) highlight creative possibilities with Scratch in a beginner friendly way and (2) provide a template example that might spark an idea for a story project and invite personalization. I created a story template project highlighting beginner friendly things to try inside of a Scratch project: add and move a character; add sounds or animate objects for the character to interact with; change scenes and more. I also designed the starter project to have a main character (a personal avatar I had generated using Imagine Yourself), inviting the user to replace it with their own!

Note: The starter project experience is designed so that replacing this template character with the user's own creation already happens by default. When the user clicks on the starter project, they are taken to the Imagine Yourself tool. If they create an image and add it to their project, that image will get added as a new costume for the currently selected sprite, replacing the template character with the user's creation while keeping the code of the story the same.

[30] Example starter projects can be found on the Scratch Website: https://scratch.mit.edu/starter-projects and on individual Scratch Lab extension pages e.g. https://lab.scratch.mit.edu/text/



Still frames from the Imagine Yourself starter project created in **Iteration B**.

Using “Surprise Me” as Default Image Generation Option

In addition to the creation of a new landing page and support materials for introducing Imagine Yourself, in Iteration B, I also switched the default image generation option in

Imagine Yourself to use the "Surprise Me" process. This decision[31] was a response to positive feedback for the "Surprise Me" process in prior workshops.

Note: When a user is iterating on a generated image through editing the prompt behind the scenes, the Dall-E 3 direct image generation process is used, so that the user has more agency in the editing process. The "Surprise Me" process is only used for the input prompts in the top-level text input box of the Imagine Yourself tool.

Experimenting with Prefixes and Suffixes

In multiple workshops and especially each of the 1-on-1 sessions—where I had opportunities to engage the creators in direct conversations about their experiences while creating—a common piece of feedback shared was about the style and mood of the images represented. Many creators commented on the cartoonish, colorful, and happy feel of the images that were being depicted in the Imagine Yourself tool—one creator even referred to the art style in terms of a "condescending tone", which she said felt characteristic of AI (Creator 2, 1-on-1 Session). She expressed the desire to see more diversity in the moods and emotions being represented in the images as well as a more serious art style, characterized by more muted colors. Many creators used prompts to modify the style of images they were getting by default, interested in depicting images with prompts including `"on a rainy dark day,"` `"foggy and stormy,"` `"muted dark colors,"` `"claustropobic feel,"` and `"dystopia."`

Prompted by this feedback, I came to reflect on a design challenge in this work. I was designing a system for a wider age range than the one represented in workshops. The typical age range on Scratch is 8-16. As noted in observations in the previous chapter, because of stereotypes in its output, guiding the generative AI technology to generate images that are appropriate for a younger audience also trends its outputs towards bright colors and smiley faces—what I've referred to as "candy-coated depictions." Creators in workshops expressed the desire to see more realistic images with muted colors or representing more of a range of emotions.

The challenges creators experienced trying to generate an art style that resonated with them led to an incomplete understanding of their experience with Imagine Yourself more broadly. Creator 2 spent some time exploring ways to create content that resonated with her. She focused on making the art seem more realistic. She appreciated the challenges of making content that was appropriate for young people, but critiqued the underlying AI system for generating images that depicted an overly-saturated, brightly colored view of the world. She shared that the art style of the images reflected the condescending tone of generative AI tools—feeling that AI reduced the concept of childhood to happy feelings and bright colors.

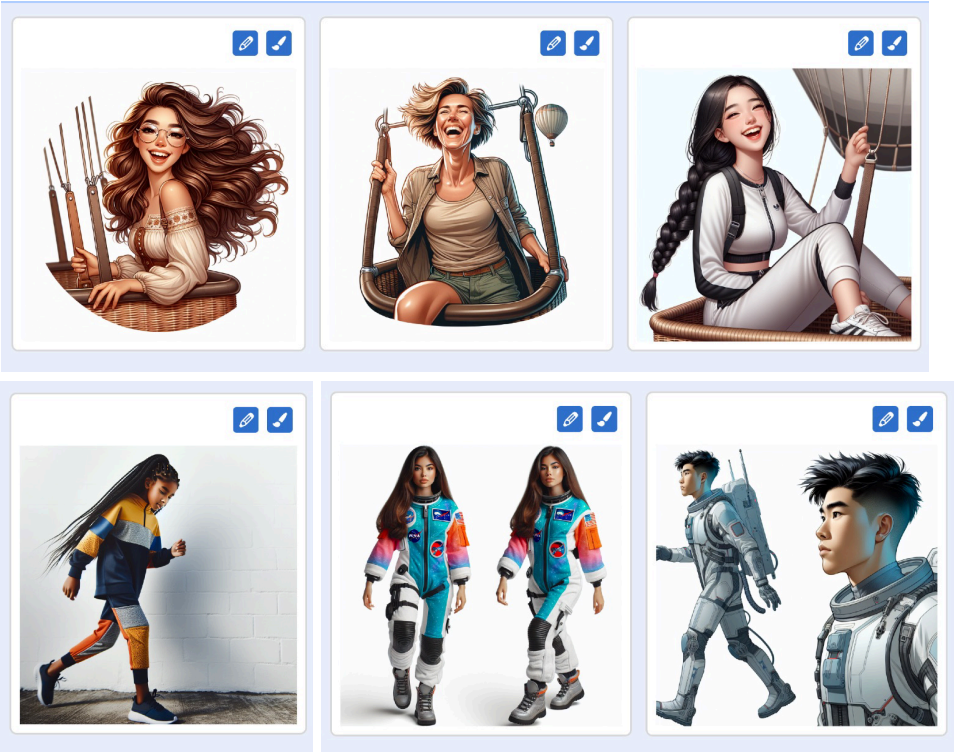
What this creator experienced is an artifact of the "appropriate for children" suffix I had added to the system. It is an artifact of how the concept of "appropriate for children" is connected to the training data that Dall-E 3 was built upon. Though this is a changeable aspect of the Imagine Yourself design, my attempted iterations towards removing or replacing this phrase were not promising.

Even in the earliest iterations of Imagine Yourself, I was not sure how I could address the issue of getting images from Dall-E 3 that felt inappropriate for my audience (see Observation #10 from the previous chapter). I never liked the phrase "appropriate for children," feeling that

[31] This decision was made possible by the increased image generation rate limits from paying OpenAI more money. This is discussed in more detail in a footnote in (Attempted) Iteration A, above.

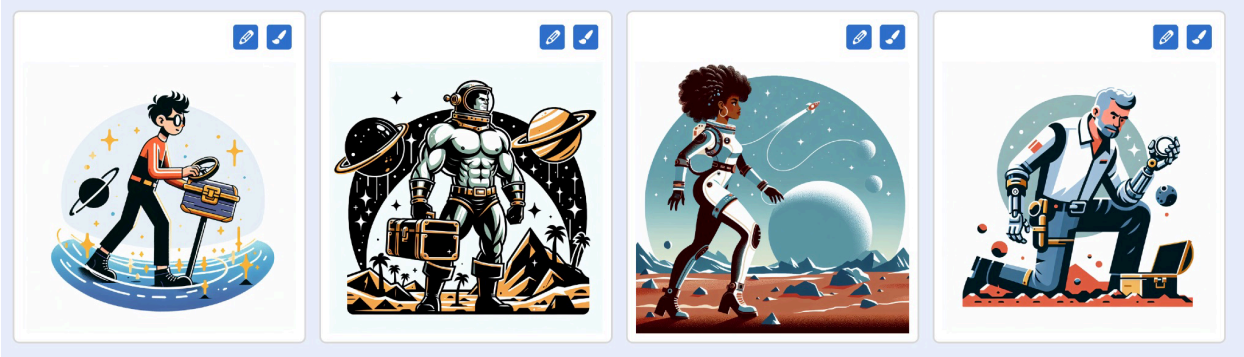
the phrase neither invited belonging nor achieved much towards the goal of making the generated images safer for young people. In my iterations I tried altogether removing this suffix or finding alternative phrases (e.g. “family-friendly”). However, I found that these changes had other unintended consequences on the image output.

Removing or replacing the suffix “Appropriate for children” led to seeing more depictions of people that felt too detailed—images that felt like they were depicting a very specific, sometimes photorealistic, person rather than a more general avatar-like image that could be representative of the creator.



Examples of generated images that felt too realistic and specific

I also tried replacing the prefix “a digital illustration of.” This prefix sometimes yielded images that had depictions of the prompted subject as a drawing on an iPad. I also wondered whether it was the combination of this prefix with the “appropriate for children” suffix that was leading to the issues with the candy-coated depictions in many of the images. Because I still did not want photorealistic images in the outputs, I tried the prefix “vector style image.” However, I found that the resulting images tended to constrain the images to a very particular art style like the star-filled artwork in the images below.



Example images generated when I tried the prefix “vector image style”

Ultimately I abandoned this exploration due to a lack of time to more carefully explore and test alternative potential solutions such as removing the prefixes and suffixes altogether or even offering ways for users’ to control which ones the underlying system uses.

While the first half of this chapter provided general context for each of the workshops as well as discussions of design iterations that occurred between them, the second half of this chapter details creators’ experiences in the Imagine Yourself workshops.

Workshop Observations and Feedback

Trends Across Workshops and 1-on-1 Sessions

Across all of the workshops, creators interacted with the Imagine Yourself tool in a variety of ways. They explored the boundaries of the new technology by coming up with increasingly specific prompts, often combining “random” concepts together. They systematically and iteratively tried to create representations of themselves, their families, or their interests. In some cases they tried to express emotions or feelings—such as `“weight of the world”` and `“Man focused on his goal in a sea of chaos.”` Through notes from workshop conversations and observations, as well as digitally collected prompt and image data from the workshops, I categorize some trends and patterns I observed in the creators’ initial explorations with Imagine Yourself below.

Exploration

Exploring the Boundaries

Some creators started off by exploring the boundaries of the new technology they were presented with. One such creator used the tool to explore what AI is. They used prompts such as `“are you human,”` `“are you real,”` and `“artificial intelligence”` to probe the system.

Others explored the capabilities of the tool, ranging from trying to visualize high level abstract concepts such as `“a billion dollar business idea”` or exploring how the tool would react with highly specific prompts. One such creator articulated,“I was trying to see if I could type a whole sentence and see if it could generate something crazy” (Creator 3, Group Workshop). This idea of trying to get the AI to generate something weird, wacky, or crazy was an avenue for exploration that many of the creators pursued. One set of creators who decided

to work as a pair were very excited to share a prompt that they spent a lot of time crafting, “[We] tried to do ‘dominos that are doing jumping jacks that are dividing fractions.’ [We] didn’t get anything that we liked. It took a while to think about this one” (Creator 4, Group Workshop).

Yet another creator described his experience and impression of AI as a whole, “I got real specific with sentences, scenarios, settings. It’s scary, in a good way. Like wow. Scary in the sense of how it all comes together, how it works, the generation of it. For example I am a reptile owner, so I put ‘bearded dragons going out shopping,’ and it did it in like 2 seconds” (Creator 5, Group Workshop).

Recreating Memes

As generative AI tools have gotten popular, early users have posted their explorations on social media. Some of these explorations have gone viral, garnering lots of enthusiasm from other online users. One such example is the 'Make it More' trend (Fear, 2023), in which users prompt image generators to create an image embodying an abstract concept and pushing the tool to generate something that is "even more" of that concept.

Screenshot of a post on X.com from user @venturetwins describing the "make it more" AI image generation trend (Justine Moore [@venturetwins], 2023)

One creator explored this same style of prompting, progressing between the prompts "`punk rock penguin,`" "`even more punk rock penguin,`" and "`the most punk rock penguin`" sequentially.

Another meme post from the site 9gag depicts an example of an image generator yielding an unexpected image featuring Spider-Man in response to a request for "`pineapple fighting a pepperoni pizza`" ([@mrtrips211], 2024). Another creator tried this out for themselves with Imagine Yourself, using the prompt "`pizza fighting a pineapple.`"

Screenshot of a post on 9gag from user @mrtrips211 depicting a generative AI meme ([@mrtrips211], 2024)

In both cases, these trends indicate general fascination with observing how image generators observe different concepts.

Specific Vision

While many creators spent time exploring the capabilities of the image generator, others tried creating representations of a specific subject they had in mind. For some it was a type of character or a creature, inspired by one of the art elements from the LikeMinds tool. Seeing one of the sky backdrops in the LikeMinds tool, one creator wanted to create an image of a `“flying armored warrior”` in a specific pose and art style.

Another creator was fascinated with learning that stars can turn blue. He spent time creating scenes with a blue sun: `“Mystical stetting in an enchanted forest, blue sun, magical creatures watching,”` `“purple planet with blue sun in the sky, astronauts floating.”`

Self-Representation

Some creators were particularly interested in exploring representations of themselves, their families, their cultures, or their homes. Creators used the tool to create depictions of their blended families, a digital “headshot” of themselves, or representations of themselves with imagined aspects — from fantastical representations such as `“platformer protagonist”` and `“mermaid”` or imagined attributes such as `“rainbow hair.”`

Below I share some trends I noted from creators’ prompts in representing themselves or loved ones.

Descriptors creators’ used in prompts for themselves :
Each of the following are examples directly quoted from creators’ prompts. In some cases I include just a portion of the creator’s prompt pertaining to the relevant categorized

observation of the common themes that were used across creators' prompting explorations. For some of the prompts I've quoted here, I've redacted age from the prompt to preserve anonymity. In those cases I've marked where the prompt referenced a specific description of age using `[AGE]`.

- Race, Ethnicity, or Nationality
    - `"Egyptian young girl"`
    - `"little Indian girl that turns into a cat when nervous"`
    - `"Really buff black guy"`
    - `"Black teenage girl"`
    - `"black princess"`
    - `"[AGE] old white woman snuggling [AGE] old son"`
    - `"witch of African descent"`
    - `"caribbean american 8-bit pixelated platformer protagonist"`
    - `"Native american lady dancing sundance realistic or anime"`
- Religion
    - `"Muslim girl wearing hijab"`
- Age
    - Many prompts referenced ages, however I've redacted these references as mentioned above.
- Gender, Gender expression
    - `"abstract art style non-binary person with green hair"`
- Sexual Orientation
    - `"lesbian family"`
- Height
    - `"6ft man with green eyes dark brown hair"`
    - relative height — e.g. `"short"`, `"tall"`
- Hair:
    - `"evil curly-haired witch of African descent forest witch with lots of hair"`
    - Hair length
        - `"shoulder length brown hair"`
        - `"black forest witch with lots of hair"`
    - Hair color
        - `"green hair"`
        - `"a blonde hair tall man"`
    - Hair style
        - `"yellow braids"`
        - `"6'0 black man with dreads and glasses"`
    - Hair type
        - `"curly hair"`
        - `"natural curly hair"`
        - `"black mermaid with big coily hair"`
- Eye color (`"blue eyes"`)
- Tattoos (`"women and man with tattoos"`)
- Description of family
    - `"a Black father with 4 sons"`
    - `"Black Family with 4 sons surreal art style"`
    - `"native american family in a drum circle"`

    - “family with mom age [AGE], children daughter [AGE], son [AGE], daughter [AGE], son [AGE]”
    - “man eating ice cream while sitting in front of a sunset while his kids play”
    -
- Skin tone
    - “light black couple”
    - “a dark skin black woman”
    - “Man with lightbrown skin and green eyes kissing light brown skin women with dark eyes. curly hair”
    - black women with light ginger hair and light brown skin
    - “caramel skin black women”
- Imagined Attribute
    - “black women with rainbow hair on a boat”
- Descriptions of other facial features:
    - “big nose”
- Weight, Body type
    - “Poison Ivy as a black woman with natural curly hair and a plus size body”
    - “A curvy black woman”
- Abstract Descriptions:
    - “girl silhouette with umbrella”
- Descriptions of interests:
    - “A curvy black woman gardening in her horticultural homestead”
- Fashion Aesthetics
    - “a goth woman with shoulder length green hair standing next to man with black hair and a curled mustache”
- Depictions of Emotions, Self Perception
    - “reflection of ones self feeling defeated”
    - “ripped apart in pain and agony”
    - “Strong Black Man”
    - “black man meditating on a beach during sunset while his sons prepare for dinner”
- Descriptions of Home
    - “brownstone apartment living room”
    - “nic e mexicanm kitchgen” (nice mexican kitchen)

As exemplified above, creators tried various techniques and descriptors for representing themselves.

I found it particularly interesting that one creator—who had been very focused on creating a personal Scratch project—chose to represent herself abstractly through a silhouette, prompting the tool for “girl silhouette with umbrella” and “silhouette of girl with balloon.” When asked to share her thought process for using a silhouette to represent herself, she shared (the quote below is transcribed from an audio recording):

> I think I just didn’t want to deal with the hassle of making something that looks like me. Not necessarily in a bad way because I think the tool can’t do it but just because I knew that it would take more time. And I think it’s difficult to—like nobody looks the

> same—so I think it's like difficult … even though it's generic … like even when we searched up Tom Brady, it didn't look like him, so I thought having a silhouette—it proves the point that I need it to. So I just went with that. (Creator 6, 1-on-1 session)

Many other creators iterated on more specific representations, and shared reflections on their experiences. The following section elaborates on creators' experiences and feedback while using Imagine Yourself.

Creators' Experiences with Imagine Yourself

Creators shared their experiences creating with Imagine Yourself in whole group discussions and conversations in 1-on-1 sessions.

Experiences with Self-Representation

One creator described both her challenge and delight in using the tool. She expressed the challenges she faced in trying to create images that depicted her as a Black, interracial woman, with lighter brown skin tone. She noted:

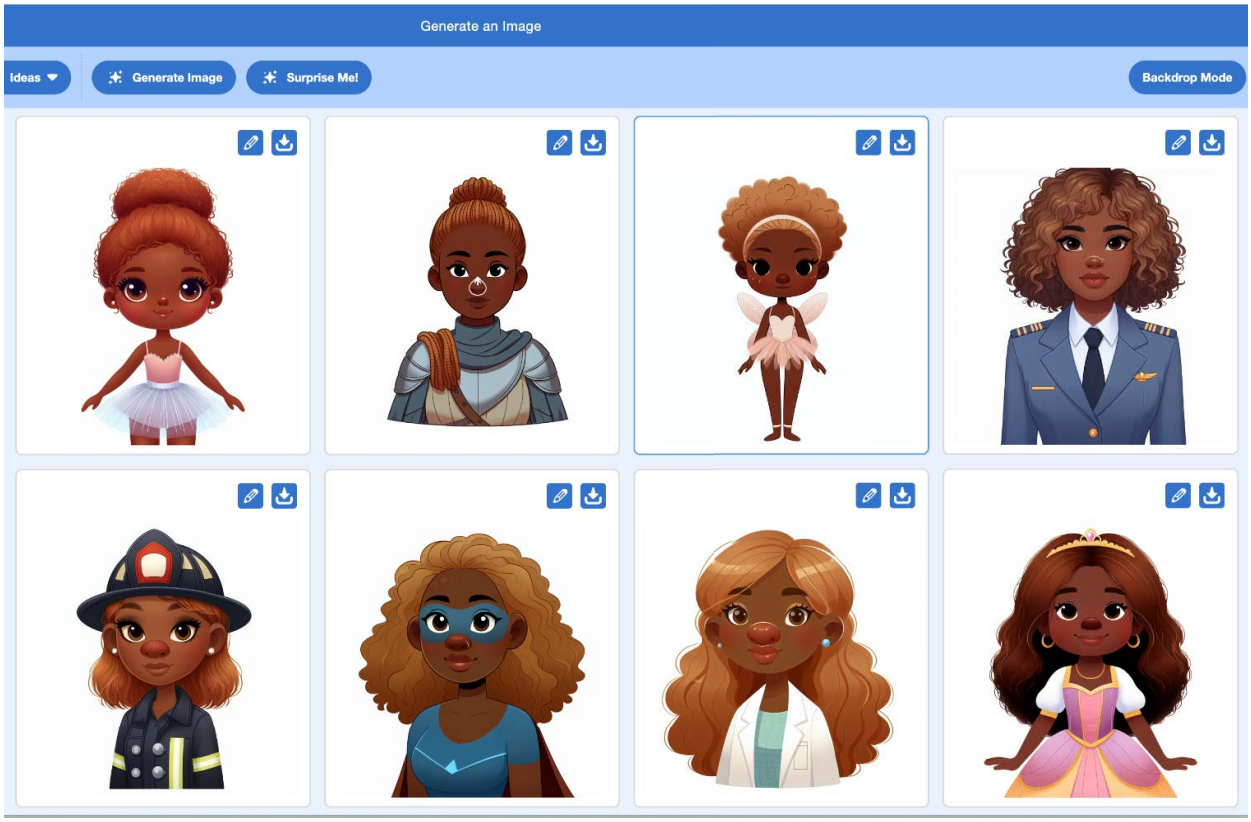
> I notice when you do a Black woman it automatically does a dark skinned woman. I thought it was pretty but I was trying to figure out how to do an interracial woman. When I tried to do light skin it made the hair straight and showed me White features, but the right skin tone. I couldn't find the balance. (Creator 7, Group Workshop)

It seemed that this creator was coming up against two types of issues, (1) she was encountering systemic biases of the underlying AI tool that Black woman implies darker skin or lighter skin implies straighter hair, and (2) she was coming up against the limitations of limited agency in the modifications she could make with this tool, e.g. wanting to change just the skin tone while keeping other features the same.

Photograph of Creator 7 sharing what she was trying in Imagine Yourself

She also highlighted her approach to experimenting with representations using the Imagine Yourself tool and her delight in seeing a wide variety of representations. She shared:

> I put the same description of a character several times, and it was really cool. I thought it was interesting that it made a superhero, and all these different career fields. With the 'surprise me' I got the different careers and hair and stuff. It was actually really cool. It looked like a lot of different people. When I did 'generate' it did a more basic image. I think 'surprise me' is better because it's more creative. If you just do the 'generate image,' it's more basic. It just did the person in a casual outfit. But now I have a doctor, a knight in shining armor, and a chef. (Creator 7, Group Workshop)



Screenshot of images generated by Creator 7

This creator's experience echoed some of the observations I had noted while experimenting with the design of the Imagine Yourself tool. In an early informal test, when attempting to use the Imagine Yourself tool with friends to create avatar versions of ourselves, some of us noted difficulties in changing specific attributes of the images depicted. My friend, who identifies as a biracial woman, had trouble creating an avatar that matched both her Indian identity as well as her lighter skin tone—which she often refers to as "beige," a mixture of brown and white. It took her over 68 image generations to land on an avatar image that she felt represented her. However, we reflected on the fact that some of the earlier generated images would have sufficed if she had a way to change the skin tone of the avatar image more easily.

These stories highlight the idea that while interacting with generative AI can make some types of creation more accessible (e.g. quickly creating images without needing extensive skills or prior experience with digital art creation tools), *text-based prompting* is not always the best way to create, remix, and personalize the visual representations we imagine. In future iterations of the Imagine Yourself tool, I would imagine an experience where text is just one of many modes of interacting and editing with the images, other tools such as within the paint editor, would make it easier to hand-modify aspects of the visual representation.

Perceptions of Creativity

While Creator 7 (highlighted above) described the behavior of the "Surprise Me" feature as interesting and "cool", she also attributed it with being "more creative". This idea of creativity (especially being attributed to the machine) came up in conversation with several creators across the workshops. Upon final reflections at the end of the workshops, many creators mentioned that they liked using the AI tool, because they felt that they were not "creative" or

“artistic,” and they found that this was a way to express themselves with art. One creator reflected, “I liked the generative AI tool because I’m not artistic. I liked being able to [keep working on] my ideas, feeling like I wouldn’t be able to make that” (Creator 8, Group Workshop). Another creator mentioned “I liked AI because I was being creative and I’m not a creative person” (Creator 9, Group Workshop).

A conversation with one creator stood out to me. This creator participated in a later workshop, working with the Imagine Yourself tool inside of Scratch. He expressed his curiosity and critical concerns about generative AI more broadly before starting the workshop. He asked questions about how his data was being collected, where the AI obtains the artwork it’s trained on, and whether AI developers have permission from the original artists of the images being used for training the AI model. We had an open conversation about these concerns, how various AI companies are handling or attempting to address these issues, where these efforts have been falling short, and how little is known about how these models fully work.

During the workshop this creator explored prompting the Imagine Yourself tool with names of artists, art styles, video games, game characters, and game styles. In some cases he tried combining different art styles in one image (e.g. `“ligne claire 80s city pop”`). In conversations with him, I learned that he was an artist himself, he liked to draw images and he was interested in exploring how the AI represented particular art styles or the work of artists he was familiar with. In the final group reflections about the creators’ experiences during the Imagine Yourself workshop, he shared conflicting feelings about using AI to make art. **Note:** I paraphrase this discussion here since I was not able to take verbatim notes during this portion of the workshop. He shared that creating with AI made him feel “insecure” and “depressed'' as an artist, because there’s a tool that can create such detailed images so quickly. He further described that, while on the one hand, he felt that there isn’t really “intention” behind the generated art, on the other hand, he could see how people may still use pieces of generated art to make things that have intent or meaning. I appreciated his openness and vulnerability in sharing his personal experience, conflicted feelings, and criticisms of generative AI technology (Reflection of experience paraphrased from conversation with Creator 10, Group Workshop).

This was far from the only instance of young people sharing concerns, open questions, and criticisms of generative AI, especially as it pertains to the unethical theft from artists. One creator shared her hesitation of making personal things with AI, and how she thought about the relationship of AI and art (the quote below is transcribed from an audio recording):

> I don’t know … because I feel like … It’s weird making personal things with AI, right? Because like, the part of being personal is human—the humanity part of it. And then with AI—and I know a lot of AI enthusiasts and I am not one of those ‘people who are like AI is going to destroy the world’ people—I use AI a lot. But I think the one thing about AI that makes it inherently AI is that it’s not human. And like using AI to create personal things kind of takes the humanity out of it and it makes me feel like really like [pause] I don’t know … it’s just me.
> …
> Because I’m also a musician. And I think a lot of people in the art, artists, musicians, writers, have different feelings towards AI than those people who aren’t specifically into those because …
> …
> I play the flute, right, and I went to camp a few weeks before this, and one of the funniest things — my teacher said, ‘your flute is your expression stick.’ It’s an extension

> of you and you use the flute to like—you're not sharing music. You're sharing emotion. And the medium is music. And like, it's human-to-human connection, composer-to-player, player-to-audience, all of those. Composer-to-audience, you're the vessel. And then it's like … we interact with that sort of stuff on a daily basis. So we're very attuned to—oh this is human. This should be reserved for humans. AI shouldn't intrude on it. (Creator 2, 1-on-1 Session)

In creating new tools for supporting young people's personal expression, it is incredibly crucial to honor, respect, and center their voices and perspectives in the design process. Young people are already thinking critically about these systems, and they are reflecting on what is important to them in terms of personal and creative expression.

It has been remarkable to see the variety of ways creators used Imagine Yourself and to learn about their initial experiences with it. Engaging in these workshops have primarily left me with more reflections and open questions about the nature of creating and engaging with AI to create personal meaning. These workshops have also led me to reflect on my role as a researcher and designer. The following chapter concludes this thesis with some of these final reflections.

Chapter 7. Reflections

Working on this project has been very different from many of my past experiences. This was my first time writing about topics that feel deeply personal and dear to me. This was my first time writing about ideas pertaining to systemic social issues. This was also my first time engaging in a *solo* process of creating, from exploring design ideas grounded by my core values to building a tool inspired by these ideas and trying it out in workshops. My past experiences have largely been collaborative projects, where I contributed to a team, but even more importantly, I got to learn from the perspectives and expertise of others. In this project, I wore many hats. I was a researcher, an engineer, a UI designer, a workshop designer, a storyteller, a facilitator, and more. Trying to balance these roles has led me to some interesting reflections about the process.

As a software engineer, I found it fun to explore natural language as a new medium to tinker with—using text-based prompts alongside code (a medium I am much more familiar and comfortable with). When considering a design challenge from this perspective, it was interesting to explore possible solutions from the lenses of both of these mediums—should I tweak the system prompt or should I come up with a code-based solution to the problem at hand, or a combination of both?

As a designer I found it a slow, murky, and at times frustrating process to tinker with the AI. While I had design goals in mind, often the design tradeoffs I was weighing were not clear. Many attempted iterations in response to feedback from creators in workshops have ultimately remained unfinished or inconclusive because iterating with AI felt both fast and slow. Iterations were *fast* in that interacting with Generative AI is conducive to rapid prototyping. However, they were also *slow* because honing and tuning the tool towards design goals felt hard to measure and evaluate. Unlike with code, it was difficult to tell—well, did it work?

As someone fascinated by the nuances of language and culture, I found myself infinitely intrigued by exploring the inputs and outputs of the image generator, curious to see how my System prompt iterations and my image prompts would be perceived and interpreted by the machine as well as categorizing the trends I was observing in the depictions. I have endless documentation on images I generated. Many of these trends are noted above. And yet, I still feel I have barely scratched the surface of understanding AI image generation.

As a workshop and learning experience designer, I was humbled to put Imagine Yourself in front of creators and be able to have rich conversations with them about their hopes and concerns pertaining to the rapid proliferation of the technology. It has been my honor and privilege to be able to share brief glimpses into their stories here.

As a writer, I have found it both challenging and rewarding to reflect on these experiences and the personal stories, perspectives, and curiosities that have led me here. I have also found it incredibly humbling and rewarding to try and share the experiences of the creators in the Imagine Yourself workshops.

While I have learned a lot from my explorations in designing Imagine Yourself, I continue to reflect on many of the same questions I started out with. Some of these point to other

interesting areas of research or exploration, while others have been the kinds of questions and concerns that require more careful consideration when working with generative AI.

I break apart this chapter into two sections. In the first, I discuss the challenges and considerations of working with AI that require careful consideration but were out of scope for the explorations documented in this thesis. These are issues and challenges that I feel are crucial to consider as generative AI technological development advances. In the second part of this chapter, I discuss potential future directions to explore towards supporting young people's meaningful personal expression with generative AI.

Broader Issues and Challenges to Consider

As I have alluded to in earlier parts of this document, while there are opportunities for exploring ways to incorporate AI into human creative practices, there are also many challenges and issues with the development of AI, working with AI as a medium, and the implications of widespread AI usage for humanity, and especially for communities who are often marginalized, excluded, or exploited by the development of new technologies.

Issues with AI range from issues of cost and access, environmental impact, effects on marginalized and non-dominant communities, and the exploitation of hidden laborers. While I chose to focus, in this thesis, on exploring particular challenges such as the issues of systemic bias and lack of inclusive representation within image generation tools, I will briefly discuss each of these other issues we must contend with as researchers, designers, and educators.

Cost

As with the advent of any new technology, cost is a primary barrier for people to access, explore, and assess the new technologies. Early adopters of new technologies tend to be those in the fields closest to the development of the technology, and tend to have more privilege and means of accessing it. In the case of Generative AI technology, cost is a barrier to entry in both the development of the technology as well as its use. This results in inequities both in who gets to develop skills with the new technology, but also in whose perspectives and voices shape the dialogue around the technology. These privileged perspectives and voices greatly influence further technological developments in this area.

As with all other technologies, generative AI tools will become cheaper to use and develop and they will become more ubiquitous in public life. However, the directions this development will take depends heavily on who has access. Developers of AI tools and technology must focus on centering communities that have been furthest from their development thus far.

Privacy & Terms of Use

In selecting the AI technology to work with, I wanted to ensure that the underlying tools I was building with were not using users' data to train and develop their AI models. I did not want young people's creations to be exploited and extracted towards the further training and development of AI models. The OpenAI API[32] does not use user data to train its models (OpenAI, 2023c). The terms of use of many of these tools also prohibit the use of these tools with young children (often restricted under the age of 13). This makes research on developing

[32] While the OpenAI API does not use user data to train its models, use of ChatGPT does.

AI tools for young people incredibly difficult. For safety considerations, in my work, I preferred to try out the Imagine Yourself tool with older teenagers and adults. However, as discussed in previous chapters, this presented a challenge between imagining and designing for a younger audience than I was able to get feedback from. These were all considerations I was contending with during the process of designing and testing the Imagine Yourself tool. Future development and AI research must consider how to make these models safer to use.[33]

Environmental Impact

Generative AI is incredibly resource intensive to train and use. Research on the environmental impacts of generative AI range from studying the comparatively high energy consumption of the special hardware used to train and run large language models as well as looking at the environmental impacts of developing widespread user-facing applications using generative AI technology (Berthelot et al., 2024; de Vries, 2023). Some researchers also examine the potential impact of image generators in particular. This is due to their significant consumer appeal—which could lead to greater widespread adoption and use—significantly impacting global energy consumption (Utz & DiPaola, 2023).

As a researcher and developer working on a generative AI tool, I am confronted by the tensions between contributing to greater energy and resource consumption by engaging in these design explorations and the responsibility and opportunity I feel to try and leverage my privilege and power as a researcher and technologist towards exploring inclusive and nurturing approaches to creating and designing with AI.

Exploitation

As I was working to support young people's personal expression with generative AI, I was constantly contending with using a technology that was furthering systemic inequities in its design and development. The development of technology, especially AI technology, impacts and exploits global communities in a variety of ways, from mining for cobalt in the Congo (Al Jazeera, 2023; Crawford & Joler, 2018; Smith & Mantz, 2006) to communities that are the site for dumping of tons of E-Waste and hazardous materials (Baldé et al., 2024), to the hidden labor used to train, hand label, and moderate the data used to train these AI systems (Perrigo, 2023). Kate Crawford and Vladan Joler's *Anatomy of an AI System* does a wonderful job of diving into the AI development pipeline for systems such as the Amazon Echo.

Beyond the broader systemic impacts of technological development more generally, Generative AI is exploitative, by nature, due to its unfettered theft of the work of countless artists, storytellers, and creators in the training of their underlying models. These models work because humans have been creating and sharing art for centuries. For many creators, their work is not only a means of livelihood but also a way for them to share and express their voices, stories, histories, and ideas. When for-profit companies take and redistribute without permission, this is a system of exploitation that furthers inequities and grossly imbalanced power structures. As a creator who seeks to center the voices of young people who have been excluded from technology and education, I do not wish to contribute to the development of systems that contribute to harms imposed on these same communities of young people. This has consistently been a tension for me to navigate, evaluate and reflect on, and I try to continue to reflect on how I can engage with these issues in ways that stay true to my values.

[33] See appendix A for more information on generative AI models and the use of inappropriate content in their underlying training data.

While I have provided a brief discussion of each of these topics, they can be explored in much more detail and are crucial considerations for responsible development with generative AI. In the second half of this chapter, however, I explore possibilities for future directions pertaining to fostering young people's personal expression with generative AI.

Future Directions

A main focus of this work has been to invite a tinkerable and iterative process of creating with Generative AI. While the Imagine Yourself tool offers some opportunities for iteration, young people have expressed difficulties in modifying specific aspects of their images. Observing parallels between young people's iterations on generated images and my own iterations with the system prompt for the "Surprise Me" behavior in the Imagine Yourself tool, I have noted that tinkering with generative AI is incredibly challenging. Language as the sole medium for expression does not allow for precise control while editing, which can make it hard to express a particular idea or to make a specific change. As creators in Imagine Yourself workshops have expressed, it was difficult to edit specific aspects of their images such as changing the skin tone of a character. While Imagine Yourself is built on Scratch, which offers a paint editor for young people to remix and draw their own images, editing and modifying the bitmap images generated by most text-to-image generators is still difficult, especially for beginners.

Future considerations for inviting iteration with generative AI should explore new ways to offer more fine grained control over the creation process—inviting a sense of agency while creating. Below are ideas to continue to explore, stemming from this Imagine Yourself design exploration. While many of these ideas pertain more specifically to image generation, similar principles can be applied to creation with other types of generative AI.

Supporting Agency while Creating

Exploring intermediate representations

One possible exploration entails building a system with generative AI that generates modular and composeable pieces instead of "complete" compositions. In the image generation context, text-to-*vector* image generation is one example of such a system. It offers an *intermediate representation*—the shapes generated as part of the vector image are separate components that can be manipulated. Editing can be done through prompting or through direct manipulation of the individual components.

Another example could be a system that offers ways to edit and track parts of the composition. The story helper tool was an exploration towards this goal, instead of generating a whole scene, I was using the AI to generate individual components of the scene to compose or modify myself. Newer developments in generative AI tools offer features like "character weights" — a Midjourney feature designed to allow the creator to keep one part of the image the same (the main subject of the image) while changing the other components across iterations (background, art style, etc.). This feature aims to address a need expressed by creators to be able to generate consistent characters across multiple images. This would be instrumental in creating animations and stories.

Opening up *more* Black Boxes

As Resnick and Silverman noted, it is important to "choose black boxes carefully" (Resnick & Silverman, 2005). This does not mean eliminating every possible black box from the system, but instead opening up the ones that enable the exploration of the relevant powerful idea. This principle brings up many open questions in the context of Imagine Yourself. For example is it okay that some parts of the system are transparent but not editable?

In the Imagine Yourself tool, young people can see and edit the revised prompt generated by the underlying system. However, the current system has two design flaws. There are parts of the prompt that are visible but unchangeable, leading to both a poor user experience and undermining the user's agency over the system. Additionally, while I aimed to open up a black box by revealing the transformed user input prompt, the parts of the Imagine Yourself system that were responsible for these transformations still remain hidden to the user, the same way the inner workings of Dall-E 3 remain hidden to me.

Imagine Yourself adds prefixes and suffixes to every prompt and the "Surprise Me" process uses an elaborate system prompt to add details to the user's prompt. These invisible processes of the Imagine Yourself system can certainly be inferred by the outputs of the process, but the process itself remains invisible. In cases where the user removes or modifies these added prefixes or suffixes, they are added back to the prompt by the underlying system, beyond the creator's control. In workshops, creators repeatedly tried to remove these prefixes and suffixes. Often removing the phrase "Appropriate for children" or modifying the prefix "A digital illustration of" to change the art style of the generated images to be more "realistic." Unfortunately even while editing these parts of the revised prompts, the same prefixes and suffixes are added back by Imagine Yourself, leading to potential frustrations or feeling a lack of agency over the iteration process.

I had a parallel experience as the designer of the Imagine Yourself system, encountering invisible processes of the underlying language model or features that I could not *fully* turn off. These experiences led to challenges, confusion, and frustrations while attempting to design towards a specific goal.

In writing about my iterations of Imagine Yourself, I have come to realize that in attempting to open up black boxes for young creators, I was turning off and rewriting parts of the underlying OpenAI system so that I could have more control over uncovering black boxes for the user. I was removing the Dall-E 3 prompt revision process to write my own "Surprise Me" revision process. I revealed parts of this new prompt transformation process to the users of my system. However, since I am not an experienced AI researcher myself, the technical quality of my early explorations do not match the caliber of the teams of people contributing to the underlying Dall-E 3/GPT system I depend on. These underlying systems should be designed to open up more black boxes to give creators the freedom and agency to imagine, play, and create.

Input into the Process, Not just the Result

Related to uncovering black boxes, young people should have input into the underlying processes of the system, uncovering some of the automatic "magic" of generative AI systems. It has felt important to me that young people—especially those who are often othered by representations within media and technology—feel seen within the Imagine Yourself tool without having to expend additional efforts as compared to their peers who are more readily

supported and represented by these same media and tools. As such, in designing the experience of generating images with the “Surprise Me” process, I faced a persistent tension between wanting young people to see themselves with Imagine Yourself and offering them the tools to invite iteration and personalization. The system prompt behind the “Surprise Me” process is a form of invisible “magic”.

Young people should be able to see not only the output of the system's revisions but also the process for making those revisions. Furthermore, they should be able to control the parameters of this transformation process. This is another black box that can be further opened up. Once again I come back to the idea that black boxes should be chosen carefully (Resnick & Silverman, 2005). Opening up some of these boxes can expose powerful ideas for young people to encounter and play with, while others could be visited in another construction kit designed for another layer of abstraction. Are there ways to expose certain aspects of the system design without compromising the underlying design values? For example young people should be able to see and tinker with parts of the “Surprise Me” system prompt, but it remains to be explored how this can be done without affecting or compromising the design of the system. An open question remains, what is the appropriate chosen black box for a system like Imagine Yourself?

Designing *with* Young People

While the above section talks about ways to improve designs of creative tools using generative AI, there are also a number of directions to explore in supporting young people as *designers*.

Critical Playgrounds

Beyond exploring ways to invite a sense of agency over the tools young people create with, young people should be supported in critiquing and designing the underlying systems they’re interacting with. In Imagine Yourself workshops, I tried to encourage discussion of the design of the AI systems young people are encountering, wanting to learn more about AI tools they have used and understand their perspectives and experiences with the technology. I found that young people participating in the workshops were eager to share their critical perspectives and concerns about the issues and challenges that arise from the development and use of generative AI technologies. Workshops with young people should not only encourage discussion and sharing perspectives, but also foster belonging with technology through involving young people in the design of the systems they’re interacting with.

I have been enamored with exploring the idea of designing critical playgrounds—open construction kits that invite young people not only to create with technology but to audit and explore the design of the technology itself.

I come back to Susie Wise’s description of Belonging:

> *Belonging* is being accepted and invited to participate; being part of something and having the opportunity to show up as yourself. More than that, it means being able to raise issues and confront harsh truths as a full member of a community. (Wise, 2022)

Critical reflection and engagement is another crucial aspect of belonging. How can we invite young people to explore the critical design challenges of the technologies they regularly interact with? Critical playgrounds cannot solely encompass a tool. They need to be supported

by learning environments that nurture young people in sharing their personal experiences and critical perspectives.

Deeper Engagements

The Imagine Yourself explorations in this project featured one-time, short workshops, with each group of people I worked with. I did not have the opportunity to form deeper connections with the creators or understand their journeys while engaging in longer form projects and creative explorations. The Imagine Yourself tool is designed to be one component of the creation process within Scratch, yet many of the workshops I hosted did not allow me to understand how young people use it as a tool among many in their toolkit. In future explorations I hope to build deeper relationships with the young people and mentors I'm working with and to engage with their perspectives as designers. The design process of Imagine Yourself has thus far been more front-loaded than I would have liked, especially because of the challenges in working with generative AI, discussed throughout this thesis. In future work with Imagine Yourself, I hope to engage in more collaborative iteration cycles with the communities I am working with and explore longer-form workshops. Deeper engagements will involve working consistently with the same community, building relationships, and exploring more engaging forms of creation—e.g. long-form story projects instead of one-off short term explorations of image generation.

Conclusion

While combing through the notes I was keeping while designing and iterating on Imagine Yourself, I came across a set of questions that I was reflecting on pretty early in my explorations. While I don't have all the answers, I have a guide post.

- Representation is a very personal process, and image generation feels *impersonal* — what are ways to think about this tension?
    - How to navigate the idea that machine generated personal representation feels impersonal, perhaps lacks authenticity … ?

- When a machine creates representations of people, is it inherently engaging in stereotype or cultural appropriation?
    - The images are not being created by or with input from the people whose cultures are being depicted

- What does it mean to create something personal or personally meaningful?

While I don't have all the answers, in this work I have tried to address these questions by engaging authentically. While writing and sharing the ideas I've been exploring in this thesis, I have felt vulnerable because I am mindful that I'm speaking from a place of privilege and power, and I want to critically reflect on my own work and on how I am engaging with the communities whose voices I want to center. In workshops, I have hoped to work towards creating a space where young people feel invited to create and share, but do not feel pressured to share personal experiences inauthentically or in ways where they don't feel comfortable. It's been important for me to share my own background and motivations behind the work in these workshops, while balancing my role as a researcher, trying not to weigh my own voice, perspectives, and ideas above those of the young people and communities I am engaging with.

Authentic engagement has been at the heart of this work, and I believe it's the most important aspect of engaging with the ideas explored here. Authentic engagement remains at the heart of finding ways to engage with AI that honor and preserve what is human.

Appendix A

What is Generative AI?

Generative AI is quite the buzz word and umbrella term encompassing many different kinds of technology. There are also similar sounding terms that are related, but have distinct meanings such as AGI "Artificial General Intelligence," "Artificial Intelligence." With many of these phrases and acronyms using the exact same words, it can get confusing to keep up. This section aims to give a general overview of generative AI and how it works.

Generative AI is a term that refers to tools that take text as input and use machine learning techniques and unfathomably large amounts of data to create brand new data as output. What do I mean by "brand new data?" This can be new text, images, videos, sounds, music, and more, assembled by the machine.

How does it work?

Most of what is currently referred to as Generative AI at the time of writing uses generalized Large Language Models (LLMs) which are trained to connect meaning between different kinds of text. This builds upon decades of research and technological advances in many fields including Artificial Intelligence and Natural Language Processing. Much of the hype around Generative AI stems from its capabilities to quickly construct new data from human language input. This new data can take the form of content that feels directly meaningful and understandable by humans such as images, videos, music, conversational text, and more.

As the technology has evolved, the queries and prompts for generating things can become quite complex and the Generative AI model can still handle it with ease. This is different from past AI systems that have been available to the public such as chat support agents on various online platforms or voice assistants such as Apple's Siri[34] or Amazon's Alexa[35]. Most of these agents are also classified as "AI" systems, but have had limited capabilities such as responding to a finite set of commands or key phrases that the AI assistant had been programmed to respond to. If you've ever had the experience of one of these models telling you something like, "sorry I didn't quite get that," then you've bumped up against some of the boundaries of these tools, whether you asked for something it wasn't programmed to handle, or you spoke in an accent that was left out of its training dataset.

In contrast with these older AI systems, Generative AI tools have a wider range of possibilities for interaction. Instead of pre-programmed commands or functionality, these tools have been trained on unfathomable amounts of data from the internet and are capable of understanding and responding to more complex and nuanced requests. While Generative AI models have their own sets of boundaries and constraints, these boundaries are rapidly being expanded at unprecedented rates with major new technological developments being released every few weeks. For example, within seconds, generative AI can be used to create a video clip or a new pop song. Of the existing LLMs, Open AI's Generative Pre-trained Transformer (GPT), the underlying model behind the ChatGPT product, is one of the most popular. GPT-4 Omni,

[34] https://www.apple.com/siri/
[35] https://developer.amazon.com/en-US/alexa.html

the latest offering from OpenAI can not only converse via text or images, but it can also understand spoken words and recognize pictures and videos from your webcam as live input while using it. Combining the capabilities of many technologies, Generative AI tools provide a powerful landscape for thinking, creating, and interacting with computers in new ways.

What is a (text-to-) image generator?

Image generators are types of Generative AI tools in which the user provides input and guidance for the AI to create a new image. The input can be in the form of text for text-to-image generators, or images for image-to-images generators, or images and text, and perhaps other input for specific tool contexts. In the most common text-to-image generation tools, the user provides a text description of the image they want to create. This text description, known as a "prompt" could be as long or as short as the user likes and include descriptors for the subject, art style, artist reference, perspective, color scheme, *anything* the user can think of that would be handy in describing their vision.

How does it work?

Most of the popular image generators today are trained through a process called diffusion aimed to provide the AI model with lots of data about images and what they look like when they are noisy (to the point where they have no distinguishing qualities of the original image — the same way old TVs displayed random pixels if disconnected from the cable box or if there was interference in the signal being received by the tv antenna). This data is used by the AI model to construct new images from the user's prompt being able to start with a grid of randomly colored pixels (again think of the tv noise channel). While this is a very crude and brief overview of how image generator models work, these systems would not be possible without LARGE amounts of data. *Unfathomable* amounts of data. In the following section I draw some attention to where this data comes from, what is known and unknown about this data, and what effects this data has on the outputs of the AI system.

Where does the data come from?[36]

Most of the image generators out there today are trained on billions of pieces of data (including text, images, videos, music, and more) from the internet. Many popular image generators use the LAION-5B dataset of images (Beaumont, 2022), one of the largest datasets of text and image pairs, derived from a general crawl of the internet. The 5B in that acronym stands for 5 Billion, meaning that there are ~5 billion text and image pairs in this dataset. The data included is of varying quality, from all across the internet. The dataset is so large that even the creators of the dataset do not know what is in the images, and in fact, research on this dataset has shown that the images include inappropriate and even illegal content or often mislabeled content (Buschek & Thorp, 2024). Research has also shown that often the text descriptions are derived from alt text captions or other text that was not originally intended to provide rich, detailed descriptions of what is in the image (Buschek & Thorp, 2024). This research has also shown that the data in the dataset also disproportionately over-represents content from the US and European countries (Buschek & Thorp, 2024).

[36] The Knowing Machines Project visual story, "Models all the Way Down" (https://knowingmachines.org/models-all-the-way), is a resource that I highly recommend to anyone interested in an approachable way to learn more about the training data used by image generators.

Appendix B

Dall-E 3 Prompt Rewriting

A somewhat hidden secret behind how Dall-E 3 and particular other image generators such as Google Gemini's image generator work (Warren, 2024) is that they rewrite the user's input prompt behind the scenes—unbeknownst to the user. This technique is often referred to as *prompt transformation*, *prompt revision*, or *prompt injection*. In Dall-E 3, the prompt transformation technique is used to add more details to create a richer prompt input for the image generator to use (OpenAI, 2023a). It is also used for safety measures or to avoid copyright infringement and related issues (e.g. by rewording references to inappropriate content or references to names of public figures or copyrighted materials). The user's original prompt may also be transformed to include additional details about the visual style of the image such as the kinds of colors to use, the art style of the image, or the visual perspective (e.g. top down or isometric). The full details of the Dall-E 3 prompt revision process are detailed in the Dall-E 3 System Card (OpenAI, 2023a).

In some cases, however, the prompt transformation / injection technique has been used to attempt to compensate for the issues of bias in the tool's underlying model. A "Know Your Meme"[37] page on "Ethnically Ambiguous AI Prompt Injection" documents reports of users on social media noting the phrase "Ethnically Ambiguous" popping up as unexpected text in their images generated by Dall-E 3 (*Ethnically Ambiguous AI Prompt Injection*, 2023). The images below are from this "Know Your Meme" page.

[37] Know Your Meme (https://knowyourmeme.com/) is an online crowd-sourced platform that documenting internet trends

Screenshots of social media posts discussing Dall-E 3’s “ethnically ambiguous” prompt injection (*Ethnically Ambiguous AI Prompt Injection*, 2023)

References

Adichie, C. N. (Director). (2009). *The Danger of a Single Story* [Video recording]. https://www.youtube.com/watch?v=D9Ihs241zeg

Adobe. (2024). *Vector Graphics Software – Adobe Illustrator* [Computer software]. https://www.adobe.com/products/illustrator.html

Afro puffs. (2024). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Afro_puffs&oldid=1227524857

Al Jazeera. (2023, September 12). *Mining of cobalt, copper in DRC leading to human rights abuses: Report*. Al Jazeera. https://www.aljazeera.com/news/2023/9/12/mining-of-cobalt-copper-in-drc-leading-to-human-rights-abuses-report

Amar Chitra Katha: Indian Mythology, History & Comic Books. (n.d.). Amar Chitra Katha. Retrieved August 3, 2024, from https://us.amarchitrakatha.com/

American Psychological Association. (2023). *Inclusive language guide* (Version 2nd ed.). https://www.apa.org/about/apa/equity-diversity-inclusion/language-guide.pdf

Baldé, C. P., Kuehr, R., Yamamoto, T., McDonald, R., D'Angelo, E., Althaf, S., Bel, G., Deubzer, O., Fernandez-Cubillo, E., Forti, V., Gray, V., Herat, S., Honda, S., Iattoni, G., Khetriwal, D. S., di Cortemiglia, V. L., Lobuntsova, Y., Nnorom, I., Pralat, N., & Wagner, M. (2024). *Global E-Waste Monitor 2024*. International Telecommunication Union (ITU) and United Nations Institute for Training and Research (UNITAR).

Beaumont, R. (2022, March 31). *LAION-5B: A New Era Of Open Large-Scale Multi-Modal Datasets*. https://laion.ai/blog/laion-5b

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.

https://doi.org/10.1145/3442188.3445922

Benjamin, R. (2020). *Race after technology: Abolitionist tools for the New Jim Code*. Polity.

Berthelot, A., Caron, E., Jay, M., & Lefèvre, L. (2024). Estimating the environmental impact of Generative-AI services using an LCA-based methodology. *Procedia CIRP*, *122*, 707–712. https://doi.org/10.1016/j.procir.2024.01.098

Betker, J., Goh, G., Jing, L., Brooks, T., Wang, J., Li, L., Ouyang, L., Zhuang, J., Lee, J., Guo, Y., Manassra, W., Dhariwal, P., Chu, C., Jiao, Y., & Ramesh, A. (n.d.). *Improving Image Generation with Better Captions*.

Black Technology LTD. (2023). *Stable Diffusion Online*. https://stablediffusionweb.com/

Buckingham, D. (Ed.). (2008). *Youth, identity, and digital media*. MIT Press.

Bulovic, K. (2022). *Designing for Tinkerability for Accessibility* [Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/144686

Bulovic, K., Bentley, Z., & Rusk, N. (2024). Designing for Tinkerability and Accessibility: Developing the OctoStudio mobile app to engage blind and visually impaired learners in creating with code. *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*, 882–886. https://doi.org/10.1145/3628516.3659411

Buolamwini, J. (2023). *Unmasking AI: A story of hope and justice in a world of machines* (First edition). Random House.

Buschek, C., & Thorp, J. (2024). *Models All The Way Down*. Knowing Machines. https://knowingmachines.org/models-all-the-way

Chou, R., & Feagin, J. R. (2015). *The myth of the model minority: Asian Americans facing racism* (Second Edition). Paradigm Publishers.

Chun, K.-T. (1980). The Myth of Asian American Success and Its Educational Ramifications. *IRCD Bulletin, A Publication of He Institute for Urban and Minority Education*, *15, No. 1-2*(Winter-Spring Double Issue).

Constine, J. (2018, January 30). Snapchat enhances Bitmoji with 1.9 septillion avatar options.

TechCrunch. https://techcrunch.com/2018/01/30/bitmoji-deluxe/

Crawford, K., & Joler, V. (2018). *Anatomy of an AI System*. Anatomy of an AI System. http://www.anatomyof.ai

Darke, A. M. (2020, April 24). Open Source Afro Hair Library. *Prettydarke*. https://prettydarke.cool/portfolio/open-source-afro-hair-library/

Darke, A. M. (2023). *Open Source Afro Hair Library*. Open Source Afro Hair Library. https://afrohairlibrary.org>

de Vries, A. (2023). The growing energy footprint of artificial intelligence. *Joule*, *7*(10), 2191–2194. https://doi.org/10.1016/j.joule.2023.09.004

Ethnically Ambiguous AI Prompt Injection. (2023, November 27). Know Your Meme. https://knowyourmeme.com/memes/events/ethnically-ambiguous-ai-prompt-injection

Fear, N. (2023, December 1). *AI art gets even more bizarre in viral 'make it more' trend*. Creative Bloq. https://www.creativebloq.com/news/ai-generator-more-trend

Garcia, A., & Lee, C. (2020). *Equity-Centered Approaches to Educational Technology* (pp. 247–261). https://doi.org/10.1007/978-3-030-36119-8

Geena Davis Institute. (2023). *Changing the Narrative: Why Representation in Video Games Matters*. Geena Davis Institute. https://geenadavisinstitute.org/wp-content/uploads/2024/01/2023-GDI-Gaming-Report-v2.pdf

Geena Davis Institute. (2024). Geena Davis Institute. https://geenadavisinstitute.org/

Goh, G., Betker, J., Jing, L., & Ramesh, A. (n.d.). *DALL·E 3*. Retrieved April 15, 2024, from https://openai.com/dall-e-3

Google. (2024). *Noto Emoji*. Google Fonts. https://fonts.google.com/noto/specimen/Noto+Emoji

India.com. (2019, July 24). *Interesting Facts About Kabhi Khushi Kabhie Gham as Karan Johar Celebrates 400M Views of Bole Chudiyan*. India.Com.

https://www.india.com/entertainment/bollywood-news-interesting-facts-about-kabhi-khushi-kabhie-gham-as-karan-johar-celebrates-400m-views-of-bole-chudiyan-3725845/

Johar, K. (Director). (2001, December 14). *Kabhi Khushi Kabhie Gham...* [Drama, Musical, Romance]. Dharma Productions.

Justine Moore [@venturetwins]. (2023, November 26). Obsessed with the new “make it more” trend on ChatGPT. You generate an image of something, and then keep asking for it to be MORE. For example—Spicy ramen getting progressively spicier 🔥 (from u/dulipat) https://t.co/UitZh8tLR0 [Images Attached]. *X*. https://x.com/venturetwins/status/1728956493024919604

Kafai, Y. B., Fields, D. A., & Cook, M. S. (2010). Your Second Selves: Player-Designed Avatars. *Games and Culture*, *5*(1), 23–42. https://doi.org/10.1177/1555412009351260

Kelly, K. (2022, November 17). What AI-Generated Art Really Means for Human Creativity. *Wired*. https://www.wired.com/story/picture-limitless-creativity-ai-image-generators/

Kumar, I. K. (2023). *Fostering Well-Being: Designing Technology to Improve the Psychological Well-being of Foster-Involved Youth* [Master’s Thesis].

Lacoma, T., & Beaton, P. (2021, September 22). What Is Bitmoji and How Do You Use It in 2022? *Digital Trends*. https://www.digitaltrends.com/mobile/what-is-bitmoji/

LoRA the Explorer—A Hugging Face Space by multimodalart. (n.d.). Retrieved August 4, 2024, from https://huggingface.co/spaces/multimodalart/LoraTheExplorer

Minero, E. (2020, August 14). Educators Turn to Bitmoji to Build Community and Engagement. *Edutopia*. https://www.edutopia.org/article/educators-turn-bitmoji-build-community-and-engagement/

[@mrtrips211]. (2024, March 17). *I asked AI for an image of a pineapple fighting a pepperoni pizza, this was one of the images*. 9GAG. http://9gag.com/gag/a4PpdN1

Museus, S. D. (2008). The Model Minority and the Inferior Minority Myths:Understanding

Stereotypes and Their Implications for Student Learning. In *About Campus* (Vol. 13, Issue 3, pp. 2–8). https://doi.org/10.1002/abc.252

Napoli, P. M. (Ed.). (2018). *Mediated communication*. De Gruyter.

Naqvi, H. (2022, September 27). Oak Tree Road Gives Cultural Comfort. *Quo Vadis - Student Newspaper of Middlesex County College*. https://www.quovadisnewspaper.com/lifestyle_culture/oak-tree-road-gives-cultural-comfort/article_63595ca6-3ea9-11ed-89f6-734914274e54.html

National Academies of Sciences, Engineering & Medicine. (2021). *Cultivating Interest and Competencies in Computing: Authentic Experiences and Design Factors* (B. M. Means & A. Stephens, Eds.). The National Academies Press. https://doi.org/10.17226/25912

Nicoletti, L., & Bass, D. (2023, June 9). Humans Are Biased. Generative AI Is Even Worse. *Bloomberg.Com*. https://www.bloomberg.com/graphics/2023-generative-ai-bias/

OpenAI. (2022, March 25). *DALL·E 2*. https://openai.com/index/dall-e-2/

OpenAI. (2023a). *Dall-E 3 System Card*. https://openai.com/index/dall-e-3-system-card/

OpenAI. (2023b). *GPT-4 System Card*. https://cdn.openai.com/papers/gpt-4-system-card.pdf

OpenAI. (2023c, March 1). *OpenAI Platform*. OpenAI Platform. https://platform.openai.com/docs/models/how-we-use-your-data

OpenAI. (2023d, September 25). *GPT-4V(ision) system card*. https://openai.com/research/gpt-4v-system-card

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., … Zoph, B. (2024). *GPT-4 Technical Report* (arXiv:2303.08774). arXiv. https://doi.org/10.48550/arXiv.2303.08774

Othering and Belonging Institute. (2024). *Home | Othering & Belonging Institute*. https://belonging.berkeley.edu/

Papert, S. (n.d.). *Hard Fun*. Retrieved August 3, 2024, from

http://www.papert.org/articles/HardFun.html

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.

Papert, S. (1986). *Constructionism: A new opportunity for elementary science education*. Massachusetts Institute of Technology, Media Laboratory, Epistemology and ….

Passmore, C. J., & Mandryk, R. (2018). An About Face: Diverse Representation in Games. *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, 365–380. https://doi.org/10.1145/3242671.3242711

Perrigo, B. (2023, January 18). *Exclusive: The $2 Per Hour Workers Who Made ChatGPT Safer*. TIME. https://time.com/6247678/openai-chatgpt-kenya-workers/

Powell, J. A., & Menendian, S. (2024). *Belonging without othering: How we save ourselves and the world*. Stanford University Press.

Resnick, M. (2017). *Lifelong kindergarten: Cultivating Creativity through Projects, Passion, Peers, and Play*. MIT Press.

Resnick, M. (2020, October 16). The Seeds That Seymour Sowed. *Medium*. https://mres.medium.com/the-seeds-that-seymour-sowed-c2860379617b

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for all. *Communications of the ACM*, *52*(11), 60–67. https://doi.org/10.1145/1592761.1592779

Resnick, M., & Rosenbaum, E. (2013). Designing for Tinkerability. In M. Honey & D. Kanter (Eds.), *Design, make, play: Growing the next generation of STEM innovators*. Routledge.

Resnick, M., & Rusk, N. (1996). The Computer Clubhouse: Preparing for life in a digital world. *IBM Systems Journal*, *35*(3.4), 431–439. https://doi.org/10.1147/sj.353.0431

Resnick, M., & Rusk, N. (2024, July). OctoStudio: Mobilising Creative Learning. *Hello World Magazine*, *24. Impact of Tech*. https://downloads.ctfassets.net/oshmmv7kdjgm/2r3hBJWLrxnGD8rvmSVo0q/7c76ac2

2fdb72bd43edf861a71862839/HelloWorld24.pdf

Resnick, M., & Silverman, B. (2005). Some reflections on designing construction kits for kids. *Proceedings of the 2005 Conference on Interaction Design and Children*, 117–122. https://doi.org/10.1145/1109540.1109556

Robertson, A., Magdy, W., & Goldwater, S. (2021). Black or White but Never Neutral: How Readers Perceive Identity from Yellow or Skin-toned Emoji. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–23. https://doi.org/10.1145/3476091

Rogers, O., Mastro, D., Robb, M. B., & Peebles, A. (2021). *The Inclusion Imperative: Why Media Representation Matters for Kids' Ethnic-Racial Development*. Common Sense Media. https://www.commonsensemedia.org/sites/default/files/research/report/2021-inclusion-imperative-report_final-release_for-web.pdf

Rosenbaum, E. (2023, August 22). Inside Scratch Lab: AI Image Generation. *Medium*. https://medium.com/scratchteam-blog/inside-scratch-lab-ai-image-generation-179f11bd921a

Rusk, N., Jain, R., & Martin, C. K. (2023). *Designing for Belonging: Partnering With Community-Based Educators to Develop a New App for Creative Expression*. 2025–2026. https://doi.org/10.22318/icls2023.808278

Scratch Foundation. (2024). *Scratch—Imagine, Program, Share*. https://scratch.mit.edu/

Smith, J. H., & Mantz, J. W. (2006). Do cellular phones dream of civil war? The mystification of production and the consequences of technology fetishism in the Congo. In M. H. Kirsch (Ed.), *Inclusion and Exclusion in the Global Arena* (pp. 71–93). Taylor & Francis.

Snap Inc. (2022). *Bitmoji—Your own personal emoji* [Computer software]. Snap Inc. https://www.bitmoji.com/

Snap Inc. (2024). *Snap Inc.* https://snap.com/en-US/

Stability AI Image Models. (n.d.). Stability AI. Retrieved August 11, 2024, from https://stability.ai/stable-image

STUDIO GHIBLI Inc. (2024). スタジオジブリ｜*STUDIO GHIBLI*. https://www.ghibli.jp/

TetraChroma Inc. (2017). *Picrew*｜ *The make-and-play image maker*. Picrew. https://picrew.me/en

The Nielsen Company. (2021). *Being seen on screen: The importance of quantity and quality representation on TV* (Diverse Intelligence Series). The Nielsen Company. https://www.nielsen.com/insights/2021/being-seen-on-screen/

The United Nations Office at Geneva. (n.d.-a). *Disability Inclusion at UN Geneva: Guiding Principles | The United Nations Office at Geneva*. The United Nations Office at Geneva. Retrieved August 3, 2024, from https://www.ungeneva.org/en/about/accessibility/principles

The United Nations Office at Geneva. (n.d.-b). *Disability Inclusive Language Guidelines*. The United Nations Office at Geneva. Retrieved August 3, 2024, from https://www.ungeneva.org/sites/default/files/2021-01/Disability-Inclusive-Language-Guidelines.pdf

Tiku, N., Schaul, K., & Chen, S. Y. (2023, November 1). *These fake images reveal how AI amplifies our worst stereotypes*. Washington Post. https://www.washingtonpost.com/technology/interactive/2023/ai-generated-images-bias-racism-sexism-stereotypes/

Turkle, S. (1995). *Life on the screen: Identity in the age of the Internet*. Simon & Schuster.

Utz, V., & DiPaola, S. (2023). *Climate Implications of Diffusion-based Generative Visual AI Systems and their Mass Adoption*.

Warren, T. (2024, February 22). *Google pauses Gemini's ability to generate AI images of people after diversity errors*. The Verge. https://www.theverge.com/2024/2/22/24079876/google-gemini-ai-photos-people-pause

Wilson, D., Shin, S. Y., Gibney, S., Coy, J., & MBD, D. (2022). *Where we come from*. Carolrhoda

Books.

Wise, S. (2022). *Design for belonging: How to build inclusion and collaboration in your communities* (First Edition). Ten Speed Press.

Women Techmakers (Director). (2022, February 2). *Black Women in Tech: A.M. Darke builds better representation in gaming* [Video recording]. https://www.youtube.com/watch?v=MqcgDnXSTSk

Zheng, R. (2016). Why Yellow Fever Isn't Flattering: A Case Against Racial Fetishes. *Journal of the American Philosophical Association*, *2*(3), 400–419. https://doi.org/10.1017/apa.2016.25

Zimmermann, D., Wehler, A., & Kaspar, K. (2023). Self-representation through avatars in digital environments. *Current Psychology*, *42*(25), 21775–21789. https://doi.org/10.1007/s12144-022-03232-6