From Mind to Machine:

An Embodied Approach to Image Creation with Generative AI

By Anusha Menon

A thesis exhibition presented to OCAD University

in partial fulfilment of the requirements for the degree of

Master of Arts in Digital Futures

OCADU CO, 130 Queens Quay East, 4th to 6th April 2024

Toronto, Ontario, Canada

April 2024

Copyright Notice

This document is licensed under the **Creative Commons Attribution-Noncommercial-ShareAlike 4.0 International License**.

You are free to:

Share – copy and redistribute the material in any medium or format.

Adapt – remix, transform, and build upon the material.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution – You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial – You may not use the material for commercial purposes.

ShareAlike – If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions – You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

To learn more, please visit https://creativecommons.org/licenses/by-nc-sa/4.0/

Abstract

This thesis explores the integration of embodied interactions within Human-to-Artificial Intelligence (AI) collaborative activity to support creative engagement and reduce the perception of AI as an uncontrollable, autonomous entity, also known as the AI “black box” rhetoric. Specifically, it investigates the utilization of kinetic sand as a sensory material and physiological data collector, facilitating the translation of users' hand motions and imprints in the sand into inputs for generative AI image creation. Informed by sociocultural frameworks of creativity, theories of embodied cognition and the positioning of AI as a statistical model, while grounded in iterative design methodologies and phenomenological analysis, the research aims to identify emergent guidelines from this collaborative creative process between humans and AI. The findings hope to contribute to the development of guidelines that inform the future design and implementation of generative AI systems for creative work. These guidelines account for embodied cognition as an essential facet of human creativity, promoting more intuitive and meaningful interactions between humans and generative AI. Ultimately, this research seeks to advance the discourse on human-AI collaboration, emphasizing the importance of embodied techniques in fostering creative synergy and mitigating the black box effect.

Keywords: Embodied Cognition, Embodied Creativity, Generative AI, Interaction, Human-AI Collaboration, Image Generation

Acknowledgements

Professors Barbara Rauch and Alexis Morris, thank you for your tireless guidance, generosity, patience, and support over the past year. I have learnt more than I could have ever hoped to through this endeavour, and I am endlessly grateful.

To the faculty at OCADU that I have had the privilege to engage with over the course of the program, being a part of your classroom gave me the confidence and resources to take on this thesis, among many other fascinating projects during the last two years. Thank you for your commitment to keeping us inspired and engaged, against all odds.

The DF cohort of 2024 has been a wealth of knowledge, support and motivation and I couldn't have asked for a better group of friends to go through this experience with.

A special thank you to my family, for continually bolstering me through every new choice in my education and career, allowing me to travel halfway across the world to achieve my dreams and never faltering in their belief in my abilities.

Table of Contents

List of Figures

1. Introduction

1.1 Motivation

The modalities of creative practice in the 21st century are continuously growing and evolving as technological innovation gives rise to new opportunities for artistic expression and audience engagement. Innovations in 3D Printing and Digital Fabrication have also influenced the fields of sculpture, architecture and product design by streamlining prototyping processes and making new forms and materials easier to access (Balletti et al. 2017; Menano et al., 2019), while a huge variety of digital tools and software have empowered creators from various fields in art, design, music and video creation to experiment with a wider range of styles thus providing greater creative freedom (Harder, 2018; Gumster, 2020; Jorda 2020). In the realm of immersive installations, Augmented and Virtual Reality (AR/VR) have revolutionised the possibilities for audience experiences, in addition to the use of sensors, touch interfaces and projection mapping (Verhulst, 2020; Larrieux and Speziali, 2020). Generative Artificial Intelligence has emerged amidst these technologies as a novel way of creating pieces of art, writing and music, with machine learning algorithms trained on enormous quantities of existing data to simulate the act of creation as we see it in human beings. An example of one of the first applications of AI in visual art is Harold Cohen's program, AARON, which could generate a variety of images ranging from line drawings in its first iterations, to eventually painting its own images through the use of robotics (Boden, 2009, p.27). Since Cohen's conceptualisation of AARON in the early 1970s, Generative AI has developed into not only an easily accessible tool through which to explore new creative possibilities, but also a controversial force in an age of automated labour and mass digitalisation (Joler and Pasquinelli, 2021), issues that will be explored in greater depth within this thesis.

However, considering the rapidity of the growth of AI as not just a useful tool, but as a competitive force in and of itself, there exists a dissonance between the kinds of creative output we are beginning to see from both human beings and AI. As mentioned above, Artificial Intelligence has historically referred to technologies that replicate the inherent intelligent functions of the human mind, but one may argue that this definition proves to be too broad in the face of current innovations in AI and encompasses a larger variety of applications that may be too simple to consider within this field (Sheikh et al., 2023, p. 15). Since there are multiple contested definitions for the term, I have chosen to base this thesis on the definition put forth by Haroon Sheikh et al. in their book *Mission AI*, which examines these numerous explanations to arrive at a singular definition that applies specifically to AI in its most current form as well as leaving room for relevance in future iterations and advancements. Their definition takes into consideration the connection between the changing definitions of AI and its evolution as a phenomenon, our understanding of both human and artificial presentations of intelligence and the need to keep pace with the exponential growth in the skills and capacities of both humans and machines. Therefore, for the purposes of this thesis, the definition of AI that I have taken into consideration is one put forth by the High-Level Expert Group on Artificial Intelligence (AI HLEG) of the European Commission (EC) in 2019, which defines it as "systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals" (Sheikh et al., 2023, p. 16). This definition leaves scope for the non-linear path of growth that AI research and development is likely to continue taking and encompasses all forms of the technology rather than just the most advanced innovations, such as deep learning, while remaining distinguishable from more relatively simplistic algorithms and general digital technology.

Artists like David Rokeby, Anna Ridler, Harold Cohen, and Rob Saunders, among many others, have used Generative AI to examine how technology can influence their own creativity and help them understand both their own practice and AI itself in new ways, despite the critiques surrounding the legitimacy of AI in art (Boden, 2009, p. 29; Zeilinger, 2021). This thesis seeks to reconcile the differences between human and machine creativity by reintroducing the embodied practices of art-making that have been lost in the process of artificially recreating the creative human mind, while still leveraging the opportunities offered by generative AI. Using kinetic sand, a sensory play material, I will facilitate embodied interactions with AI to collaboratively create art, by engaging the body's senses, movements, and instincts. By collecting visual and motion data through a camera, the AI acts as an active collaborator in the making process with the human body, allowing for a higher level of control and participation in the process by the user, thus potentially mitigating the perception of AI as an autonomous force. (See Figure 1).

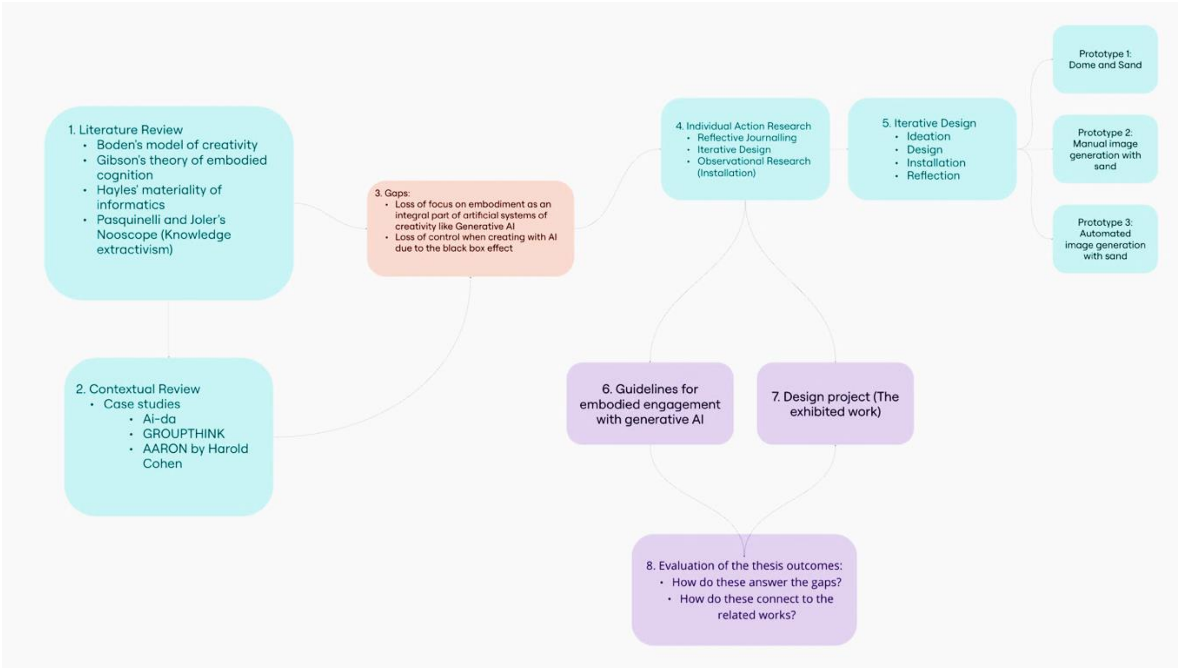


Figure 1: Map of the Research Summary

1.2 Current AI Interfaces for Image Generation

In recent years, there has been a surge in the availability of both open and closed source software for commercial as well as recreational image generation. Some of the most popular models like Stable Diffusion, DALL-E and RunwayML operate using text-based inputs with numerical parameters like random seed, iterations, guidance scale and strength at their most basic level (Stable Diffusion Models, n.d.; Betker et al, n.d.; Runway, n.d.). More advanced users may, for example, use knowledge of machine learning and greater computational power to build more nuanced models tailored for specific use-cases, from the base models that are already available, such as Stable Diffusion's checkpoint models that have been specifically trained to become Anything v3 for anime styled images, Realistic Vision for photorealism and Dreamshaper for painted styles (Stable Diffusion Models, n.d).

Recent advancements have further broadened the capabilities of most of these models to include image inputs in addition to the text prompts to provide greater control over the desired output, such as the Stable Diffusion model used in this thesis project. Clipdrop.ai (created by stability.ai, the developers of Stable Diffusion) offers this feature through existing image uploads as well as a digital drawing canvas that can be interacted with using a mouse or trackpad (Clipdrop, n.d). While these modes of image generation offer a variety of options for input, there is a lack of existing programs that can be used to directly draw information from an embodied creative process, in the same way that the body might inform how one interacts with physical materials (See Figure 1). This made it necessary to create an interface using kinetic sand and computer vision for this thesis, providing a medium through which to gain a richer breadth of physiological data for the computer to "perceive", including imprints of the hands and fingers in the sand, the motions and gestures performed by the user and environmental factors like light and shadow.

Model	Input	Parameters	Programming Skill Requirement	Level of embodied control
Stable Diffusion	Text	Seed, iterations, guidance, strength	Beginner to Advanced	Low
DALL-E	Text	Seed, iterations, guidance, strength	-	Low
Runway ML	Text, Image	Seed, iterations, guidance, strength	Beginner	Low
Stable Doodle	Text, Digital Drawing	Seed, iterations, guidance, strength, visual style	-	Medium
Midjourney	Text	Seed, iterations, guidance, strength	-	Low

Figure 2: Comparison table of current AI Interfaces for image generation

1.3 Research Summary

1.3.1 Problem Statement

The commodification of AI within a capitalist society contributes to knowledge production gaps due to the prioritization of efficiency and profit margins over a holistic understanding of the biased knowledge structures within AI and ethical considerations of generative AI use, resulting in a fragmented knowledge landscape and huge societal ramifications. Considering these challenges, there arises a pressing need to examine the shifting dynamics of embodied cognition within creative practice. The transition from human-centred to AI-driven creativity within a techno-centric world presents a paradigmatic shift with profound implications for artistic expression and human experience. Being in its early stages of development, discussions of AI are also prone to falling within the "black box rhetoric", which cites a lack of understanding of the processes of reasoning within the AI as a reason for lack of control over it. Current interfaces for image generation with AI present little opportunity for understanding the decisions made by the algorithm to influence the output it creates, particularly due to the computational nature of working with text and numerical parameters in contrast with the physical materials that artists have traditionally worked with. As a result, the lack of sensory exploration afforded by these interfaces amplifies this feeling of a loss of control and further alienates creatives from engaging with this technology.

Thus, there emerges a critical imperative to recentre the human body and embodied practices into creative making, particularly within the realm of generative AI. Due to the closely interlinked nature of creativity and intelligence, attempts to replicate these traits computationally has resulted in a focus on logical reasoning and decision-making processes that can be mathematically coded into a system. This has resulted in a loss of focus on embodiment, which, as a crucial facet of creativity, is highly relevant to the development and use of Generative AI tools for creative processes, particularly visual art and image-making.

1.3.2 Hypothesis

To maximise the efficacy of AI tools in affording us novel methods of creating, a shared focus on embodiment of the artificial system and embodied knowledge of the artist is necessary, which could both reconcile the differences and bridge the gap between human and machine creativity.

1.3.3 Research Questions

Primary

1. How might we transform our understanding of the affordances of generative AI by investigating the role of the body in the creative process to reconcile the differences between human and machine creativity?

Secondary

a. How might outsourced computation through computer vision enable effective embodied interactions with generative AI?
b. How might tactile interfaces for generative AI lead to creative image generation by involving multiple senses?

c. How might an exchange in perspectives between humans and generative AI lead to equal creative collaboration in the process of image generation?

1.3.4 Objectives

This thesis aims to create a set of guidelines that highlight the value of embodied cognition in the creative process of image generation. It does so by reintroducing physical interactions with tactile materials during the process of creation with generative AI and by embodying the computer with sensors that allow it to gather a greater degree of information on the environmental context in which the human body and the physical material are operating. The purpose of focusing on embodiment is twofold:

1. To provide the human collaborator with a greater degree of control over the creative process and therefore, the creative output, effectively reducing the impact of the AI black box effect
2. To contribute to discussions of embodiment in creative computation research, which eventually guides the development of more effective and ethical image generation models in the future

1.3.5 Methodology

The research methodology of this thesis borrows methods from Individual Action Research (IAR) (Willis J.W. et al, 2014, pp. 96-97) and the iterative design process of Human-

Computer Interaction (HCI) design as it incorporates self-reflective as well as observational evaluation methods. The research is carried out through the following steps:

1. Literature and Contextual Review
2. Creation of the Interactive Kinetic Sand-AI system
3. Exhibition of the system for user observation
4. Journaling and Reflection
5. Creation of Guidelines for embodied engagement with Generative AI
6. Discussion of the artifact and guidelines

The iterative process of creating the Sand-AI system presents opportunities for designing a functional proof of concept of the theories examined in this paper as well as a space for reflection on the relationship between the human body and generative AI in the process of image-making. Through each prototyping step, new factors such as sensory stimulation, human agency, AI agency and embodied control are introduced, to examine the influence of these factors on the interactions with the prototype. The findings are then used in combination with the analysed theory and related work to create a set of guidelines. Finally, the guidelines and prototypes are contextualised within the existing research to discuss the results of the thesis research and possibilities for future work.

1.3.6 Contributions

The primary contributions of this thesis are as follows:

1. The creation of a set of guidelines for embodied engagement with generative AI.

2. Prototypes demonstrating embodied Human-AI collaboration in the context of image generation. These prototypes will demonstrate how tactile interfaces facilitate multisensory interactions that enable a deeper dialogue between the user and the AI model, using computer vision.

This thesis will also provide additional contributions over the course of the research and further context and information about the prototypes and guidelines being created. These include:

1. A literature and contextual review covering relationships between creativity, intelligence and AI, the significance of embodiment in computational creativity and human-AI collaboration.
2. Analysis and discussion of theoretical frameworks such as structural coupling, a sociocultural model of creativity and theories of affordances, which support the research.
3. Opportunities for future development of the research, detailed through evaluation and analysis of the created guidelines and prototypes.

1.3.7 Scope and Limitations

There are several challenges relating to scope that have been anticipated within this thesis, the most significant one being the focus on image generation models alone and not all forms of generative AI. This was informed by my personal background in multiple forms of visual art and image-making, ranging from drawing and painting to graphic design, causing image generation to be the logical first step of my own journey as a researcher studying creativity and generative AI. Secondly, the focus of this thesis on embodiment makes it necessary to incorporate certain forms of hardware and software into the computational systems being studied. This will be limited to relatively simple, inexpensive and accessible technologies such as cameras and computers with a moderate level of processing power. Finally, it is worth noting that advancements in generative AI are progressing at an increasingly rapid rate and the models investigated within this thesis may not be on par with the complexities and capabilities of the newest forms of generative AI at the time of finishing this paper. This also limits the applicability of the principles established at the end of this thesis, causing the research to be conducted on a broad-level assessment of some of the shared aspects of current image generation models.

2. Literature and Contextual Review

Creativity and its relationship to AI has been a highly researched and debated subject for decades, across numerous disciplines. The rising success and exponentially growing complexity of AI every year appears to foreshadow an eventual surpassing of human capabilities. Although the in-depth explorations of creativity within this thesis are limited to creative processes in visual art, this literature review seeks to define and contextualize the key concepts within the fields of digital humanities, using theories from cognitive science, philosophy, art and design and creative computation, to focus on the theoretical potential of generative AI as a creative collaborator through embodied user interactions. These concepts include the relationships between creativity, intelligence and AI, the significance of embodiment in computational creativity and human-AI collaboration (See Figure 3).

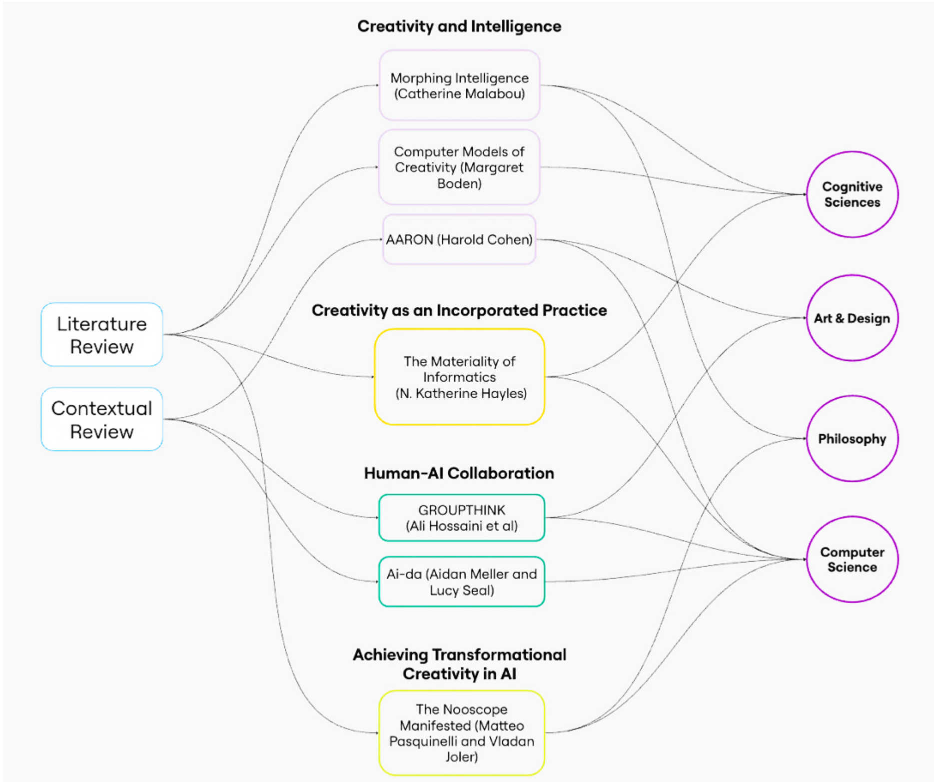


Figure 3: Summary of literature and contextual review

2.1 Creativity and Intelligence

When first approaching the concept of creative intelligence in both natural and artificial systems, examining the historical evolution of the term “intelligence” itself offers some insights into our perception of artificial intelligence today. From a philosophical perspective, Catherine Malabou examines the evolution of the definition of intelligence itself, from a quantifiable genetic trait to a processual means to an end, a method by which we gain and share knowledge that may vary across lived experiences (Malabou 2019). In her introduction to *Morphing Intelligence*, she categorises this historical evolution within three phases, or “metamorphoses” of intelligence (2019, p.13). The third or final metamorphosis, which is yet to come is defined by Malabou as automatism and the eventual disappearance of the lines between natural and artificially intelligent beings. Here, the definition of intelligence through biological or psychological factors proves to be almost redundant in the face of intelligent machines that have evolved past the point of "simple robotization" (Malabou, 2019, p. 15) With this framing of intelligence as a relationship between the body and the mind, both natural and artificial, it can be said that the definition of creativity, as a facet of intelligence, also continues to evolve in a techno-centric world. Margaret Boden acknowledges this relationship between creativity and intelligence, recognising how, despite the continued mystery surrounding a biological and psychological understanding of where creativity comes from and how it functions as an aspect of the human condition, it is a shared trait among all human beings, manifesting in everyday life as much as it does in historical innovation (Boden 2009, p. 24).

Boden’s model of creativity dissects the three ways in which it can take shape; combinational, exploratory, and transformational. These classifications are defined based on how the creator arrives at their novel idea and the implications this process has for the conceptual space that they are operating within. Combinational creativity can be seen in the formation of previously unrealized combinations of indirectly linked concepts and is the easiest form of creativity to achieve. Exploratory creativity is found in the discovery of new ideas within the socio-culturally accepted norms of that space because of extensive inquiry and pushing or examining the limits of that space. Transformational creativity is achieved through the breaking down of the generally accepted rules of a space and changing or removing them to redefine what exists within it. This allows for the generation of ideas that previously could not have existed within that space because of its limitations and is therefore touted as the most aspirational form of creativity (Boden 2009, p. 25). While these distinctions allow for a critical analysis of examples of artificially creative agents, Boden notes that there is no clear-cut distinction between exploratory and transformational creativity and that there is no straightforward answer to which of the three types would be the easiest for a computer to achieve. She then continues contextualizing these three forms of creativity through numerous examples of artificial intelligence, citing the strengths and limitations of each case as a way of laying out their complexity, not only in terms of their creation but also in their functioning and creative output.

Boden's analysis lays out a strong case for the various ways in which it has been established that creativity is an extremely difficult aspect of human intelligence to mimic, let alone develop from scratch. This creates an argument for the necessity of unique traits of the human condition that will be studied through this thesis, while also emphasizing the potential of machines to empower human beings to achieve more with their creativity than ever before.

This was exemplified in as early as the 1970s, by Harold Cohen with his creation of AARON, a program designed to autonomously generate artwork in Cohen's own style (Boden, 2009 p. 27). The evolution of AARON began with simple black and white line drawings that reflected Cohen's own knowledge of artistic principles, techniques, history, and aesthetics which led to the program working with a robotic arm to manually paint the images, finally culminating in a version that creates multicoloured prints (Boden, 2009 p. 27). Boden references Cohen's project to emphasise the potential of computers to excel at tasks that are typically challenging to human beings themselves, in this case, AARON's ability to exceed its creator's own expertise, as stated by Cohen (Boden, 2009 p. 27). While AARON was a program designed exclusively for Cohen's own practice and functioned as a product of his own creativity, current models of generative AI still offer opportunities for examination of how creativity in the visual arts functions as a product of interpersonal exchanges of perspectives. This manifests through each agent's distinct understanding and perception of the world around them, as will be further discussed through theories of embodied cognition.

2.2 Creativity as an Incorporated Practice

In examining the relationships between human and AI creativity, there are several similarities and differences that emerge, particularly when considering the entanglement of the human experience with technology in our present context. N. Katherine Hayles offers key insights into ideas of what the relationship between human and machine could look like, in a posthuman context. In consideration of liberal humanism taking the view that consciousness and cognition is what defines humanity and not its embodiment, even the notion that machines could be made to contain or augment this consciousness is analysed through a carefully critical lens. Hayles' revision of the existing narrative takes it from critiquing the notion of Posthumanism as producing tropes of immortality or exaggerated power to, instead, tapping into the potential to redefine the perception of the body and its ability to process and use information (Hayles, 1999, p. 5). This evokes discussion on the divide between 'bodies' and 'embodiment' or materiality and information and which takes precedence over the other. In the chapter "The Materiality of Informatics", Hayles further deepens the discussion on the body's capabilities as a vehicle for inquiry. Embodiment and the body are presented as a pair of interlinked yet separate conceptual entities, where the relationship between the two is comparable to that of inscription and incorporation.

According to Paul Connerton and Maurice Merleau-Ponty as referenced by Hayles, inscription refers to the process of encoding information into a medium (1999, pp. 199-201). In the digital realm, this could involve typing text on a keyboard, coding a program, or creating digital images or sounds. Hayles argues that inscription involves not just the act of encoding information but also embodies the choices, intentions, and cultural contexts of the person doing the inscribing. Each act of inscription reflects the author's subjective decisions and biases, shaping the meaning and interpretation of the resulting artifact. Inscription highlights the active role of the human agent in creating digital artifacts, underscoring the importance of human agency and intentionality in digital production. The act of creating visual art with generative AI in its predominant form today, using text-based prompts and adjusting parameters could also be categorised as an example of inscription.

Incorporation, on the other hand, refers to the process by which humans absorb and integrate technologies into their lives and selves. It involves adapting to and internalizing the affordances, constraints, and ideologies embedded within technological artifacts. Hayles suggests that as humans engage with digital technologies, they not only use these tools but also assimilate them into their cognitive processes, habits, and social practices. This incorporation of technology influences how individuals think, perceive, and interact with the world around them. Incorporation highlights the reciprocal relationship between humans and technology, where technology shapes human behaviour and cognition, while humans, in turn, shape and reinterpret technology to fit their individual needs and desires.

In summary, Hayles' concepts of inscription and incorporation provide a framework for understanding the dynamic interplay between humans and technology in the digital age. Inscription emphasizes the active role of human agents in creating digital artifacts, while incorporation highlights how technology becomes an integral part of every human experience, shaping cognition, identity, and social practices.

This rejection of the idea that embodiment is universal to all beings with a body provides an important guiding principle for this thesis and the importance of embodied creativity in engagement with generative AI. For the purposes of this thesis, when considering human-AI collaboration within this framing of inscription and incorporation, it can be said that bringing the nuances of the embodied human experience is critical to art-making practices and could serve as a key entry point to redefining the position of Generative AI in an industry that relies on that social, cultural, and environmental context. Hayles further articulates the necessity of nuanced narratives in crystallizing the knowledge gained through embodied practices as it is often not formalizable and cannot be “programmed into explicit decision procedures”, making it near impossible to code into a machine (Hayles, 1999, p. 202).

2.3 Human-AI Collaboration

Agency is essential to creativity, as it denotes an individual's capacity to make choices and decisions informed by their own knowledge and experiences, creating unique avenues for expression. This aspect becomes more intriguing in the context of artificial intelligence, as AI systems are becoming increasingly capable of autonomously generating artistic content, as previously discussed through Boden’s analysis of computational creativity. The interplay between human agency and AI-generated ideas raises questions about the nature of authorship and the role of humans in the creative process. Contextualising this idea within responsive live performance, the project GROUPTHINK by Ali Hossaini et al. brings to light the notion of creative collaboration between audience and performer through AI (2022). Echoing Malabou's speculative third metamorphosis of intelligence, characterised by a shared, automated intelligence (Malabou, 2019), the project focuses on the responsiveness of live audio-visual performance to the physiological reactions of the audience, proposing a shift from the traditional cues of gaze, breath, dancing and clapping to neurophysiological interfaces as a method of experiencing live performances on digital channels. In the performance, the remote audience's collective heart rate is monitored and transmitted to a triptych of monitors which translated the data as a visual score that served as the backdrop for the live-streamed musical performance. Here, the heart rate was monitored through the audience's webcam and reflected to the performers through the visual score and AI-generated graphics. The performers could then change the tempo and mood of their performance in response to the audience's reactions and the piece was created through a multi-level collaboration between the audience, performers, and AI (Hossaini et al., 2022). The piece was perceived as successful if a level of "physiological entanglement" and shared agency could be achieved by the audience and performers.

Overall, the project served as a metaphor for the potential future of collaboration between human and machine in art-making practices, emphasising the speculative concept of an "Internet of Neurons" inhabited by AI and human beings using neurophysiological interfaces to interact with each other (Hossaini et al., 2022). In the associated paper, the authors reflect on the ability of the piece to raise discussions around the benefits and dangers of harnessing technology to create art, questions on authorship and individuality and the invasiveness of the media (Hossaini et al., 2022). Aspects of creativity are also brought into question by this project; when the tools for creation possess their own intelligence, who is the creator?

Ai-da, the world's first humanoid robot artist, created by Aidan Meller and Lucy Seal in 2019, further exemplifies the complex nature of the question posed above ("Ai-da", n.d.). Named after Ada Lovelace, a mathematician and writer considered the world's first computer programmer, Ai-Da embodies the fusion of technology and artistic expression. Her robotic form is equipped with a range of sensors and technologies, including cameras in her eyes, which enable her to observe the world around her. She possesses the ability to analyse scenes and respond to visual stimuli, giving her a unique capacity to create art that is influenced by her perceptions.

Ai-Da is not only capable of drawing and painting but also engaging in conversations about her artistic process. Her creators programmed her with the ability to interpret and respond to the emotions of those interacting with her, arguably adding a layer of emotional depth to her artistic endeavours. The robot's artistic style is influenced by a combination of algorithms and data input, creating a blend of computational precision and creative intuition. Her works often explore the intersection of technology, humanity, and the environment, reflecting the creators' intention to provoke thought and discussion about the role of AI in society. According to their website, this intention stems from Donna Haraway's articulation of the Cyborg and Yuval Harari's request for a cross-disciplinary involvement in responding to the AI and biotechnology revolutions ("Ai-da", n.d.). Given the embodied form that Ai-da occupies, not only in terms of artistic capability through her robotic arm and sensors, but in gendered and cultural identity as well, the level of autonomy displayed by Ai-da is arguably higher in comparison to more commonly accessed forms of generative AI like Stable Diffusion and

Midjourney. Although the creators themselves acknowledge that this project still functions through human-machine collaboration and relies on a significant level of human control ("Ai-da", n.d.), their intention to blur the lines between human and artificial intelligence raises important questions on who or what is allowed to assume creative authorship in a world of rapid technological advancement. Ai-da represents a milestone in the journey of AI development towards the eventual goal of complete autonomy in AI, or General Artificial Intelligence (GAI).

2.4 AI as a Tool for Creativity

As computer scientists continue to strive for higher performing machines through advancements in AI, so does the level of debate around machine intelligence and its comparison to human capabilities increase. The scepticism surrounding artificial intelligence tends to echo the belief that computers can only create within the limits of what they are programmed to do and the kind of data they are fed (Boden, 2009). Pasquinelli and Joler approach this idea by using the Nooscope map to challenge the premature positioning of AI as an autonomous entity and critique the processes by which knowledge is inputted and processed by machine learning algorithms (2021). They posit that AI in its current form operates mathematically, with no conscious decision-making capabilities which not only results in a randomised, yet seemingly coherent regurgitation of information, but also a compression of a huge variety of knowledge systems and domains into vectorised latent space (2021, p.1268-1270). The gravest implications of such a system are knowledge distortion and data bias, which emerge from not only the algorithm itself, but also in the manual labour through which information is categorised and labelled. In the context of the image generation programs engaged with through this research, this is visible in the presentation of specific imagery and stylistic choices by the AI in response to certain prompts.

Why then, would a model that not only replicates human flaws, but also magnify and further distort them in an effort to simulate intelligence, be useful to humans as a creative collaborator? In the age of mass digitisation and automation of labour, the positioning of AI as an entity that can operate not only on par with, but superior to human capabilities has perilous consequences for the role of human creative practitioners. The necessity of the role of the critical human collaborator in the operation of AI programs is only made clearer by the identification of these faults, a necessity that arises from the mere existence of this technology itself. Given the precariousness of the issues being dealt with on a larger level in the development and use of AI systems, this thesis project positions the image generation model being used as a tool only, a lens through which to examine human creativity, embodied cognition and the dynamics of human-technology relationships, without centring the AI as the primary creator but still acknowledging the larger context in which users approach AI systems as an exciting and novel means of producing creative output. Drawing inspiration from Boden's analysis of the relationship between Harold Cohen and AARON as well as the cognitive capabilities of physically embodied computers as exemplified in Ai-da, this thesis seeks to carefully and critically examine how the affordances of generative AI can be leveraged to enhance human creativity by engaging the user in an embodied dialogue with their medium.

2.5 Summary

This literature and contextual review discussed the key texts and related works informing this thesis research, through the concepts of creativity, intelligence and AI, the significance of embodiment in computational creativity and human-AI collaboration. Beginning with a discussion of creativity as a facet of intelligence and the implications this has for AI, the ability of technology to empower human beings to maximise their creative potential is examined through Cohen's generative AI program AARON. To further explore the role of machines in augmenting human consciousness and capabilities, Hayles' analysis of inscription and incorporation is discussed as a lens through which to view the dynamic interplay between humans and AI as mutually influential entities in an increasingly digitised world. The role of agency in this relationship and its impact on human-AI collaboration in creative endeavours is discussed through an analysis of the projects GROUPTHINK and Ai-da, which use highly contrasting approaches to evoke discussions around the benefits and dangers of harnessing technology to create art and questions on authorship and individuality. Finally, the role of AI as a creative tool rather than an autonomous entity is discussed through Pasquinelli and Joler's use of the Nooscope map to critique the larger issues of data bias and the "black box" rhetoric prevalent in contemporary applications of AI. Through this analysis, questions of creative control and human agency are brought to light to inform the approach of this thesis towards creating a collaborative space between humans and AI in the process of image generation. Having set the theoretical context and reviewed the existing applications of creative practice through human-AI collaboration, the next chapter highlights the key theories that will be used to inform the execution of the thesis research through the ideation and evaluation of the created prototypes and suggested guidelines for engagement with generative AI.

3. Theory and Frameworks

The research conducted in this thesis is based on a combination of theoretical discussions around human and computational creativity and embodiment, some of which have been discussed through a literature and contextual review in the previous chapter. However, establishing a key set of frameworks through which to inform and examine the practice-based research will help maintain a clear and consistent position throughout the research journey and create a foundation from which to build the final protocol. This chapter focuses on elucidating the three main theoretical frameworks governing the approaches to examining the creative processes taking place between humans and AI in this thesis (See Figure 4).

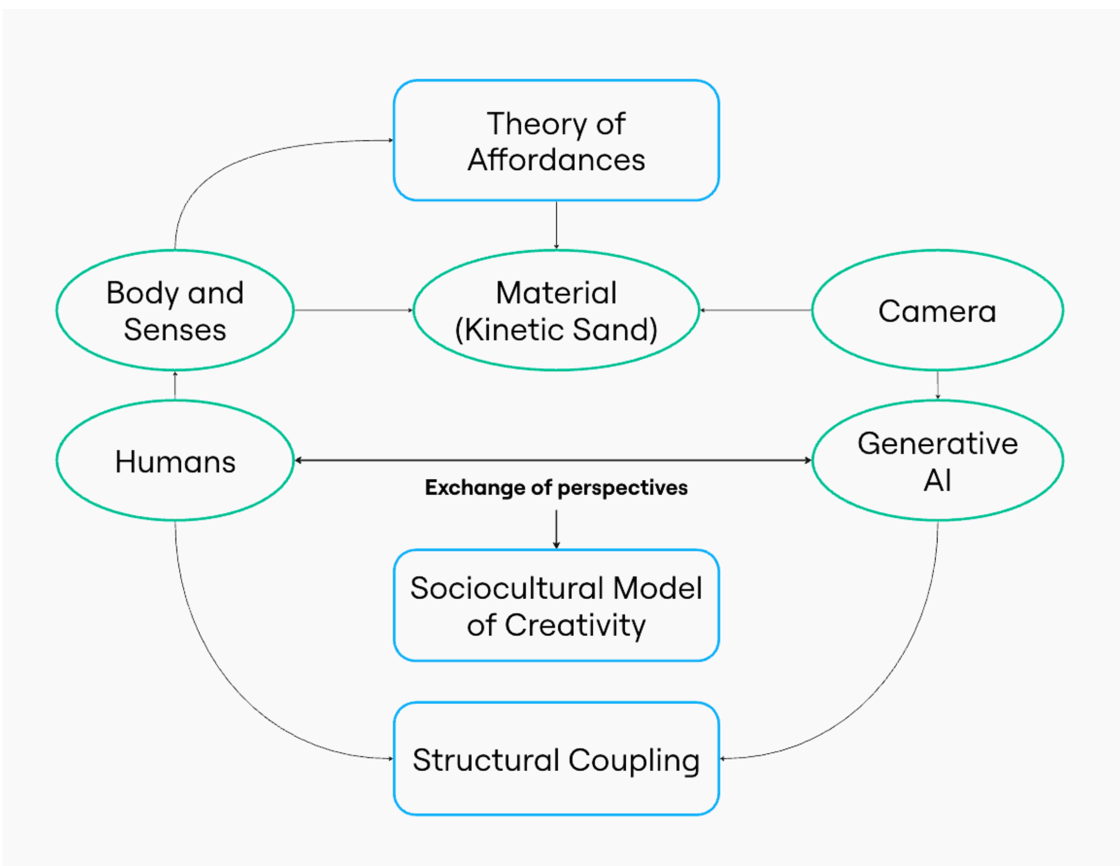


Figure 4: Connections between key theories

3.1 A Sociocultural Model of Creativity

To frame the conditions in which the creative processes of engaging with AI will take place, this thesis will draw on a sociological model of creativity, designed to focus on the collaborative nature of creative practice and the benefits of bringing a diverse set of perspectives to the process. Vlad P. Glăveanu emphasizes the role of multiple perspectives and influences on the creative process (2020). Glăveanu's four propositions about the creative process are as follows:

1. Differences of perspective increase creative potential,
2. Exchanging perspectives, within and between individuals, fosters creative processes,
3. These exchanges result in perspectives that reveal previously unperceived affordances,
4. It is the affordances of material objects or of unique idea combinations that guide the development of novel perspectives in creative work (Glăveanu 2020, p. 9-12)

In the context of computational creativity, these exchanges in perspectives could manifest as a collaborative creative process between humans and AI. Theoretically, if differences of perspective are an indicator of increased creative potential, what better way to maximize one's creative capabilities than through collaboration with non-human agents like Generative AI? This perspective-affordance sociocultural theory of creativity is grounded in relational notions of perspective, position and affordance that emphasise the significance of diverse viewpoints, the necessity of communication and co-creation in these processes and space for new discoveries and reflective thinking (Glăveanu, 2020, p. 13) This framework proves essential to the research conducted in this thesis as it foregrounds the importance of providing better avenues of communication between humans and AI to foster more effective creative processes and as a result, better creative output. The perspectives brought forth by the AI, are highly processed and compressed forms of knowledge that have been randomised and rearranged by the model to mimic human capabilities, which in this thesis is the creation of images. While recognising the ethical concerns of knowledge extractivism and data bias present in current forms of generative AI (Pasquinelli and Joler, 2021), it is worth nothing that the randomisation of the data contained in the model, as presented in the output, could lead to previously unrealised perspectives when perceived by the user. Considering that the highlighted issues are large-scale, structural problems in the overall development of AI in a capitalist society, this thesis seeks to focus on the possibility of novel creative ideation through human-AI collaboration, with the hope that these larger issues will eventually diminish as AI comes out of its nascent stages.

3.2 Theory of Affordances

To expand on the concept of affordances posited by Glăveanu, Gibson's model of embodied cognition delves into the complexity of perceiving and engaging with the world, not just from a human perspective but as all living beings do. Gibson's Theory of Affordances looks at how spatial, material, and environmental affordances are perceived by beings not just based on visual cues but through all the senses and on a genetic level, an inherent instinctual understanding of the world. Visual perception as a whole-body system rather than simply the eye's connection to the brain is articulated through this discussion of affordances and provides a look into how embodied cognition manifests in everyday life and knowledge-building.

Through this post-cognitivist lens, Gibson emphasizes that the environment is rich in information, and individuals are attuned to this information through their sensory systems. Perception is seen as an active exploration of the environment, not a passive reception of sensory data. As mentioned in the text, "the possibilities of the environment and the way of life of the animal go together inseparably. The environment constrains what the animal can do and the concept of a niche in ecology reflects this fact" (Gibson, 1979, p. 143). In consideration of material affordances alongside Hayles' posthuman analysis of embodied cognition, a being's ability to be creative depends on this embodied engagement with the world and a rejection of the notion that creativity exists solely in the mind. While experimentation with sensor-based data that allows artificially intelligent agents to gather more complex information on their environment allows for richer meaning and motivation behind AI-generated art, the necessity of human collaboration at this stage in the development of generative AI and its use as a tool is evident. As further posited by Gibson, the information, or even misinformation, that can be gathered through visual perception can create a significant impact on the response one has to their environment, and in the context of this thesis, their creative output. It may be said that creative collaboration between human and machine relies on this symbiotic perception and response to the affordances provided by each other's actions – where Gibson spoke of the uniqueness and dynamism of our perception of other living beings, AI now steps into a previously unoccupied space between object and agent.

3.3 Structural Coupling

To further detail the premise on which this research is being undertaken, I will look into the concept of structural coupling. This is especially relevant given the focus of this thesis on the study of human Ai collaboration as a creative process and the role of a mutual exchange of knowledge and perspectives in this process, the concept of structural coupling comes largely from the work of Humberto Maturana and Francisco Varela. It broadly refers to the dynamic relationship between two systems where each system influences the structure or behaviour of the other through their interactions (Maturana, 2002, p. 16). Some of the components of structural coupling include the involvement of autopoietic systems, which are self-creating and self-maintaining systems that are typically living organisms, both of which are mutually influential in the way they contribute to each other's adaptation and evolution (Maturana, 2002, p. 7-10). This relationship is also typically non-linear, meaning the relationship between the system is not necessarily proportional or predictable. Small changes in one system may lead to disproportionately significant changes in the other contributing to the complex and dynamic nature of the coupling. Although Maturana and Varela originally formulated the concept in the context of biological systems (Maturana, 2002, p. 16), it is commonly applied in the study of embodiment in creative computation as discussed in chapter 2 of this thesis. For the research conducted in this thesis, structural coupling offers a lens through which to analyse Human-AI collaboration through the embodiment of both agents. It is important to note that the AI models used in this thesis are not self-governing or self-preserving, but to improve the applicability of the principles derived from this research to future developments in AI, which are striving to achieve autonomy in AI agents, considering this relationship through the lens of structural coupling forms a theoretical foundation from which future applications can be examined. Some of the concepts within structural coupling that will be reviewed through this research project are the reciprocal nature of the human-AI interactions, embodied using kinetic sand as a tactile interface. The design of the interactive system will focus on elucidating the physiological nature of the human’s responses to the AI and the visibility of the connections between the physical environment and the AI's output. The AI also functions as a cognitive extension of the human body, by retrieving the data gathered from the interactions with the kinetic sand and responding to it by iterating or modifying the output it produces.

3.4 Summary

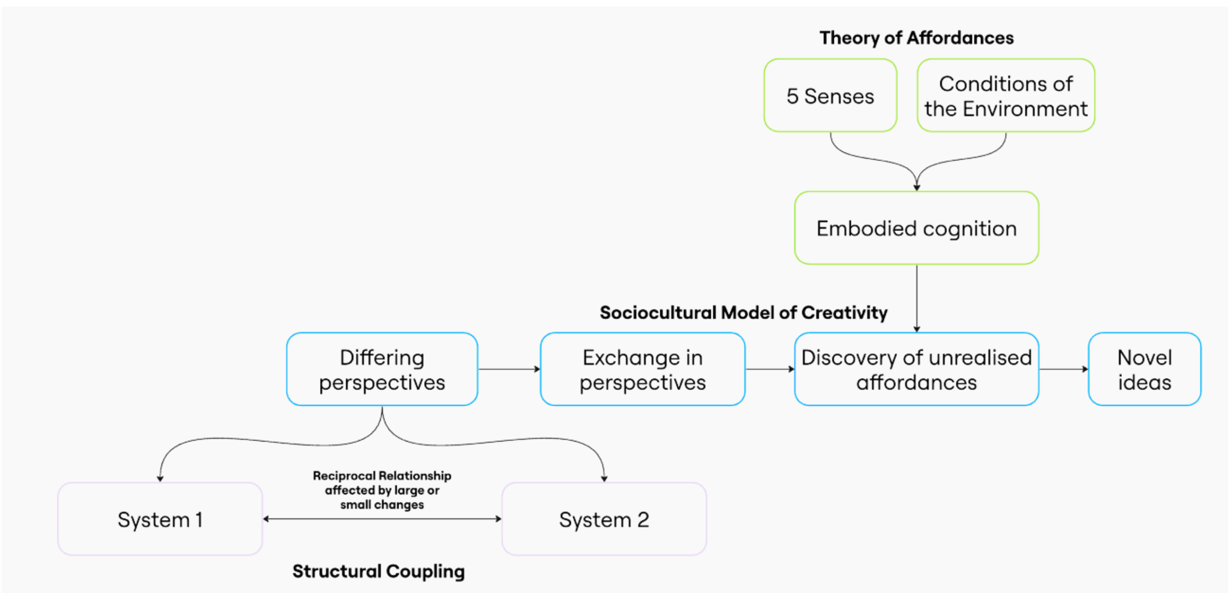


Figure 5: Connections between key theories

The theories of a sociocultural model of creativity, affordances, and structural coupling ground the research conducted in this thesis through their emphasis on the role of embodied reciprocal interactions as facilitators of creative dialogue (See Figure 5). Glăveanu's perspective-affordance approach to discussing creativity highlights the necessity of communication in fostering the discovery of new perspectives to increase creative potential. The role of embodiment in supporting the perception of these affordances is further detailed in Gibson's theory, which is rooted in embodied cognition, or the ability of organisms to use multiple senses to gauge the utility of their environment and the materials in it. The relationships formed between the users and AI through the application of these theories over the course of this research will be discussed through the concept of structural coupling, which posits the mutually beneficial role that interdependent organisms play in each other's functioning in specific environments. These theories were selected for their ability to create a link between creativity and embodiment in the context of both humans and computers, given that generative AI is designed to simulate human intelligence and creativity. Chapter 4 will discuss the methodology through which this thesis research is conducted and how these theories are applied and further analysed through iterative prototyping, observational research and literature and contextual reviews.

4. Research Methodologies

This thesis takes a mixed methods approach, using a detailed literature and contextual review, as covered earlier along with Individual Action Research and Iterative Design as practice-based research methods. Through a cycle of reviewing theoretical frameworks, ideation, design, reflection and evaluation, the product of the thesis, which is the discussion of suggested guidelines for embodied engagement with Generative AI, will be conceptualised and iterated. Although this final product is theoretical in nature, an iterative design practice is a necessary part of the research for a critical and detailed observation and evaluation of applications of such a protocol in real world scenarios. These observations take place through my own engagement with the prototype as well as installation and exhibition of the prototypes in gallery environments, to maintain the artistic context that this research takes place in (See Figure 6).

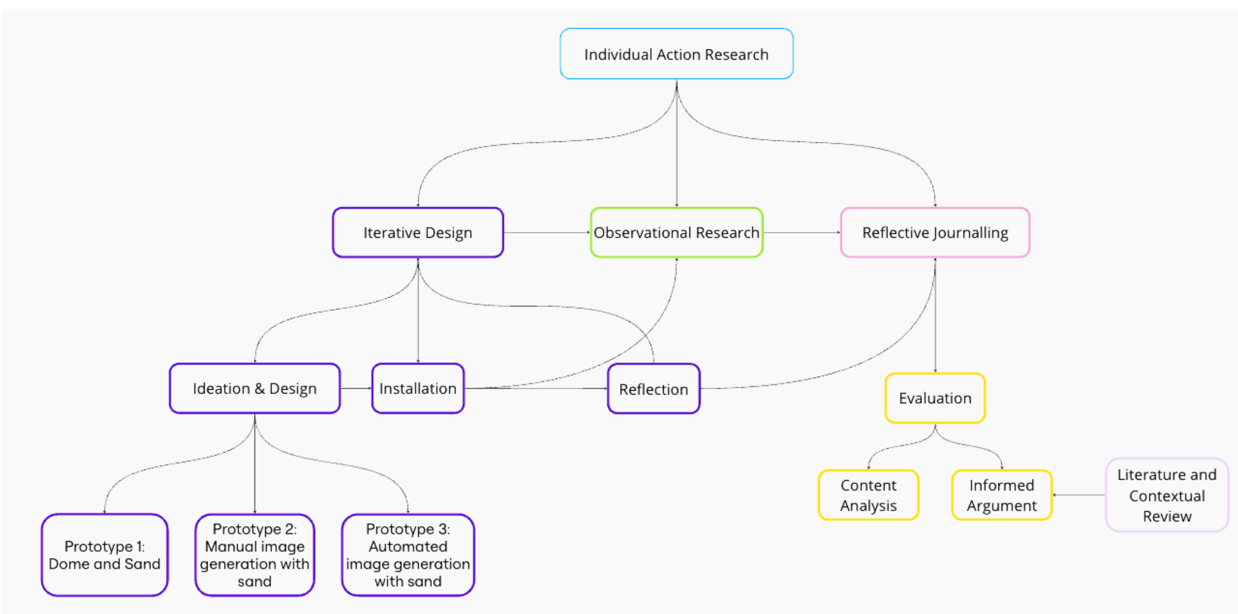


Figure 6: Map depicting the steps of the methodology

4.1 Individual Action Research

This thesis will be using Individual Action Research (IAR) as its primary methodology, using iterative design and Personal Action Research Journals (PARJs) (Willis J.W. et al, 2014). IAR takes an interpretive and self-reflective approach to action research, using PARJs to comprehensively document each iteration of the research process and critically analyse and evaluate the findings to inform the next stage of research (Willis J.W. et al, 2014, pp. 96-97). According to La-Kicia K. Walker-Floyd, self-study is central to IAR, using subjective qualitative data with an understanding and acknowledgement of socio-cultural contexts and biases, while still aiming to achieve researcher objectivity (Willis J.W. et al, 2014, pp. 96-97). Identification of a specific issue in its historical and social context, designing of experiments and journal-based reflection that borrows from phenomenological methods of analysis are the key steps in conducting IAR within this thesis. Journal-based reflection can be analysed through the coding of information to identify patterns that lead to valuable insights. These insights inform the next iteration of the potential solution, which in this thesis, is the complete process of creating with generative AI using the kinetic sand. As using generative AI requires a highly iterative process in its existing form, with targeted areas of improvement being identified and reworked to achieve the best possible outcome, translating this methodology to the research conducted in this thesis proved to be a pragmatic approach to take.

4.2 Iterative Design

The Iterative Design methods in this thesis are grounded in the theoretical frameworks explored earlier in the literature and contextual reviews, which are then used to prototype each interactive installation for user observation. Iterative design begins with the identification of a goal within a problem space, followed by a cycle of ideation through reasoning, empirical observation, and strategizing, to inform possible solutions, that are then revised within the cycle again (Adams and Atman, 1999). Iterative design is a popular choice among engineers and designers in the field of Human-Computer Interaction (HCI) (Adams and Atman, 1999), making it an appropriate choice for the design practice through which I am conducting my research, due to its dependence on efficient and intuitive human-AI collaboration. To effectively study this collaboration in real-world applications to generate well-evaluated protocols, designing effective systems of interaction will be crucial to the research. Additionally, as this collaboration is centred around proposing tactile interfaces for human-AI collaboration, ensuring the system is functioning in its best possible state will require an iterative approach.

4.3 Evaluation

To evaluate the prototypes at each stage, I will be using a combination of observations through the prototype installations as well as continuous self-reflection through my own interactions with the Generative AI system, which will both be documented through descriptive journaling. This will be analysed to evaluate each system for its ability to create an embodied system of interaction, using five criteria:

1. Sensory Stimulation: Based on Gibson's theory of affordances, an organism's understanding of a material's utility is linked to its sensory perception of that material, relative to its own needs at the time. This

requires more than just a visual understanding of the material's properties, engaging with the body as far as possible, to evoke memory or enable new discoveries (Gibson, 1979, p. 143). This is also in line with Glăveanu's perspective-affordance theory of creativity, where the affordances of material objects or of unique idea combinations guide the development of novel perspectives in creative work (Glăveanu 2020, p. 9-12)

2. Level of Control: Based on Pasquinelli and Joler's articulation of the need to demystify AI and mitigate the rhetoric of the black box effect (2021, p.p. 1267), the prototype should support and signify the user's agency in the process of image generation. While the AI model functions as a tool of perception, the user needs to be the one in control of the creative process through their embodied actions and manipulation of the physical material.
3. Level of Creative Satisfaction: Based on Glăveanu's model of creativity, the perception of affordances and unique ideas, guides the development of new creative perspectives (p.p. 12), as mentioned before. Therefore, it is essential for the prototype to fulfil the user's aim to create a unique artwork with the physical material as well as the AI model, to demonstrate how embodied engagement with generative AI could increase the user's creative satisfaction with the process as well as the generated images.
4. Level of Collaboration: Glăveanu's model of creativity also emphasises the importance of an exchange in perspectives, which needs to take place between the user and the Sand-AI system in this research. This provides a more collaborative space for the creative process of image-making to take place in and signifies the effectiveness of the user's embodied interactions

with the sand in communicating a desired image to the AI model. This also points to the concept of structural coupling in humans and AI, and the reciprocal nature of the interactions taking place between the user and the Sand-AI system (Maturana, 2002, p. 16). Emphasising the level of collaboration through the prototypes could present insights for the future of relationships between humans and AI as generative AI becomes a more tool prevalent in everyday life.

5. Fidelity of the Output: This metric is indicative of the functionality of the AI model in communicating its understanding of the user's perspective of the kinetic sand back to the user, through the generated image.

These five criteria form the observational parameters of my prototype evaluation, which finally inform the set of guidelines for embodied creative practices in image generation. To validate and recontextualise the prototypes and guidelines within their fields, I will be using an informed argument (Hevner, 2004) to evaluate these two artifacts against other related work in Generative AI and Computational Creativity research.

4.4 Summary

The iterative and highly reflective process of this thesis research revolves around using the insights gained from the literature and contextual review and the three key theories to inform the process of designing and engaging with the prototypes and the observations derived from their installation. Although the application of IAR in the case of Walker-Floyd's research focused on designing effective teaching methods, which is a largely different context, I was inclined to borrow steps from this methodology as it provided a space for subjective and critical observation in this research, which is essential to the process of conceptualising the set of guidelines at its culmination. Detailed, systematic documentation of both my own experiences with the interactive system as well as its functionality in a gallery environment provide a way to deeply analyse how the creative dynamic between humans and generative AI is impacted by embodied interactions, which is a highly individual phenomenon. Therefore combining the more subjective, theoretically informed and individual-focused methods in IAR with the iterative design processes of HCI provided a systematic approach to conducting this research. This is demonstrated in the following chapter which details the process of ideation, execution, reflection and evaluation for each of the three prototypes created in this thesis.

5. Creation of the Interactive System

Due to a lack of existing accessible systems for embodied engagement with Generative AI, I chose to create a set of prototypes that could be used to examine the impact of various factors on the user's perception of the image generation process. My material of choice for this system is kinetic sand, which possesses distinctly unusual material properties of viscoelasticity, making it mouldable yet fluid in its form. This would enable a strong sensory experience for the user, with intuitive tactile interactions like touching, pressing and sculpting that could be translated into data for the AI to perceive. The first prototype explores the feasibility of the sand in achieving this effect while the second tests at the communication of the visual properties of the sand to the AI model and its impact on the output. Finally, the third prototype streamlines this process by automating the communication of this data and adding a hand tracking component to increase the dialogue and therefore, the level of collaboration between the user and the AI model. (See Figure 7)

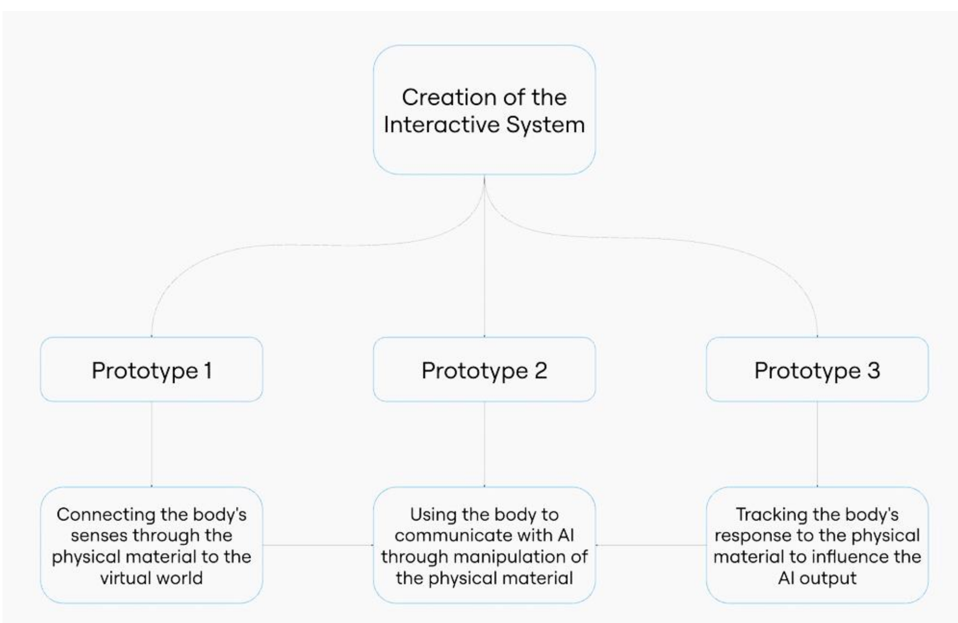


Figure 7: Overview of prototyping process

5.1 Kinetic Sand as Creative Interface

Play-based learning is a globally popular method of early education, deemed to empower the child to learn through their own experiences and imagination. From psychologists like Eleanor K. Gibson who are proponents of embodied cognition as a form of inquiry, to educators like Maria Montessori who encourage independent learning and the empowerment of children in their developmental stages, play as education is a long-researched and well-supported domain. Today, businesses and institutions alike are engaging with the field to bring new methods of play to the world, through products and systems that sustain multidimensional forms of play. Sensory play is one among these trending topics of interest, believed to be highly stimulating to the brain and promoting increased cognitive development and fine motor skills (Gainsley, 2012). This emphasis on the importance of multi-sensory interaction as a mode of learning gave rise to innovative new materials containing huge varieties of textures, visual features, scents, sounds and opportunities for creative use. Kinetic sand is one popular play product in this area, being a hydrophobic material that looks and feels like wet sand but is dry to the touch (See Vol. 2, Fig. 8). Additionally, when submerged in water, the sand can continue to be moulded and shaped, much like regular natural sand, and remains dry even immediately after removal from the water. Spin Master, one of the most popular sellers of the product, utilises these affordances to its advantage, creating toolkits of various accompanying accessories designed to promote play with the sand in different ways, through themes like "Digging for Dinos" or "Ice Cream Station" as well as simple tools to manipulate and explore the texture of the sand itself (spinmaster.com, n.d.).

Figure 8: Kinetic Sand

Given the distinctive properties of this kind of sand, it is worth noting that it has not been extensively explored as a medium of interactive art. While there have been a few examples of its use in exhibits such as *Please Do Touch The Art* by Dan Lam, where viewers were invited to break the tension between the art and the audience by touching and interacting with the original sculptures made by Lam (waterfall-gallery.com, n.d.), explorations with its one-of-a-kind visual and tactile appeal have been limited. My explorations with kinetic sand as a potential material for my thesis experiments, examined it not only as a medium of multi-sensory interaction for human beings but as a sensing object itself.

My understanding of its affordances was further developed purely through my own meditative interactions with it, which included:

1. Running my fingers lightly across the surface of the sand to observe the ridges that formed and the differences in these ridges based on how densely or loosely packed the sand was kept.
2. The surprisingly slow speed with which the granules fell apart when prodded or released from the fingertips, owing to the chemical composition of the material that gave it its viscoelasticity.
3. Closing my eyes to feel the texture without any visual cues led to associations with gooey, slimy materials that are typically unpleasant at first yet compelling to continue exploring.
4. Forming the sand into a dense block and cutting it into slabs with a knife, to notice the unusual structural integrity it held when sculpted into shapes.
5. Pressing down on the surface as though it were clay, to release my own pent-up energy while also creating organic forms.
6. The scent released by certain varieties of kinetic sand that clung to my fingers afterwards, that added to the soothing effects of the experience.

When placed in the context of a digital environment, as this thesis is focused on engagement with generative AI, the kinetic sand transformed from a play object to a combination of creative medium as well as interface. My analysis of the sand shifted from observation of its material qualities to a search for potential data collection points as it took on the active role of a sensing entity. As my aim now was to find ways in which the sand could collect data from my body to send to the computer, I took note of the specific points of interaction, such as pressure; the coordinates of my fingers and palms as I moved my hands through it; the speed with which I moved; the areas where the sand was more densely or loosely packed; the shapes that the sand formed and so on. The distinction between the organic nature of the sand itself, and the programmed, quantitative nature of the computer, created a space for me to better consider the possibilities of communicating with machines in new ways, by better utilising alternative forms of conversation with the human body. The sand offered opportunity for non-textual cues that are commonly used in traditional artistic practice to be sent to the computer, allowing for greater creative potential from both the human and the machine.

5.2 Prototype 1: Kinetic Sand and Sensory Play

The main aim of this prototype was to demonstrate the connections and disconnections between the mind and the body when engaging with a medium that combined analogue and digital materials. As my thesis explores the impact of embodied interactions when engaging and creating with generative AI, I wanted to begin the research process by gaining a deeper understanding of unusual human-computer interactions and the impact such interactions could have on a creative experience. To do this, I designed an object that would bridge the gap between analogue and digital interfaces to maximise the use of specific senses, natural human instinct, and playfulness (See Figure 9).

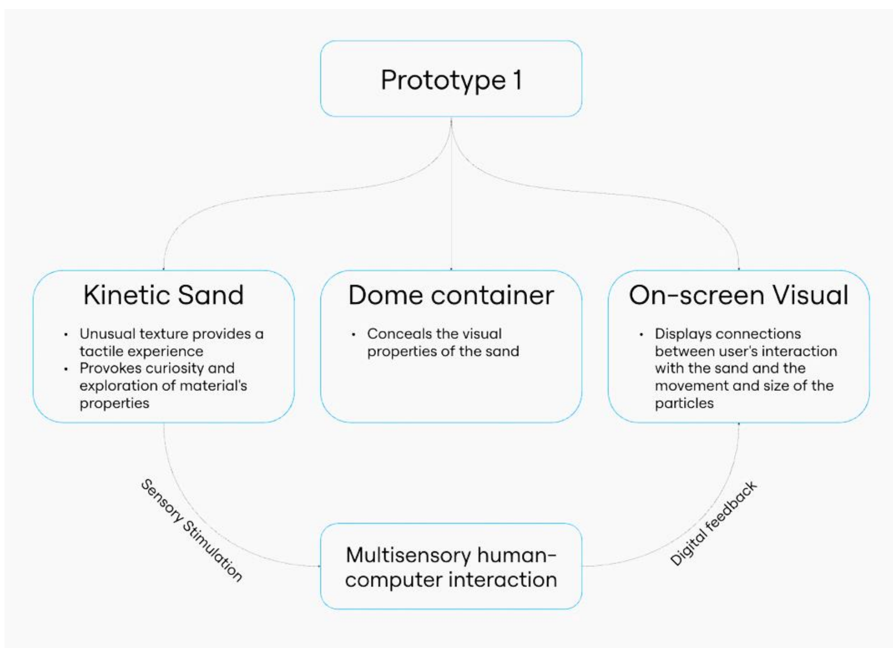


Figure 9: Overview of Prototype 1

I decided to design a dome to contain the sand, as it offers a more organic, less harsh shape than a simple box and would be slightly less intimidating to the user, considering they would be unable to see the contents. I opted for a geodesic dome structure that could be built out of cardboard as this offered good structural integrity as a base for papier mâché and paint to be added for detailing. I decided to change the contents of the dome from traditional art materials to kinetic sand because, in my experience working with children's creative toys and games, I had found popular sensory play materials to be much more exciting and unexpected in their textures, as they are designed for cognitive and motor development in early childhood.

5.2.1 Process

The dome was constructed using found cardboard, of a single-wall box. This made it easier to cut while retaining strength from the corrugation. I used approximately 30 isosceles and equilateral triangles altogether, and mounted the dome on an asymmetrical curved base, to resemble natural forms like caves and dunes. The entire structure was then covered in papier mâché for additional support and texture and painted blue and brown. I allowed the paint to apply in uneven strokes and shades to intensify the effect of underwater light and shadow and as inspiration from colours that can be found in nature (See Figure 10).

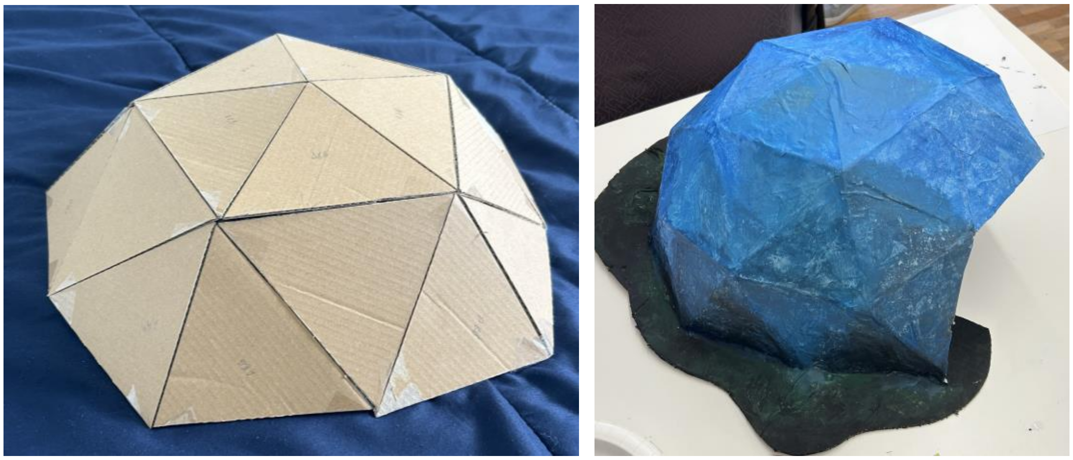


Figure 10: The geodesic dome, before (left) and after (right) painting

For the digital interactivity, I used an Arduino Nano IoT device with an ultrasonic sensor, to collect data from the user's interactions in the dome and send it to a P5.js file that displayed visual feedback on a projection. The setup was extremely simple with only one sensor that measured distance using ultrasonic "pings" - this was to minimise the amount of computational power required to run the installation, to ensure the visual feedback could be easily configured and to simplify the pipeline of information. I realised that the participant's interactions did not have to be directly or "accurately" measured in terms of pressure, shape, or gestures, which were some of the initial ideas I had considered, but rather needed to simply be reflected as visual feedback in some dynamic form that appeared to be influenced by their actions. Placing this distance sensor at the back of the dome allowed me to carry this out effectively and focus on the graphics of the visual and how that might impact the experience, rather than the specific type of data collected.

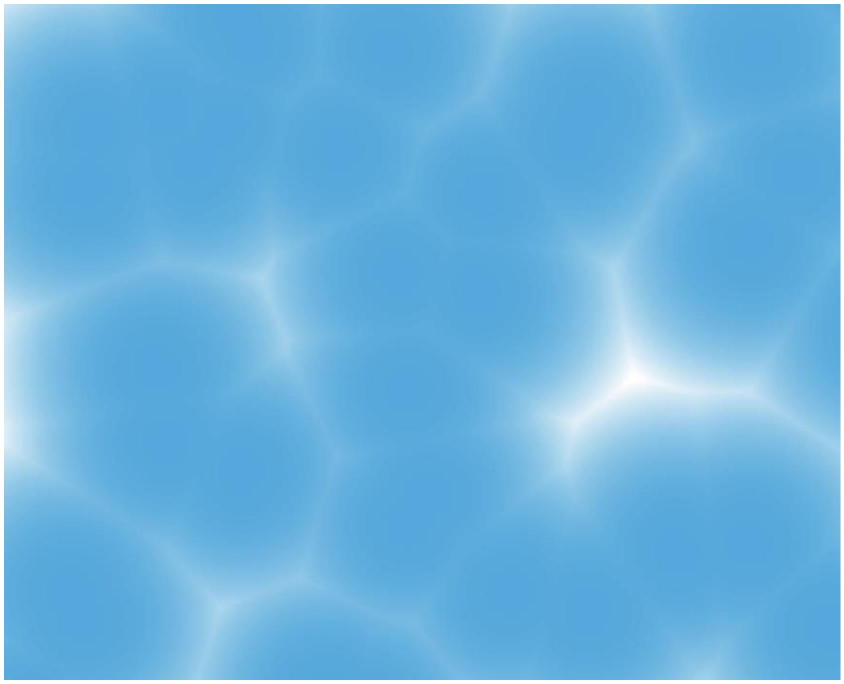


Figure 11: First iteration of the water surface visual

I took inspiration from the texture, reflectiveness and movement of water, which offered a contrast to the grainy, unusual feel of the sand yet conceptually matched it well, especially in combination with the cave-like appearance of the dome - this narrative was inspired by the notion of deep sea exploration, where one ventures into the unknown, relying on senses, existing knowledge and intuition to carry out research and inquiry. I also appreciated how the ambience created by the projection of water-like visuals was calm and soothing in contrast to the potential discomfort caused by the texture of the sand (See Figure 11).

It took several iterations using trial and error to use the data from the sensor with the displayed visual, as the file grew heavy and was difficult to process on my laptop. I decided to go with dynamic particles to reflect the grains of sand and mapped the origin points of each particle group to locations determined by the distance of the hands from the sensor (See Figure 12).

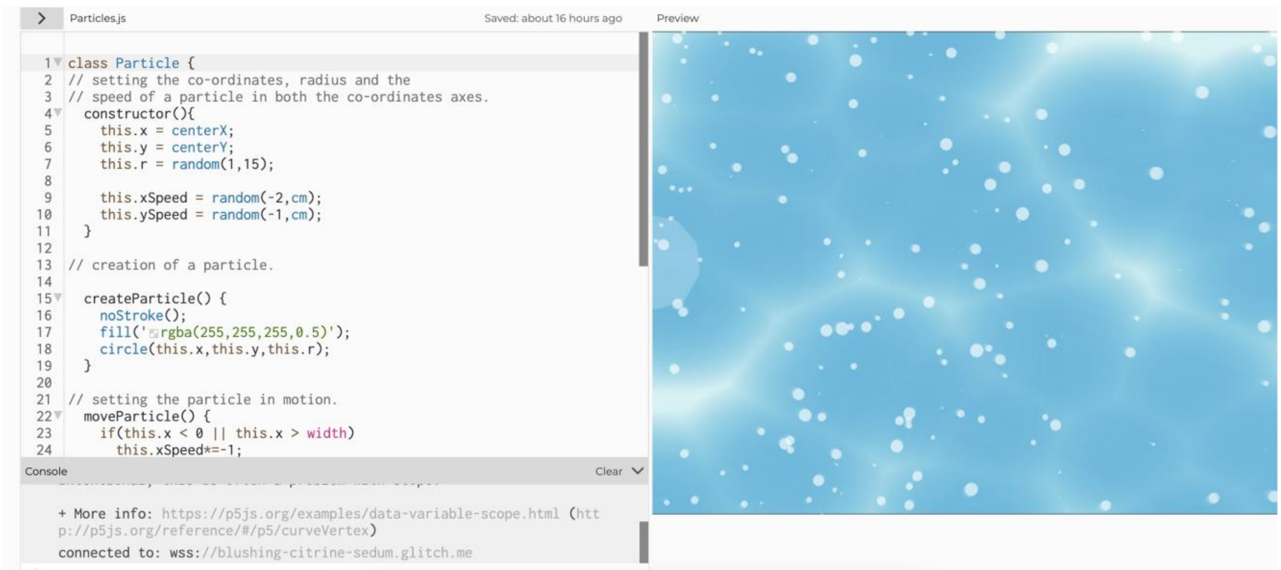


Figure 12: Second Iteration of the water surface visual

5.2.2 Reflection

Upon the initial presentation of the prototype, a setback arose as the sensor data failed to be interpreted by the P5.js file, rendering the prototype non-functional. Despite this technical limitation, participants were able to engage with the dome interface through available visual cues, subsequently expressing various reactions to the kinetic sand, including interest, provocation, and a mixture of attraction and aversion. This provided an opportune moment to raise inquiries regarding the agency perceived by participants during interaction, particularly whether they felt aligned with the computer or the tactile medium.

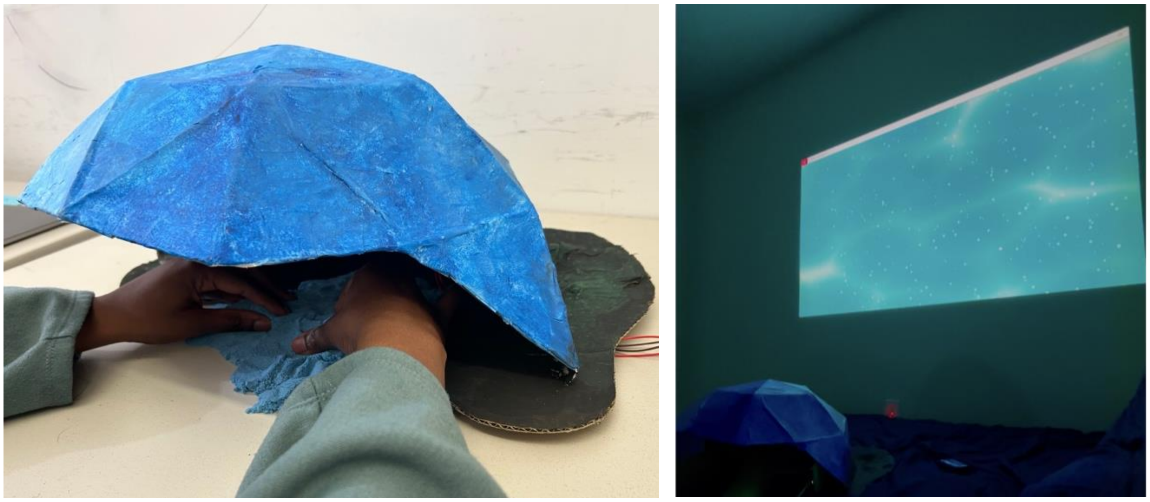


Figure 13: First iteration of Prototype 1, with the interactive particle system

Additionally, there were discussions around whether any repulsion experienced was intrinsic to the sand itself or a deliberate facet of the experiential design. While responses to these inquiries varied, their significance in shaping the subsequent prototype iteration, pivotal to the ongoing thesis research, was duly acknowledged. Noteworthy insights gained from the prototype experience primarily pertained to the intention behind the selection of materials and the manner of interaction design. While initial focus had centred on aspects of human-computer interaction and environmental perception, a deeper exploration into the nuanced aspects of analogue interfaces emerged as imperative for clarifying how participants are drawn towards the digital realm through the inherent affordances of its interactive components. In applying these insights to a text-to-image AI model, it could be said that while a computer keyboard allows the user to input specific word-based guidance, kinetic sand could provide space for intuitive and embodied knowledge to emerge through the user's perception of the material. This indicated the necessity of incorporating more sensor data such as computer vision and hand tracking in future iterations.

5.2.3 Evaluation

The two installations of this prototype primarily highlighted the efficacy of the kinetic sand in physically engaging the user in a sensory experience (See Figures14 & 15). Emotional responses to the sand were highly descriptive and physical responses indicated observably strong reactions, both positive and negative. Some of the repeated descriptors used were “disgusting”, “weird”, “gross”, “satisfying”, “fascinating” and “scary”, all of which indicated strong physiological aversions or affinities for the material. I conjectured that the negative reactions were intensified by the inability to see the material, which was in line with my goal of using the dome set up to be able to better augment and observe the sensory experience of the user.

However, almost all the users, expressed a feeling of disconnectedness from the visual on the screen. While the ambience of the installation was soothing and provided respite from any negative feelings towards the sand, there was no other visual or textural connection between the tactility of the physical experience and the vectorised digital visual. The link between the movement of the hands in the doom and changes in the visual were perceived by the user but the actual cause of these changes could not be identified leading to a lower sense of control over the digital space.

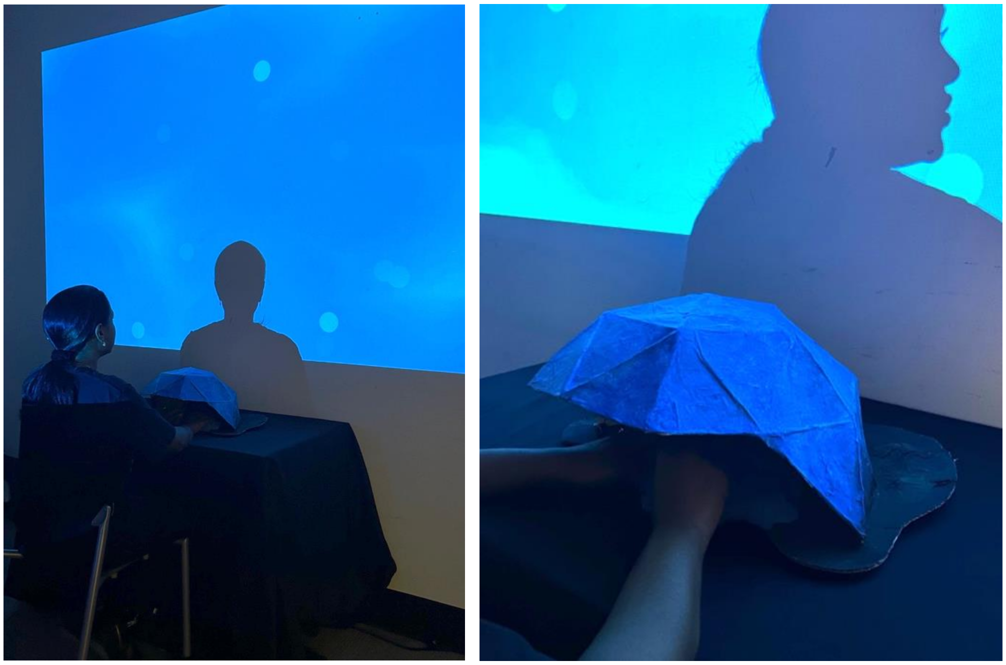


Figure 14: Final version of Prototype 1

Criteria	Level
Sensory Stimulation	High
Level of Control	Low
Level of Creative Satisfaction	Low
Level of Collaboration	Medium
Fidelity of Output	Low

Figure 15: Evaluation of Prototype 1

5.3 Prototype 2: Kinetic Sand and AI

The first two installations of the experiment revealed that merely relying on material affect did not suffice to foster engagement with the digital interface. While the absence of a visual connection between the kinetic sand and the computer visual presented an intriguing tactile experience, it also resulted in a perceived disconnection from the digital realm in terms of medium and emotional response. Nonetheless, the material demonstrated efficacy in captivating users' interest in the digital sphere, facilitating exploration of the dome space through intuitive motions and responsive interactions tailored to users' perceptions of the sand. Some users' positive response towards the unconventional fusion of technology and bodily capabilities underscored the importance of maintaining this effect in the subsequent prototype iteration while mitigating any sense of disconnection. Considering these pros and cons, the decision was made to centre the interaction around kinetic sand as the primary medium, with AI serving as a complementary, yet valuable collaborator. Furthermore, to enhance the ability of the computer to analyse the users' interactions with the sand, I planned for the integration of computer vision and an image-to-image pipeline within the AI framework (See Figure 16).

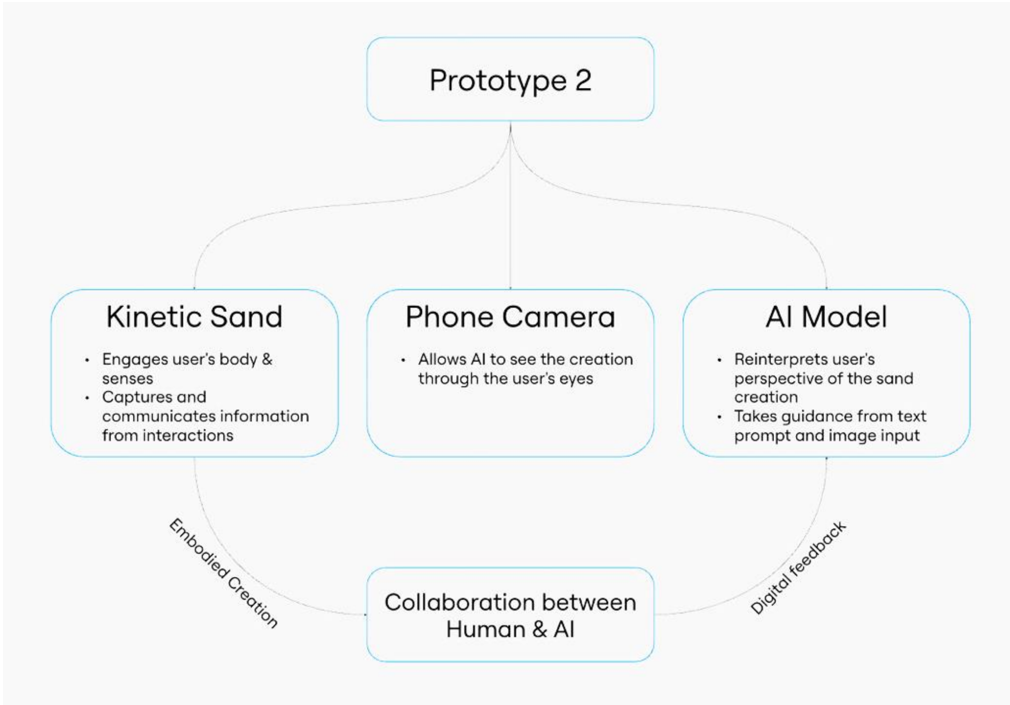


Figure 16: Prototype 2 Overview

5.3.1 Process

I used Touch Designer to set up a camera feed linked to a computerender API (compute(r)ender, 2023), a service that allowed me to integrate Stable Diffusion into my prototype, with all basic parameters like seed, guidance scale, strength, and iterations available for use within the Touch Designer workspace itself (Blankensmith, 2002). This allowed me to connect a video feed of the sand, which I set up on my phone as a testing device, to Stable Diffusion, providing the required image input. Initial experimentation with the basic functionality of the setup revealed a creative process that I found to be extremely interesting, simply by virtue of the way the AI interpreted the images I captured. Enabling the AI to see the environment it operated in, opened an avenue for the computer to seemingly respond to its environment while also following the text prompts I put in. For the initial test prompts, I opted for a level of specificity that guided the AI towards a certain art style and subject matter while still leaving room for random choices in the details of the images. Analysing the first few iterations of the images revealed a pattern in the way the AI filled in spaces, which I assumed to be the knowledge of aesthetics and composition that was incorporated within the training data (See Figure 17).

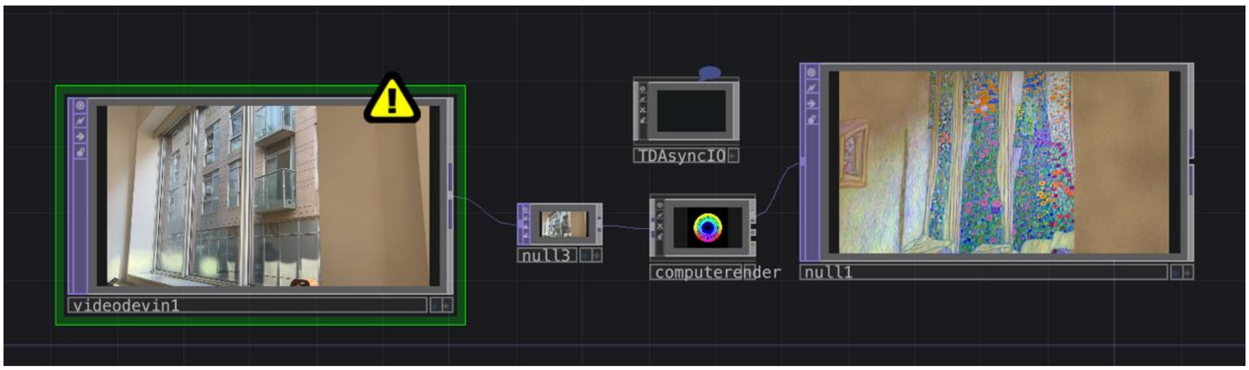


Figure 17: Image-to-image generation, Test 1

The next step was to bring back the kinetic sand, which offered further opportunity to control both the image I fed into the program and the prompt itself. Through my interactions with the sand, I found myself noticing specific shapes and forms that emerged from the imprints I left, inspiring prompts that were less specific with details but relevant to the image input, in my eyes. As this flow seemed to naturally emerge from the interaction, I decided to generate more images, testing how well the AI would replicate what I saw in the sand and how much it would either surprise or frustrate me in the process. Based on the prompts I put in, it seemed to achieve both effects, following the broader, landscape-based, or abstract prompts with better results and failing to create a coherent portrait based on my crude image input. When prompted towards abstraction and more stylistic terms like “playful” or “textured”, the generated images were interesting, but did not feel connected to my initial process, or the sand itself and were conceptually dissonant (See Figure 18).

Figure 18: Image-to-image generation, Test 2

Finally, I tested how the AI would respond to the image input when guided by extremely vague prompts, to study the impact of the lighting, textures and prints in the sand and the colours on the final output. Using ambiguous prompts, such as "freeform, any style," led the AI to select subjects such as wavy hair or loosely draped cloth, which exhibited no recognisable correlation to the sand medium, beyond the AI's translation of the interplay of light and shadow on the surface of the sand into the final image. In that moment, the visuals elicited a sense of excitement, attributable to the element of surprise at the specificity of the subjects of the images. The AI's apparent utilization of "creative" interpretation to output imagery that transformed abstract concepts into recognizable subject matter could be compared with the creative techniques employed by artists to overcome creative block or approach a new artwork. Techniques such as sketching freehand squiggles and subsequently reinterpreting them with an open-minded and discerning eye into specific subjects epitomized this process.

Upon receiving slightly more directive guidance via the text prompt "freehand, artistic style," the output was less imaginative, but still visually compelling in its composition and aesthetics. While it contained more abstract visuals that were still derived from the forms created in the sand, it was interesting to note a recurrence in more fluid or smooth styles of visuals rather than a highly textural subject like the sand itself (See Figure 19).

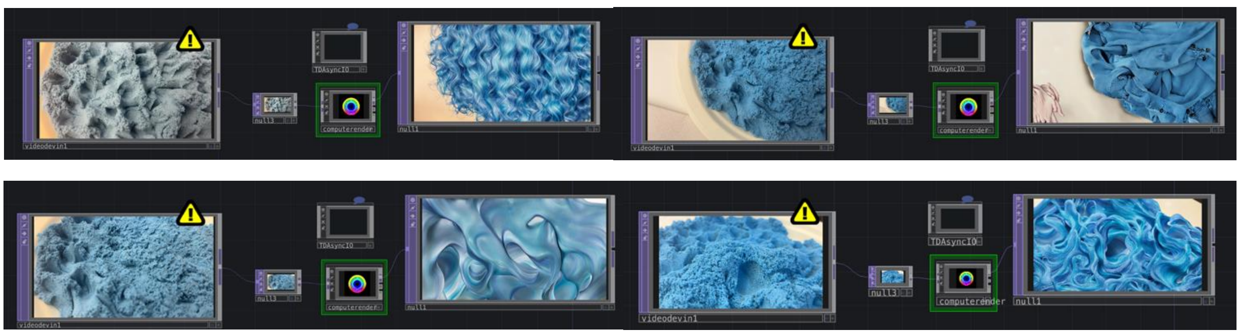


Figure 19: Image-to-image generation, Test 3

The former phase of this round of testing highlighted the importance of maintaining greater control over the text prompts. While the element of surprise imbued the interaction with excitement, it revealed a potentially problematic perception of the AI as a sentient, deliberative entity, veiling its fundamental nature as an algorithmic system simply trying to find a coherent match for the input using a premeditated set of training data and rules and a mathematically randomised selection process. This approach also takes away a greater level of control from the user, relinquishing the final decision-making process to the AI model itself, and making the question of creative ownership even more complex. This approach to creating the interactive system would only further exacerbate the black-box effect and go against the intention of the research, which is to find ways to collaboratively create with the AI, using it more as a tool through which to gain additional perspective in making images.

Based on these 3 rounds of testing, I opted for a predetermined text prompt that was both detailed and style specific as the initial point of interaction. This prompt could later be adjusted based on what the audience perceived in the sand as they interacted with it. This approach introduced an additional layer of complexity to the collaborative process by integrating my own interpretations of the sand into the user-AI collaborations. It also delineated my role as an artist within the gallery setting, inviting users to partake in my collaborative endeavour with the Sand-AI system. This approach would still only function as a starting point as users retained the freedom to draw inspiration from the sand during their interactions and propose their own prompts if desired. However, a notable drawback was the necessity for my substantial involvement in each user's experience to adapt prompts or parameters as necessary and manually generate the images. Nevertheless, constrained by the limitations of the timeframe available for exhibiting the system in a gallery context, I resolved to proceed with this version of the system. I saw this as an opportunity to observe users' reactions to the collaborative Sand-AI experience rather than conducting further user testing for functionality.

Figure 20: Final version of Prototype 2

The gallery installation of this system comprised a straightforward setup, maintaining focus on the essential elements: the sand, the camera feed, and the AI-generated output. To do this I used a single screen containing the video input and AI output windows from the Touch Designer program itself positioned in front of against a moving particle system depicting the last generated output. The main purpose of the particle system was to keep the visible screen in an idle state, inviting viewers in to interact with the system. Four colours of kinetic sand were placed in individual containers to avoid colour mixing, which offered a larger variety of options to the user compared to the single blue used for initial testing. Two trays were also provided for greater surface area to explore the sand with. I continued to use my phone as the camera, affording users the freedom to frame the sand according to their own visual perceptions, giving them control over the computer's view of the imprints left in the sand through the framing, lighting, and camera angle of the image input. This installation experienced a large amount of footfall with participants from a variety of backgrounds, offering opportunities for more nuanced insights (See Figures 21 & 22).

Figure 21: Exhibited Prototype 2

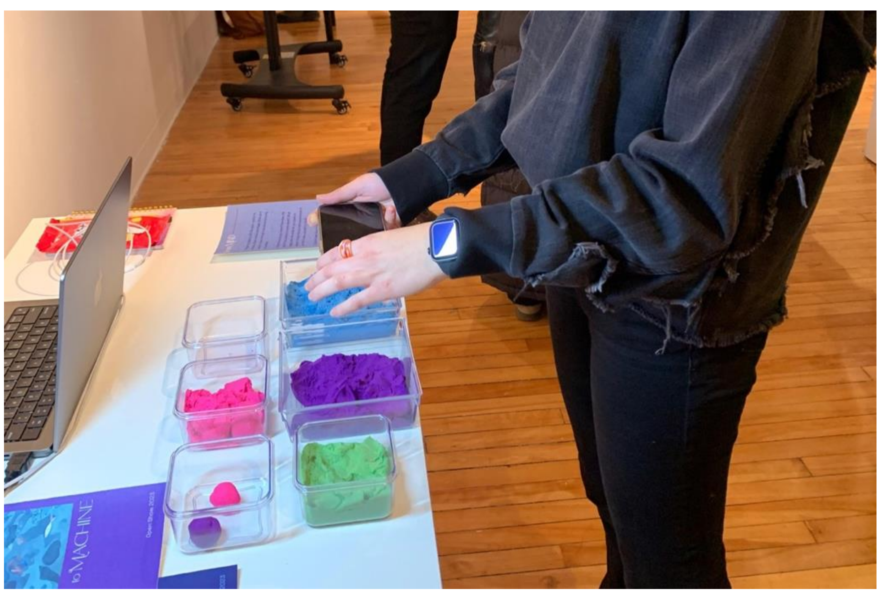


Figure 22: Visitor interaction with Prototype 2

5.3.2 Reflection

This prototype was installed using a very simple setup that maintained focus on the colourful varieties of kinetic sand available for the users to interact with and the large screen that provided a visual depiction of the creative process to onlookers in the space. I did want to provide some context for the set up through a set of small posters that gave the audience a glimpse into what could be created with the AI program as well as a set of broad guidelines on how to engage with the installation. I found that given the high traffic of visitors to the gallery space there was little to no time for reading the posters or engaging in a slow and meditative interaction with the installation like I had intended. Almost all of the interactions were extremely quick, with users first being drawn in by the appearance of the kinetic sand on the table, further emphasising the success of the tactile, tangible material in drawing the user towards the digital. The users then appeared more excited to generate the image once they found out that the process culminated in the creation of an artwork through the use of AI.

A significant drawback of requiring a high degree of manual control over the program was that I had to be present at the installation as both the designer of the program as well as an additional facilitator in the creative process. This involves guiding users towards playing with the kinetic sand, motivating them to seek out visuals in the sand itself that I could input as text prompts, and then finally generating the actual image while the user held the camera over the sand. In combination with my own prolonged interactions with the sand and considering the relationship I had built with the program over the course of the research so far, exposing the set up to a more diverse audience in a gallery context generated a much richer breadth of insights that could contribute to the overall principles I seek to establish at the end of this thesis.

5.3.3 Evaluation

There were several key insights gained by identifying patterns between the preliminary rounds of testing and the observations made during the gallery installation of this prototype. In comparison to the first experiment, the connection between the digital and the physical was significantly higher. The sand successfully maintained its tactile qualities, and with the added ability to see the sand while working with it, created an engaging sensorial experience. The material still seemed to be foreign to the users which caused some hesitation in uninhibitedly interacting with it, but those that chose to explore it fully appreciated the strangeness of its properties in the same way that users did in the first experiment, if not to a greater degree due to the added visual perceptions of the material (See Figure 23).

Criteria	Level
Sensory Stimulation	High
Level of Control	Medium
Level of Creative Satisfaction	Medium
Level of Collaboration	Medium
Fidelity of Output	Medium

Figure 23: Evaluation of Prototype 2

While the connection between the physical and digital in this set up was significantly higher, this was mostly due to the direct connection between the visual of the sand and how it was translated into the AI output. My role as a facilitator in all the user experiences and the rushed environment of the gallery limited the time that users could spend with the set up. Despite this, users were still compelled to identify interesting prompts or imagery in the sand after they had interacted with it and appreciated the manifestation of these visuals in the AI output, revealing a level of creative satisfaction in the outcomes of the process.

5.4 Prototype 3: Automated Interactions and Hand Tracking

5.4.1 Process

With this prototype I wanted to streamline the interaction with the machine and allow the user to focus on their embodied experience and creative process by reducing their dependence on traditional computer controls like clicking to generate the image or typing to tweak parameters like the seed with every iteration. To achieve this, I used the same set up as prototype 2, with multiple colours of kinetic sand available for use.

The process of creating the image was to be carried out in two stages:

1. The user interacts with the kinetic sand in any manner they like, allowing the unique affordances of the material to guide them towards creating any form or pattern.
2. The user then picks up the phone camera and frames their creation in any angle that they prefer. The act of picking up the phone triggers the AI to begin generating images based on the image and pre-determined text input it receives. The user continues modifying either the sand or the camera angle until they are happy with the images (See Figure 24).

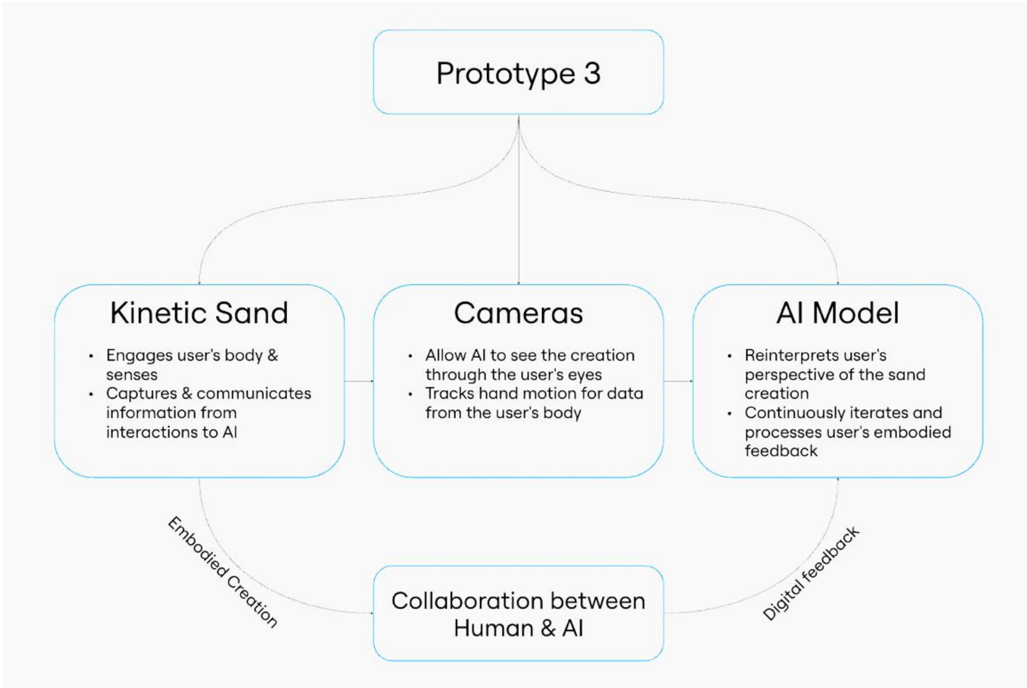


Figure 24: Overview of Prototype 3

Within the existing Touch Designer program I had set up in the second prototype, I used MediaPipe, a computer vision plug-in created by creative technologists Torin Blankensmith and Dom Scott (Blankensmith, 2023) to enable hand-tracking within the program, which allowed me to collect real-time data from the user's hand movements. I chose to link the "velocity" data with the seed parameter in Stable Diffusion, based on my observation of rapid motion being linked with an urgency to greatly modify the generated image, while slower movements focused on exploring specific areas of the material and adjusting details. The rapid motion would cause the seed to change to a greater degree, resulting in an image that would differ more significantly from the previous iteration and vice versa (See Figure 25). I also set up a "blob" recognition element within the program that would use the detection of light in the camera to automatically prompt the system to generate an image, thus removing the need for a facilitator, or the user themselves to click a button to create the final image while still holding the camera (See Figure 26). This was also intended to bring the user's focus to the repeated image generation, indicating to the user that the creation process would be iterative and inviting them to continue modifying the sand.

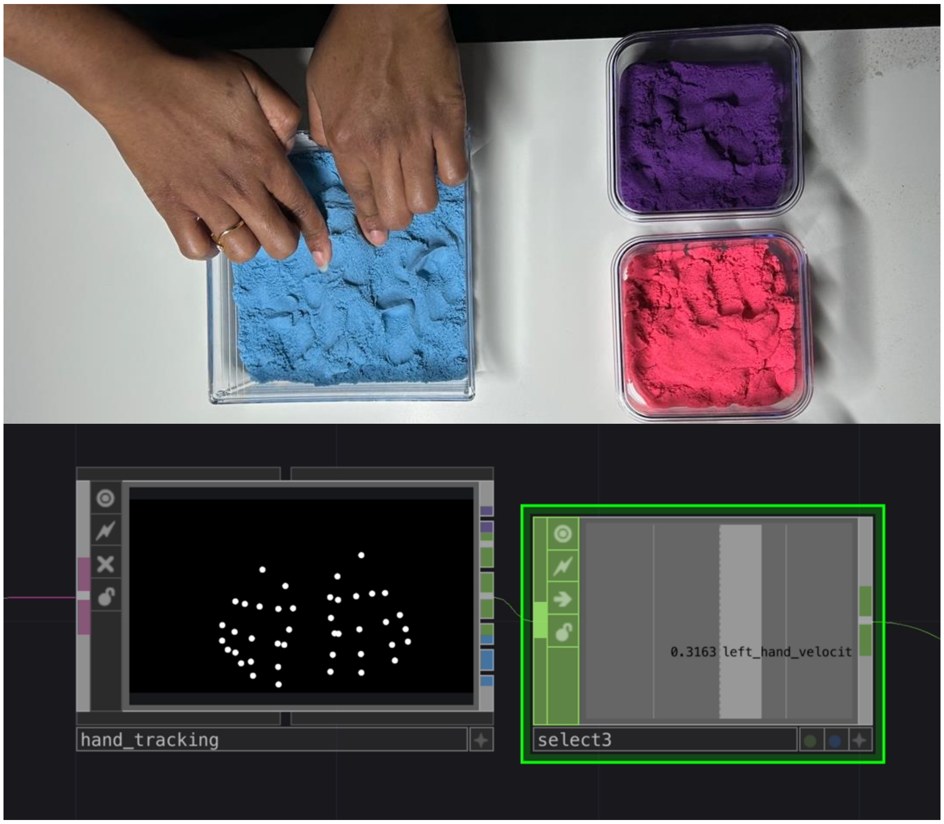


Figure 25: Hand Tracking setup in Prototype 3

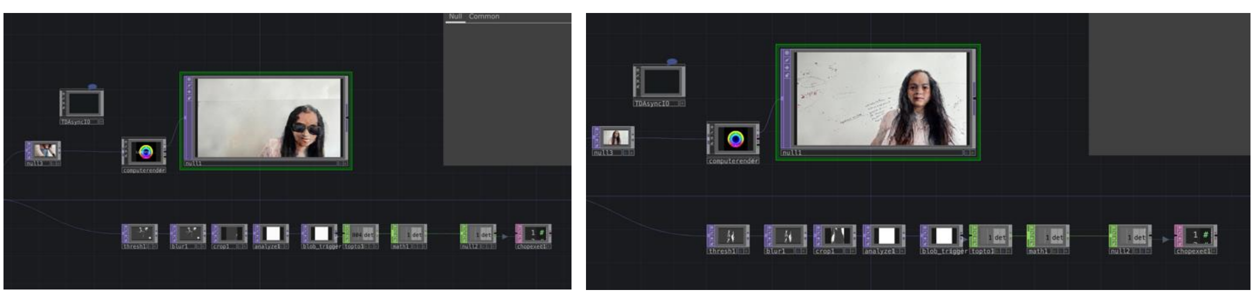


Figure 26: Using image detection to automate image-to-image generation

The first time I tested the prototype as an installation for user observation, my main aim was to test the functionality of the program and to observe whether the user could seamlessly follow the whole process as intended. The installation was a lot more simplistic and less stylised than I intend for the final set up to be. I planned to eventually create a more inviting atmosphere with the set up to provide a stimulating artistic environment for the user and clear points of guidance to communicate what the user is meant to do when visiting the installation. These would involve hanging designed posters depicting my own artworks created with the program and printed labels and signs attached to the table, indicating the order of steps to follow, without distracting the user from the main features. I would have also preferred to provide a larger quantity of sand, to create a visually proportionate set up and to allow greater freedom of creativity and movement during the interaction. I also plan to have a fixed webcam for the hand-tracking element of the program, so that it is not dependent on the user's handling of the camera for it to work.

5.4.2 Reflection

I was largely satisfied with the ease of setting up the full installation as it felt like an easy arrangement for any artist to achieve in their own workspace, without being an overwhelming, inaccessible gallery-specific piece, which I believe would make it more accommodating to an audience with varying levels of expertise. The audience also seemed to engage well with the generated content on the screen and the sense of curiosity, wonder and enthusiasm I hoped to evoke through the sand and the image generation was observable.

The feedback I received was targeted across enhancing several aspects of the installation, to create a better user experience as well as communicate my concept more strongly. One suggestion was to consider the way the camera would be presented - preferably a device that would be held with two hands or contained in a structure similar to a 'GorillaPod', to reduce the banality of handling a regular phone camera. I intend to test a flexible or easily adjustable tripod first as it would prove to be a less time-consuming solution at this stage of the process. However, I also recognised the value of keeping the set up as neat and free of cumbersome equipment as possible, which a case with two handles would achieve.

Another suggestion was to display the generated images, either as a tiled screen or as a slideshow, continuously updating as more users interacted with the program. The tiled screen resonated with me as it would aesthetically present as its own collaborative artwork and the journey from the first user's interaction with the sand to the last's could be traced. I plan to execute this within the Touch Designer program itself and display it on a second screen, perpendicular to the interactive screen, for it to be visible to an audience without being distracting to the user. It was also recommended for the interactive screen itself to be designed to present the generated image more clearly, with the camera feed given less prominence, for both aesthetic purposes and to highlight the standout feature of the installation.

I was also advised to test projecting the generated artwork onto the kinetic sand itself, for the audience to get a sense of how the artwork was derived from their creation with the sand. While the thought of layering the two mediums spoke to me as a strong conceptual and aesthetic approach, I felt like it would work against the idea of focusing one's senses on the physical medium itself before collaborating with the AI and it would limit the camera angles available to the user while capturing the image of the kinetic sand as it would require a fixed camera angle that matched the angle of the projection to achieve the full effect.

5.4.3 Evaluation

Observations made during user interactions in the gallery space revealed several valuable insights that contribute to the experience design of the interactive system itself as well as broader implications for the Human-AI relationships being studied through this research, far more than previous experiments had revealed. A common thread across responses to the creative process was a much higher focus on the AI output itself, rather than a continued engagement with the kinetic sand, especially compared to previous experiments. This shift in focus manifested in different ways, with some users simply spending less time looking at the sand or moving their hands through it, while other users abandoned the sand entirely to investigate what might happen if the camera captured a different visual to be translated into an artwork.

Some larger issues with the interaction design included a disconnect between the text prompts guiding the artwork and the physical materiality of the sand itself – I had elected to prompt the system to generate painted artworks in the style of baroque or renaissance masters, as a nod to traditional artistic practices being amalgamated with new technology (See Figure 27). However, this kind of artistic interpretation was too symbolic and far-removed from what the installation intended to achieve and resulted in more fragmented interactions. Another obstacle in achieving a seamless sand to AI pipeline was the banality and impracticality of the phone camera being used to record the sand and the hand movements simultaneously. This caused difficulties in framing the image of the sand easily and gathering enough motion data from the hands to fully translate the user's physical experience to the AI program.

Figure 27: Images generated by users with installed Prototype 3

Despite these challenges, the installation provided a rich set of observations on the continued relationship-building occurring between the users and the AI program, still successfully mediated by the kinetic sand. There was an improved level of control over iterations when the setup functioned appropriately and the automation of the process allowed for a shift in attention from the numerical parameters (associated more closely with the AI), to the user's own bodily engagement with the system, creating a better balance in the level of collaboration between the human being and the AI (See Figure 28).

Criteria	Level
Sensory Stimulation	High
Level of Control	High
Level of Creative Satisfaction	Medium
Level of Collaboration	High
Fidelity of Output	Medium

Figure 28: Evaluation of Prototype 3

5.5 Exhibition

5.5.1 Process

When preparing for the final exhibition of the interactive system, I used the insights from the previous prototypes to inform certain decisions in the interaction design and set up (See Figures 29-31). One of the main changes made to Prototype 3 was the replacement of automated image generation with a button for the user to click when they were ready to create their image. This was done by changing the camera to a webcam fitted to an iPad, which served as the viewer for the user and programming one of the keys on the iPad to act as a button to initiate the generation process. This provided the user with more control over when they felt ready to create the image and more time to observe the connection between their interactions in the physical sand and its impact on the digital image. I also found it necessary to incorporate a loading screen while the generation process was underway, to provide some form of feedback to the user while they waited for the generated image to appear. The text prompts for this exhibit were pre-set by me for each of the three days, beginning with a retro-futuristic cityscape, then a Van Gogh-style painting of flowers in a garden and ending with underwater scenes filled with marine life. This was done with the intention of demonstrating how each user’s interactions with the sand enabled them to create unique images, despite having the same subject matter in each image. A carousel of the images generated that day was projected onto the wall next to the setup for reflection on the collection of artworks created by all previous users and the similarities and differences between them.

Figure 29: Exhibition setup with Interactive System and Image Carousel

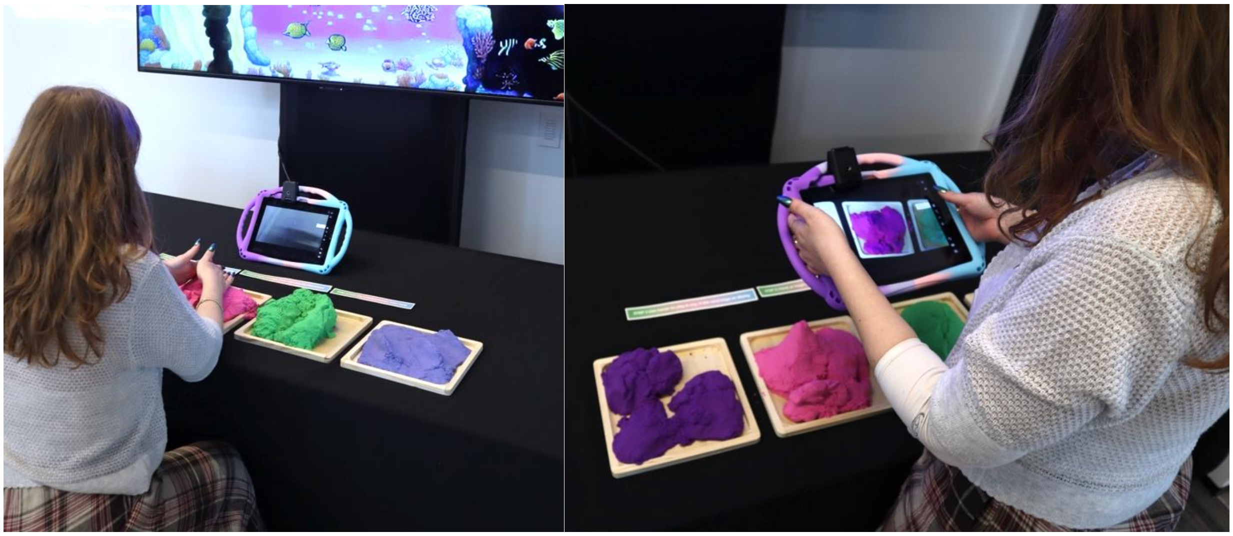


Figure 30: Physical interactions using the kinetic sand and camera

Figure 31: Response from the AI model upon initiating image generation

5.5.2 Reflection

The enhanced interaction design within this setup worked very well for users to independently engage with the exhibit, with minimal guidance. Many users approached the system with apprehension upon finding out that it incorporated an AI model, due to their own preconceived notions of AI as a dangerous technology but were generally positive towards the applications of AI in such a context. Recognising that AI is on track to be an inseparable part of daily life and could also be a useful tool played an important part in this ideological shift, which contributed to the success of the system in mitigating the fearfulness and mystery surrounding AI, without requiring much explicit theoretical discussion.

5.6 Summary

The three prototypes created in this thesis project individually focused on three distinct aspects of the embodied approach to image generation, highlighting the influence of each of these aspects on the overall success of the prototype in having the desired impact. The first interactive system prioritised the sensory experience of engaging with the kinetic sand and its ability to function as a medium of controlling a digital visual. By setting aside the role of AI in this experience, and restricting the visual impact of the sand, it was easier to observe the effects that its tactility and unusual material properties had on the user experience. Despite the lack of clarity in the connection between the sand and the computer image, it prompted interesting questions around preconceived expectations of human-computer interaction, by creating a conceptual and physical contrast between the softness of the mouldable sand and the starkly sharp digital visualisation. The kinetic sand was then carried forward into the second prototype which incorporated generative AI into the system and equipped the computer with a camera, allowing for an increased dialogue between the physical and the digital. This was demonstrated by the focused testing of the AI as an image generator relying on the camera feed of its environment for input, in combination with the routine system of text-based prompts and numerical parameters. In the gallery environment, users experienced a higher level of control and creative satisfaction as compared to prototype 1, and the sand successfully invited increased embodied engagement, prompting the user to touch and mould it in various ways. However, the interaction was too long and cumbersome for an installation and restricted the time each user spent with the system, limiting the observations that could be made and the number of generated images. This resulted in the user relinquishing control to the AI itself and the parameters it typically functioned on, which exacerbated the very black box effect that I was trying to mitigate. To better facilitate the interactions and allow the user to focus on the sand, I adapted the system to automatically and continuously generate images once the user began filming the sand and linked the motion data from their interactions with the sand to the randomisation of the content used in the images. This allowed for more control over the process of iterating the images based on the rapidity and urgency with which the hands moved to manipulate the material, resulting in a more streamlined, mindful, and embodied creative process. This also facilitated increased communication between the user and the computer by enabling the AI to respond to the perceived actions of the user as they attempted to achieve their desired result. Each of the three prototypes provided a rich base of insights through their successes and drawbacks which, in combination with the key theories of affordances, structural coupling and a sociocultural model of creativity, informed the suggested guidelines in the following chapter.

6. Guidelines for Embodied Creativity in Image Generation

Through the iterative process of prototyping and installation, it became clear that a set of parameters needed to be established for a successful working prototype that provides an embodied experience of creative image generation. In combination with the theories of sociocultural creativity, material affordances, structural coupling, and the black box effect, I used the observations from my own engagement with the prototypes as well as the gallery installations to propose a set of guidelines for embodied creativity in image generation, detailed below. These guidelines are intended for the development of future interfaces for image generation, applications of generative AI in creative practice and to contribute to discussions of critical making with generative AI (See Figure 32).

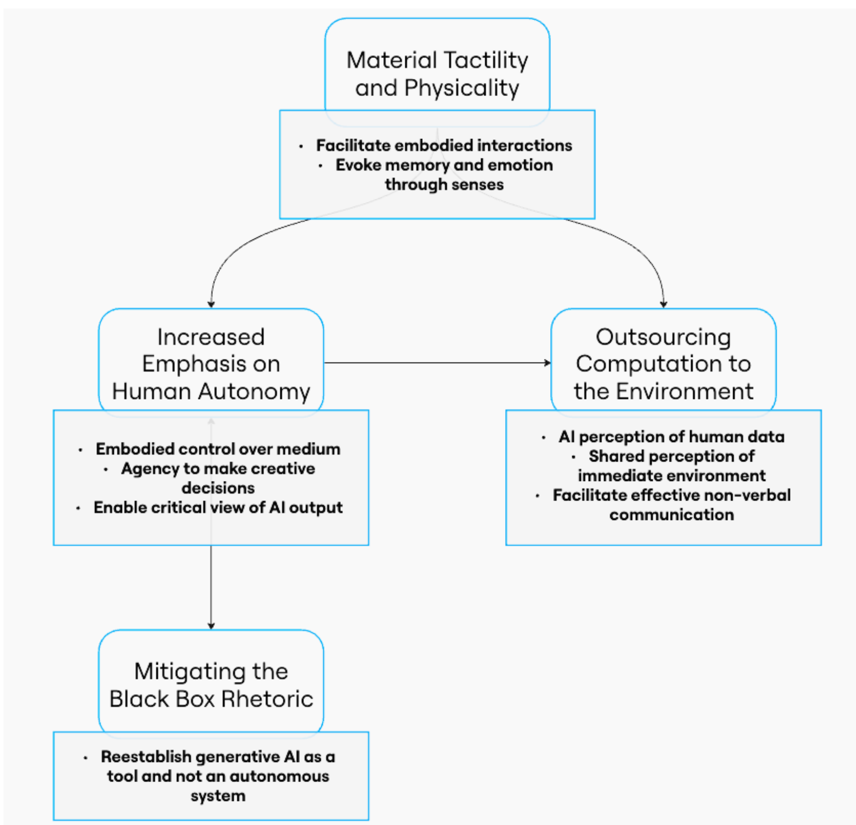


Figure 32: Recommended Guidelines for Embodied Creativity in Image Generation

6.1 Material Tactility and Physicality

In the experiments, kinetic sand served a dual role as both a medium for physical manipulation and as a sensor. It presented opportunities for users to physically create with it, manipulate it, and respond to it through multi-sensory engagement and intuitive interactions based on emotional responses like curiosity, disgust, or fascination. Simultaneously, it communicated these interactions to the computer through visual imprints denoting areas of tactile contact, as well as unique textural variations corresponding to the intensity of pressure applied and the manner of manipulation. This tactile interface offered an experience that was distinct from yet deeply intertwined with the digital realm, fundamentally reshaping the user's approach to creating with the AI program. In today's models, the invisible, heavily abstracted and compressed data contained within the model, also known as the latent space (Pasquinelli and Joler, 2021), is navigated to generate an image by using text prompts and numerical parameters. With an embodied system of interaction however, the medium of exploring the latent space of the AI is now physical rather than digital - faster movements of the hands take the user on wider journeys through the latent space by guiding the AI to randomise its output to a greater degree. The physicality and tactility of the sand guides the user towards a visual that can be observed before the AI generates it, rather than the AI attempting to replicate what the user has in their mind's eye.

The material for creating this tactile experience has been limited to kinetic sand in this thesis, to maintain consistency across the prototypes, given the limited time and scope. However, the primary function of the sand is to provide a physical medium for the user to create with, while facilitating a greater degree of sensory stimulation than typing on a keyboard, touching a screen or using a mouse can provide. This effect can be replicated across a variety of materials, particularly those associated with sensory play, such as textured clays or slime.

6.2 Outsourcing Computation to the Environment

While the physical material creates a space for the user to seek inspiration and creative satisfaction through a sensory experience, it also exists an environment with other factors that affect the user's cognitive abilities, while influencing the perception of the material itself. To maximise the translation of this sensory experience from the physical to the digital, embodying the computer with sensors that can communicate these non-verbal cues as data allows it to take advantage of these factors to create output that can be better received by the user. This also assists in blurring the boundaries between the user's physical environment, which influences the user's own decisions, and the latent space, which governs the decisions made by the AI, resulting in a process that allows both collaborators to share an embodied experience, while maintaining their own knowledge systems. This could result in a greater degree of collaboration and better contextual specificity within the output, which in turn raises the potential for a higher degree of creativity. This is also in line with the idea put forth by Guckelsberger et al. (2016) which posits that "co-creative and social creativity systems are only meaningful if each agent has a different perspective on a shared world, allowing them to complement each other and for creativity to emerge from their interaction" (Guckelsberger et al, 2021, p. 197). Within this thesis, the AI was embodied primarily through programmed computer vision, with a brief focus on ultrasonic sensors. However, there is scope for the use of a much larger variety of embodying technologies, as can be observed in the cases of Ai-da and GROUPTHINK, discussed in Chapter 2.

6.3 Mitigating the Black Box Rhetoric

A common thread across user experiences with every prototype and my own self-reflective journey with the Sand-AI program, was a sense of marvel and wonder at the decisions made by the AI in interpreting the visuals and prompts provided by the user. The novelty of the technology plays a role in evoking positive reactions like this, however, the critical approach to generative AI maintained throughout this thesis underlines the problematic nature of such perceptions. An essential guiding principle of this thesis has been the positioning of AI as a statistical model, as described by Pasquinelli and Joler (2021), and therefore a tool with a higher degree of affordances than those available in other technologies and mediums. The black box rhetoric lays emphasis on AI as an intelligent machine that cannot be understood and therefore cannot be controlled. However, an increased sense of control over the creative process, aesthetic decisions and final output is necessary for a balanced collaboration between humans and generative AI, as has been exemplified through the results of this research project. AI literacy and increased discussions on the complex nature of knowledge extractivism and data bias contained within algorithms, especially from varied cultural contexts, could improve the effects of the black box. In consideration of the sociocultural model of creativity followed in this thesis, relying on the unique perspectives offered by generative AI depends on considerations of what kinds of biases and problematic categorisations are held within the system. Until the development of machine learning algorithms moves beyond its nascent stages and its internal mechanisms are better understood, there is a need for critical dialogue and transparency in generative AI interfaces, to lessen the effect of the black box on the sense oof creative satisfaction and control achieved when making images with generative AI.

6.4 Emphasis on Human Autonomy

Building further on the notion of shared control over decision-making in creative processes, establishing and emphasising human autonomy is essential to fostering effective collaboration between humans and AI. A focus on human embodiment as an equally integral part of the Human-AI relationship as the embodied AI is necessary to create a space of exploration where perspectives can be *mutually* beneficial. Through this research, the kinetic sand served as a tactile interface through which both the human and the AI could derive inspiration and knowledge, whether physiologically or through abstraction of the sensorial experience. While still being imperfect, this human-sand-AI system exhibited greater scope for communication by removing the limitations imposed by the text-based and numerical guidance systems prevalent in current image generation models. This demonstrates the potential for creating higher-quality human-AI interactions that centre the nuances of the individual human experience without being overshadowed by the magnitude of information, particularly biased or distorted data, that is contained within the AI's latent space. In theory, this would foster improved communication of perspectives and understanding of affordances on the user's end, meeting the criteria of Glăveanu's model to maximise creative potential.

6.5 Summary

The research conducted in this thesis aimed to identify how our understanding of the affordances of generative AI could be influenced through an embodied approach to creative image generation, which culminated in the conceptualisation of four suggested guidelines. These guidelines are intended to inform future applications of image generation models in interactive systems by emphasising the role of the human body in the process of creation. Beginning with the significance of tactility and materiality of the interface, the kinetic sand's ability to evoke emotional and physical responses in users is cited as a way of engaging the human senses in their interactions. Similarly, the AI model's perception of its environment and user is considered integral to the process as well, and the incorporation of sensors into the computer enables it to respond to non-verbal cues from the user, potentially leading to more collaborative and contextually specific outputs. The guidelines also address concerns surrounding the "black box" perception of AI, advocating for transparency and critical dialogue to empower users and mitigate feelings of loss of control. Finally, the emphasis is placed on human autonomy in the creative process, suggesting that effective collaboration between humans and AI requires a balanced approach that values the unique contributions of each party. In the following chapter the guidelines and prototype are recontextualised within the key theories and related work to further detail their implications on future applications of this research and to evaluate them against similar work in computational creativity research.

7. Discussions

Having proposed a set of guidelines for an embodied creative process with generative AI, exemplified using kinetic sand in the three prototypes, this chapter will investigate the connections arising between the theories of structural coupling, material affordances and creativity discussed earlier and the results of the conducted installations. First, the nature of the interactions taking place between the user, the sand and the computer will be analysed through the concept of structural coupling. Next, the implications of the guidelines on approaches to embodied creativity in human-AI collaboration will be studied. Finally, the prototypes will be compared with other related approaches to embodied interactions with generative AI art tools.

7.1 Human-AI Structural Coupling

An ongoing challenge not only within this thesis, but also in the theoretical contextualisation of AI in a human-centric society, is the positioning of AI as either an autonomous system, capable of independent decision-making or a statistical model through which knowledge is simply processed and recirculated (Buongiorno, 2022; Varela & Maturana, 2002; Joler & Pasquinelli, 2021). While this thesis maintains a cautiously critical view of AI and through its research, attempts to position it as the latter, analysing the human-AI relationships facilitated through this thesis project is necessary, considering Maturana and Varela's own acknowledgement of the possibility of autopoietic artificial systems (Buongiorno, 2022, p.1038) As detailed in previous chapters, the presence of the kinetic sand between the user and the AI, as a medium of making art as well as a sensor and interface for the computer, is intended to provide a space for the user to incorporate part of the creative decision making of the computer within their own cognitive capabilities, while also responding to the computer's perception of the human's decisions. This is based on Hayles' notion of incorporation as a process of embedding processes of interaction with our environment within our existing bodily capacities, making them habitual rather than formalised steps (Hayles, 1999).

Through the research conducted, it can be observed that the kinetic sand could potentially facilitate a new form of structural coupling between humans and generative AI, while disrupting the existing systems of communication with these models as we know them today, by taking away the need for typing, clicking buttons or using verbal commands. The sand provides opportunities for multi-sensory interaction, tapping into the user's own perception of the material and its affordances through tactile exploration, while also conveying this information to the AI program to be interpreted and responded to in the form of a visual output. In an ideal scenario, the project would involve a similar embodied method of communicating specific text prompts to the computer as well, to cover the full scope of control that is available to the user to guide the AI, completely erasing the need for direct interaction with the computer. Unfortunately, the necessity of the text-based prompts in the conducted experiments limits the user from potentially engaging in a fully embodied relationship with the AI during the creative process, thereby also limiting the applicability of a more restrictive definition of structural coupling among artists and AI. However, the scope and timeframe of this thesis have allowed for the creation of a process that is still largely dependent on embodied interactions, despite its shortcomings, as a preliminary guideline for the future development of Generative AI systems as creative mediums. If the intended goal of AI development is Artificial General Intelligence, it is imperative to recognise the role of an extended cognitive systems and the relationships formed through these systems, particularly in the context of creativity, which relies on the specificity of one's own relationship with the world around them gained through embodied cognitive processes.

In summation, while we are still several developmental milestones away from autopoietic artificial systems, this thesis presents a starting point for further analysis of potential symbiotic relationship between artists and generative AI, through mediation by tactile interfaces like kinetic sand.

7.2 Potential for Embodied Creativity in Human-AI Collaboration

The scope within which embodiment as a facet of creativity has been explored within this thesis is limited, due to several factors including timeframe, access to more advanced technology, in terms of both costs and skill set and the availability of gallery settings for increased user interaction. However, the results of the research conducted demonstrate a clear theoretical and practical connection between the framing of creative processes as an embodied act and the process of creating valuable work with generative AI. Communicating the nuances of individual human experiences through only text and numbers to a system that largely relies on algorithmic modes of decision-making is an extremely difficult task that presents obstacles in what should be a seamless process of creation. This is not to say that embodied interactions could make the process linear or predictable, as the very exercise of being creative requires iteration, dialogue, exploration, and opportunities for discovery, as posited by Glăveanu (2012). However, providing more varied avenues for a mutual exchange in perspectives can be facilitated through embodiment of both the user and the AI in the creative process, by making use of a shared context like physical space or material. Here, kinetic sand served the role of medium as well as sensor, translating moments of physiological reactivity into data that could be interpreted by the AI. However, executing such a system successfully remains a challenge, as optimizing the integration of the sand's effects within the program itself posed significant difficulties in Human-Computer Interaction (HCI) design, which, within the confines of this thesis, remained partially unresolved. This included issues with effectively capturing the textural qualities of the sand to facilitate better connections between the physical and digital imagery, successfully integrating pressure sensors into the base of the sand to map specific areas within the image that could be improved and leveraging the ability to create 3D objects with the sand. Hence, the evaluation of the sand's role is primarily informed by personal meditations on its material affordances and observing user responses across the iterative development of three prototypes.

7.3 Reflections on the Interactive System

This thesis project drew inspiration from several related works that have been mentioned and analysed throughout this paper. Each stage of prototyping focused on a specific aspect of these theories or projects to create a space for critical and detailed observation, eventually informing the recommended guidelines in Chapter 6. I will be going through each of the prototypes in this section, reflecting on the insights gained and comparing them with the theories discussed in Chapter 3 and the related works analysed in Chapter 2. These comparisons are based on the evaluation criteria detailed in Chapter 4, as derived from the discussed theories (See Figure 33).

Work	Level of User Control	Level of Collaboration	Level of Creative Satisfaction	Sensory Stimulation	Fidelity of Output
AARON	High	High	High	Low	High
Ai-da	Medium	Medium	-	Low	High
GROUPTHINK	Medium	High	Medium	High	High
Prototype 1	**Low**	**Medium**	**Low**	**High**	**Low**
Prototype 2	**Medium**	**Medium**	**Medium**	**High**	**Medium**
Prototype 3	**High**	**High**	**Medium**	**High**	**Medium**

Figure 33: Comparison Table of the 3 Prototypes with Related Work

7.3.1 Prototype 1

Prototype 1 was focused on creating a multisensorial experience when engaging in a dialogue with a computer system, setting aside the aspect of generative AI to fully evaluate the effectiveness of the kinetic sand as a medium of interaction. My own experiences with the kinetic sand, through a meditative analysis of its tactile and visual properties, highlighted its versatility as an art material, indicating the role it could play as a facilitator of embodied image-making processes. Given its mouldability and viscoelasticity, it provided a three-dimensional space for exploration with the user's hands, which showed its potential to be translated into a rich source of data from the user. The novelty of the material also made it playfully engaging to the user, inviting a greater degree of interaction and provoking curiosity. When combined with a digital image however, there appeared to be a disconnect in the soft, malleable texture of the material and the starkly solid computer visual. While users observed the connection between the movement of their hands and the changes in the digital image, it was difficult to understand how the kinetic sand tied the experience together. Nevertheless, it created a unique and impactful sensory experience which I took forward into my next prototype.

The idea for a multisensorial dialogue between a user and the computer came from the project GROUPTHINK, which incorporated the physiological data of an audience with a visual and musical performance. This project focused on bringing the reciprocal effect of a live audience on a concert to a remote performance by turning data from the audience, like breath rate and heart rate, into AI generated visualisations. This influenced the tempo, rhythm, and sound of the musical performance which in turn had an effect on the audience (Hossaini et al., 2022). While GROUPTHINK succeeded in creating a conceptually sound experience for the user, with a clear relationship between their embodied presence and the performance, this was not entirely successful in Prototype 1. However, the prototype did succeed in engaging the user's senses to a high degree and provoking questions around preconceived ideas of human-computer interaction, which was a key aspect of the GROUPTHINK project as well.

7.3.2 Prototype 2

Drawing on the success of the sand in creating an embodied experience for the user in prototype 1, this was applied to the process of image generation in prototype 2. Here, I focused on the role of the user in the generation process, providing an accessible, tactile space for creative control when communicating with the AI model. Harold Cohen's AARON was the source of inspiration for this prototype, not in terms of the ability of the model to generate high-quality artwork, but in its functioning as a lens through which to gain a new perspective on the creative process. AARON was fully coded and designed by Cohen over a span of several years and was trained exclusively on the artist's own work. The model demonstrated the creative capabilities of artificial intelligence in its technological context at the time and Cohen executed a high degree of control over both its functionality and the training data that guided its output, through his skills as both a programmer and an artist. This emphasis on creative control was a key feature of prototype 2, providing the user with the agency to direct the composition of the AI's output and use their embodied knowledge to communicate with the AI through the sand, maintaining a focus on the physical aspects of artistic making processes and removing the need for any knowledge of programming. AARON was also a highly collaborative process between artist and computer, as it presented a translation of Cohen's own knowledge, skills, and creative practice, despite Cohen's opinion that AARON could outperform its own creator (Boden, 2009, p. 27).

The tactility of the kinetic sand interface provided the user with the sensory experience of using their hands to make art, a process that was replaced by robotic arms in Cohen's programme. It is worth noting that Cohen's aim was to create a computer that could make art with as little human input as possible, which is ideologically oppositional to the goal of this thesis. However, in its time, AARON was a standalone product of Cohen's unique skills and interest in the amalgamation of technology and art. Now, the context has shifted to prioritise the automation of labour and the prevalence of AI in image-making necessitates a focus on embodied engagement with generative AI to highlight human agency and creativity in this process.

7.3.3 Prototype 3

Prototype 3 was largely similar to the previous iteration in its functionality, but incorporated elements that further highlighted the aspects of collaboration, agency and embodied interactions in prototype 2. In this version, the AI model automatically continuously generates images when the camera is picked up and uses computer vision to collect motion data from the user's hand movements, which affects the randomisation of the generated image's content and aesthetic. This prototype draws on features of all 3 related works discussed in this thesis, namely Ai-da, AARON and GROUPTHINK. Ai-da is a humanoid robot artist created by Aidan Meller and Lucy Seal, with the ability to create physical artworks using a combination of robotic arms, generative AI models and sensors. Ai-da is programmed to not only ideate and create art, but also respond to and interpret the emotions of those interacting with her. This ability to perceive the world around her and translate it into artworks adds a layer of depth to her creations that arguably may not be observable in other popular generative AI models like Stable Diffusion and DALL-E. Like AARON, Ai-da demonstrates a higher degree of autonomy than the prototypes explored in this research, since Meller and Seal aimed to provoke critical discussions on the role of AI in society by embodying Donna Haraway's cyborg ("Ai-da", n.d.). However, Ai-da still requires a degree of human control and functions through human-machine collaboration, exemplifying how a reciprocal relationship between an AI system and human users could exist. This is especially relevant to prototypes 2 and 3, which similarly use cameras to perceive the kinetic sand as it is manipulated by the user and respond to this perception appropriately, albeit with less autonomy. Prototype 3 displays a higher degree of machine autonomy by automating the image generation, yet still maintains a high level of user control and user-AI dialogue by facilitating the same embodied interactions and reading these interactions for additional data. The user is then allowed to respond to the AI's outputs and the feedback loop continues.

8. Conclusion

This research was conducted with the objective of gaining a better understanding of the affordances of generative AI, by reintroducing moments of embodied cognition to the collaborative process of making images with AI to reconcile the differences in how humans and machines perform acts of creativity. As a result, this research intended to culminate in a set of guiding principles for more effective creative engagement with generative AI, to offer a solution to challenges posed by the black box rhetoric surrounding AI and the loss of embodied creativity in the current process of image generation. Regaining control over the creative process, identifying ways to communicate embodied and contextual knowledge to the AI and achieving creative satisfaction from this process were some of the issues tackled throughout this research.

The research culminated in the discovery of four key guidelines, recommended to maximise the creative potential of Human-AI collaboration:

1. Material Tactility and Physicality
2. Outsourcing Computation to the Environment
3. Mitigating the Black Box Rhetoric
4. Increased Emphasis on Human Autonomy

The applicability of these guidelines is currently limited to Human-AI collaboration through image generation models only but is intended to present a starting point for future research into this relationship and the dynamics it presents. Considering structural coupling as a concept that could be applied to creative practitioners and generative AI in the future, as AI becomes an increasingly entangled part of day-to-day life, this research aims to contribute to the discussion on ethical and effective use of generative AI in creative applications. Additionally, the scope of creativity has also been deliberately constrained to certain fields in order to focus this research within a short timeline. Broadening the definitions of creativity, in both human and computational contexts would be extremely useful towards studying the applications of generative AI in different interdisciplinary and cultural contexts.

8.1 Insights Gained

There were a significant number of challenges encountered throughout this process, both large-scale and small. Some of the larger obstacles were the very issues this thesis attempted to resolve such as critically navigating the black box rhetoric accompanying the use of AI and demystifying algorithmic processes. In self-reflective engagement with the Sand-AI system as well as observational studies of users interacting with the setup, a natural wonder and surprise at the possibilities of image creation with AI outweighed the desire to re-examine personal autonomy in the creation process. HCI design posed a challenge on a smaller scale as iteratively navigating effective systems of interaction to maximise the insights gained occupied a large part of the research process. Yet there were lessons to be learned on the value of critical making and the translation of these design processes to the actual act of generating images with AI, providing an unexpected point of comparison in how I approached my own creative capabilities as opposed to my approach to collaborating with AI. While reflection and observation provided key components of the guidelines detailed at the end of this research, the iterative design process refocused my attention on the nature of an actual creative process itself, slightly mitigating the very effect that I was striving to overcome through the installations. The establishment of a foundation of core theoretical frameworks of creativity, embodied cognition and affordances that guided the experimental setups was also key to formulating the guidelines that emerged from them, which assisted in overcoming some of the limitations of the experiments themselves.

My own relationship with generative AI also transformed significantly over the course of this thesis. As a novice in the space of image generation, and coming from a visual arts-centric background, the process of learning about the workings of AI, from the high-level parameters to more advanced features like inpainting and styleGANs, I found my own view of AI changing. I initially viewed AI as a collaborator with levels of intelligence and abilities on par with human creativity, but merely taking a different form, as a computer. As the experiments progressed and my literacy increased, I found myself positioning it far more as a statistical model with a much lower level of agency than the human user. It became a conscious decision to acknowledge the role of AI as a co-creator despite being more like a tool, as its abilities outshine any creative medium we have encountered yet, and it met the criteria of being able to present a perspective different to the one the user approached it with.

The research also emphasised the possibility to retain the humane aspects of engaging in creative processes even in the use of AI. The models of Generative AI discussed in this paper popularly function as a tool of automated image creation, with little perceivable human intervention. This often causes users to overlook the data fed into the models during the process of machine learning and the human labour behind it, as discussed by Pasquinelli and Joler (2021). Unfortunately, the research undertaken here could not fully address the sociocultural ramifications of this process due to the complexity of and scope of the subject and the time constraints of this thesis. However, reintroducing the human body into the user's own interactions provided a space for mindful and individualised image creation, which in some ways decentralised the AI model from the creative process and emphasised its role as a tool. Reintroducing the body also brought out the emotional component to engaging in acts of creation, personalising each user's experience with the system. This was especially highlighted through the playful nature of experiencing the kinetic sand's unusual and distinct properties and the novelty of generative AI tools as they rapidly continue to increase in capability and complexity. By combining these two highly contrasting experiences, I was able to derive an incredibly rich breadth of observations that could inform various applications of an embodied image generation system in future research.

8.2 Directions for Future Work

Over the course of this research project, there arose several opportunities for future work, which include:

1. Positioning a tactile interface-based AI system in a creative or art and design space, to observe its impact on inspiration, creative block and overall design process. This could involve highly speculative approaches to creative spaces or focused user testing within the art and design community.
2. Revisiting the proposed guidelines in this thesis as the AI "black box" effect is reduced over time. At the time of research, this rhetoric is still prevalent given the novelty of image generation models, limiting the success of this project in mitigating the effect completely. As the mechanisms with the AI latent space are better understood across a wider audience, it will be valuable to note its impact on the guidelines proposed here.

3. Investigating the impact of traditional art materials like paint, pencils, paper and so on within similar image-to-image generation processes.
4. Expanding the sensorial computation used in the interactive system to track and perceive additional environmental factors and data from the user, to facilitate further modes of communication between the user and the AI model. This might include pressure sensors, speech-to-text functionality and gesture tracking among others.
5. Scaling up the environment in which the user can create, to facilitate body tracking and larger spaces of movement. This could allow for applications of AI image generation in multidisciplinary artworks or performance art that incorporates a high level of embodied creativity like dancing or theatrical work. This also has the potential to discover how principles from Somaesthetics (The Journal of Somaesthetics, n.d.) could inform the designing of embodied interactions with generative AI.
6. Applications of an embodied image generation process in Health Design, particularly among non-verbal patients who may rely on body language, posture, facial expressions or eye movement to create in art therapy programs.
7. Sentiment analysis through voice and facial recognition, as can be seen in models like Hume.ai (hume.ai, n.d.), as AI continues to develop further. It would be useful to incorporate computation that translates affect and emotional responses from the user into the creative output.

8.3 Closing Thoughts

This thesis project was highly ambitious in its scope, taking on multiple large areas of research in its effort to answer the question of reconciliation between human and machine creativity through embodied interactions. I believe it was successful in part, in identifying ways to reintroduce the role of the individual's unique social, emotional and physical experience to the process of generating images with AI, by examining the gaps in this system and proposing a way forward to critically re-evaluate how generative AI is both used and developed. The incorporation of a tactile interface like kinetic sand highlighted these aspects of the individual experience, and through the incorporation of computer vision, furthered the possibility for more accessible systems of communication in human-AI collaboration. However, generative AI is still a subject of novelty and fetishization and it is challenging to broaden the community of users that understands the mechanisms that allow it to function, particularly in the face of mass digitalisation and automation of labour. The guidelines recommended in this paper cannot fully address the question of the replaceability of the human in creative processes and will require further expansion and iteration, particularly as AI continues to rapidly grow in its capabilities. Increased research in this space could benefit greatly from examining the nature of embodied creativity in humans as an important facet of generative AI development and it would be interesting to see how varying materials and aspects of computation impact the relationship between the user and the AI system.

9. Bibliography

Ai-Da. (n.d.). *About*. Retrieved December 4, 2023, from https://www.ai-darobot.com/about

Adams, R., & Atman, C. J. (1999). Cognitive processes in iterative design behavior. 1, 11A6/13-11A6/18 vol.1. https://doi.org/10.1109/FIE.1999.839114

Runway (n.d.). *AI Tools*. Retrieved February 16, 2024, from https://runwayml.com/ai-tools/

Balletti, C., Ballarin, M., & Guerra, F. (2017). 3D printing: State of the art and future perspectives. *Journal of Cultural Heritage*, *26*, 172–182. https://doi.org/10.1016/j.culher.2017.02.010

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*.

Boden, M. (2009). Computer Models of Creativity. *AI Magazine*, *30*(3), 23–34. https://doi.org/10.1609/aimag.v30i3.2254

Body of Knowledge: Art and Embodied Cognition. 27-29 June, 2019. (n.d.). Retrieved February 27, 2023, from https://blogs.deakin.edu.au/bok2019/

Buongiorno, F. (2023). Can Algorithms be Embodied? A Phenomenological Perspective on the Relationship Between Algorithimic Thinking and the Life-World. *Foundations of Science*, *28*(4), 1035–1045. https://doi.org/10.1007/s10699-022-09855-z

Choi, S. K., & DiPaola, S. (n.d.). *AI as other: An art-as-research approach to generative AI art practice*.

Clipdrop. *(n.d.) Clipdrop–Stable Doodle*. Retrieved February 16, 2024, from https://clipdrop.co/en-US

compute(r)ender. 2023. Retrieved February 16, 2024, from https://computerender.com

David Rokeby: Sorting Daemon. (n.d.). Retrieved December 8, 2023, from http://www.davidrokeby.com/sorting.html

Embodied Creativity: A Critical Analysis of an Underdeveloped Subject. (2015). *Procedia - Social and Behavioral Sciences*, *187*, 312–317. https://doi.org/10.1016/j.sbspro.2015.03.058

Erkut, C., & Dahl, S. (2019). Incorporating Virtual Reality with Experiential Somaesthetics in an Embodied Interaction Course. *The Journal of Somaesthetics*, *4*(2), Article 2. https://doi.org/10.5278/ojs.jos.v4i2.2399

Face, Hand, Pose Tracking & More with TouchDesigner MediaPipe GPU Plugin. (2023, September 8). Derivative. https://derivative.ca/community-post/tutorial/face-hand-pose-tracking-more-touchdesigner-mediapipe-gpu-plugin/68278

Fdili Alaoui, S., Schiphorst, T., Cuykendall, S., Carlson, K., Studd, K., & Bradley, K. (2015). Strategies for Embodied Design: The Value and Challenges of Observing Movement. *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, 121–130. https://doi.org/10.1145/2757226.2757238

Gainsley, S. (2012, December 16). *Look, Listen, Touch, Feel, Taste: The Importance of Sensory Play*. Yumpu.Com. https://www.yumpu.com/en/document/read/6120313/look-listen-touch-feel-taste-the-importance-of-sensory-play

Gaver, W. (2012). What should we expect from research through design? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 937–946. https://doi.org/10.1145/2207676.2208538

Generate AI Images with Stable Diffusion using Image to Image Generation with any TOP. (2022, December 21). Derivative. https://derivative.ca/community-post/tutorial/generate-ai-images-stable-diffusion-using-image-image-generation-any-top

Gibson, J. J. (2014). *The Ecological Approach to Visual Perception: Classic Edition*. Psychology Press. https://doi.org/10.4324/9781315740218

Glăveanu, V. P. (2020). A Sociocultural Theory of Creativity: Bridging the Social, the Material, and the Psychological. *Review of General Psychology*, *24*(4), 335–354. https://doi.org/10.1177/1089268020961763

Harder, J. (2018). *Graphics and Multimedia for the Web with Adobe Creative Cloud: Navigating the Adobe Software Landscape*. Apress.

Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. The University of Chicago Press. https://hdl.handle.net/2027/heb05711.0001.001

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Hossaini, A., Gingrich, O., Rahman, S., Grierson, M., Murr, J., Chamberlain, A., & Renaud, A. (2022). GROUPTHINK: Telepresence and Agency During Live Performance. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, *5*(4). https://doi.org/10.1145/3533610

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. MIS Quarterly, 28(1), 75–105. https://doi.org/10.2307/25148625

Hume AI Publication in Nature Human Behavior: Deep Learning & Vocal Bursts in Different Cultures • Hume AI. (n.d.). Retrieved March 13, 2024, from https://www.hume.ai/blog/hume-ai-publication-in-nature-human-behavior-deep-learning-and-vocal-bursts

Jordà, S. (2001). *New Musical Interfaces and New Music-making Paradigms*. https://doi.org/10.5281/zenodo.1176366

Larrieux, E., & Speziali, S. (2022). Augmented Objects as Portals into Virtual Worlds:Using Audio to Create Immersive Experiences in Extended Realities. *AudioMostly 2022*, 44–51. https://doi.org/10.1145/3561212.3561243

Making and interpreting: Digital humanities as embodied action | Humanities and Social Sciences Communications. (n.d.). Retrieved February 16, 2024, from https://www.nature.com/articles/s41599-023-02548-3#Sec1

Menano, L., Fidalgo, P., Santos, I. M., & Thormann, J. (2019). Integration of 3D Printing in Art Education: A Multidisciplinary Approach. *Computers in the Schools*, *36*(3), 222–236. https://doi.org/10.1080/07380569.2019.1643442

Mosaic Virus, 2019. (n.d.). Anna Ridler. Retrieved December 8, 2023, from http://annaridler.com/mosaic-virus

Publication in iScience: Understanding what facial expressions mean in different cultures • Hume AI. (n.d.). Retrieved March 13, 2024, from https://www.hume.ai/blog/iscience-facial-expression-different-culture

Ritchie, A. D. (1946). The Creative Mind. *Nature*, *157*(3991), Article 3991. https://doi.org/10.1038/157535a0

Romesin, H. M. (n.d.). *Autopoiesis, Structural Coupling and Cognition: A history of these. And. Other notions in the biology of cognition*.

Romic, B. (2022). Negotiating anthropomorphism in the Ai-Da robot. *International Journal of Social Robotics*, *14*(10), 2083–2093. https://doi.org/10.1007/s12369-021-00813-6

Saunders, R. (2012). Towards Autonomous Creative Systems: A Computational Approach. *Cognitive Computation*, *4*(3), 216–225. https://doi.org/10.1007/s12559-012-9131-x

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. In H. Sheikh, C. Prins, & E. Schrijvers, *Mission AI* (pp. 15–41). Springer International Publishing. https://doi.org/10.1007/978-3-031-21448-6_2

Stable Diffusion Models: A beginner's guide - Stable Diffusion Art. (2022, November 25). https://stable-diffusion-art.com/models/

The Journal of Somaesthetics. (n.d.). Retrieved June 7, 2023, from https://journals.aau.dk/index.php/JOS#:~:text=Somaesthetics%20describes%20an%20integrative%20field,significances%20in%20culture%20and%20societies.

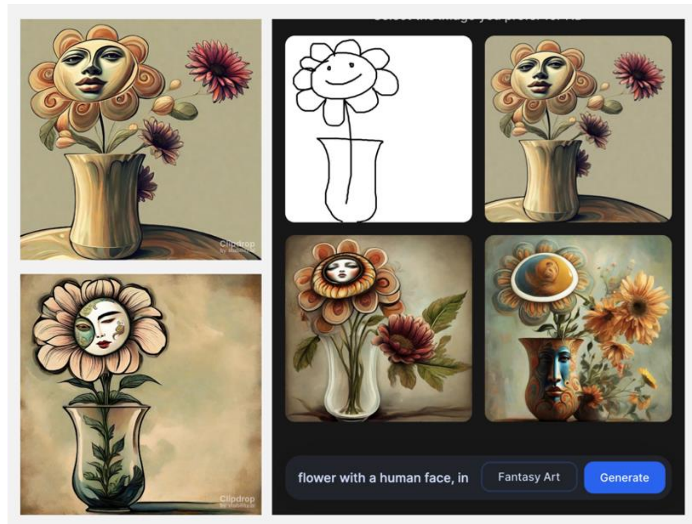
Verhulst, I., Woods, A., Whittaker, L., Bennett, J., & Dalton, P. (2021). Do VR and AR versions of an immersive cultural experience engender different user experiences? *Computers in Human Behavior*, *125*, 106951. https://doi.org/10.1016/j.chb.2021.106951

Willis, J. W., & Edwards, C. L. (Eds.). (2014). Individual Action Research. In *Action Research: Models, Methods, and Examples* (1st ed., pp. 95–109). Information Age Publishing, Incorporated. https://ebookcentral.proquest.com/lib/oculocad-ebooks/detail.action?docID=3316052

Zeilinger, M. (2021). Tactical entanglements: AI art, creative agency, and the limits of intellectual property (p. 186). meson press.

10. Appendices

Appendix A: Experiments with Stable Doodle and Clipdrop.ai



I created a physical artwork using acrylic paint and canvas in a collaborative process with a group of three friends that I used as a creatively stimulating, embodied experience of making art (Pictured above, left). I compared this with the process of generating the same kind of image using both image-to-image (Pictured above, right) and text-to-image models to reflect on the differences in the process, noting the disconnect between my vision and the generated output, the lack of space for intuitive iterating and the repetitive and monotonous nature of the process of tweaking the prompts and parameters to get closer to my envisioned artwork.

Appendix B: Template for Journal Entries

Prompt 1: Did the sand have any effect on the experience? If yes, was it positive or negative?

Prompt 2: Do any of these words relate to my experience with the sand?

- Inspiration
- Motivation
- Satisfaction
- Control
- Creation

Prompt 3: Did it feel disconnected from the visual? In what ways?

Prompt 4: Did I feel like looking at the sand as I played with it?

Prompt 5: Did I feel in control of the image on the screen?

Prompt 6: Was the connection between the sand and the image clear through the interaction?

Prompt 7: Was it frustrating to not reach a desired outcome for the visual through the interaction? Did moving the sand help channel this frustration?

Prompt 8: Was the overall experience:

- calming/peaceful?
- frustrating/disgusting?
- Neutral?

Appendix C: Additional Reading

Artificial intelligence as relational artifacts in creative learning. (n.d.). Retrieved December 11, 2023, from https://www.tandfonline.com/doi/epdf/10.1080/14626268.2023.2236595?needAccess=true

Buongiorno, F. (n.d.). *Embodiment, Disembodiment and Re-embodiment in the Construction of the Digital Self*. *36*.

Choi, S. K., & DiPaola, S. (n.d.). *AI as other: An art-as-research approach to generative AI art practice*.

Eddy, Martha. Mindful Movement: The Evolution of the Somatic Arts and Conscious Action. Intellect Books, 2016.

Gibson, E. J., & Pick, A. D. (2003). *An Ecological Approach to Perceptual Learning and Development*. Oxford University Press.

Park, J. H., Li, Y., & Niu, W. (2023). Revisiting creativity and critical thinking through content analysis. *Journal of Creativity*, *33*(2), 100056. https://doi.org/10.1016/j.yjoc.2023.100056

Rajko, J. (2018). A Call to Action: Embodied Thinking and Human-Computer Interaction Design. In *The Routledge Companion to Media Studies and Digital Humanities*. Routledge.