Developing Embodied Familiarity with Hyperphysical Phenomena

Gray Crawford

Dan Lockton primary advisor

Daragh Byrne secondary advisor

A thesis submitted to the School of Design, Carnegie Mellon University,
for the degree of Master of Design in Design for Interactions

© 2019 Gray Crawford

Abstract i

Acknowledgements ii

Introduction 1

Spatial Interfaces and The Siren of Intuition 4

Overview · Literary Foundation 6

Expectations from Prior Physics 7

Coherence and Coordination of Phenomena 8

Engagement with Hyperphysics 10

Affordances 12

Embodiment · Homuncular Flexibility 12

Tools for Thought 13

Provocations · Prototypes 15

On Notation as Physics, and the Human Capacity to Learn Environments 16

Overview of Prototypes 20

1. Bodily Fusion with Objects via Raymarched Signed Distance Fields · Blend Body 21

2. Correlating Agency while Acting at a Distance · Hidden Force Mappings 24

Prototype · Hidden Force Mappings 25

3. Slightness of Bodily Representation · Dust Bodies 1-3 28

4. Parameter Space Exploration via Play · Bulk Particle Series **31**

5. Hypermaterial Validity via Multisensory Integration · VR Wrist-Finger Haptics **33**

Prototype · VR Wrist-Finger Haptics 34

6. Embodiedly Controlling High-Dimensional but Non-Spatial, Non-Physical Structures · Xoromancy **36**

BigGAN and High-Dimensional Space 36

Xoromancy 38

Reflections on Prototyping **42**

Discussion · Conclusion **43**

References **48**

Appendices **50**

Xoromancy **50**

Abstract

Spatial computing (VR, etc) reveals an expansive and underexplored possibility-space of interactions in which the physics subtending affordances and phenomena can itself be designed, rewarding novel approaches to interaction design.

Through reviewing literature and prototyping spatial interactions, this thesis explores the impact of previously-unencountered physical dynamics upon the development of familiarity with systems and identifies significant representations of external objects and the body itself, with an eye towards the larger goal of transformative tools for thought.

While identifying promising avenues from the convergence of literary sources, this thesis synthesizes research and investigates applications of interactional dynamics by designing prototypes of spatial interactions given different materialities and computed physics, gaining insight through direct engagement with novel spatial phenomena. These VR prototypes illustrate design considerations for now-accessible interactional and material unorthodoxies, recognizing consequences and applications for embodiment spanning body-environment fusion (in depiction and interaction), multisensory integration, and high-dimensional dataset traversal.

Historically the mediums and notations available in a time period have contributed to and somewhat delineated the set of available thoughts. As a medium, spatial computing, by such nuance of bodily input and output and such rigor of calculation, likely affords domains of thought and experience barely conceivable at present.

Acknowledgements

I'd like to thank

Dan Lockton and Daragh Byrne for their supportive advisement and the numerous conversations that revealed new avenues of inquiry —

Golan Levin and the STUDIO for Creative Inquiry community for fostering an atmosphere of incisive exploration and kindred mirth —

the people of the spatial design + art communities on Twitter who continually discover what is possible and further the shared knowledge-space, a true hivemind of creativity and learning —

and of course my late mother, Sarah Nelson Crawford, who cherished and enacted the true nuance and beauty of the hand, and who would have loved to know her influence upon this work.

Introduction

Spatial computing[1] offers a unique and mostly unexplored design space where the body can spatially engage with a vast parameter-space[2] of novel, interactive, computed / simulated physics systems ("*hyperphysics*"). Since most of these computed phenomena diverge greatly from the laws of physics we grow up immersed in, there is opportunity to grow familiar via direct, embodied engagement.[3]

Hyperphysical materials and structures' wide set of behaviors and coordinations enable not only powerful tools and external representations, but also novel representations and functions of the body in space. This is largely unexplored territory as design has classically only operated on external structures. In spatial computing, the environment and the body are equally available to design, enabling new species of depiction and agency relevant to a spatial interface designer's conception of perception and neuroplasticity.

Through the course of this thesis, I worked primarily with Leap Motion[4] hand tracking in VR, documenting how familiarity develops from direct, embodied engagement with novel hyperphysics, through a series of prototypes spanning body-environment fusion (in depiction and interaction), multisensory integration, and high-dimensional dataset traversal.

Historically the mediums and notations available in a time period have contributed to and somewhat delineated the set of available thoughts. As a medium, spatial computing, by such nuance of bodily input and output and such rigor of calculation, likely affords domains of thought and experience barely conceivable at present.

---

[1] *virtual reality*, *augmented reality*, etc

[2] which I deem *hyperphysics*, includes our universe's "mundane" laws as a coordinate of its component parameters (assuming that our universe's laws are in principle simulatable, or that some perceptively-equivalent analogue is), and includes subsets of 2D and 3D (and 4D, etc) physics. Hyperphysics is the set of all possible environments and computable structures that can be reactively coordinated with the body.

For the purposes of this thesis that means the motion of the body in space, though comparable dynamics would apply to future, completely mental computed experiences in Brain-Computer Interfaces that bypass muscular movement to operate directly in the seat of experience.

[3] But what is the threshold of novelty beyond which familiarity cannot be developed?

[4] Leap Motion is an optical hand tracking solution using two infrared cameras to detect the 3D coordinates of the hands.

. . . . . .

A designer's ability is correlated with their understanding of materials. A carpenter needs to know about wood's properties and behavior before they can accurately imagine how the wood might be applied or expertly guide their hand into shaping it. This also applies to hyperphysical phenomena and materials. It follows, then, that designers would benefit from direct experience and bodily engagement with these new hyperphysics to map the possibility-space and better conceive of applications.

In the widest scope, these new media afford not only more powerful tools, but also more powerful ways of thinking, specifically because they involve direct bodily engagement.[5]

Sensitivity and attunedness to phenomena involves the development of familiarity with the contents of perception such as external environmental phenomena, the body itself, and mental structures that scaffold perception and thought. I'm especially interested in how the body can be represented now that its movement and structure are more easily received, modulated, and exported back into the sensory field. The body itself now becomes a designable structure when classically only the external environment and mental frameworks were designable.

I'm interested in how multisensory integration reinforces the development of familiarity repeatedly encountering mutually-coherent inputs across the senses. Modern computed environments can now receive and provide a wider set of sensory channels so as to support stronger familiarities with more nuanced tools.

My thesis serves to exhibit a set of concerns and possibilities for spatial interactions that may not be immediately apparent to entrant designers, with the hope that such awarenesses might provoke/instill novel frameworks for conceiving of spatial interactions and the role of embodiment in spatial computing.

[5] in the same way that electronic calculators separate the user from the mathematical operations and thus atrophy their arithmetical ability, whereas abacuses directly engage the user in every step of the mathematical operation, instilling a fluency that transcends the need for the artifact to be present at all by allowing the operations to be invoked purely mentally and via muscle memory.

Spatial Interfaces and The Siren of Intuition

The process of design classically operated solely in the domain of the physical, affecting only material objects and systems. Only relatively recently did design begin operating on computed systems, involving visual structures depicted on screen rather than structures subtended by material elements that give them their properties. That classical digital depiction was nevertheless a generally two-dimensional affair, with flat representations and comparably limited user input systems (in the form of discrete buttons, and exclusively planar mouse- or finger-movement). With the advent of modern motion-tracking systems, computers' depictive capabilities vastly expanded, gaining the ability to 1: provide coherent optical flow with precisely-tracked head movement and 2: represent the motion of the body in space, mapped to the tracked motion of limbs.

Here, now, came the ability to represent entire 3D environments and, critically, the subjective experience of being bodily immersed in an environment rather than the previous experience of manipulating control surfaces of a computer and seeing the effect isolated in a tiny fraction of the visual field.

Designers can now design immersive embodied experiences of different environments with different and arbitrary physical laws, whereas previously they were confined to everyday materials and the familiar set of physical laws subtending their behavior. The entirety of a designer's previous experience with immersive environments involved a singular set of physics, and now the design space is opened to entirely new classes of physics and phenomena, and thus new classes of affordances and designed objects.

The dichotomy here involves the dynamic of intuitions about the behavior of objects and their affordances, and the novel space of design that is opened up by spatial computing. Much modern design follows the trope that the best UI is as intuitive as possible, but how should this evolve within the domain of novel environments where no intuitions have as of yet been developed? An early answer might be to bring in and therefore perpetuate current UI and interactional grammars, as both users and designers are already intimately familiar with their dynamics and are already attuned to the structure and behavior of classical affordances. However, this serves to dilute the unique aspects of spatial computing, and replaces what can inherently support novel and expansive behaviors with a mere simulation of the mundane. This cycle is continually established with each successive new medium, this inertia of interactional grammars, and has calcified, for example, the inefficient QWERTY keyboard

layout well past the point when typewriters mechanically jammed with any more optimal layout, or the paper-based document model even when screens could represent 3D forms, opting for the prior and the familiar over the new.

My caution here is that, even presented with this supremely expansive and unexplored domain of spatial interaction design and physical phenomena in immersive computing, designers will hamper themselves and their users by perpetuating old UI mechanics, turning VR etc into a rough simulacrum of the constrained physical world rather than the means for its transcendence.

Further, since so many of the established UI norms are constrained to the 2D input mechanisms of computing systems previously described, their reapplication into spatial computing fails to fully leverage the immense nuance of physical input that motion tracking affords.

If every designer and craftsperson aimed to only make objects that were intuitive and immediately apparent, could virtuosity exist? Does virtuosity not arise through practice with high-dimensional-input artifacts, 'tuning' the brain and body to the specific dynamics of the available affordances? If we constrain our movements to the limited scope of previous media, we might make intuitive artifacts, at the expense of allowing users to develop new intuitions for the new spaces they are gaining access to. The draw to provide already-intuitive interfaces hampers the full exploration of interface-parameter-space. More critically, if designers reside only in bubbles of their current intuition, they constrain their conception of what can be designed, and what might be.

Overview ·
Literary Foundation

Expectations from Prior Physics

People build familiarity with ordinary materials and objects, the "interactional grammar" of physical affordances. This presents a challenge if computed environments can diverge from that familiarity, and users expect certain behaviors from the start that confine the designer's hand (and mind) to provide only what aligns with expectation. On the other hand, leveraging these expectations while selectively breaking them with confined novel behaviors provides opportunities to slowly ween users away from their ossifications.

. . . . . .

The physics system that we find ourselves in at birth determines the nature of our bodies, the phenomena we encounter, and the affordances available in the environment. We become familiar with and develop intuitions about the 'interactional grammars' that we repeatedly come into contact with. Or, as Sutherland (1965) puts it, "We live in a physical world whose properties we have come to know well through long familiarity. We sense an involvement with this physical world which gives us the ability to predict its properties well." This is the default state that designers have operated within since antiquity.

With the advent of computer-rendered dynamic media, phenomena could be represented that diverged from the phenomena driven by physical laws classically confining designed artifacts. This larger space of possible physical dynamics, of which the physics of our universe is but a subset, I refer to as *hyperphysics*. Since these phenomena are observed and interacted with by users who developed in ordinary physics, most users are presumably attuned to the nuances of phenomena (or do "not enter devoid of expectations that come from their previous experience" (Blom, 2007)) and may be immediately aware of the similarities, recognizing that "content that is familiar to the user from the real world will be initially and automatically considered the same as a real object" (Blom, 2010). Or, as Golonka & Wilson (2018) state: "When we encounter a novel object or event, it will likely project at least some familiar information variables (e.g., whether it is moveable, alive, etc), giving us a basis for functional action in a novel context". The challenge is how to communicate hyperphysical affordances that do not have exact analogues in ordinary physics.

For example, many objects in rendered environments (such as depicted in "virtual reality" fields of view or superimposed on the

outside world in “augmented reality” fields of view) are capable of being grasped and moved around, no matter their apparent mass, extent, smoothness, etc., even non-locally. Yet Gibson (1979)’s conception of what is graspable (granted, conceived prior to the availability of spatial computing) requires “an object [to] have opposite surfaces separated by a distance less than the span of the hand”. This requirement is now seen as being compartmentalized to only ordinary physics, but should designers of spatial user interfaces (SUIs) abandon it completely? Surely it’s useful to leverage the already-developed familiarity with ordinary physics’ interactional grammars, but at what expense? How tightly should SUIs be coupled to ordinary physics? What is conserved in intuitiveness is lost in the full exploration of the hyperphysics capable of being simulated, as “there is no reason why the objects displayed by a computer have to follow the ordinary rules of physical reality with which we are familiar”(Sutherland, 1965).

Coherence and Coordination of Phenomena

This familiarity is built up via repeated exposure to consistent observed physical behavior, where covariance of stimuli unifies the parallel streams of input into singular percepts. Relevantly, this incentivizes designers to provide multiple sensory responses for a given phenomena or user action, fleshing out the validity of the subjective experience. A difficulty, however, is that without coordination between designers across experiences, the preponderance of divergent interactional grammars and hypermaterial[6] depictions might inhibit users from developing overarching familiarities.

. . . . . .

We find much of mundane physics intuitive because we develop and spend our whole lives fully immersed in it. When the environment offers a consistent set of phenomena and consistent responses to input, the brain becomes accustomed to the perceived patterns and builds a set of intuitions about the phenomena. Piaget (1952) notices that “adaptation does not exist if the new reality has imposed motor or mental attitudes contrary to those which were adopted on contact with other earlier given data: adaptation only exists if there is coherence, hence assimilation.” This consistency comes from the fact that ordinary physics do not change over time

[6] ...a *hypermaterial* being a material behaving according to some set of hyperphysics...

or location, and the perception of the unity of events arises from multiple senses receiving coordinated impulses. In Gibsonian (1979) parlance, “when a number of stimuli are completely covariant, when they always go together, they constitute a single ‘stimulus’”. Piaget (1952), in noting that “the manual schemata only assimilate the visual realm to the extent that the hand conserves and reproduces what the eyes see of it”, communicates the unification of tactile and visual sensory input, that merely “the act of looking at the hand seems to augment the hand's activity or on the contrary to limit its displacements to the interior of the visual field.”

Usefully, since our bodies are themselves physical, we can directly impact the environment and observe the effects in realtime, becoming recursively engaged with the phenomena in question. Chemero (2009) describes this recursive engagement thusly:

> *Notice too that to perceive the book by dynamic touch, you have to heft it; that is, you have to intentionally move it around, actively exploring the way it exerts forces on the muscles of your hands, wrists, and arms. As you move the book, the forces it exerts on your body change, which changes the way you experience the book and the affordances for continued active exploration of the book*.

This is assisted by the fact that our senses are not located exclusively in the head. “...in perception by dynamic touch, the information for perception is centered on the location of the action that is to be undertaken” (Chemero, 2009). Thus we can correlate the visual feedback of where, for example, the hand is amidst the environment, the proprioceptive feedback of the hand’s orientation relative to the body, and the tactile and inertial feedback provided by the environment upon the hand.

Being confined to the laws of ordinary physics, the parallel input sources agree, providing a consistent “image” of the environment. The fewer senses available, the less well-defined the final percept is, and partial disagreement between senses can allow “anomalous” sense-inputs to be overridden. This can lead to perceptual illusions like when, at a stoplight, a large bus in the adjacent lane begins to move forward, and provided it occupies an adequately large section of the visual field, the sensation of yourself moving backwards is induced, even if there is no vestibular agreement with the optical flow. Thus, to provide as rich and internally-coherent experience as possible, spatial computing systems need to provide many parallel sources of sensory input that agree, forming a unified sensory field. Sutherland (1965) agrees that “if the task of the display is to serve

as a looking-glass into the mathematical wonderland constructed in computer memory, it should serve as many senses as possible."

Two difficulties arise. The physical behavior of rendered environments depicted in spatial computing need not align with ordinary physics (the alignment in fact being a difficult if not impossible feat), and the rendered environments need not be internally consistent either (especially given that 1. simulated physics can change in realtime at the whim of the designer (something that ordinary physics is by definition incapable of) and 2. independent rendered-environment-designers can make environments available that have vastly different physics and thus different interactional "grammars"). Thus the lived experience of the user, navigating between ordinary physics and the variant and likely inconsistent physics of rendered environments, involves shifting between inconsistent interactional grammars. Will this have a negative affect on the brain? Will expertise with unorthodox physics developed in a simulated environment have a zero-sum relationship with the embedded expertise navigating ordinary physics?[7] Is the brain plastic enough to contain and continue developing facility in an ever-increasing number of interactional grammars?

Engagement with Hyperphysics

The usefulness of an environment is a function of its physical capacities, and thus the expanded set of hyperphysics within simulated systems supports, in principle, a proportionally-expanded usefulness. Direct bodily engagement is possible not only with simulations of micro- and macroscopic phenomena, but even more esoteric and unorthodox phenomena not directly realizable within

---

[7] Examples of "bleed" between interactional grammars are fairly common, as in the case of offhandedly swiping to scroll vertically on a magazine page or non-touchscreen laptop screen. These involve specific interactions with manipulable objects. This "bleed" scales to even more fundamental methods of interaction, such as locomotion / traversal, as I have personally experienced. A short anecdote:

Ready At Dawn's VR game *Lone Echo* takes place in the microgravity of a space station, and locomotion involves using your hands to pull or push off of objects and walls, or moving over surfaces hand-over-hand. I played the game for about eight hours over the course of my first two days, and within a few hours became fairly adept at maneuvering myself in zero gravity. Using only my hands to move around became second nature. Upon waking on the third day, after I put in my contact lenses, I *pushed off the wall* to exit the bathroom. No mere object interaction, this bleed involved my body's fundamental way of moving through the world.

My familiar reflexes saved me, but it was a wild moment where I more accurately recognized neuroplasticity's reach, and spatial computing's. If such bleed is possible over only eight hours of experience, what will months or years of aggregate immersion produce?

our universe's laws. This vastly expands the space of interaction design, and rewards open and explorative mindsets and design approaches. Our neuroplasticity enables us to attune ourselves to the nuances of whatever our senses happen to provide, and this expanded space of computer-mediated experience supports untold applications of that plasticity.

. . . . . .

The opportunities afforded by bodily engagement with hyperphysical simulated systems, however, are numerous. The usefulness of the environment is a function of its physical capacities, and thus the expanded set of hyperphysics within simulated systems supports, in principle, a proportionally-expanded usefulness:

> *Concepts which never before had any visual representation can be shown, for example the "constraints" in Sketchpad. By working with such displays of mathematical phenomena we can learn to know them as well as we know our own natural world.* (Sutherland, 1965)

> *We lack corresponding familiarity with the forces on charged particles, forces in non-uniform fields, the effects of nonprojective geometric transformations, and high-inertia, low friction motion. A display connected to a digital computer gives us a chance to gain familiarity with concepts not realizable in the physical world.* (Sutherland, 1965)

It is fundamentally an accident of birth to have been born into ordinary physics, but the mind is in principle capable of becoming fluent in many other physics:

> *Our perceptions are but what they are, amidst all those which could possibly be conceived. Euclidean space which is linked to our organs is only one of the kinds of space which are adapted to physical experience. In contrast, the deductive and organizing activity of the mind is unlimited and leads, in the realm of space, precisely to generalizations which surpass intuition.* (Piaget, 1952)

A key constraint then becomes the ability of designers to envision novel physics to then manifest, as

> *...computers are so versatile in crafting interactive environments that we are more limited by our theoretical notions of learning and our imaginations. We can go far beyond the constraints of conventional materials...* (diSessa, 1988)

Affordances

Hyperphysics supports novel behaviors that have no necessary analogue in ordinary physics. Thus the entire structural, visual, and dynamic "language" of ordinary affordances is inadequate to fully cover all possible transformations and behaviors that hyperphysics supports. Even fundamental material behaviors like collision are not in principle guaranteed. Dourish (2004) describes how collision can be an essential property for certain useful arrangements:

> *Tangible-computing designers have sought to create artifacts whose form leads users naturally to the functionality that they embody while steering them away from inconsistent uses by exploiting physical constraints. As a simple example, two objects cannot be in the same place at the same time, so a "mutual exclusion" constraint can be embodied directly in the mapping of data objects onto physical ones; or objects can be designed so that they fit together only in certain ways, making it impossible for users to connect them in ways that might make sense physically, but not computationally.*

However, the greater space of possible physical behaviors offers opportunities to create new affordances with new interactional grammars that can take advantage of the specificity of computing power and the precise motion tracking of the body.

Embodiment · Homuncular Flexibility

The body's relationship to tools is often quite fluid, where prolonged use allows tools to be mentally fused with the body, and engagement with the world is perceived at the tool's interface with the world rather than the body's interface with the tool. Blind people can build a relationship with their cane such that "the cane is … incorporated into [their] body schema and is experienced as a transparent extension of [their] motor system" (Heersmink, 2014). The opportunities for spatial computing are even more potent here, where the medium's capacities for tracking the body's motion allows an even greater mapping between the rendered environment's behavior and the user's motion than ordinary dynamic media constrained to two-dimensional screens and rudimentary inputs.

The ability to depict the body in novel and hyperphysical ways, while still mapping the depicted body's movement to the base movements

of the user[8], enables profoundly transformative computer interfaces such as increasing the number of limbs,

> *Participants could hit more targets using an avatar with three upper limbs, which allowed greater reach with less physical movement. This was true even though motions mapped from the participants' tracked movements were rendered in a different modality (rotation of the wrist moved the avatar's third limb in arcs corresponding to pitch and yaw). Use of more intuitive mappings might enable even faster adaptation and greater success.* (Won et al, 2015)

or changing the physical form of the hands to better interface with a task, as explored by Leithinger et al (2014): "...we can also morph into other tools that are optimal for the task, while controlled by the user. Examples include grippers, bowls, ramps, and claws — tools with specific properties that facilitate or constrain the interactions". The question then becomes how many familiar aspects to include so as to conserve intuition, framed by Won et al (2015) as "...what affordances are required for people to use a novel body to effectively interact with the environment?", especially when "such realism may reinforce the user's desire to move as he or she would in the physical world." Though, critically, the brain's plasticity allows for novel environments to eventually become quite literally second-nature, as in the classic Heideggerian example of the hammer, articulated by Heersmink (2014): "When I first start using a hammer, my skills are underdeveloped and the hammer is not yet transparent. But gradually my hammer-using skills develop and the artifact becomes transparent which will then alter my stance towards the world."

Tools for Thought

Ideally, the increased adoption and bodily engagement with hyperphysics will prove us with new tools to understand and represent not only the world around us at scales heretofore inaccessible (as Sutherland (1965) envisions about subatomic particles: "With such a display, a computer model of particles in an electric field could combine manual control of the position of a moving charge, replete with the sensation of forces on the charge, with visual presentation of the charge's position", but also purer forms of knowledge such as mathematical relationships), and will lift our minds to new heights as previous notations for thought have already done. Gooding (2001) articulates it well:

---

[8] A mental phenomenon referred to as *homuncular flexibility*

> *Computer-based simulation methods may turn out to be a similar representational turning point for the sciences. An important point about these developments is that they are not merely ways of describing. Unlike sense-extending devices such as microscopes, telescopes or cosmic ray detectors, each enabled a new way of thinking about a particular domain.*
>
> *The sciences frequently run up against the limitations of a way of representing aspects of the world — from material objects such as fundamental particles to abstract entities such as numbers or space and time. One of the most profound changes in our ability to describe aspects of experience has involved developing new conceptions of what it is possible to represent.*

As the scale and complexity of problems experienced by humanity grows, it is critical to augment our problem-solving ability, a large part of which involves the creation of new forms of representation, ideally giving us a better grasp on the most fundamental questions. Gooding (2001), again, articulates it well:

> *But the environment is increasingly populated by artefacts which function as records and as guides for reasoning procedures that are too complex to conduct solely with internal or mental representations. In this way we are continually enhancing the capacity of our environment for creative thought, by adding new cognitive technologies.*

These tools are still in their infancy, and only through an open exploration of the frontiers of their possibility-space will we find the most powerful means to augment our intellect.

Provocations · Prototypes

On Notation as Physics, and the Human Capacity to Learn Environments

I aim to frame simulated environments, such as those found in virtual reality etc, as comparable to (or in the same spectrum as) notations (like mathematical notation) in that they are "environments" with rulesets that can be internalized. This is a framing I haven't explored completely, and I aim to use this essay (in its etymological basis) as an attempt to assay this framing's consonance with other areas of interest within my overall thesis, namely the behavior of hyperphysical affordances and the opportunities for embodiment with simulated objects.

Our bodies have aspects that afford certain approaches to the world. By default these are determined by our physiology, and then circumstance edits that phenotype, augmenting our bodies with environmental objects that can be embodied. We are provided by birth with a body of environmental objects bound to classical physics that we gain facility in maneuvering around, that we feel identified and embodied with. When I use the term *internalize*, I mean to evoke this ability of the body and brain to gather information about the dynamics and behavior of objects sensed from the world via observation and, crucially, direct manipulation and physical engagement, such that the brain restructures to better model the perceived phenomena ("schema") as a sort of internal model[9]. This is the development of familiarity, the transition from explicit guiding commands for limbs and attention to a more automatic, fluent engagement.

Objects in the world can be found or fashioned and be incorporated into the body and that now-changed body encounters the environment in different ways. Critically, as the environment is encountered repeatedly, the (perhaps newfound) capacities of the body collide with and interface with the environment, simultaneously giving the user/owner opportunities to internalize the dynamics of their body and the dynamics of the environment (particularly useful in the ways the environment is newly-accessible or perceivable specifically from the body's new capacities through embodied object augmentation).

---

[9] Finely-calibrated mental models afford a rough level of simulation and prediction of the internalized phenomena or ruleset. Mental math is one such model, as is the ability of the imagination to picture the dynamics of the material of a scarf versus a bag of chips.

The available elements within the environment need not be exclusively materials to be manipulatable. I see the manipulation of elements on/with

pages (as with algebraic notation),
screens (as with the interactional "grammar" of a certain software), or
materials (as with the pattern of operation/manipulation of beads on a soroban/abacus)

as being the manipulation of what the brain treats as an internally-coherent environment whose rules and parameter space can be explored and learned.

To take algebraic notation as an example, its spatial, modular structure of coefficients, variables, and operators has specific rules the user must follow when rearranging elements to maintain equality and mathematical truth[10]. Crucially, the spatial operativity of modern algebraic notation engages the user in ways decidedly unavailable with prior notation-attempts. Though earlier algebraic notations are capable of describing the same mathematical relationships as the more modern algebraic notation, the paragraphic notation[11] prevalent in Ancient Greece, while accurately articulating the mathematics, is notationally unavailable to spatial rearrangement in the way that algebraic notation engages spatial intuition. As a tool for thought, it does not afford manual manipulation of elements in a way that algebraic notation allows the user to explore the system through manual rearrangement. It is this manual operativity that I see as a quality of explorable environments, whether manifested notationally on static paper, or dynamically on screens, or with spatially-interactable objects (be them material or simulated).

These notations are, in a sense, internally-coherent environments created by humans, able to be partially inhabited (through the affordances of their supporting medium, classically though perhaps too often paper) by the body and thus the mind. The most powerful thing about some notations is that their ruleset is more

---

[10] If the user successfully follows the rules of algebraic operations, every mathematical statement produced will be correct, and therefore "true". For example, if the user were to rearrange $2x = 6$ to $x = 6/2$, a correct rule-following, both statements are equally true encapsulations of the mathematical system thus described. However, if the user is mistaken about the rules or unwittingly disregards a rule mid-operation, perhaps rearranging $2x = 6$ to $x = 2/6$, the two arrangements are incompatible statements about an ostensibly identical mathematical system, and thus the errant rearrangement is false as pertains to the original formulation.

[11] A propositional formulation, a sentence-by-sentence description and logical proof of certain geometric elements and relations

internalizable, that their mode of operation can become purely mental, not requiring the initially-supporting medium, and their internalization scaffolds a mental model / simulation of that "environment's" ruleset/laws in the same way that we develop a mental model (or models) of our classical environment's laws, our brains able to simulate hypotheses and without even desiring so, pursue causal chains so ingrained in our set of expectations that it doesn't even feel like thinking or analysis, but something far more direct.

I now wonder if these schema, these mentally-internalized models of experienced environments (be them the classically spatial or more notational) form a sort of Gibsonian ecology in our own minds that via repeated engagement arranges itself into alignment with our external circumstances, whatever they may be (this is where I see simulated, virtual environments' hyperphysics entering into relevance). Might this be relevant for the development of expectation/preparedness/familiarity? I've wondered how Gibson treats prediction, as that does seem to require a sort of internal model/representation independent (though at basis directly dependent on prior sense impressions) of current sense data.

This power of the brain, to plasticly incorporate objects into itself when given enough time to wield them as it learns to wield the genetically-provided object of the body, becomes especially powerful when the objects to wield and embody have a range of behaviors beyond what classical physics allows, as is the case with computer-depicted-and-simulated objects as are interactable in VR etc. This connects back to my framing of notations as alternate "environments", with the key difference that the rules for (for example paper-based-) notation are maintained/forwarded by human mastery of that ruleset, and failures of "accurate depiction" if the rules are forgotten or a single operation is made incorrectly break the environment, whereas the computer ostensibly is rigidly locked into self-accuracy, not to mention the orders of magnitude greater depth of ruleset simulation possible by digital computation.

This greater range of possible behaviors boggles the mind, which makes the job of the designer difficult, and will likely be a cultural, likely generational project, to explore the parameter space of possible "universes" of behavior rulesets to find the most useful (and embodiable) simulated objects/phenomena.

A role of many designers has involved tool design within the classical physics of our lived environment. As computers became ascendant as tools, their simulating ability allowed the design of

phenomena (UI) that could behave in ways other than classical physics, specifically allowing novel tools for thought and thus novel ways of situating/scaffolding the mind. However, the depictive media (e.g. screens) available to represent computed phenomena available were too often exclusively two-dimensional, with only two-dimensional input, failing to leverage the nuanced spatial facility of the body. Now there exist computing systems capable of tracking the motion of the head (thus orientation within possible optical arrays) and the motion of any prehensile limb, capable of simulating three-dimensional phenomena and providing a coherent and interpretable optical array as if the user was themself present in amidst the simulated phenomena.

Critically, a role of the designer no longer purely involves the design of phenomena within physics, but has come to also encompass the design of the physics themselves, exploring how different system-parameters can sustain different phenomena, different notations, and thus new modes of behavior, productivity, and thought.

Overview of Prototypes

In the following sections I am going to discuss prototypes[12] I built to investigate the development of familiarity with novel embodied interactions.

Blend Body uses raymarching to represent the hands such that they fuse with the environment, provoking novel experiences of tool assimilation and UI identification.

Hidden Force Mappings investigates embodied modulation of an intricate particle system via non-obvious control mappings as a testbed for rapid neuroplasticity, provoking experiences of nonlocal agency.

Dust Bodies 1-3 use the hands as force fields to attract and be represented by intricate particle systems, provoking consideration of the implications of slightness when representing the body.

The ***Bulk Particle Series*** document iterations of a rigidbody particle system attached to the body, provoking consideration of play as a design research method for parameter space exploration.

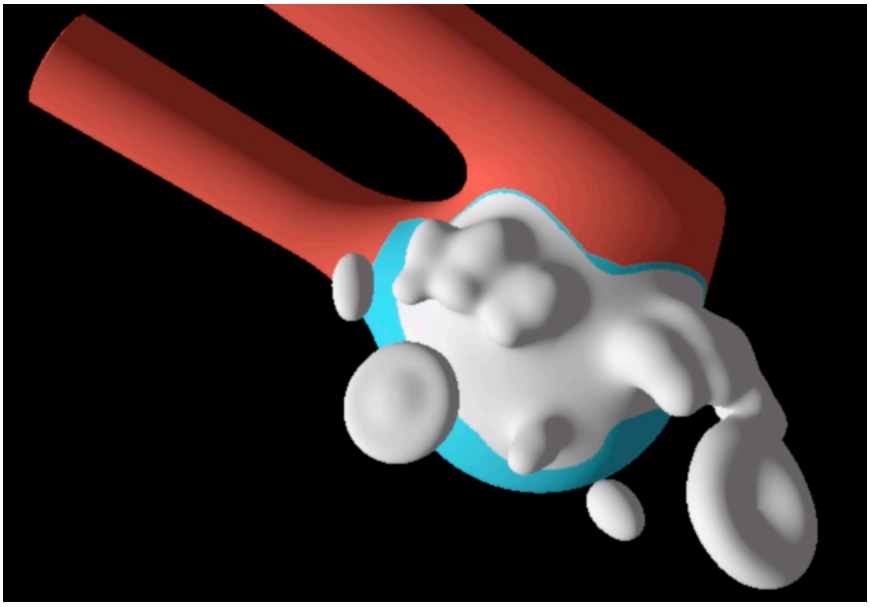
VR Wrist-Finger Haptics implements VR haptics with Apple Watch, provoking conversation on the dynamics of multisensory integration in embodiment and perception of computed objects.

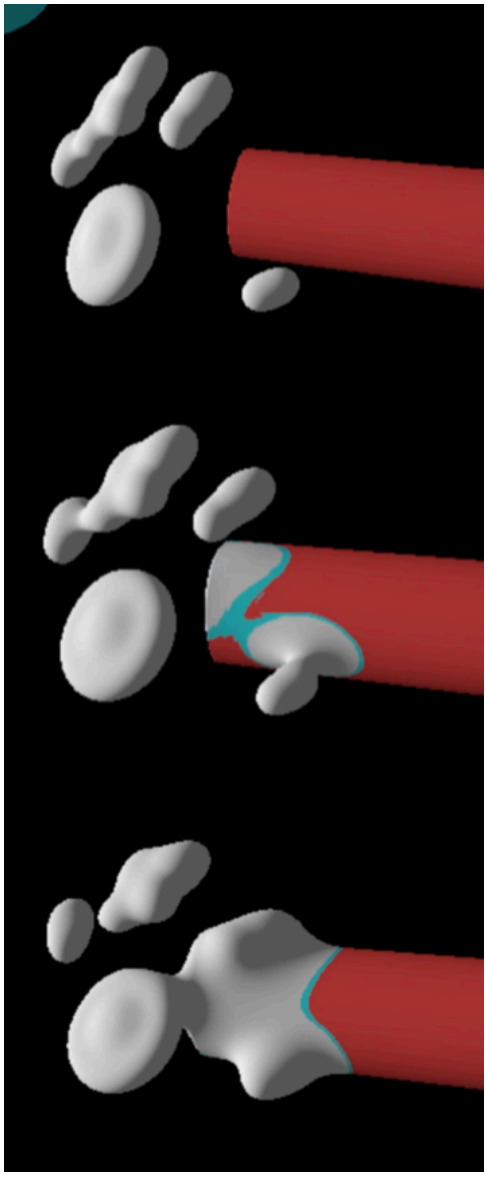
Xoromancy investigates embodied control of high-dimensional structures, using hand tracking to control the generated visual output of a massive neural network.

[12] Available at github.com/graycrawford/masters-thesis-VR

1. Bodily Fusion with Objects via Raymarched Signed Distance Fields · *Blend Body*

Raymarching signed distance fields (SDFs) is a method of rendering 3D shapes without using polygons. It defines each object as its geometric primitive, each influencing a shared surrounding concentric "distance field", and renders an isosurface at a given radius away, thus visually fusing any objects that are within 2r of each other. This property produces very organic forms, where any collision smoothly joins the objects into a melted, singular mass.





I had seen this technique[13] used for external objects, but never for the rendering of the hands themselves, and I suspected it might be quite compelling.

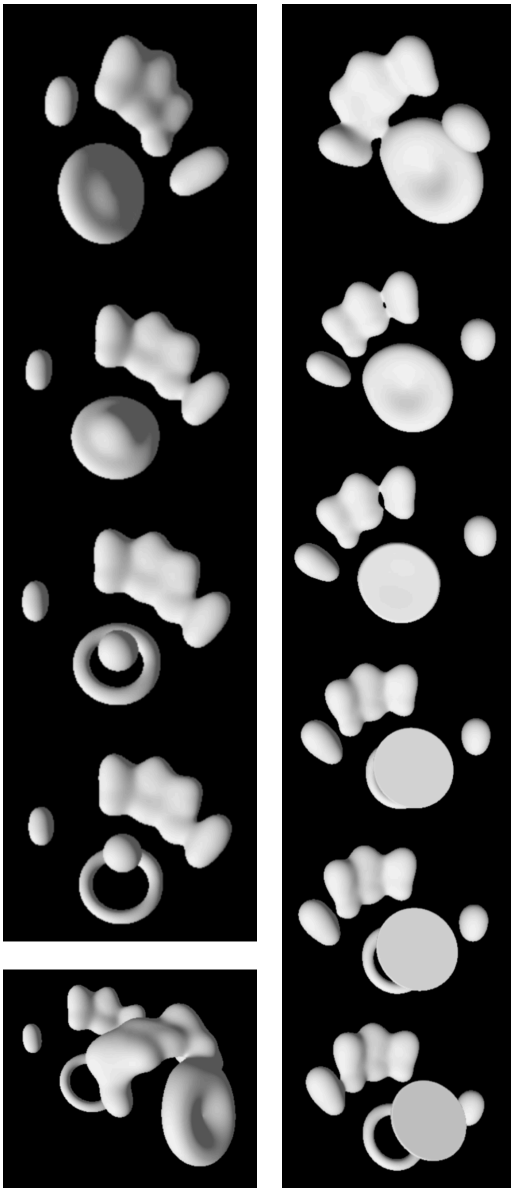
After placing two raymarched spheres on my thumbtip and index fingertip, upon pinching the spheres smoothly transitioned from separate objects to a rounded hourglass to a single ellipsoid. I added the other eight fingertips and populated the world with a sphere and a couple cylinders to observe raymarched fusion with external objects. This was immediately mesmerizing, and changing the effective isosurface radius changed my hands from separate spheres only overlapping within close proximity to a singular doughy mass where the underlying proprioceptive motion remained intact if not only slightly masked.

I added spheres for the rest of my finger joints and knuckles, and found that it felt slightly more dynamic to only include the joints that

[13] Kevin Watters and Fernando Ramallo's Raymarching Toolkit for Unity

I could move separately. My knuckles weren't contributing to the prehensility and only added to the lumpiness's visual extent, so I removed them.

I envisioned that this rendering technique might allow hands wherein the UI was fused with or emitted out of the body directly, elements stored in the palm until their activation.



I used a torus as my palm, as it leaves a circular hole that a spherical UI element could fit in. Upon activation when the sphere floats above the palm, the torus offers negative space behind the sphere which provides extra visual contrast, heightening the appearance of the floating UI. By rising above the palm, the sphere delineates itself from its previously-fused state, spatially and kinetically demonstrating its activeness and availability. This materiality prototype operates more as a wireframe, as a chance to engage with the dynamics of these species of meldings without an immediate application. The sphere is pokable and pinchable, perhaps the type of object that could be pulled away from its anchor and placed somewhere in space (expanding into a larger set of UI elements).

On my right hand, instead of a prehendable object, I wished to see how something closer to a flat UI panel might behave amidst the hand. To remain consistent, I again chose the torus as the palm, and embedded a thin disk in its center that, when the palm faces me, rises above the palm a few centimeters. While docked, the restrained real-estate of the torus again provides the panel breathing-room such that the pair do not, in their fusing, expand to occupy a disproportional volume. In its current implementation, the panel remains the same size through its spatial translation. Future development will change its size during translation such that in its active state it is much larger and might perhaps be removable, existing apart from the hand as a separate panel.

These experiments begin to touch on this novel materiality, and point at ways that UI might be stored within the body, perhaps reinforcing an eventual bodily identification with the UI itself. Further, the ways that grabbed objects fuse with the hand mirrors how the brain assimilates tools into its body schema, and begins to more directly blur the line between user and tool, body and environment, internal and external.

What are the implications of such phenomena? Could a future SUI system be based around body-embeddedness? What would distinguish its set of activities from surrounding objects? What body parts are most available to embeddedness? The arms are arguably

the most prehensile part of the body, and most often within our visual fields, so their unique anchorability is easy to establish.

In future explorations of this rendering technique, I aim to expand on the direct mapping of user motion to behavior of objects in the visual field. How might the sphere behave as an icon of a tool that adheres itself to the fingertip directly, becoming the tool itself (rather than merely a button to enter that tool mode)? How might scaling of objects afford svelter embeddedness before scaling to useful external sizes? Might distant, environmental surfaces communicate their ability to be interactive by showing a partially-submerged hand when gestured toward?

2. Correlating Agency while Acting at a Distance · *Hidden Force Mappings*

It is common in mundane[14] physics to act upon objects directly or via intermediating structures that mechanically extend us. Thus we become familiar with relying on our bodies and tools to interact with external structures. However, computers can simulate alternate physics (hyperphysics) that support action at an (apparent) distance, the simulation itself being the subtending structure allowing for mediation across more than physical movement, but any pairing of parameters. How might this capability affect perception of causality between events and user agency?

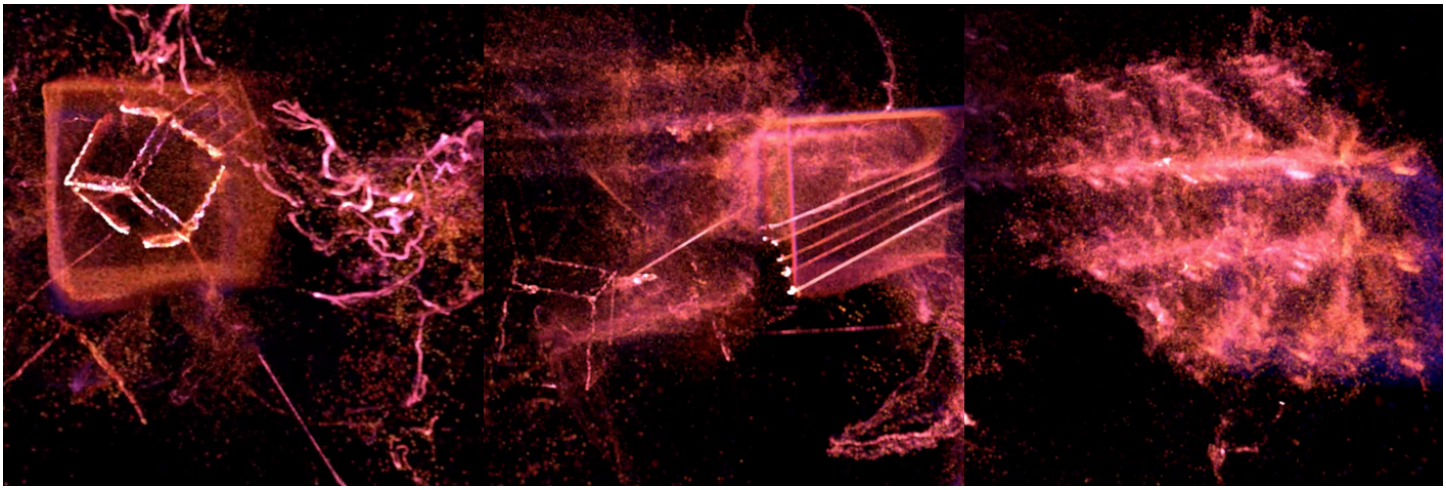
In the simple case of a hyperphysically simulated, purely movement-based causal link, the computer keeps track of every detail of the interaction, translating the movement of the acting object to the affected object even if such objects are not mechanically in contact as we would expect in mundane physics (generally — of course magnetism and gravitation act at a distance but I refer to more common human-scale interactions). This leads to hyperphysical interactions without any necessary and apparent visually intermediating structures, potentially confusing users expecting consistency with mundane physics.

However, there is action-at-a-distance precedent with non-spatial interfaces, in the form of the mouse/trackpad and cursor on desktop operating systems. Though there is no visible structure transferring the motion from the hand/finger to the cursor, the connection is apparent because the causally linked aspects of action and reaction are of the same kind (2D motion), and are mapped 1:1 or a via some scaled ratio. Horizontal input movement is mapped to horizontal output movement, etc. This mapping is easy to break, for instance, if the mouse is rotated 90º, switching the causal link between input and output dimensions, and requiring some period of repeated engagement and failure before the new mapping becomes familiar. This, though, is likely yet more easily learned because the remapped input and output parameters remain of the same type, that of 2D movement.

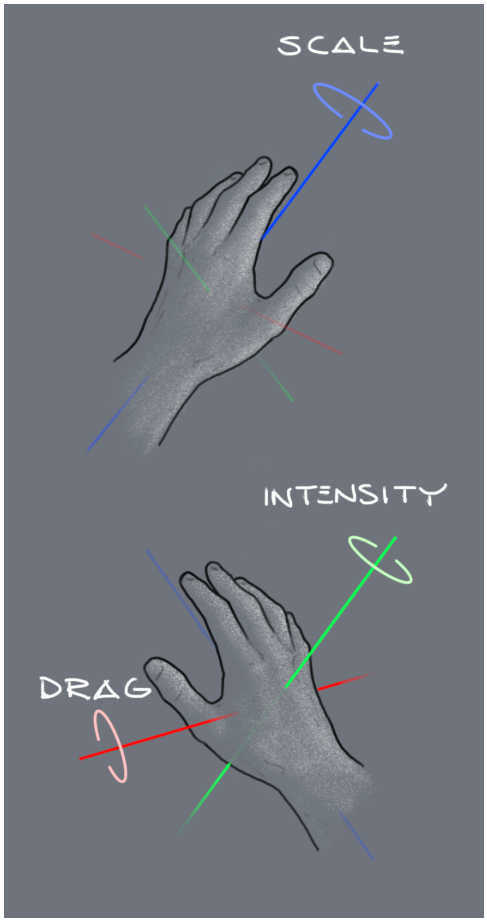
Difficulties abound, however, if the mapped parameters differ in type — if, for instance, input altitude is mapped to output color, or input acceleration is mapped to output volume. Many mappable parameters lack spatial extent or location to “connect” from or to.

---

[14] in-universe / established / classical / common / foundational / base / “real”



Conceiving of this area, I wondered how apparent initially non-obvious and differing-in-kind parameter mappings might become given some amount of practice. To directly experience such a mapping-set, I situated myself in a hyperphysical environment (using Unity's VFX Graph particle compute shader authoring environment) where multiple world-scale forces acted upon the millions of instantiated particles. The forces were themselves the output parameters to my hands' movements' inputs. I intentionally chose non-obvious mappings such that there were no "necessary", (mundane-physics-informed) correspondences that might provide extra traction/bootstraps for my neuroplasticity to correlate.



I mapped my left hand's rotation around the z-axis to the scale of a tessellated 3D block of vectors (vector field) that the particles are drawn into turbulent alignment along. By rotating my hand to the left, I would decrease the vector field scale, causing the tessellation to shrink to the size of centimeters, and to the right, grow to the size of many meters. On my right hand, I mapped the rotation around the y-axis to the intensity of the vector field, such that an increase in the input value would increase the strength of the vector field, and a rotation around the x-axis to the drag coefficient in the volume the particles were instantiated within.

All these mappings set and attached to my body motion, I entered into this hyperphysical space with the goal of seeing if such arbitrary and nonlocal mappings were still familiarizable. By rotating my hands, the output mappings of the vector fields invisibly changed, the only visual changes occurring due to the second-order movement of the mass of millions of particles attracted to and following the state of the vector fields affecting them. Through direct engagement of my body in space with these computed

parameters, and perceiving the visual effects, I was, over the course of a dozen minutes, able to accurately maneuver my hands to produce specific particulate behaviors and structures[15], never before having had any experience with such mappings.

The experience initially was one of more open exploration, where each motion produced a response that was not abundantly clear. Clarity arrived when repeating a motion, returning to a prior coordinate in parameter space that consistently recreated the prior environmental state, the particles falling into lockstep with their prior state and triggering an association in memory between the proprioceptive sensation at that parameter space coordinate and the content in the visual field.

This demonstrates that the mind is capable of correlating disparate phenomena and familiarizing itself with novel mappings that lack direct mundane analogues, and that agency is felt when affecting nonlocal phenomena even when lacking structure that mundane physics requires for similar agency. Likely if such mappings were displaced not only in space but also in time, such perceived agency would dissipate as the sensations, while inputting, grew increasingly temporally disconnected from the perceptible output.

This test could be taken further with increasingly abstracted mappings, seeing what thresholds of parameter type and connection break the perceived agency.

As I became accustomed to my agency over these particles, I began thinking about the subjective experience of having a body, and what the body means in the space of hyperphysics. In our ordinary physics, with our genetically-determined physiologies, our connection is continuous with the objects (body) that we have agency over. This connection can be augmented, in the case of embodiment described earlier, with prehended external objects that we incorporate into our body schema. However, in this particle system, there is no analogous, humanoid representation of the body, yet there is a direct correspondence between the proprioception and certain visual structures that behave coherently and whose behavior can be correlated with proprioceptive feedback. This reactivity is explicitly nonlocal, and lacks even the rudiments of 1:1 mapping of hand position to the position of coherent structures in the visual field. Yet I came to develop a sense of identification with these phenomena, as if they *are* me. I suspect that one key missing

---

[15] ...such as in the third frame, where the tessellated vector field's intensity and drag are held static at high values, causing the particles to be guided into an ordered grid of turbulence...

factor in these mappings is that it is almost one-way. I send output motions, and I receive by default my proprioception, and the visually rendered behaviors in the scene. If I were able to receive sensory feedback from the phenomena happening where I am enacting my agency, perhaps then would those visual structures be cemented as subjectively my body, and not mere objects being puppeted.

3. Slightness of Bodily Representation ·
Dust Bodies 1-3

When phenomena are driven by body motion, and those mappings grow increasingly distant or slight, what is the sensation when discerning the connection between our proprioception and the visual structures?

For example, a direct 1:1 mapping of a sphere's position to the palm's position allows immediate identification with the driven structure, as visual movement is immediately correlatable with proprioception, even if the structure differs from our physiology. However, a mapping where [body part location :: sound frequency or amplitude] is not a sensorially-equivalent mapping of like with like, but the agency is still apparent (if the scaling is such that the sound changes are discernible), and that recognition only falls out of direct engagement and perception of the effects.

It can become increasingly difficult to identify one's self with the observed phenomena as the amount of abstraction between the input body data and the output environmental behavior increases, as seen in *Dust Bodies*. The Unity VFX Graph particles conform to the signed distance field of my hand, and the raw sensory input to my body varies greatly though is nevertheless causally mapped to my bodily motion. Some particles may be so close to my hand SDF that they are drawn quite closely to my hand's motion and are thus positioned essentially 1:1 with my hand surface, being immediately correlatable and thus identifiable as me, whereas further-flung particles receive weaker forces drawing them towards the SDF and enter ever-changing and decaying orbits that somewhat mask the structural basis for their motion.[16]



[16] Video of Black Dust Body

All this is dependent on the parameters chosen in the physics simulation (which could couple the particles very tightly to the body, or make them even more weakly driven and thus less representative of the body's motion), which itself affords a level of real-time modulation to be another set of dimensions of realtime expressivity and environmental responsiveness.



That the body's representation is not a relatively-unchangeable polygon mesh but instead a *field* affords many reactive behaviors that would be difficult to implementably polygonally, where the particles can be additionally affected by newly-introduced field structures that warp the particles' trajectories to flow into visual structures that can interface with the environment or change the shape of the underlying hand.[17]



Our base physiology makes us almost uniformly opaque, which can partially obscure the structures that we work with, and we implicitly learn entire "dances" to maneuver ourselves out of the way so as to view the work unobstructed because we inherit this opacity. If our

[17] Video of Fiery Dust Body

bodies were entirely invisible, we would only have a sense of our motions proprioceptively and when directly interacting with visual structures from which we infer our own motions. This is successfully used in many VR games where the hand is by default visible, but upon grabbing an object the hand is made invisible and the 1:1 moved object communicates the body's motion, correlatable with one's proprioception.

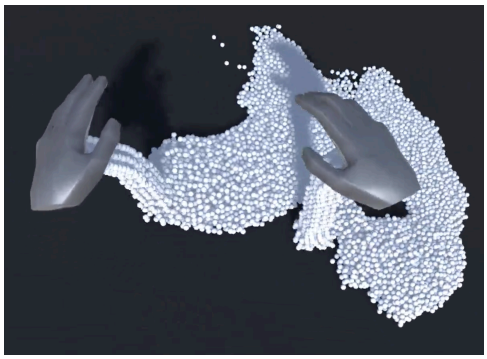
If the body is given only a subtly perceptible amount of opacity, while remaining generally translucent, much more of the background environmental content is available to view, deferring to the content. Even though the visual representation is so slight, there is sufficient data to correlate with proprioception and the structure is easily identifiable as oneself.

4. Parameter Space Exploration via Play ·

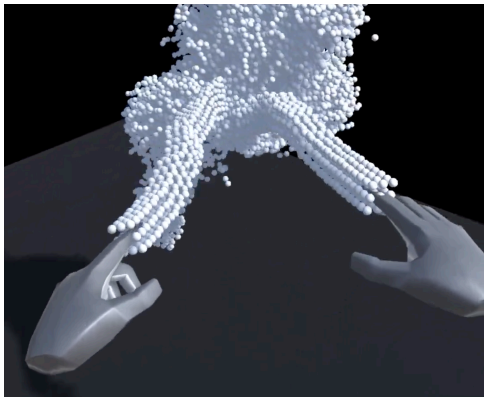
Bulk Particle Series

Whereas the previous particle systems were reactive on a massive scale, they were unable to collide with each other. In the search for a more physically reactive system, I discovered NVIDIA's Flex particle simulation library for Unity. Flex allows for many thousands of colliding particles, and as long as they are all the same radius, they can be meshed into flexible fabrics, enclosed volumes, rigid- and soft-bodies, or remain free-flowing fluids.

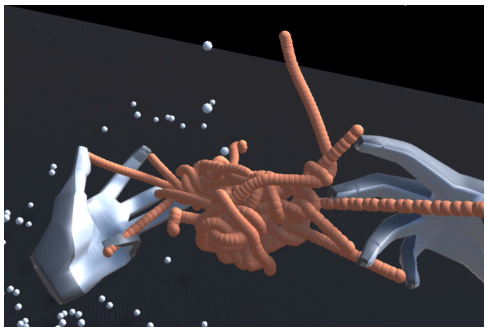
Without a strict plan of attack, I began placing particle emitters on my head and hands, any place that offered me manual control over the placement of the particles in realtime. Immediately I found the physical dynamics captivating, spending multiple hours tuning the available parameters. As the particles were emitted from my palm and came to collide with a flat surface I erected, I changed the parameters of the simulation such that there was supremely increased friction, causing the particles to bunch up immediately upon colliding with each other or the surface, behaving more like a highly viscous goo.



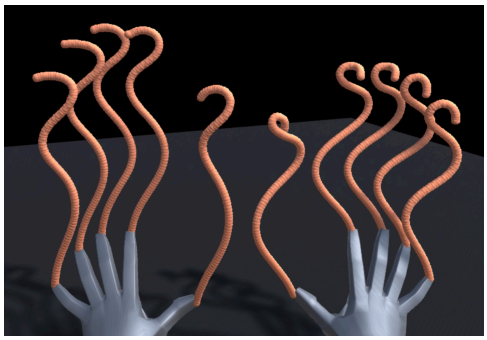
Turning the friction back to zero and the dampening up high produced an effect comparable to if the surrounding "medium" of the air was highly viscous itself, rapidly slowing down any fast-moving particles. In combination with turning the gravity to zero, the intense drag caused the particles to bunch up at a distance perhaps half a meter away from my palm, just floating in mid-air. However, bringing the palm closer caused slightly-less particles with velocity remaining to impact the already-stationary pileup, scattering the static chunk at the point of impact. Applying this effect to different portions of the chunk enabled me to sculpt its form, as I gained finer control and sensitivity to the nuances of its physical behavior.



Wanting to experience the fabric dynamics possible, and at the suggestion of Golan Levin, I attached ropes to the tips of my fingers, initially hoping that I could connect one fingertip to the other. Being unable to do that, I found that, even with one end of the rope loose, the dynamics were nevertheless immediately captivating. With zero gravity, turning the friction up quite high caused the ropes to inescapably tangle with each other, and with such long loose ends floating freely, their pendulous, tentacular behavior was available to play with and, again, was so captivating as to be a distraction from its own development, which is likely indicative of some innate positive aspect.



Turning the gravity to be negative, the ropes were pulled vertically away from my fingers, and with dampening turned up, immediately looked like kelp floating in the slow currents of the sea. By “submerging” my hands under the surface/ground, only the “kelp” was visible, and allowed its behavior and reactions to my movement to be the only visible objects in my field of view. Thus I was able to puppet the kelp around, imagining myself the currents and driving the sway and twist of the vertically-pulled ropes. Thus the physical reactivity revealed an application in the maneuvering and puppeteering of objects when the hands are hidden from view.



It was curious to notice how merely the direct engagement with the phenomena was enough to elicit possible applications, in combination with the progressive altering of the parameters driving the physics simulation. In implementing large-scale novel physics simulations, is the only way to identify possible applications by directly interacting with the dynamics and observing the emergent reactions? If so, it points to the usefulness of play as a research method, especially when mapping out undocumented parameter-space, and how building intimate familiarity with materials helps to better conceive of their possible applications.

5. Hypermaterial Validity via Multisensory Integration · *VR Wrist-Finger Haptics*

As mentioned in the *Coherence and Coordination of Phenomena* section, simultaneous stimuli across the senses reinforce each other to form more concrete perceptions of given phenomena.

Physical controllers such as the Oculus Touch provide haptic feedback with embedded vibration motors, and their haptics are used very successfully in many VR experiences. Valve's *The Lab*'s *Longbow* provides haptics when the user draws the bowstring back, replicating the subtle clicks when tensioning a string. When combined with stereo audio feedback spatialized from the location of the string, the visual feedback of the string, and the proprioception of the pulling finger, these sensory channels fuse into a relatively rich perception of a bow. Curiously, sensations not directly provided by the VR output surfaces can be somewhat provoked or confabulated out of the perceptions from other senses. The combination of sense inputs from *Longbow* serve to also hint at the sensation of tension between the arms, forces a real bow is able to impart but that are impossible with the provided technology. Similarly, passing the hands through the stack of physically-reactive cloths in my 2018 VR music piece *Strata* has for myself and others independently produced the sensation of cobwebs or some light tactile texture, even though there is only visual indication of the dynamics and materiality of the physical structures.

The drawing cursor in the VR CAD program *Gravity Sketch* uses haptics to provide sensations from a distance. The amplitude of the haptic motors in the Touch controllers are mapped to the velocity of the cursor in space, such that moving the cursor slowly produces a subtle buzz in the moving hand, and a quick motion produces a stronger buzz. This helps ground the cursor as a physical object capable of imparting forces, however subtle. This is useful when the cursor comes into contact with a surface. Since rendered surfaces can't provide resistive force to a limb pushing on it, the rapid haptic amplitude discontinuity between free movement and collision immediately communicates the moment of collision to the hand, which in combination with visual input reinforces the unity of the event and the mutual connection between you and the computed environment. The cases where the cursor is projected a distance away from the body but continues to communicate its velocity and collisions via haptic sensations speaks to the flexibility of the mind to incorporate distant objects into the body schema, and influences an extrapolated conception of the body.

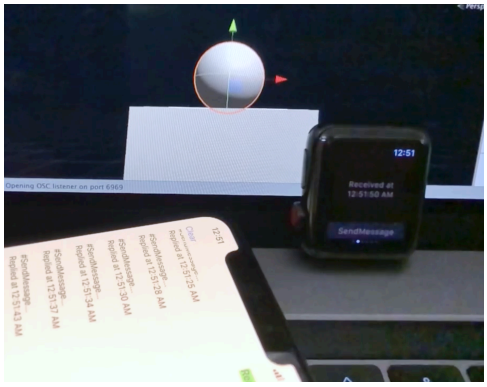
Prototype · VR Wrist-Finger Haptics

My prototypes having previously involved the combination of only vision and proprioception, it seemed useful to prototype for other senses too. I created a setup using an Apple Watch as a haptic source in VR.

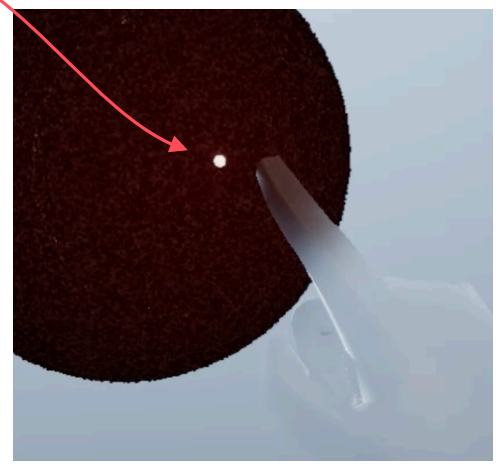
While controllers can usefully provide haptic feedback, if one is using optical hand tracking solutions like the Leap Motion, there is no held hardware to house a haptic motor. I knew the Apple Watch had a robust and subtle haptic (or in Apple's parlance, Taptic™) generator, and could be strapped to my wrist without impacting my hand movements. I searched for previous work using the Apple Watch in VR to no avail, and recognized that I could build it myself.



In Unity I took a collision trigger associated with my fingertip and wrote a script to, upon collision, send a message over the Open Sound Control (OSC) protocol to my iPhone, which ran an app I built to send a trigger to the connected Apple Watch, which I programmed to play a haptic pulse upon reception of the trigger.

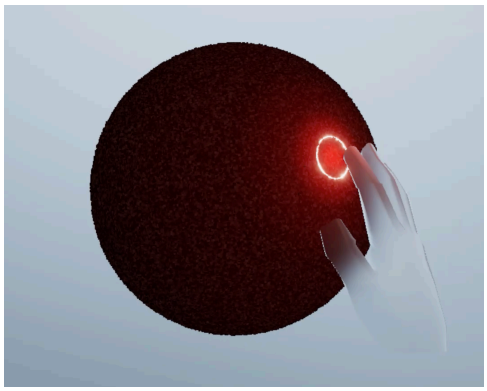


Sadly, the latency of each step in the chain produced a total delay of between 100 and 300 milliseconds between fingertip collision and haptic pulse, way outside the ~50ms window of perceptible simultaneity. To combat this I wrote a script to project a point in front of my fingertip, in the direction of the fingertip's velocity vector, the distance correspondent to the fingertip's velocity magnitude. With the projected point as the true collision trigger, the signal sent to the Apple Watch occurs prior to the fingertip collision itself, and generally synchronizes with the visual indication of fingertip collision.



Without public precedent for Apple Watch haptics in VR, my implementation is likely the first outside of (presumably) Apple itself. This opens the field for researching consumer VR haptics with optical hand tracking without custom haptic hardware.

Obviously, the main drawback of this implementation is the distance between where the collision takes place and where the haptic feedback is provided. I hypothesize that, given sufficient time with the implementation, some more unified percept would arise, similarly to how the texture of and collision with paper is felt localized *at the tip* of the pencil even though the forces are being provided to the body along the sides of the fingers. I also hypothesize that mapping the magnitude of the impact velocity to the magnitude of haptic sensation would provide another dimension of correlation to assist the perceptive fusion of the events. Could an extrapolated effect



work for the action of structures visible external to the body, such that they are felt as part of the body? Further exploration into the flexibility of this implementation is needed.

More work is needed, generally, to see what sorts of sensations can be produced out of the presence of other sensations in VR. How aligned do hyperphysics have to be to normal physics to be coherent?

In the case of audio, many location cues of audio sources come from how the sound interacts with our ears' cartilage and the resonant properties of the cranium. The head-related transfer function (HRTF) describes how frequencies from different directions reflect, refract, and are absorbed by our heads before we hear them. VR implementations simulating someone's custom HRTF with virtual audio sources produce startlingly realistic audio that, combined with visual depiction matching the simulated audio location, are nigh indistinguishable from the real behavior.

The physics behind HRTF are relevant to only a subset of all hyperphysics, but given that our physiology is so determinant to our subjective experience, is it optimal to create spatial experiences that lie only in that subset, to leverage the inbuilt sensitivity and association with certain frequency responses? Might there be a set of requirements for spatially-computed physics that have them align with our current physiology and sense organs for most rich results? Or does our sensitivity to, for example, our personal HRTFs come only from a pure set of experiences unaltered by spatial computing, and such inbuilt sensitivities might be editable if a different mapping of audio placement and resultantly perceived frequencies were provided from birth?

6. Embodiedly Controlling High-Dimensional but Non-Spatial, Non-Physical Structures · *Xoromancy*

Spatial, embodied familiarity can be applied to the control and exploration of non-spatial structures, such as might exist purely mathematically, with no explicit physics applied to any sort of objects with spatial extent. Though these non-spatial structures aren't themselves spatial in the sense of how our bodies and minds experience mundane space (or most spatial computing experiences (VR, etc)), since the (spatial) body is tied to the control of computers generating the non-spatial structure, a comparable embodied familiarity is available, though unique in its divergence from the body's motion's conventional unification with explicitly spatial outcomes. Similar dynamics arise involving the development of embodied familiarity with non-spatial structures.

BigGAN and High-Dimensional Space

In the midst of this thesis a new image generation technique was published: BigGAN[18]. This species of neural network, when trained on massive dataset of photographs, can come to generate strikingly photoplausible images via the coadaptation of two opposing (or Adversarial) networks. The "generator" network attempts to increasingly accurately reproduce from scratch images to appear as though they were taken from the training dataset, and the "discriminator" network attempts to increasingly accurately differentiate the generated images from the ground truth dataset images. The pixel differences between generated and dataset images are then used to backpropagate through the neural networks and adjust the mathematical weights amidst the neurons to, over time, attune the generator and discriminator to generate and discriminate increasingly subtle visual and structural characteristics in the original dataset.

In practice, BigGAN (as trained by Google on fourteen million photographs from ImageNet, those neural weights made freely available) can be made to generate unique images from the startlingly realistic, to the curious and absurd, to the extremities of human conceivability. BigGAN develops a structure which categorizes the visual patterns relevant to identify the provided

[18] Brock, A., Donahue, J., & Simonyan, K. (2018). *Large Scale GAN [Generative Adversarial Network] Training for High Fidelity Natural Image Synthesis*. arXiv preprint arXiv:1809.11096.

categories of images into clumped localities where, for instance, all the lion-categorized images occupy a locally adjacent space, the more similar of each lion images themselves clumping in more local clumps, etc. A different clump far distant would contain all the images similar to, for instance, bedrooms. And a spot equidistant between those two clumps would contain a smooth interpolation of visual structure and color to intermix bedroom and lion.

This structure, or "embedding", has one thousand dimensions / sliders / parameters controlling the relative influence of each of the one thousand image categories chosen by Google to train on, taken from ImageNet.

BigGAN contains a second type of dimension, called the z-vector, a 128-dimensional vector, with 128 components / sliders / parameters. The z-vector affects primarily compositional and structural aspects of the image, whereas the 1000-dimensional category vector affects the colors and textures of the image. BigGAN creates an image when it receives a command containing the 1128-dimensional vector that is the coordinate of the specified image.

For example, the first image is at a point in this 1128-dimensional space in between the collection of alligator-categorized images and the collection of dishrag-categorized images.

The last image is at a point in between volcano, jellyfish, macaw, and lorikeet categories.

The middle image is at a point in between the two 1128-dimensional coordinates of the prior images.

The "space" that BigGAN makes available to explore is not a "space" in the classical or mundane sense of the word. In it there are no "objects" that have spatial extent and that interact according to a set of physical laws. But BigGAN does provide a coherent "space" wherein one can modulate their "position" along any of 1128 dimensions (true, mathematical dimensions, not explicitly spatial dimensions) and produce visual output consistent with that traversal.

This space of possible images is dauntingly massive, and cannot be rapidly explored if the control scheme is trapped into the discrete and stepwise. The online BigGAN explorer *ganbreeder.app* enables productive exploration but also reveals the mismatch between the limited web UI and the vast and nuanced "latent space" within BigGAN.

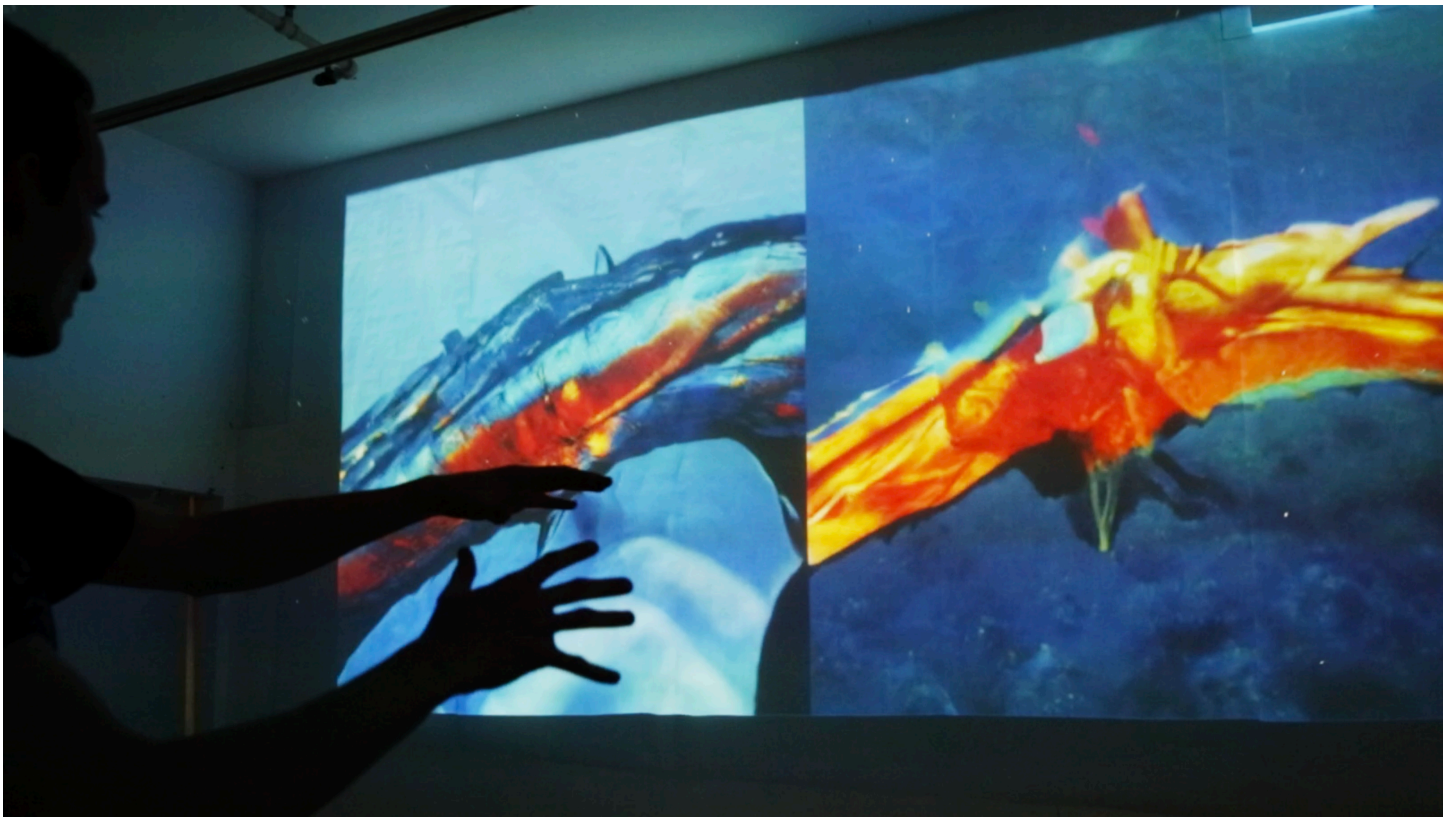






With continued experience with the set of BigGAN-generated images, I became increasingly curious about how an embodied spatial control method might mesh more directly with the continuous and high-dimensional space of image traversal.

Xoromancy



I collaborated with fellow student Aman Tiwari (CMU, 2019) over the course of one month to create an interactive BigGAN explorer controlled by body motion entitled *Xoromancy*. By surfacing fourteen components of the z-vector to be controlled by the motion of the hands, seven per hand, *Xoromancy* provides opportunities for proprioception to sync with the visual output from BigGAN, increasing the rate of exploration and reinforcing the perceived agency over and unity with the dynamics of the output.

Though the body classically builds fluency with spatial, physical interactions, *Xoromancy* demonstrates that the spatial extent of the body can be tied to explicitly non-spatial outputs while nevertheless retaining the species of familiarity-development common to spatial interactions, as the output responds in concert with specific body motions.

The central challenge of *Xoromancy*'s interaction design involved the mapping between input movements and output dimension

modulation. We agreed early on to use the Leap Motion hand sensor as the input source, and though there is a massive set of possible orientations and contortions that the hand can mechanically undergo, the choices of structuring the input/output mapping are far fewer, in part due to the limitations of Leap Motion's tracking. The final mapping involved, for each hand, maps translational movement along the x, y, and z axes to three components of the z-vector. The rotation of each hand was mapped to four more components of the z-vector using the 4D quaternion method of decomposing 3D rotations providing *Xoromancy* an extra z-vector dimension of control per hand than if the classical Eulerian xyz decomposition were used.

Work had to be done to tune the scaling factor of each transform to fit within the range of z-vector component magnitudes that produced coherent output imagery. Outside of a certain range, z-vector values produce visuals that initially seem to break out of the (unchanged) category vector, intermixing subjects, before extreme z-vector values dissolve into harsh visual artifacting.

The choice of the seven dimensions per hand was most constrained by the desire/requirement that each dimension be traversable orthogonally to any other — that is, that no action to modulate a given vector component would necessarily modulate any other, that they be independent unless otherwise desired. For example, one could rotate one's hand (affecting the quaternion-controlled dimensions) without affecting the dimension mapped onto translation along the y-axis.

Attempts to map all 128 z-vector components onto the hand is a near-impossibility given current hand-tracking fidelity, and much work needs to be done to explore/determine what the maximum number of orthogonal dimensions might be mapped onto the set of contortions of the hand. Conversations with computational artist Memo Akten illuminated the possibility of treating the high-dimensional space of all possible comfortable and mechanically valid conformations of the hand as a subsidiary, embedded "manifold" of all possible 3D arrangements of the 20-odd joint vertices (ignoring mechanical limitations). A neural network could be trained to project between the mechanically-valid manifold and the 128- or 1128-dimensional space of BigGAN images, optimizing for maximum visual difference of output per given input movement. Such an approach highlights the size of the design space once neural network optimization is used to descend towards mappings that might be impossible to arrive at to manually situate, but that a computer would have little trouble in balancing.

Upon mapping hand motion in those fourteen total dimensions to the first fourteen components of the z-vector, the nuance possible was immediately apparent. Although there is no explicit meaning behind the ordering or operation of each z-vector component, once each component is mapped to a body movement, merely moving the body instantly displays how the visual form is controlled. Visual patterns of growth or sliding or transformation or inversion are apparent, and are reinforced when the proprioceptive awareness of the hands' movements are correlated to visual output. Since hand movements are reversible, and the visual feedback reflects that reversion in realtime, coordinates in BigGAN's "latent space" are available to be coherently returned to.

Xoromancy is distinct amongst the other prototypes in that 1. it is not VR-based, and 2. it is the only one deployed into the world and tested beyond myself, having had multiple public exhibitions.



First premiering at CMU's *Frame Gallery*, February 22-24, 2019, *Xoromancy* enabled the audience to explore curated BigGAN images. Participants had a range of reactions, some of whom used exclusively gross movements, whereas others were quite meticulous and measured in their BigGAN traversal. Multiple participants expressed the sense that they were "mastering" it over time and this matched my and Tiwari's personal experience, where repeated

engagement gave opportunities to increasingly attune oneself to the intricacies with the BigGAN model.

At the time of writing, *Xoromancy* is installed at New York Live Arts for their *2019 Live Ideas* festival from May 8-22.

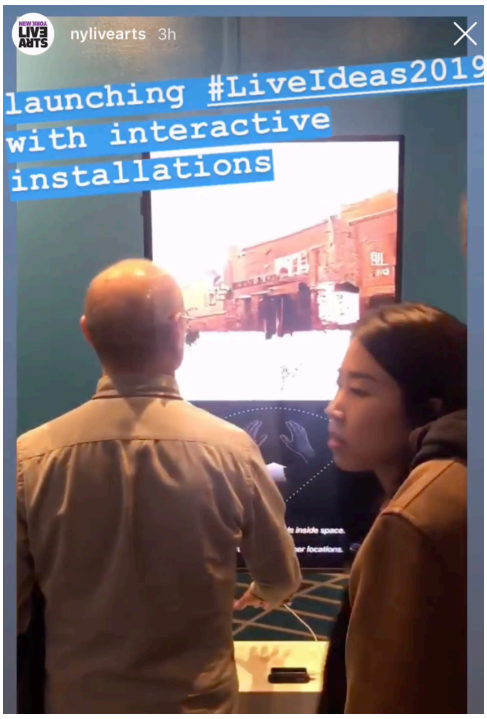
In light of its originality, *Xoromancy* was also accepted to *IEEE-GEM* [Games Entertainment & Media] *Conference 2019* as *Xoromancy: Image Creation via Gestural Control of High-Dimensional Spaces* (Tiwari & Crawford, 2019)[19], and will show from June 19-22 at Yale University.

. . . . . .

BigGAN's latent space, while not being literally a "space" containing physical dynamics or affording embodied movement within, is nevertheless a space that the body and mind can grow direct, embodied familiarity with, provided that some species of motion tracking enables mapping between the body and the output. Qualitatively distinct experiences of this latent space are possible when the latent space is itself attached to you, where the only visible results of your motions are explicitly non-embodied and non-spatial, instead being 2D images that change in deeply nuanced ways that coherently reflect your motions.

The scale of the GAN latent space situate GANs and other generative neural networks as a pivotal new opportunity for spatial, embodied interaction, where the wide ranges of motion of the body can interface with such nuanced and nigh-inexhaustible visual content. As machine learning continues its inexorable climb into greater relevance, I predict many rendering methods will come to incorporate it.

Though *Xoromancy* is not a spatial representation, I see no reason why entire coherent stereo optical arrays can't in principle be generated by neural networks in realtime, with all of the capabilities that current spatially-computed renders offer. This opens up an entirely new class of hyperphenomena, and deserves much investigation.





[19] See the Appendices for submission texts to IEEE-GEM and SIGGRAPH

Reflections on Prototyping

Though varied in form and implementation, my prototypes are the products of an approach to exploring spatial computing that provokes questions surrounding the nature of the body and its relationship with computed phenomena. As the field of spatial computing develops and evolves, new methods of generating sensory phenomena are developed and become available, in effect continually widening the available hyperphysical design space.

Many of my prototypes stemmed from seeing Twitter posts showcasing or announcing new rendering methods such as BigGAN, or Unity's Visual Effect Graph. Implementing them in an embodied context produced novel outputs and provided insights that I would likely not arrived at had I only used my imagination. At the moment, Twitter houses the thrust of the field's public conversation around spatial computing, and the visually satisfying novelty of creative reapplications serves to amplify the collective awareness of the more expansive possibilities. As a unified learning and sharing platform, it accords quite well with the dynamics of the field's burgeoning medium.

By evaluating my prototypes solo and in real time, I was able to iterate rapidly. Assisted by runtime editing in Unity, this often took the form of associating a hyperphysical structure with the location or movement of my hand, tweaking parameters in the simulation as I actively modulated the structure with my tracked hand's movements. This expressivity of bodily motion brought out nuances of physical reactivity that I might not have arrived at had I only watched the reactivity rather than having been its causal source.

What I explored was a minute section of the available exploration space, nevertheless revealing many insights into new dynamics within spatial computing. It is my hope that designers operating in this field take opportunities to experiment with novel materialities and mappings, as collective effort to discover and map hyperphysical frontiers furthers the field and deepens the collective understanding. Even more critically, the direct, embodied experience of these new phenomena gives a nuanced understanding that is impossible to fully convey with paragraphs, so in many ways the best method of propagating these understandings is through the creation and sharing of the experiences themselves.

Discussion · Conclusion

The framing of hyperphysics urges designers to conceive of the classical physical and computational context as only one of a massive set of possible physics, ripe for exploration. Critically, the body's relevance in hyperphysical exploration cannot be overstated. Within spatial computing, the representation and behavior of the body is as available to designers as any other computed structure, and this heralds a transformation in how we perceive and conceive of our bodies and our relationships with the external.

Many invariants taken for granted when designing in the physical world fall away when opened up to the computer to simulate, and this space of design supports a range of phenomena that we haven't had access to before, requiring methods to explore and assess the available design space.

The series of prototypes produced in this thesis investigate the opportunities of representing the body with novel hypermaterials that can undergo transformations impossible for our physiology, and methods of mapping bodily motion into computed environments, enabling direct, embodied control over unusual phenomena. In both cases the flexibility of the mind in adapting to the novelty of hyperphysics is observed, raising questions of what the threshold of neuroplasticity is for increasingly divergent physics.

When the body has agency over external objects, without any explicit depiction of the body in its classical physiology, questions arise over what the boundary is between the body and the environment. Is any structure that responds to your action "you"? Post- the material delineation between body and environment, is your body all the structures that you have control over? The incorporation of external objects into the body schema is an established phenomenon, and this path likely holds many fundamental curiosities for future designers to investigate.

The easy approach to depicting hands in spatial computing is to drive analogously-structured polygonal meshes of the hands from one's hand motions. But with essentially complete freedom of depiction, what are the minimum set of visual cues needed to establish identification with a hand-structure? What do rendering techniques like particle systems afford for the transformative quality of future bodies?

Outside of traditionally-spatially-rendered solutions for spatial computing, what opportunities are there for neural networks to be driven by body motion, the embodied control of explicitly non-spatial structures?

Spatial computing is a context for us to better understand and personally experience states of being that we've never before had access to. It is the context for new tools for thought leveraging the power of modern computation with bodily attunement.

In its current state, a robust spatial computing barely covers sight, sound, and haptics, but an extrapolated spatial computing completely informs the sensory field, which is a daunting and delicate context. Exploring the largely unintuitive possibility-space of current spatial hardware and software provides opportunities to experience in the short term where concerns might come into effect, informing a deeper extrapolation of where computer-mediated experience will lead.

Designing hyperphysical spatial interfaces allows us to explore novel interaction methods that are not possible yet with atoms, but that in the future might be. Though not materially instantiated, computed implementations approach the level of utility as the future material instantiation, and provide opportunities to prototype interactions not possible today.

The value latent in this expanse of now-available hyperphysics is the set of its possible applications toward augmenting human intellect. This situates embodiment in spatial computing and this thesis amidst a lineage of topics explored by thinkers like Douglas Englebart, Ivan Sutherland, Alan Kay, and Bret Victor.

Our ability to represent and manipulate external information and to output our own is deeply influenced by the capacities of the supporting medium and interactional notation/grammar. Thus the set of thoughts available to come out from a static, paper-based medium differs from the set available from a 2D dynamic, computed medium, which both differ from the set from physical objects which differs from the set from 3D dynamic, computed, embodiable structures. With more powerful and dynamic representations, we can think and create new things.

By communicating this range in the factors influencing and falling out of perception of bodily engagement with hyperphysics, this thesis hopes to instill a deeper conception of the fundamental powers and nuances that embodiment with spatial computing offers, such that readers might widen their conception of what is possible to create.

I hope that out of this thesis, readers will recognize the flexibility of the brain, consequently recognizing the available parameters in a spatial design and incorporating their modulation via the body,

leveraging direct manipulation and embodied engagement. All that is required to get a direct sense of a new phenomenon's dynamic properties is to attach it to or associate it with the body, manipulating it to reveal its responses, and develop familiarity.

As spatial computing is still young, many of its fundamental UI patterns have yet to be solidified, making this time ripe for exploration before patterns become entrenched. I hope this thesis documents approaches to authoring embodied interactions in such a way that inclines readers to experiment with hyperphysical possibilities before importing an established UI pattern from a more constrained space of mundane physics or 2D computation.

I hope this thesis indicates that the brain is capable of attuning itself to phenomena far beyond the range of phenomena available in nature, such that future embodied tools leverage that innate neuroplasticity to transform what it means to be human.

References

Blom, K. J. (2007). *On Affordances and Agency as Explanatory Factors of Presence*. Extended Abstract Proceedings of the 2007 Peach Summer School. Peach.

Blom, K. J. (2010). *Virtual Affordances: Pliable User expectations*. PIVE 2010, 19.

Chemero, A. (2009). *Radical Embodied Cognitive Science*.

Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.

Disessa, A. A. (1988). *Knowledge in Pieces*.

Dourish, P. (2004). *Where the Action Is: The Foundations of Embodied Interaction*. MIT press.

Engelbart, D. (1962). *Augmenting Human Intellect: A Conceptual Framework*. Stanford Research Inst. Menlo Park CA.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Psychology Press.

Golonka, S., & Wilson, A. D. (2018). *Ecological Representations*. bioRxiv, 058925.

Gooding, D. C. (2001). *Experiment as an Instrument of Innovation: Experience and Embodied Thought*. In Cognitive Technology: Instruments of Mind (pp. 130-140). Springer, Berlin, Heidelberg.

Heersmink, J. R. (2014). *The Varieties of Situated Cognitive Systems: Embodied Agents, Cognitive Artifacts, and Scientific Practice*.

Leithinger, D., Follmer, S., Olwal, A., & Ishii, H. (2014, October). *Physical Telepresence: Shape Capture and Display for Embodied, Computer-mediated Remote Collaboration*. In Proceedings of the 27th Annual ACM Symposium on User interface Software and Technology (pp. 461-470). ACM.

Maravita, A., & Iriki, A. (2004). *Tools for the body (schema)*. Trends in cognitive sciences, 8(2), 79-86.

Piaget, J., & Cook, M. (1952). *The Origins of Intelligence in Children* (Vol. 8, No. 5, p. 18). New York: International Universities Press.

Rybarczyk, Y., Hoppenot, P., Colle, E., & Mestre, D. R. (2012). *Sensori-Motor Appropriation of an Artefact: A Neuroscientific Approach*. In Human Machine Interaction — Getting Closer. InTech.

Smith, R. B. (1986). *Experiences with the alternate reality kit: an example of the tension between literalism and magic*. ACM SIGCHI Bulletin, 17(SI), 61-67.

Sutherland, I. E. (1965). *The Ultimate Display*. Proceedings of IFIP Congress, 506-508.

Won, A. S., Bailenson, J., Lee, J., & Lanier, J. (2015). *Homuncular Flexibility in Virtual Reality*. Journal of Computer-Mediated Communication, 20(3), 241-259.

Further Reading

Bret Victor

The Humane Representation of Thought (2014)

Media for Thinking the Unthinkable (2013)

A Brief Rant on the Future of Interaction Design (2011)

Jaron Lanier

Lanier, J. (2017). *Dawn of the New Everything: A Journey through Virtual Reality*. Random House.

Appendices

Xoromancy

IEEE-GEM submission text

Xoromancy explores methods for gestural control of high dimensional spaces, applied to the control of generative adversarial networks. Correlation between proprioception and visual feedback builds an intuitive understanding for navigating the highly-nonlinear mappings between input dimensions and generated output imagery.

Generative adversarial neural networks (GANs) can produce uncanny, highly diverse, and pseudo-real images that are otherwise impossible to produce in real-time. However, the input spaces of modern neural networks are vast, and stepwise or random exploration is a slow traversal method. Xoromancy addresses this by combining Leap Motion hand tracking with realtime GAN-generated images. With 128 dimensions of bigGAN's z-vector input and 1000 dimensions of its category vector input available to control, the central challenge is mapping as many dimensions of hand movement onto a subset of the available vector components. Each dimension affects some property of the generated image, whether it be what the subject of the image is or the specific instantiation and visual composition of that subject within that image.

This implementation maps seven vector components to each hand for a total of 14 controllable vector components. Each hand has X, Y, and Z translational axes mapped, as well as I, J, K, L quaternion rotational axes mapped. In addition, there are gestures to navigate to the next pre-set point if one wishes to explore another corner of the input space. These 14 components are then mapped onto the 128 + 1000 dimensions of the input space.

The subjective experience of moving one's hands through space and receiving real-time visual feedback of the generated image enables rapid learning of how to modulate the hands to generate specific images and explore in specific directions, building up intuitions for the underlying structure of the latent space that would be essentially unexposed via a purely stepwise traversal.

SIGGRAPH submission text

Many tools for exploring the image latent space of generative adversarial networks (GANs) provide only stepwise, image-by-image traversal. Xoromancy's use of hand tracking for controlling many latent space vectors simultaneously with realtime image response enables rapid and fluent exploration.

It is the first interactive real-time tool leveraging the human body's proprioception and fine motor control to control the generation of images by neural networks. These images, produced by bigGAN (Brock et al, 2018) are high-resolution and can often near-photorealistically represent objects and scenes that would be otherwise time-consuming to render through other means.

Artistically these bigGAN images exist in spaces that can range from the highly realistic to the distinctly surreal, often containing perceptual cues conveying realistic lighting, texture, and forms but with decidedly unreal subjects.

However, the input spaces bigGAN and similar networks are extremely high-dimensional and are prohibitively difficult to explore and conceptualize. This makes exhaustive exploration intractable and random exploration slow and often unrewarding.

We recognize that the human body is itself a continuous and high-dimensional control system, and modern body tracking such as the Leap Motion sensor enables the body's motion to drive the high-dimensional input vector controlling the GAN's visual output.

Although the possible dimensions of bodily movement are many, only a subset are mechanically comfortable to control and conceptually intuitive. The input mappings chosen for Xoromancy support both fast exploration through a variety of images and specific refinement towards a goal once an image has been found.

Xoromancy includes distinct gesture categories to independently control the subject of the generated image and its specific composition and form.