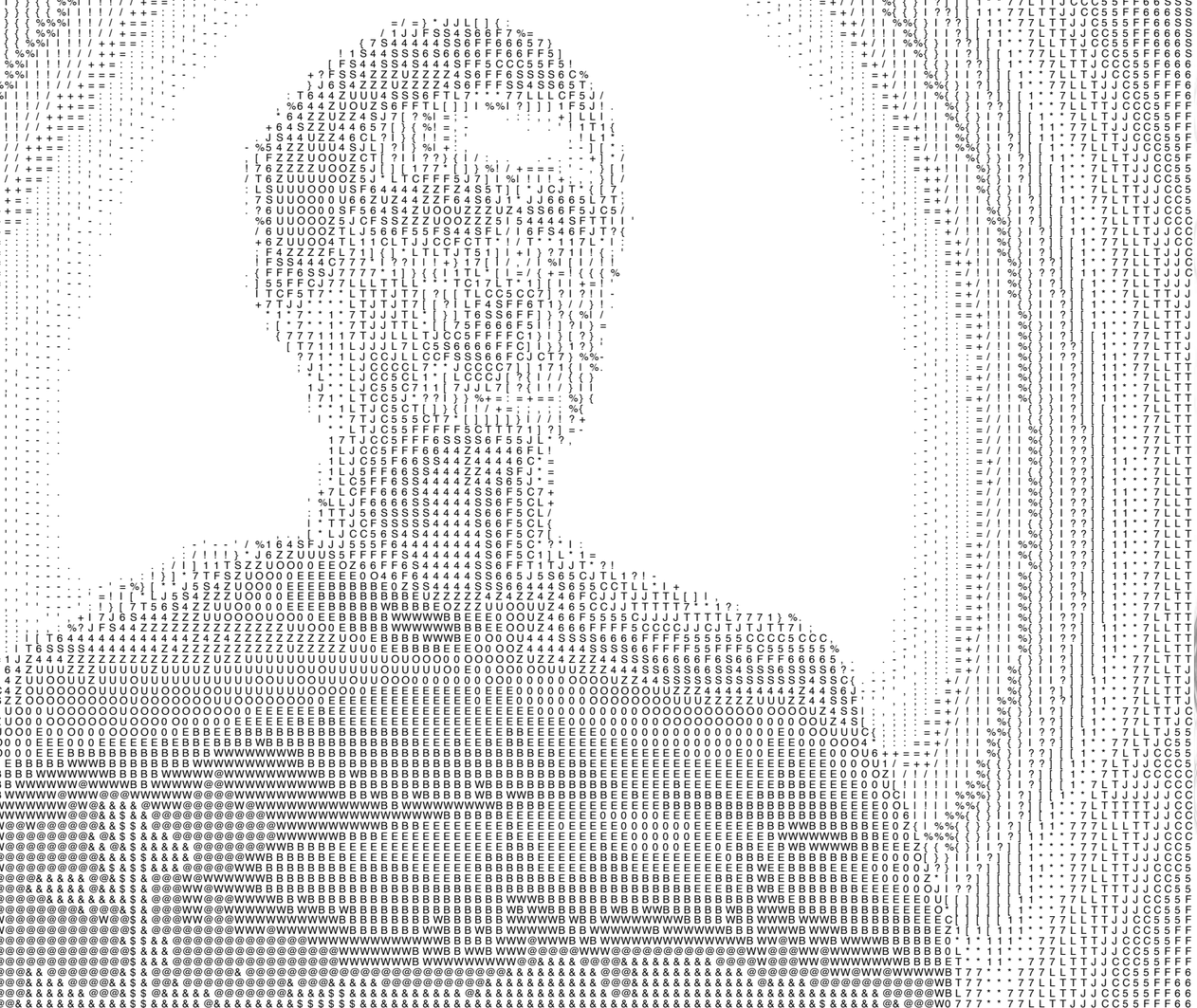


Dynamic Systems of Engagement

Scott Murray

Dynamic Systems of Engagement

Scott Murray

Dynamic Systems of Engagement

Scott Murray

This thesis is submitted in partial fulfillment of the requirements for the degree of Master of Fine Arts in Design and is approved by the MFA Design Review Board of the Dynamic Media Institute at the Massachusetts College of Art and Design.

May 2010 — Boston, Massachusetts

Gunta Kaza, Thesis Studio Advisor
Professor of Graphic Design

Joe Quackenbush, Thesis Seminar Advisor
Assistant Professor of Design

Jan Kubasiewicz, Program Coordinator
Professor of Design

Abstract

We now have the tools and technology to create just about any interactive system imaginable. But how can we ensure that our designs are engaging? *Dynamic Systems of Engagement* illuminates how dynamic, interactive, computationally based systems offer new opportunities for engagement with participants and third-party observers. Through numerous case studies, I explore three core themes: data visualization, dynamic systems, and engagement.

I consider data visualization broadly as a process of interpreting and expressing data of all kinds, not just numbers and text. I explore principles of systems design to illustrate how dynamic systems differ from works of static, pre-composed media, like painting, film, and television. Finally, I connect these themes to methods of interaction and engagement.

My past projects illustrate a range of design possibilities grounded in these ideas. From *Gesture Project*, which responds to physical gestures with patterns of rotating, color-changing discs, to the *ASCII Photo Booth*, a high-tech, low-fi interpretation of a traditional photo booth, these interactive studies illuminate nontraditional uses of data visualization, systems design, and interface concepts.

Although the concepts are valuable, more important is how real people *respond* to the designs. That is, what is the *experience* like? I conduct extensive user research with each project, the findings of which are used to refine the designs and inform future projects.

I adopt a framework of challenge and reward for sustaining engagement, which I then employ for two primary thesis projects, *Practice* and *Cheeky*. Although each project has its own distinct content and approach, both elicit engagement by employing visual mirroring, establishing tension and ambiguity, and finally resolving that ambiguity, providing closure to the experience. Both projects address the questions: How can we challenge someone while keeping them engaged, and how can we incentivize participants to overcome the discomfort of the challenge?

Practice is a new interactive video piece that employs metaphors of stillness (physical and psychological) and reflection (visual and personal). While most interactive video installations reward motion, *Practice* rewards stillness, and in so doing tests participants' tolerance for physical discomfort and emotional ambiguity.

Practice employs computer vision methods of face detection and face tracking to identify participants' presence and level of engagement, so that mere visual stillness, without engaged users, elicits no reward. Visual and aural cues incentivize users to overcome the discomfort of the challenge, by establishing anticipation of the rewards to come. And through it all, the system collects data on participation, which is analyzed and visualized.

Cheeky, a second interactive video piece, is introduced and shown to apply the same principles of experience design to engaging and humorous ends.

Contents

Thanks

Thanks to Gunta Kaza for her keen emotional eye and invaluable visual and psychological insights.

Thanks to Joe Quackenbush for numerous valuable critiques and for so often just “getting it” and articulating why an idea was great, even when I wasn’t so sure.

Thanks to Jan Kubasiewicz for his philosophical guidance and bureaucratic acrobatics, without which I would not have been able to complete my time at DMI from San Francisco.

Thanks also to all the other DMI family members for their support over the past three years, including Agata Stadnik, Audrey Fu, Brian Lucid, Colin Owens, Cici Li, Dan Johnston, Dennis Ludvino, Elaine Froehlich, Jason Bailey, Lou Susi, Mike Golembewski, Simon Bae, and many others. You all influenced my work, sparked new ideas, and battled the MassArt network in order to bring me to class via video. (Now please stop calling me “Computer.”)

Special thanks to Aimée Bruederle for her close review of this document, and to Huy Le for his ongoing encouragement and the use of his salon for my crazy art installations.

Extra-special thank-yous to my amazing partner, Nora Heins, without whose prodding, encouragement, and support, I never would have pursued an MFA in design, let alone discovered this new life path. Thank you for nudging me in the right direction.

A view of Rainbow Pools.

Preface

On the road up to Yosemite, there is an unmarked turn-off that leads to a dirt road. Around the bend, a small parking area comes into view, and — in the summer — so do hundreds of swimmers seeking respite from the heat: Teenage couples (or couples-to-be), moms and dads, and hundreds of little kids easily scramble over the granite boulders as though they weren't at high altitude. And in the center of all this is Rainbow Pools, an oasis of cool mountain water in the high foothills. A 20-foot waterfall cascades down into the largest pool, a large hole in the ground originally dug as a mine shaft, going who-knows-how-deep into the Earth. After the mine was abandoned, a kitschy restaurant was built on the granite shelf above the pool, so visitors could admire the top of the falls and the swimmers below as they ate. The restaurant burned down years ago, and now, only the water remains.

My partner Nora introduced me to Rainbow Pools for the first time last summer. I stripped down to my swimsuit and paddled around with a local terrier, before hopping back out on the rocks. As a kid, I learned to swim in chlorinated pools, and never stepped into open water until my junior year abroad in Australia. Without the safety of a designated, humans-only swimming area, I am flooded with anxiety about all the critters whose space I'm entering, and how much or which parts of me they are interested in eating. In chlorinated water, Band-Aids and clumps of hair don't bother me; at least I can see the bottom, and verify for myself what is (and is not) in the pool with me. After my very first non-chlorinated swim in Australia, my phobia spiked when we were told that there could have been estuarine crocodiles nearby, but in that time of year, it was "unlikely." Well, I guarantee that no crocodile, estuarine or otherwise, has ever taken a dip in the Rinconada Public Pool in Palo Alto (where I was a Junior Lifeguard for three summers) — they wouldn't be able to afford the $2.50 admission.

See, I joke to diffuse the tension. Back at Rainbow Pools, I was comfortably out of the potentially infested waters, and noticed that children were swimming up to, then through the waterfall itself — and disappearing. Then, a minute or so later, they reemerged, laughing and splashing. I jumped back in, determined to push past my fears and swim behind the waterfall.

It was an easy decision to make from 100 feet away, at the far end of the pool. But once I was in the water, very near the base of the waterfall, it became clear how much water was falling off of that granite shelf, because it

A peace lantern in the Japanese Tea Garden in San Francisco.

A bridge in the Japanese Tea Garden in San Francisco.

was all pushing me back, away from my goal. So I swam faster. And I'm a decent swimmer, but, as I have said, I swim best in filtered, pH-balanced, still water. Beneath the falls, I needed more air, since we were at 5,000 feet, and the thundering force of the water was fighting me every step of the way. Plus, I didn't have goggles, flippers, and a kick-board — I couldn't see a thing, I couldn't hear Nora cheering me on (was she?), not anything but the pounding roar of the beautiful, indifferent falls.

I remember maneuvering to the left side of the waterfall, so I could both swim and grasp at the slick, vertical rock wall, hoping for a handhold. My heart was already racing but sped up when I realized that the rock existed only above the water line; when I felt for a foothold underwater, there was nothing there. (Probably where the electric eels hang out, I thought.) And then, in that final push beneath the cascades, there weren't any thoughts at all. I was pure action and reaction — swimming, breathing, grasping, grabbing, and then, all at once, I was out of the water, sitting on a small, moss-covered granite ledge, feeling beaten up but very alive, and looking back, up, and out at the underside — the inside! — of the waterfall.

My racing thoughts began again, picking up where they left off, but for a moment, they had been silenced by a combination of ambiguity, discomfort, and focus. Would I make it? I didn't know. What could I expect once I arrived? No idea, as it couldn't be seen. In that moment, my immediate future was ambiguous and unpredictable. The next moment could have played out in any number of unforeseeable ways.

I was also extremely uncomfortable. My fear of deep, dark water was in play, as was the fact that I couldn't see or hear anything useful. I was literally grasping in the dark, in search of direction and stability, while being hammered with a turbulent stream of water from above. Oh, and I was out of breath, too. I wasn't in pain, exactly, but there was heightened physical, emotional, and intellectual discomfort. I couldn't reason my way out of the waterfall. The only way to move forward (or even go back) was to focus.

In Japanese Buddhism, *satori* is sudden enlightenment, a rare and brief glimpse of truth. I understand satori as those brief moments of thought-less-ness in which one's day-to-day mind falls away, and the whole being operates at such a level of intense focus that time seems to slow down, without room for conscious thought. In reflection, it may not be that time "slows down," but that in satori, we lose awareness of time (and all other abstract, intellectual constructs), as we focus only on the present moment. It is only in retrospect, when remembering our experience of no-time, that we can analyze it with the benefits and entrapments of abstract thought. We attempt to bring that special moment forward into the present one, in order to relive and reexamine it. But in satori, when thought leaves us, all that's left is reality. For the Buddhist, perceiving reality correctly is, of course, the ideal achievement. Buddhism calls the result enlightenment, and satori is like a tiny, bite-size taste of enlightenment.

In the Zen tradition, satori can be induced by koans, seemingly nonsensical stories provided to students by their Zen masters. A student will receive the koan and meditate on it, for hours, weeks or even months, struggling to comprehend the incomprehensible. Koans are designed to essentially snap the student "out of his mind" and into reality, thereby achieving satori. Thus, they are intentionally illogical and cannot be "solved" like a riddle or even "understood" using the mind. (Consider the now-proverbial "sound of one hand clapping.") The koan is a tool for short-circuiting the student's mind, and bringing her back to the present moment.

One of my favorites is from *Mumonkan* ("the gateless gate"), an early collection of Zen koans:

> The wind was flapping a temple flag, and two monks started an argument. One said the flag moved, the other said the wind moved; they argued back and forth but could not reach a conclusion. The Sixth Patriarch said, "It is not the wind that moves, it is not the flag that moves; it is your mind that moves." The two monks were awe-struck. (Sekida 1996)

This story is more receptive to intellectual interpretation than most koans. I understand it as saying that we each create our own realities — that all "things" are in the eye (or mind) of the beholder. That is, that which we perceive "exists" only because we perceive it to exist. (If a tree falls in the forest, and no one hears it, does it make a noise?)

Most koans, though, are more esoteric:

> When the monks assembled before the midday meal to listen to his lecture, the great Hōgen of Seiryō pointed at the bamboo blinds. Two monks simultaneously went and rolled them up. Hōgen said, "One gain, one loss." (Sekida 1996)

And some are just completely nonsensical:

> Nansen Oshō saw monks of the Eastern and Western halls quarreling over a cat. He held up the cat and said, "If you can give an answer, you will save the cat. If not, I will kill it." No one could answer, and Nansen cut the cat in two.
>
> That evening Jōshū returned, and Nansen told him of the incident. Jōshū took off his sandal, placed it on his head, and walked out. "If you had been there, you would have saved the cat," Nansen remarked. (Sekida 1996)

By presenting these stories here, in an academic document written in English by a Californian of European ancestry, I am taking these koans far from their native contexts. Anyone reading these koans here will interpret them through a lens not informed by a lifetime

Early Intake, a.k.a. Rattlesnake Gulch, a remote swimming hole on the Tuolumne River just below Hetch Hetchy.

of Buddhist upbringing and years of monastic training. By which I mean, it is easy to argue that these stories are nonsensical, because they don't make sense to me (and likely also to you).

Nonetheless, we can find the koans instructive, and even humorous. Jōshū's response to Nansen's story is downright comedic, although perhaps not intentionally so, and thereby illustrates the engagement value of the "awkward moment." Jōshū's action can be deconstructed and analyzed for clues to the nature of reality, but not before it catches us off guard, surprising us with its apparent absurdity. The sandal-on-the-head routine seems an inappropriate non sequitur following the serious dilemma of a halved feline. We have been primed with a dramatic conundrum, and are expecting a clean, morally acceptable resolution, but one is not offered. Koans, by design, offer our minds no such easy out; the tension of the unresolved awkward moment keeps us intrigued.

Comedy (another field in which I have no expertise, yet cite freely here) also relies heavily on the "awkward moment." First, a situation is defined in which no obvious outcome could resolve the tension in a satisfactory way without violating one or more social norms. Second, the situation is diffused, modified, or corrected, by way of a punch line. In between those two steps lies comedy's potential for engagement. In that moment when our minds are so uncomfortable and unable to predict what could happen next, we focus completely on the scene before us. As the ancient koan authors certainly knew, satori can be induced by humor.

Little did Jōshū know that he would be succeeded hundreds of years later by an industry of comedians and television programs that vie for our attention (and laughs) by making those moments increasingly awkward. *The Office*, *Curb Your Enthusiasm*, and *It's Always Sunny in Philadelphia* make us as uncomfortable as possible, albeit in the name of ratings, not enlightenment. I find some episodes of *The Office* physically painful to watch. (On subsequent viewings, those episodes are more enjoyable, since I already know how the story is resolved.) What comedy calls the awkward moment, what Zen calls satori, what we experience as the slowing of time and enhancement of focus: these are all the same phenomenon by different names.

My personal experiences of satori have arrived not through mental exercises, however, but intensely physical ones, such as my swim at Rainbow Pools. In those moments, struggling against the current, the immediate future was ambiguous, my physical state uncomfortable, and my attention intensely focused. My satori at the pools was triggered by a heightened state of anxiety, coupled with the fact that I couldn't predict how my current state would unfold into a future state. What would I find on the other side of the falls? Would there be a ledge convenient for me to sit on? Or just a solid rock wall, forcing me to turn back? Would the current pull me under before I could find out? Thankfully, those fleeting moments of intense focus concluded safely. But for me, Rainbow Pools was a koan, challenging my mental state, tricking me into a new level of awareness.

I wonder, as designers, can we deliberately induce this uncomfortable, yet rewarding state? Can interactive, dynamic media engage and challenge people as effectively as koans and comedy?

Introduction

Although the meaning of the term "design" shifts when applied to different disciplines, broadly speaking, it is a process of problem-solving within constraints. A successfully designed umbrella keeps its owner dry. A successfully designed bridge bears the weight of transport moving over it. A successfully designed pacemaker keeps ticking, year in and year out.

Those objects, when well-designed, are more or less invisible. Yet, if the designs fail, they suddenly become highly visible. Works of graphic design function in just the opposite way: visibility is key to their success. They must draw attention to themselves in order to communicate and serve their intended function.

Interactive design, a relatively young amalgamation of several fields — graphic design, film, video, sound design, computer science, and human-computer interaction — must also draw attention to itself in order to succeed. But visibility is not enough; participants must be *engaged* and motivated to interact, to communicate back to the system. Without engagement, there is no interaction between human and system, the communication flows in only one direction, and the design remains isolated, unexperienced. If the audience is not engaged, the design is a failure.

Engagement, therefore, is essential to an interactive project's success. The degree, duration, and depth of engagement required depends upon the project's goals. For an online advertisement, a fleeting moment of engagement may be all that's needed in order to register one brand impression. In that case, the message is brief, task-oriented, and not meaningful ("Click me!"). I am interested in creating meaningful and memorable experiences, the digital equivalents of a coastal hike on a beautiful day, or a great party with your closest friends. These experiences stay with us as happy memories that are kept alive through periodic retellings: "Remember that time when…?" Designing and facilitating such significant experiences requires deep, sustained levels of engagement with participants.

Traditional media have many tools for engagement available to them, including motion, audio, and captivating narrative elements. But computationally based media offer opportunities for new levels of engagement through interaction. In this context, the "design" encompasses not just a system of communication, but also a *system of interaction*. Just as the structural engineer can design a bridge, the interaction designer can consciously and deliberately design *systems* that successfully engage participants.

Fortunately, unlike bridge-builders, we have the luxury of working mostly with software, not steel, which allows us a great deal of flexibility in our methodology. My approach is to first define the problem, then sketch out possible solutions, and test them with potential users as early and often as possible. Testing provides me with data — both raw usage data and anecdotal data about the user's personal experience with the project — which then informs subsequent refinements to the design.

My thesis projects, *Practice* and *Cheeky*, are the results of one year of intensive design research — plus two additional years of graduate study, and a lifetime of fascination with dynamic media. To appreciate where I am now, I must first explain where I came from.

Personal History

At age four, I was given my first interactive, dynamic, digital object: a Speak 'n' Spell. The clunky orange box pronounced words that I then learned to spell on its keypad. This is my earliest memory of a device with true interactivity: it would prompt the user, accept input, and then respond with feedback.

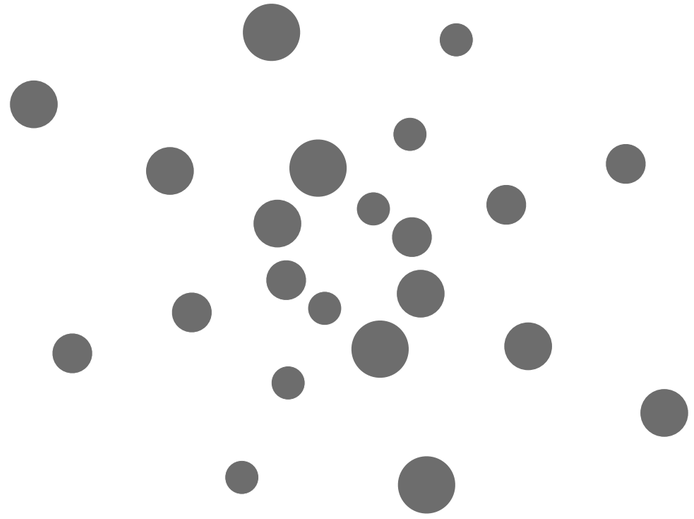
At age five, I faced a monumental ethical dilemma. I wanted to attend a class on LOGO, the early computer graphics programming language, but it was for ages six and up only. At my parents' suggestion, I lied about my age to the friendly staff at the Junior Museum in Palo Alto. This deceit caused me a great deal of distress, but perhaps the anxiety motivated me to learn as much as possible before being found out and removed from the classroom.

```
FORWARD 100
RIGHT 90
PENUP
FORWARD 100
PENDOWN
```

So began one of my first computer programs, although I didn't realize it was programming at the time. All I knew was that I could tell the computer what to do, and it would do it. Also, it would do it correctly, every time. If a mistake was made, it was my own, and I could correct it. My simple circles, triangles, and squares weren't beautiful, but they existed because I had learned to speak the computer's own language. There was something attractive and intriguing about that. Maybe part of it was feeling like an insider, knowing a secret language that only I, the computer, and my classmates spoke. Part of it, too, was satisfaction from solving the puzzle — figuring out how to write a program that would execute my vision. But it was also about simply providing instructions and watching the machine carry them out. The more complex the instructions, the more complex the final image. Using recursion (a word I didn't know at the time), I could even repeat parts of the instructions, building up patterns with spirograph-like complexity.

That same year, my family bought its first computer. Our Macintosh 512KE was very friendly to regular users, but not to aspiring programmers. My neighbor's Commodore 64 was the opposite — just switch it on, and begin entering BASIC commands at the prompt. I remember returning home, excited to try BASIC on my own computer, only to be sorely disappointed by the Mac's non-responsive and indifferent blinking question mark.

A few years later, I discovered HyperCard, an application that made it easy to create buttons, text fields, and pictures, and write code so that behaviors could be attached to those elements. I played with HyperCard for hours. I made *Cipher*, a program that could encrypt or decrypt text using a basic substitution cipher. (This was perhaps my first project to incorporate audio. When encryption was complete, which could take minutes, a prerecorded sample would play: "*Sa-sa-sa-sa-cipheeeerrr...*") In sixth grade, I made an interactive, choose-your-own-adventure style adaptation of the book *Killing Mr. Griffin*,



in lieu of a traditional, linear, written book report. Ms. Dolan didn't totally "get it," but she understood enough to know that this was a new form of creative expression. (Many creative liberties were taken, as the hidden cave and massively explosive finale don't actually occur in the book.) Even though it didn't resonate with her, she encouraged me to pursue similar projects, and she still gave me an A.

Following that, I created *Adventure*, a non-linear exploration through an underground cave, where clicking on doors and passageways led to other screens, all hand-drawn with the mouse, each one containing some hidden message or scary sound. HyperCard's built-in transitions (basic dissolves, slides, and wipes) gave the piece a cinematic quality, and I could even animate transitions — making a bookcase swivel or flame flicker — albeit slowly. Some people loved it, while many couldn't understand it — my first taste of interactive design's power to engage others (or fail to do so).

Around that time, I acquired my first modem. I had found a magazine on Bulletin Board Systems at the local newsstand, loaded with pages of phone numbers that could be dialed *by computer* — what a novel idea! Within fifteen minutes of opening the box, I had connected the modem to the phone line, issued my first AT commands, and signed on to a BBS — my first experience with networked computing. In the following weeks, I tested every BBS in the 303 area code, and my frequent requests of "Don't pick up the phone!" reminded my family that I was "online," that word's first entry into our daily lives. After the next phone bill arrived, I was told to scale back my time online, since I had racked up a hundred dollars in so-called local-long distance calls. I waited until late at night, when the calls were cheaper, and explored the worlds of FidoNet, Usenet, text filez, and many other things I can barely recall.

My middle school had a computer lab full of Apple IIEs, preloaded with BASIC, just like the old Commodore 64. For some reason, Ms. Danzer, the physical education instructor who made us run laps around the field, was also the computer and typing instructor. Her degree of disappointment with me on the track reversed as soon as we came inside to the lab, where she was impressed with my quick grasp of procedural instructions, logic, and variables. While my classmates poked around the Oregon Trail and searched for Carmen Sandiego, I made ASCII art and interactive knock-knock jokes.

Also that year, I completed and mailed a form to the Boulder Valley School District. They wrote back a couple of weeks later to say that my first email address was ready: scmurray@bvsd.k12.co.edu. I dialed into BVSD's server with my modem and typed "mail" at the prompt — my first Unix command. This definitely was not a Macintosh, but a hardcore *mainframe*, just calling out to be explored. I experimented, though gently, as I didn't want to break anything and have my account revoked. I wrote emails to my only other friend with email about how cool it was that we could write messages to each other without stamps or interference from nosy teachers.

In eighth grade, I somehow begged a hundred dollars to buy *Morph*, a software package that mimicked the visual effects made famous in Michael Jackson's "Black or White" music video a year earlier. A friend and I set to work, using his new LC II's video capture card and Adobe Premiere 1.0. I crafted a sequence of our friends smiling and laughing as they morphed into one another. We used the clip in our History Day submission, an 8-minute video on the history of communication technologies, rendered entirely on the computer. Our piece made it to the state-level competition, where it won only fourth place, so we didn't get to take our groundbreaking video to nationals in Washington, DC. That was how I learned that non-geeks will never appreciate the technology alone — the whole experience has to win them over. Although our project was cutting-edge technologically, it occupied only 320 by 240 pixels in the center of the TV screen. The judges complained that it was too small, and thus hard to see, disregarding my explanation that making a full-screen video would have required twice as much expensive RAM and 48 hours to render, instead of only 24.

I made my first website in 1993 at the age of fifteen. My high school had received a grant of $100,000 from Pfizer, which covered the cost of about fifty Macintosh LC IIIs, an Apple QuickTake digital camera, and our own T-1 connection to the Internet, which was still

An early career dream: designing letterforms.

unknown to most people. To thank Pfizer and demonstrate what we did with all that money, my teacher asked me to put together a small website. I taught myself HTML and filled a floppy disk with several inter-linked pages of student work and photographs of our fancy computers with color displays.

At college, I worked at Vassar's computer center, recovering theses from near-death experiences on demagnetized floppy disks. I continued making websites, both for myself, as a creative outlet, and for student groups. Upon my graduation, the alumni association hired me to revamp and expand their tiny site. This was my first professional, paid experience in design, and I did a decent job, considering my inexperience. I remember acknowledging that the redesign had to welcome and accommodate a specific *audience*, including older alumni with limited web experience. This was in sharp contrast to the work I had done until then, which was purely for my own satisfaction.

On September 11 of that year I built an online check-in system, so New York-based alumni could reassure others that they were safe. Since phone lines were down, people could post notes like "I spoke to so-and-so this afternoon, and she made it out of midtown safely..." We emailed the link to about 20,000 alumni. The server nearly ground to a halt, but that project probably had a greater emotional impact on its audience than anything else I've worked on. Alumni wrote and called us in tears, overflowing with gratitude.

Six years and three jobs later, I had learned a great deal about web standards and usability research, but I left work every day with a headache, bored in a position that was about ten percent creativity and ninety percent bureaucratic struggle. So I applied to graduate programs in design. My focus was on graphic design programs (I had a fantasy of being a type designer, which has since passed), but I knew that technological expertise was my primary asset. I couldn't speak the language of letterpress, but hoped that Illustrator and InDesign would get me in the door.

Thankfully, I was encouraged to look into MassArt by a friend, and the language on the DMI website made perfect sense to me, in contrast with other art schools' self-important copy. So I applied, and at the interview I could see that we were speaking the same language. But I still thought that DMI was a graphic design program. I didn't understand the distinction between that and *dynamic media* design until halfway through my first semester. Time-based, interactive, non-static, databased, generative, systems... These terms floated through my brain until I suddenly got it: This isn't about posters, books, business cards, or even websites! This is about designing the point of interface between people and their machines. The machine needs a system of rules to follow, and the people need a pleasant, satisfying psychological experience. I began to see how all my past experiences had prepared me for this new direction.

During that first semester, though, I grew restless creating only conceptual designs — I had come to school to get creative and make new things. I wanted them to actually *work!* Keynote was not bad for faking it, but then I heard about Processing, an open-source, free application then emerging from MIT. I downloaded it, bought a book (the best forty dollars I've ever spent), and got cracking.

What a surprise: This was the modern-day LOGO! The triangular turtle had been superseded by far more advanced visual capabilities, but it was essentially the same idea: a programming language for creating visual art, digitally. Processing, unlike LOGO, could also be programmed to accept user input and react in kind. So the process of using LOGO may have been dynamic, but its output was not. Processing, however, could "output" both static images and *dynamic experiences*. The system designer interacts with Processing by writing code, while the participant (e.g., end-user, client, visitor) interacts with the finished, compiled work. As the designer, I get to not only engage in the creative puzzle of translating a concept into computer-speak, but I also get to define the terms of the puzzle itself. I decide which elements and influences will be included in the system, and I then enjoy watching others interact with that system and figure it out.

Methods of Engagement

We now have the tools and technology to create just about any interactive system we can dream of. So how can we make them engaging? Dynamic systems offer three new avenues for engagement: motion, interaction, and content.

Visual motion is a powerful tool for attracting and maintaining attention. Motion is processed pre-attentively, meaning that we perceive motion before we are consciously aware of it (Ware 2004). That's why it can be so hard to tune out the visually busy animations of online advertisements; your eye will jump straight to them, even though, consciously, you are already aware of the source of motion and have decided it's no longer relevant to you. Motion can be used to capture (and re-capture) attention, even against the participant's conscious will.

Motion can also be used to communicate information, such as through data visualizations or interface feedback, within certain limits. Yet our perceptual abilities are limited in how much motion they can comprehend at once (Ware 2004). While large numbers of independently moving visual objects may be beautiful, they cannot be interpreted meaningfully. (Later, I will review my *Gesture Project*, which captivated participants on a visceral level through hundreds of moving shapes.)

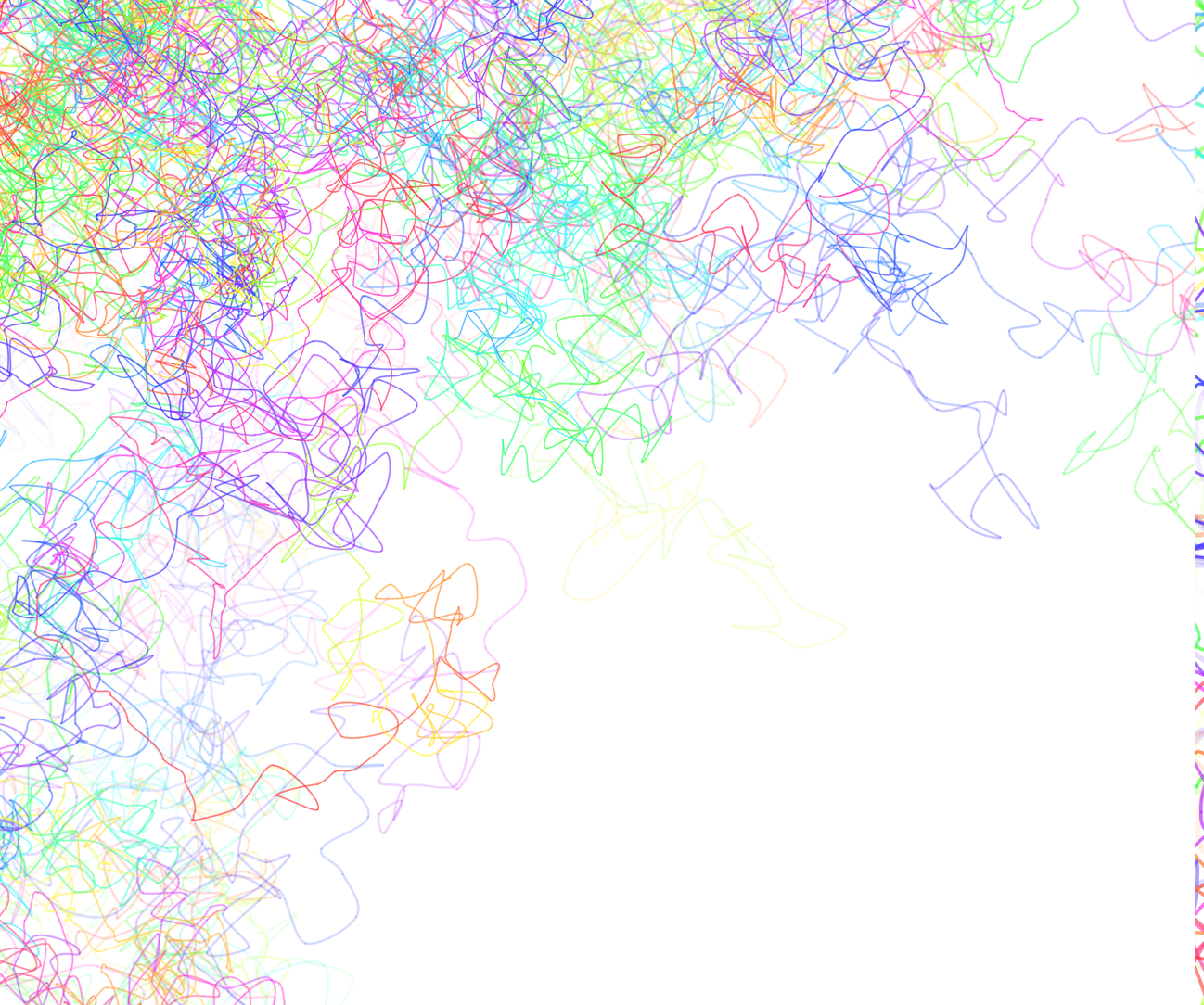
Once users are paying attention, they can be engaged by interactivity. I define interactivity as the two-way feedback loop between system and user. The faster that feedback is provided, the better. A more responsive system is always more engaging.

Dynamic systems can employ any number of interaction methods, but I find one to be most effective: mirroring. Visual mirroring is a medium-specific method as it requires live video input. By capturing an image of the participant, flipping it horizontally, and projecting it back to them, the system acts as a mirror enhanced with computational power. The "mirrored" image, then, can be augmented with additional visuals or replaced entirely with some algorithmic interpretation, as done by the *ASCII Photo Booth*.

According to artist Camille Utterback, a pioneer of interactive video installations and 2009 MacArthur Fellowship awardee, people intuitively understand how to move and interact with mirrored video by virtue of their experience in the physical world (i.e., using real mirrors) (2009). Psychological research supports and expands on that view. English psychoanalyst Donald Winnicott was one of the first in the field to study the powerful role of mirroring in child development. Winnicott summarized his findings in the statement, "When I look I am seen, so I exist" (1967). Mirroring is such a powerful technique for engagement in part because it shows us that we are being seen. Our actions and very existence are validated and proven to be real. I explore the psychology behind mirroring in more depth later in this document.

After motion and interaction, the third new opportunity for engagement is through data as content. Backed by computational power and network connectivity, dynamic systems can draw data from practically anywhere, either the user's mouse movements or a database on the other side of the planet. The potential for engaging users through unexpected, current, time-sensitive content is enormous. And this, more than anything else, is why we must think of these projects in terms of *systems*: Unlike highly composed paintings, sculptures, photographs, and even films, the content of a dynamic system is not fixed and may not be defined in advance. Rather, the designer defines the *scope and structure* of the system's contents. Then, only once the system is executed does it collect, transform, and express that data.

The use of quasi-random content from networked sources is an exciting opportunity, for it practically guarantees that each experience with the system will be unique. But designing a system to handle a limitless variety of content is not possible; some restrictions must be put in place. And creating such a system requires a new design approach.



Projects Overview

Although my thesis process revolved primarily around one major project (*Practice*) and a minor one (*Cheeky*), these projects cannot be understood in a vacuum, as they were preceded by so many other projects. All of my prior projects at DMI have influenced my thesis projects, so I have elected to include descriptions of the most relevant ones here.

Data visualization has been an important theme in my work, and I discuss several projects, ranging from *Search Explorer*, an interactive tool for voyeuristic exploration of web search queries, to *Relationship Visualizer*, a tool for visualizing connections in network data. My *iTunes Library Visualization* was also a valuable learning experience, as was the *BART Trains Visualization*, which taught me that innovative designs are not necessarily successful designs. I briefly explore how data might *sound* — call it data auralization — with two projects, *Audible Particles* and *Aural Data Plot*.

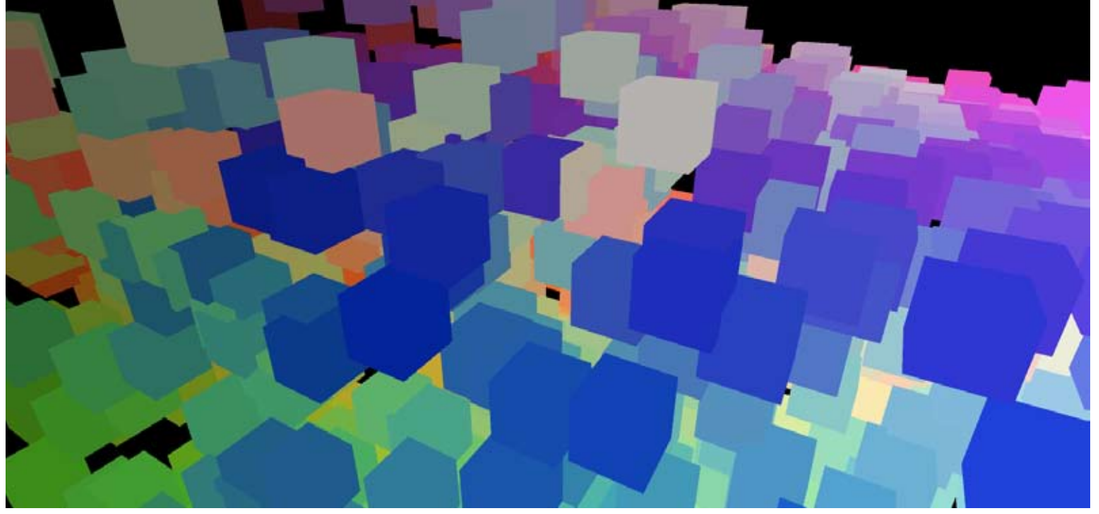
With *Dictionary Words* and *Questions & Answers*, I explore elements of interactive, nonlinear narrative structures, to varying degrees of success. The *Gesture Project* and *ASCII Photo Booth* introduced me to the engagement potential of visual mirroring. *Anticipation Study*, my very first project created with Processing, continues to inform my thinking about how to communicate without words, and how to imply what is not yet seen.

Goals of This Document

In the pages that follow, I document and reflect on my personal process of engagement with dynamic media, concluding with in-depth analyses of my thesis projects. By the end, I hope to have shared the most pertinent and valuable insights and discoveries made during this journey at DMI.

My hope is that this document will be accessible, enjoyable, and valuable for anyone interested in design and interactivity.

Conceptions of Interface and Engagement



I began this year of research with one primary goal in mind: to create one project of significance. It should be interactive, with visual and aural elements, but above all, it must be engaging. Ideally, it would also invoke a rich emotional experience for many users. But engagement is requisite to experience, so I begin by exploring recent thought on user engagement within the context of human-computer interfaces.

Traditional

The term "interface" is loaded with meanings, but it is typically used to describe the collection of elements that mediate between a system and its human operator. Today, that commonly means a set of physical input devices (mouse, keyboard, video camera) and output devices (visual display, speakers, printer). Most of any computer's parts are not accessible to the average user (central processing unit, memory), but are at the core of any computer-based system. In the early days of computing, humans had to go to great cognitive lengths to communicate to these so-called "system internals," via interface methods like physical switches, punch cards, and keyed-in commands. Today, we type, tap, and scribble, but years ago, computer "users" were computer "operators," a term which reflected the skill required to operate the complex and physically intensive interfaces of the day. These early interfaces were highly abstract, requiring hundreds of hours of dedicated learning time to achieve proficiency. I suspect that they were engaging only for the few who had invested the time and energy to understand them.

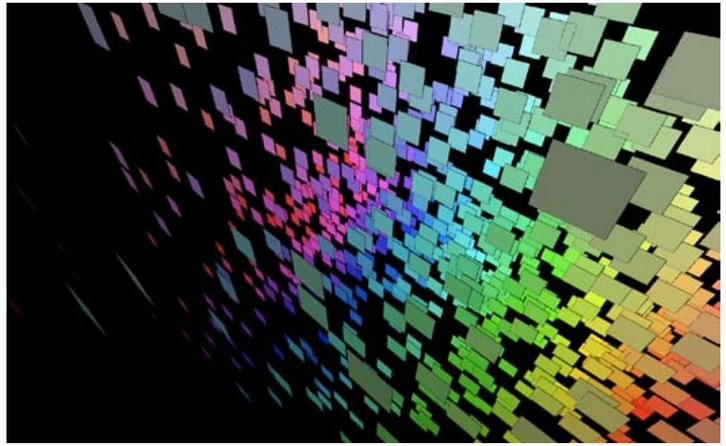
The graphical user interface (GUI) is a relatively recent phenomenon that evolved to address the need for more accessible methods of interaction. Visual imagery is relatively data-intensive, and it was not until the late 1970s and early 1980s that digital memory and computation power was sufficiently advanced to support visual interfaces. As the cost and size of components diminished, interest in the new concept of "personal computers" grew. But if computers were to be adopted by the masses, the interface would have to be far more intuitive.

The Xerox Star introduced the first true GUI in 1981, using the now-familiar metaphors of a "desktop," "files," and "folders," along with the mouse and keyboard input devices. The Star was soon succeeded by Apple's Lisa (1983), the Macintosh (1984 onward), and Microsoft's Windows (1985 onward).

Although there was some variation in each of these systems' visual representations, the fundamental interface structure was the same. By enabling users to "move" visual objects with actual, physical motion through mouse motions like drag-and-drop, these GUIs enabled direct manipulation of abstract, ephemeral data. (We owe the term "direct manipulation" to Ben Shneiderman, a pioneer in the fields of interface design and information visualization.) Consider the difference between (A) sliding a mouse to the right while observing a visual object on-screen move a corresponding amount to the right, and (B) typing in a command that describes the desired action: "move object right." In the latter example, the process of articulating the command burdens the user with both a foreign syntax and yet another layer of conceptual abstraction. The user must first imagine the desired outcome (object should be on the right), determine what needs to be done to achieve the outcome (object is moved), and express that command in syntax that the machine can understand.

As Brenda Laurel, researcher and theorist, explains in her book *Computers as Theatre*, "direct manipulation interfaces employ a psychologist's knowledge of how people relate to objects in the real world in the belief that people can carry that knowledge across to the manipulation of virtual objects that represent computational entities and processes" (1993). Thus, if done successfully, such an interface would be more accessible (more "user-friendly") simply by taking advantage of its users' preexisting cognitive, perceptual and expressive abilities. It may still be necessary to explain that a particular icon represents a "folder," but most people do not need to be told what a folder is. At a certain point, however, GUI metaphors break down. (Real-world "windows" rarely overlap each other, for example, and "menus" do not spawn generations of "sub-menus.") Nonetheless, the traditional graphical interface is greatly more accessible to the general public than its predecessors.

Theatrical



Laurel is dissatisfied with the traditional model of interface — that of merely a layer mediating between human and machine. She puts forward an alternate conception of interface as theatre, which she summarizes this way:

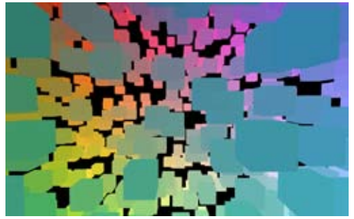
> In a theatrical view of human-computer activity, the stage is a virtual world. It is populated by agents, both human and computer-generated, and other elements of the representational context (windows, teacups, desktops, or what-have-you). The technical magic that supports the representation, as in the theatre, is behind the scenes. Whether the magic is created by hardware, software, or wetware is of no consequence; its only value is in what it produces on the "stage." In other words, *the representation is all there is*… (Laurel 1993, italics original)

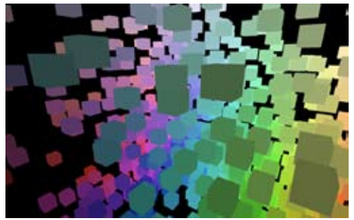
Laurel conceives of interface as a "stage" in which both the computer and its users "perform" together. The interface, then, is not really just a collection of visual representations, but *an active space of collaboration*. The stage is a "virtual world" that doesn't exist in physical space, but only as perceived by human participants. Just as theatre, film, and novels can evoke whole "virtual worlds" that, in our imaginations, expand beyond the physical constraints of the stage, screen, or page, so, too, can computing be enlisted to create intellectually and emotionally immersive experiences. If anything, a computationally based system could produce even more immersive experiences; unlike static compositions (such as films and books), we can become actors or co-creators within this active space called "interface."

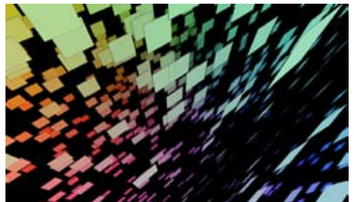
Reconsidering interface in this light, so-called "direct manipulation" interfaces may not be so direct after all. They tend to require physical tools (mouse, keyboard) to function as intermediaries between human action and the computer, so some level of cognitive abstraction is still involved. (Does it really make sense to push the mouse forward on a horizontal plane in order to move the on-screen pointer "up" on a vertical one?) The manipulation in traditional GUIs is not really direct, but indirect. It is only "direct" relative to the highly indirect command-line, punch-card, and manual-switch interfaces that preceded it.

Despite those limitations, researchers (including Donald Norman, formerly of Apple's Human Interface Group) have argued that the process of using direct manipulation interfaces can induce a *feeling of direct engagement* (Hutchins et al. 1986). Attempting to define an emotion is problematic, but direct engagement can be described generally as a positive, satisfying experience. Yet, in observing this phenomenon, the researchers write: "Although we believe this feeling of direct engagement to be of critical importance, in fact, we know little about the actual requirements for producing it" (Hutchins et al. 1986).

I believe that we can answer this enormous question — how can we engage people? — by merging Laurel's conception of interface with psychological research on attention, emotion, and "interface," in the pre-computing sense of the term.







Stills from Color Mapper, a tool for visualizing different color models.

Psychological

I should not write off so-called "direct manipulation" interfaces so quickly. There are scores of such interfaces that are, in fact, highly engaging, and can maintain a user's interest for hours at a time (e.g., computer games, web browsers, email and other messaging programs). But even the simplest of these applications still involves some degree of learning, since each involves abstractions ("indirect manipulation"). Users must explore the interface long enough to develop a mental model of how it works, what the system is doing (or not), what kinds of input it can accept, and what ranges of output it may produce. This exploratory process may take only a few minutes or seconds, but it is always there. We are not born understanding how to surf the web.

What if there were an entirely different model of interface, based on true "direct manipulation" yet without *any* learning curve whatsoever? Such an interface would eliminate conceptual abstractions (interface elements) simply by catering only to known human abilities. This new kind of interface would be intuitive and instantly engaging, since users would "just know" how to interact with it — no new learning required. Of course, the potential applications for such an interface might also be limited to performing actions that we already "know." For example, a highly skilled action such as producing technical schematics may fall outside of the realm of possibility for a strictly intuitive interface. Nonetheless, the idea is still worth exploring.

Camille Utterback uses video "mirroring" in many of her projects. The mirrored image may show the user's likeness, yet also be augmented with additional visuals or algorithmic interpretations of the image. During a recent talk at UC Berkeley, Utterback observed that people intuitively understand how to move and interact with mirrored video, by virtue of their experience in the physical world (2009).

Yet psychological research indicates that our familiarity with mirroring is not learned through life experiences (such as using real mirrors), but is biologically innate. In a paper summarizing the field's research on mirroring and child development, psychologist Malcolm Pines observes that "human infants start reacting to a mirror at 18 weeks and at 42 weeks, over 60 percent of them try to look at the back of a small mirror, as if looking for the other infant, or kiss the mirror but do not seem to recognize themselves. Laughter was reported to be elicited by a mirror image at 17 weeks" (1985). Visual mirroring is a powerful force very early in our lives.

An infant's response to an actual mirror is interesting, but more relevant is the role of interpersonal mirroring, especially that between parent and child. It is well-established

Social

(and easily observed) that parents *reflect* their infant's behavior back to the child.

> By speaking to infants in baby talk with its higher pitch, repetitive phrases, and the free use of nonsense utterances, by the use of exaggerated facial expressions which offer clear patterns to the infant, clear indications of emotional state and offerings of joy, greeting, approval, encouragement, or even of amused disapproval, the parent acts as a benign reflection to the baby of its presence within the context of a social relationship. (Pines 1985)

Called biological mirroring, because it is observed in cultures worldwide, this process is a near-universal means of communicating with a pre-verbal infant. Quite literally, our very first social interactions are of other people mirroring our own actions.

For the infant, parental mirroring sets the groundwork for more complex social relationships, and it is also the first step of individuation, of illustrating to the infant that she is an individual being. Through imitative behavior, the parent acts "as a psychobiological mirror, an active partner in the infant's developing capacity for social relations and the beginning awareness of self-representation" (Pines 1985).

By reflecting the infant's behavior, with minor modifications, parents use the interaction to further engage the child, and to instruct, encouraging positive behaviors and discouraging negative ones. When performing this process of augmented reflection, "the mother's intention seems to be that it helps her to develop and to sustain a more meaningful *dialogue* with her baby and, furthermore, to facilitate the development of a more deliberate imitation by the baby. The mother's answering gesture provides the infant with an interest-holding event which is temporarily contingent upon his own performance of a similar event" (Pines 1985, italics original).

In other words, the parent-infant *interface* is an ongoing feedback loop, a two-way communication between the learner and the learned. Just as with a dynamic, computational system, the output does not directly mirror the input. Rather, the parent takes the infant's actions, interprets them, and performs them back, but slightly modified so that the difference is infused with instruction. This modification of the performance also serves as an "interest-holding event" — a method of sustaining engagement. Augmented mirroring, then, is a means of both engagement and communication, and it is fundamental to the human experience. "Using a mirror," "mirroring," and "reflecting" are not learned skills, but biologically innate and necessary human abilities.

Traditional interface models imagine a layer that mediates between one user and one system. But in the context of interactive video installations, there is potential for multiple concurrent users. It is essential, then, to consider not just a one-to-one, user/system interaction, but also the interaction dynamics between the people themselves.

In her recent lecture, Utterback related her observations on how video installations are always multi-user experiences. Any work using video as the input medium automatically becomes social. Even if intended for only one user, it must be designed to accommodate more, simply due to the practical consideration that multiple people can enter the visual space in front of the camera. Utterback's work takes advantage of the inherently social nature of interactive video art. Many of her systems track multiple participants, and encourage collaborative interactions. For example, *Untitled 5* generates paint-like strokes on-screen in response to user motion. But when those strokes cross paths, they change form, and grow out in new directions. The end result is visual output that cannot be invoked by one user exploring alone.

In a single-user system, the user can focus entirely on his relationship with the system. Having multiple users, though, requires that a social negotiation occur alongside each individual's relationship with the system. Traditional, direct manipulation GUIs are not designed to accommodate more than one user at a time. As a result, attempts to collaborate with multiple people sharing

Participants experiencing Practice in San Francisco.

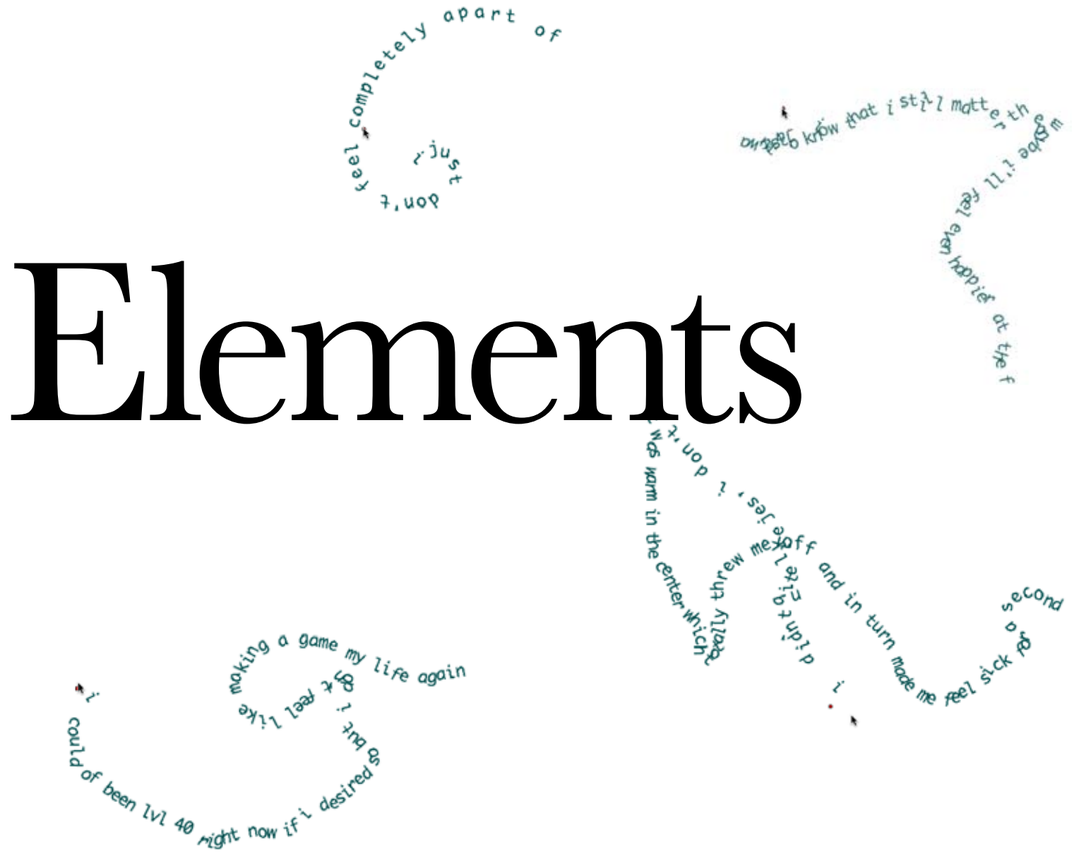
one machine typically requires taking turns — e.g., "Okay, now you take the mouse" — effectively acknowledging the interface's inability to accept input from more than one user. A successful interactive video piece will acknowledge each user's participation and accommodate them gracefully.

Any such installation will occupy a space, just as a mirror occupies a space, whether in a bathroom or a hallway. People within the space (in view of the video camera) become active participants, since their actions influence the output of the system. But what about the people outside the space, looking in? Or the people moving into or out of the space? What do the people in these transitory spaces contribute to the interface?

As people enter and exit the space, they function as observers — third parties not yet interacting with the system, yet able to perceive both the system's output and its active users. These observers are not influencing the system directly, but they may influence the active users either directly (e.g., by verbalizing instructions, suggesting actions) or indirectly (e.g., simply by observing, making the users self-conscious). In that liminal space, the observers are unwitting participants in a larger stage that encompasses the system, its users, and the observers themselves. This adds complexity to the interactive dynamic — it is no longer just an interface between a system and a user (or multiple users). We must also account for additional "interfaces" between the users and the observers. To use database jargon, the relationship is not one-to-one, nor even one-to-many, but many-to-many.

If this larger stage is itself observed, it can be considered a form of unintentional performance art, in which the participants are wholly unaware that they are playing roles scripted for them in advance by the systems designer. Laurel's theatrical interface offers many opportunities for understanding how to consciously craft these roles, since her model considers each individual to be an "actor" on the stage.

In my first year at DMI, I wrote a research paper on flash mobs, a phenomenon derived from the "happenings" of the 1960s and whose creator considered it both performance art and a social experiment. Yet most of the mobs' participants (or "performers") were unaware that they were acting out roles on a predefined stage. Each individual was simply responding to instructions and making decisions informed by the social context, yet that context was very deliberately structured ("designed") in advance. But since these performances occur in public space, they are typically surrounded by and infiltrated with third-party observers, who, in turn, become part of the show. So all the actors have varying levels of awareness and appreciation for the larger system at work, but they are all engaged, whether by choice or social coercion.



Stills from a sketch testing the physics of emotion statements for Practice.

Medium

Reflecting on my past projects, my first and most obvious observation is that they share a common medium: the screen. Nearly all of my projects are screen-based, and use the mouse or keyboard to take input. (A few print-only pieces are exceptions.) I haven't done any work with so-called tangible interfaces (although mice and keyboards are quite tangible and tactile), nor sensor-based input methods or haptic feedback.

This focus on screen-based work was born partly out of convenience, but also from my desire to leave DMI with semi-marketable skills. It's hard enough to explain "dynamic media" to people — why complicate my elevator pitch with distance sensors and hacked webcams? I thought that the familiarity of screen-based work would be more accessible to potential clients. Plus, it's easy to deploy a screen-based application on a large scale via the web — less so for physical objects and custom installations.

In terms of form, my early work was all black-and-white, beginning with my first projects in HyperCard. At DMI, too, I chose to work mostly in black-and-white (and gray) for about six months, while I learned the basics of Processing. Color was an additional dimension that I just couldn't handle on top of my other new learning. But eventually I got there, and studied how to work with color computationally. I now try to make my projects as colorful as possible (when appropriate), simply because colors tend to make the interaction more beautiful, and often more engaging.

These reflections led to an insight that a project's medium or media must be considered in a more structured context. Every dynamic system employs at least three media: one for input, one for output, and one that coordinates the two and executes the rules of the system itself. In a traditional GUI, the mouse and keyboard serve as input media, the display is the primary output, and the system rules are executed by the software's logic on the CPU. (Of course, this is an oversimplification, as networks, storage devices, and graphics cards are also involved.)

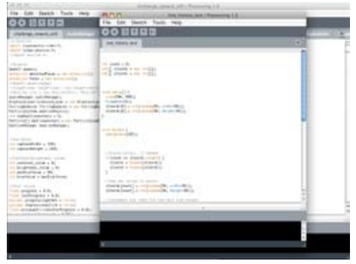
One may argue that these final items — the software and central processing unit, or main chip in the computer — are invisible to the user, and, therefore, do not qualify as media. But I argue that we should consider it as such, for the process of medium selection, more than anything, defines the range of possible inputs and outputs for any system.

Medium selection is critically important. If Picasso had used marble instead of paint, and Rodin opted for pastels over bronze casts, would they have created comparable masterpieces? Yes or no, their work would have been very different from the paintings and sculptures we know today.

The decisions we make about which digital tools to use are just as important. If I will be performing live music that has to be pitch-perfect and on-beat, then Max/MSP or Pure Data are probably my best bets. For fluid, interactive work on the web, Flash, Flex, or any number of JavaScript libraries may fit the bill. When selecting an appropriate medium for our system, we need to consider what will work best with the desired input and output media, as well as our own technical abilities and the timeline for the project.

For my purposes and skill level, Processing has been a great medium for developing and deploying projects. It's a free, open-source programming language for artists and designers. Because it's built on the Java language, anything you create in Processing

Sketching in code: stills from twelve iterations, culminating in the desired line quality.



Processing's sketching environment. The front window contains code that generated the line drawings on the opposite page. Once the line sketch was complete, its code was integrated back into the main program for Practice (rear window).

can be posted to the web as a Java applet, or exported to a native program for Mac, Windows, or Linux. It uses a simplified syntax that makes programming simple visuals relatively easy, and it automatically tracks key interaction data (like mouse position and keystrokes) that are more difficult to manage in other languages. Processing is particularly well-suited for data visualization projects, since it has extensive built-in support for capturing and processing data from a variety of sources.

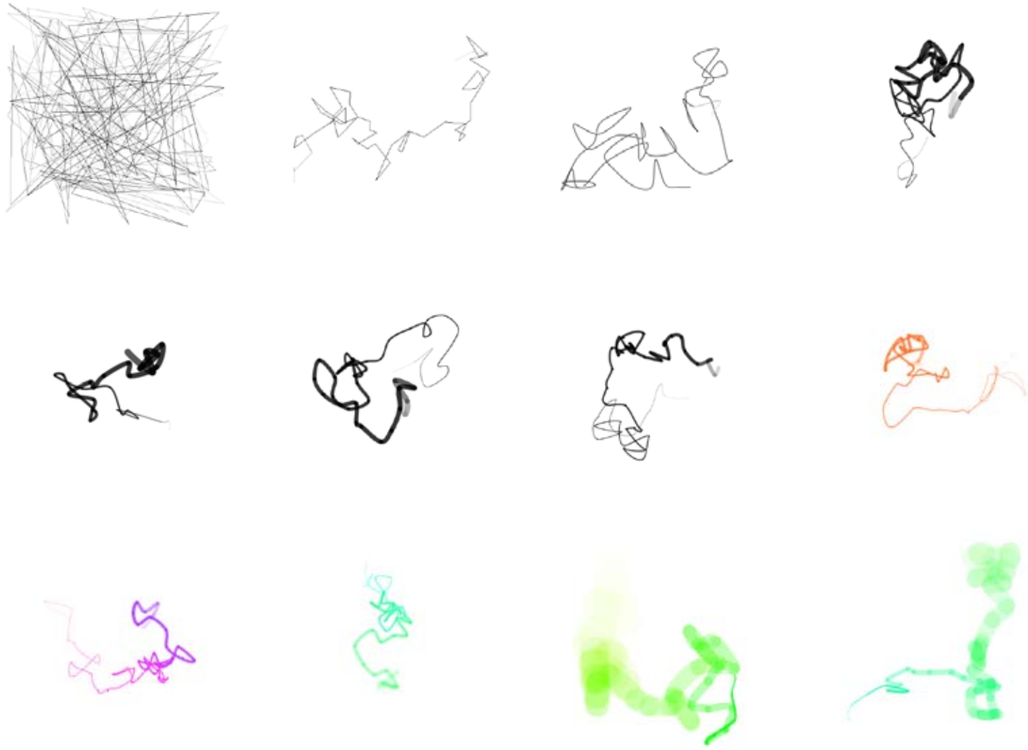
I was introduced to Processing during my first semester at DMI, when I bought the book *Processing* by Ben Fry and Casey Reas, the programs co-creators. Learning to program seemed the only way to employ true interactivity in my projects. So I sat down with the book, and slowly taught myself how to use this new tool. *Anticipation Study* was my first finished project, an interface that positions abstract shapes in response to mouse movements. For me, the learning process was a smooth one, since the Fry and Reas book related many elements back to fundamental concepts of systems design.

While I sometimes hate to admit it, I truly am a "systems guy" and I've always loved systems thinking. My undergraduate major was a self-designed exploration of "environmental theory" — essentially a study of how social and biological forces (like religion and ecology, itself a systems-based science) interact with each other. In my daily life, I have all kinds of systems in place to do everything from washing the dishes to reminding me to pay the rent. When working, I would rather learn a new tool and figure out how to incorporate it into my workflow than perform a task in a less efficient way.

So Processing fits the way I already think. To make a cluster of circles that orbit each other, you simply define what one circle looks like and how it behaves. Then you make ten, a hundred, or thousands of them, and let them interact with each other. That is a systems approach: to watch complexity emerge from simple rules.

Processing is also intended to be a sketching environment — like pencil and paper, but with code. Everything about the interface is designed to help you quickly write code and execute it as quickly and often as possible. Draw a line, done. Make it thicker, done. Make it a curve, colorize it, adjust the transparency, and so on. Sketching with code in Processing is analogous to using paper — you start rough, and then gradually hone in on your final image through numerous miniscule refinements.

Upon entering my thesis year, I continued to use Processing for two reasons. First, because of my growing familiarity with its language. Second, because I had an extensive amount of code written that I could re-use which would speed my development time for future projects. Spending less time on coding meant I would have more time for conceptual work, user research, and design refinements.



Data

I've long understood data intuitively: what "data" are, and how it (or they) can be captured, stored, and manipulated. But once I started using Processing, I had to dig deeper and build some structure around that vague intuitive sense. Programming requires explicitness. Computers are truly binary: they perceive only *this* or *that*, and nothing else. It is only because of the tremendous speed with which they move between *this* and *that* that complexity arises, and thus, the perceived magic of computation. At their root, computers are very simple machines with only rudimentary abilities to receive instructions.

To program — to write in the computer's language — one has to think on the computer's own terms, especially where data are concerned. A number must be identified as a specific *kind of* number (whole number or decimal point). A chunk of text must be identified as a specific *kind of* text (one single character or a sequence of many characters). The format and structure of the data must be defined before the data themselves are even known.

My relationship with data has evolved over the last few years as I've learned how to consider and document it explicitly. The process has called for more structure in my own thinking when planning projects. I've gravitated toward this process and gained a great deal from it, although at times I worry it has kept my process and work *too* structured. There is certainly value in letting go and creating more intuitively, but it's difficult to work with data that way. The content constrains the process.

Lev Manovich has argued that transcoding and transformation are both fundamental phenomena in new media (2002). After being reduced to ones and zeroes (the lowest common denominators for all binary data), any data can be adapted to any purpose. An uptick in the stock market can translate into a taller line on a chart, and mouse coordinates can be used to position a shape. To the machine, it doesn't matter where the data comes from — the market, the mouse, a video camera, or a song. They are all just numbers, and all numbers can be mathematically manipulated. My colleague Jason Bailey's recent visualization projects exploit this phenomenon in its most abstract form. Jason has taken a Miles Davis recording, captured the literal numeric values from the digital recording, and then interpreted those values in visual form. So the result is indeed a data visualization, although its data source is so far abstracted from our aural experience of the music that there is no meaningful, human-perceivable mapping. To me, the unexpected beauty that results only adds to its value.

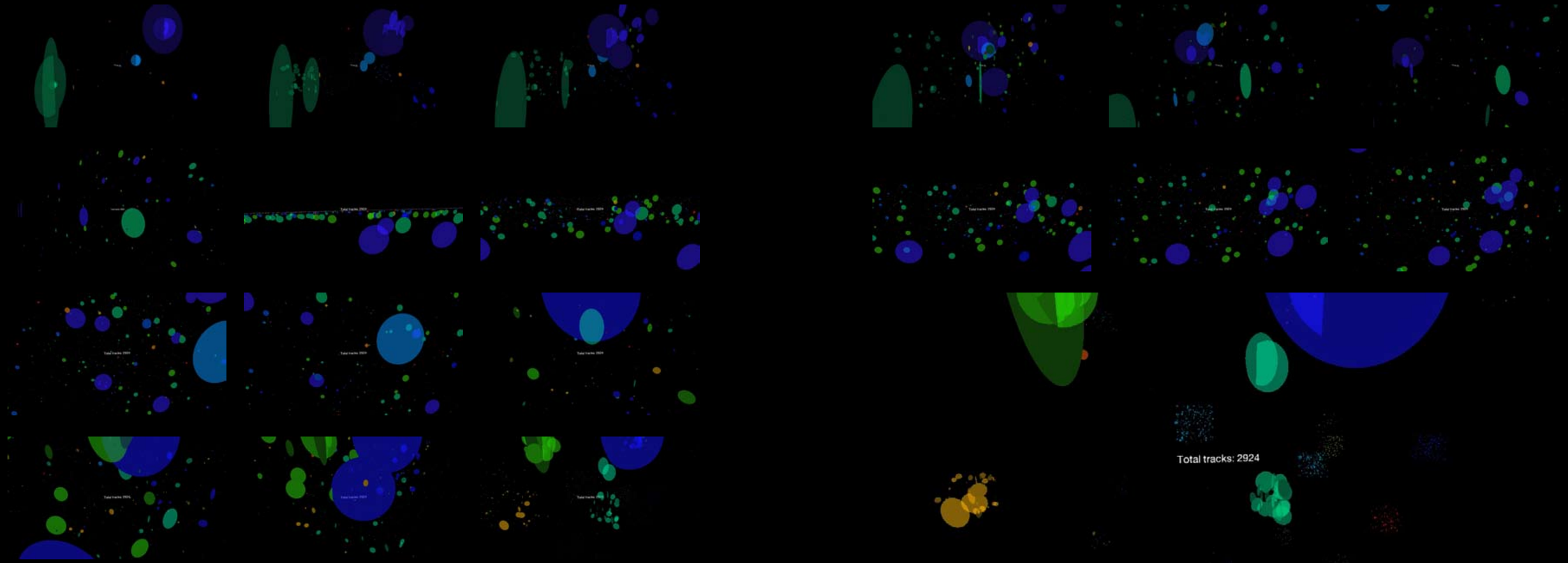
Yet this abstract mapping raises questions about human perception (What kinds of visualization are meaningful?) and the philosophical divide between art and design (Is this useful, or beautiful, or both?).

The process of giving visual form to data necessarily merges aesthetics with statistics, two traditions that rarely see eye to eye. So what constitutes a meaningful representation of data, and what is an artistic interpretation? When is each appropriate, and how can we achieve both at the same time?

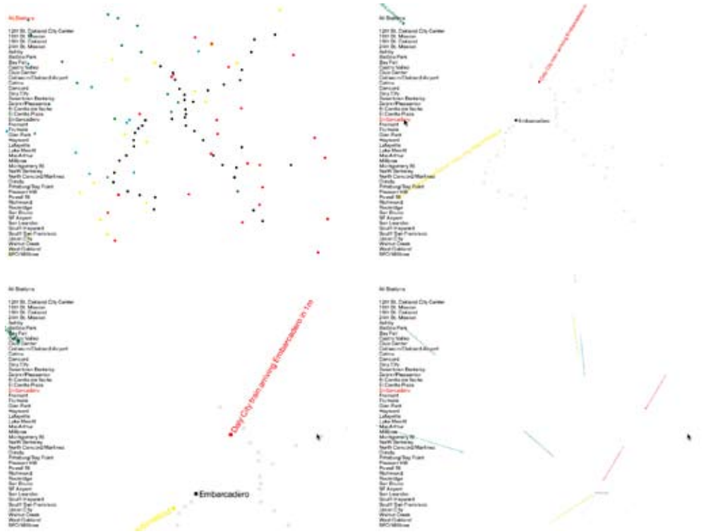
I spent the fall 2008 semester pursuing these questions at UC Berkeley in a course with Maneesh Agrawala, who recently received a MacArthur "genius grant" fellowship for his groundbreaking work in the field. I had already developed a few visualization projects of my own, including the *iTunes Library Visualization*, which represented songs on my computer as circles in 3D space. Longer songs were larger circles, and they could be grouped and sorted in space by different attributes like genre or number of times played. The iTunes project generated great eye candy, but I quickly realized that cool visuals are not necessarily meaningful visuals. At Berkeley, my plan was to learn what makes a successful visualization.

In the process, I built two interactive projects for that course. First, the *BART Trains Visualization* represented Bay Area subway trains' arrival times on a map. The data was pulled from BART's website, so the visualization was real-time. As an experiment, this project was invaluable, although my visual approach was not perceptually sound.

Stills from the iTunes Library Visualization project.


Total tracks: 2924

Stills from the BART Trains Visualization project.

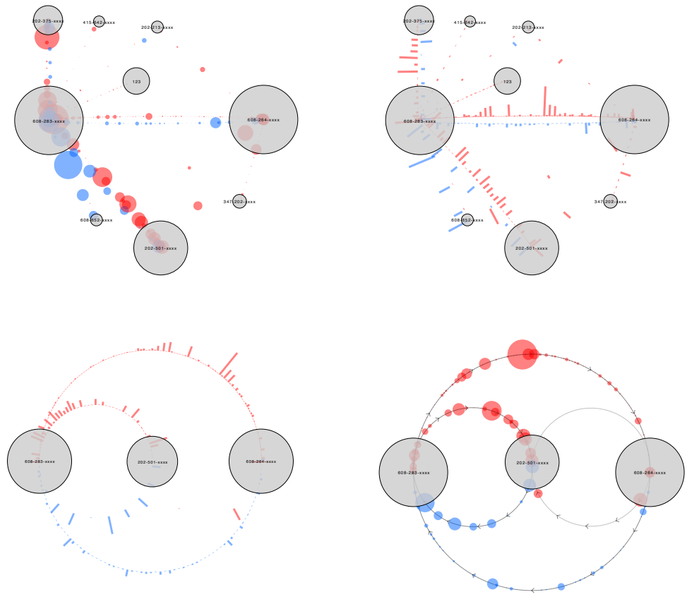


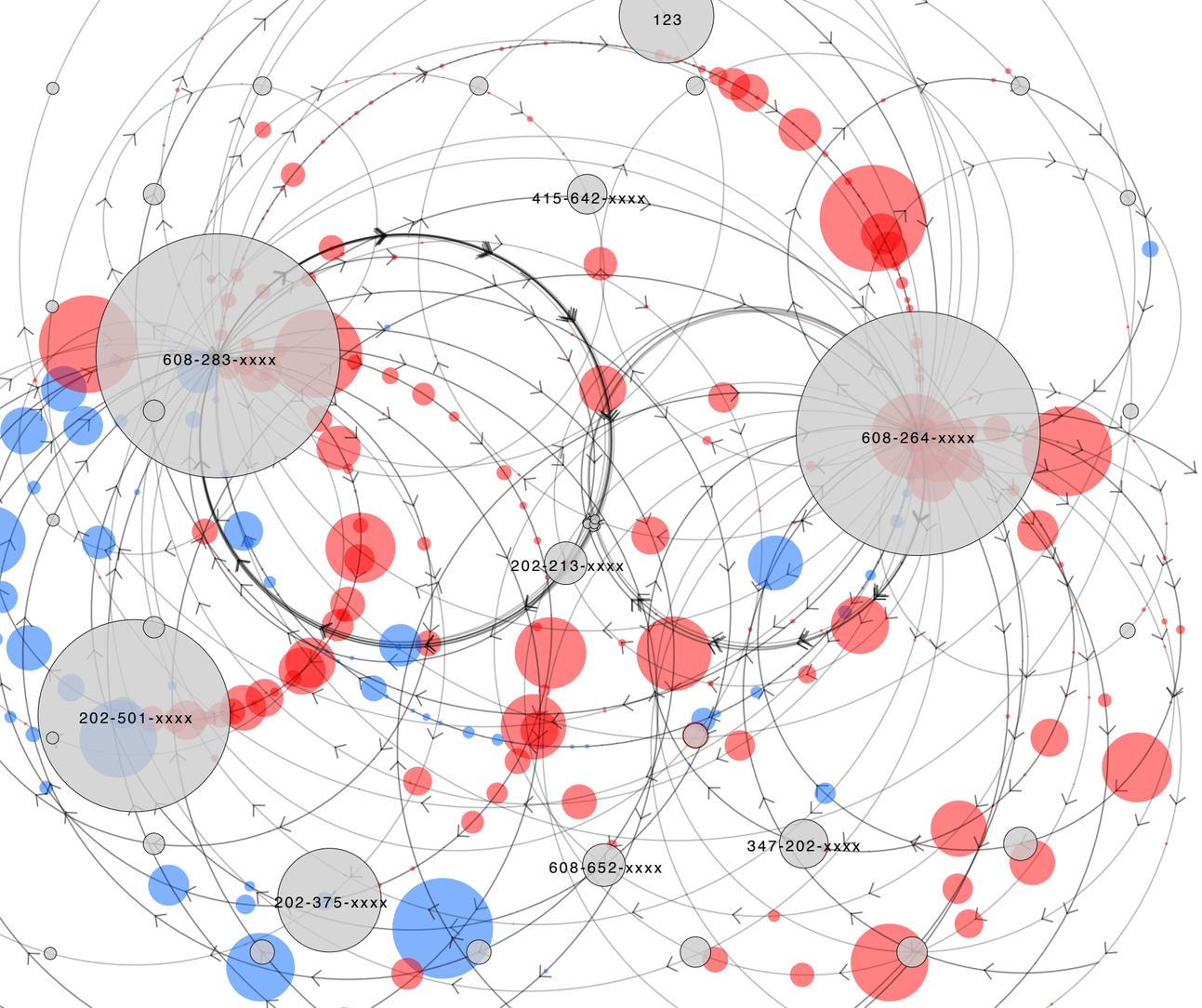
Users were confused about why the trains, represented by color-coded circles, would appear to be "off the tracks," displaying in the middle of the bay, for example, or as far south as San Jose. The reason was that I made a conscious decision that trains should not follow the track alignments, but instead display simply a certain distance away from the target station. Since all the trains were in motion at a constant speed, the distance between any train and its station represented the amount of time until it would arrive. My concept was that, perceptually, users could perceive the distance and the rate of motion, and from that deduce a sense of how much time they had to run to the subway station. (The minutes and seconds remaining were also shown on-screen.) In any case, this approach was neither visually clear nor perceptually sound, but it was a great technical achievement for me, and a valuable lesson came with it: innovative designs are not necessarily successful designs.

My final project for the course, *Relationship Visualizer*, was a visualization of networked relationships, for which I used my own telephone records. This project showed not just connections between "nodes" (phone numbers, in this case), but the frequency, direction, and duration of those connections (actual telephone conversations). Professor Agrawala was impressed by the originality of my concept and design, but warned me to be careful with using so much motion. We had spent a few weeks of class on human visual perception, and while humans are great at identifying the presence of motion, we're not able to track the motion of more than a few visual objects at once. Therefore, a screen full of moving objects is, perceptually, a big mess — visually stimulating and beautiful, perhaps, but not meaningful (or, not a visualization *from which* we are able to extract meaningful connections about the data). Following that feedback, I kept the existing motion, but added a view with arrows to indicate directionality, so the visualization could be interpreted clearly whether in motion or not. This project is one of the most-visited on my website, which I think testifies to the need for tools to help us make sense of networked information.

During his summer 2009 course on sound design, Colin Owens and I started talking about how to pair data visualization with audio. How could sound be used to reinforce and augment visual representations of data? I immediately imagined a swarm of data points floating around in space, and position-based audio that would indicate where the points were in space, like the hum of a swarm of bees. I created exactly that with *Audible Particles*, a

Different views of the sample data set from Relationship Visualizer.

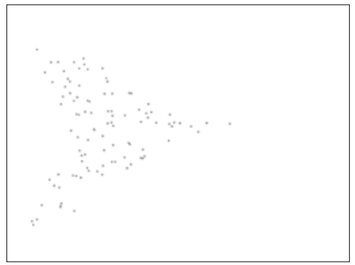


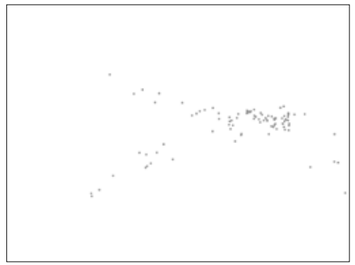


Pushing the limits of human perception: too many shapes and too much motion results in an overstimulating, illegible visualization.

```
608-283-XXXX,123,1
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,978-476-XXXX,2
608-283-XXXX,608-264-XXXX,2
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,202-375-XXXX,10
608-283-XXXX,202-501-XXXX,46
608-283-XXXX,608-264-XXXX,8
608-283-XXXX,608-264-XXXX,2
608-264-XXXX,608-283-XXXX,26
608-283-XXXX,608-264-XXXX,2
608-264-XXXX,608-283-XXXX,1
608-283-XXXX,202-501-XXXX,20
202-375-XXXX,608-283-XXXX,1
608-283-XXXX,608-264-XXXX,2
608-264-XXXX,608-283-XXXX,1
608-283-XXXX,202-375-XXXX,5
608-264-XXXX,608-283-XXXX,1
661-478-XXXX,608-283-XXXX,2
608-283-XXXX,661-478-XXXX,1
608-283-XXXX,661-478-XXXX,3
608-283-XXXX,415-642-XXXX,2
608-283-XXXX,123,2
608-283-XXXX,415-642-XXXX,2
608-283-XXXX,415-642-XXXX,1
608-283-XXXX,415-642-XXXX,1
608-283-XXXX,202-501-XXXX,1
608-283-XXXX,123,1
608-283-XXXX,415-876-XXXX,1
608-283-XXXX,800-266-XXXX,3
608-283-XXXX,888-327-XXXX,1
202-501-XXXX,608-283-XXXX,45
608-283-XXXX,800-592-XXXX,15
608-283-XXXX,800-266-XXXX,15
608-283-XXXX,301-639-XXXX,25
608-283-XXXX,608-264-XXXX,10
608-283-XXXX,202-766-XXXX,1
608-283-XXXX,202-375-XXXX,5
608-283-XXXX,202-375-XXXX,3
608-283-XXXX,202-501-XXXX,1
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,608-264-XXXX,2
202-375-XXXX,608-283-XXXX,9
608-283-XXXX,202-501-XXXX,1
608-283-XXXX,202-501-XXXX,23
608-264-XXXX,608-283-XXXX,6
608-283-XXXX,415-354-XXXX,1
608-283-XXXX,415-354-XXXX,6
608-283-XXXX,415-571-XXXX,31
608-283-XXXX,415-505-XXXX,3
608-283-XXXX,608-264-XXXX,9
608-283-XXXX,608-264-XXXX,1
608-264-XXXX,608-283-XXXX,1
608-283-XXXX,202-375-XXXX,1
608-283-XXXX,415-492-XXXX,2
608-283-XXXX,510-628-XXXX,2
608-283-XXXX,415-354-XXXX,3
608-283-XXXX,608-264-XXXX,3
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,608-264-XXXX,1
608-264-XXXX,608-283-XXXX,1
608-283-XXXX,202-501-XXXX,9
608-283-XXXX,123,3
608-283-XXXX,608-264-XXXX,1
202-375-XXXX,608-283-XXXX,26
202-501-XXXX,608-283-XXXX,1
608-283-XXXX,202-501-XXXX,21
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,608-264-XXXX,2
510-628-XXXX,608-283-XXXX,3
608-283-XXXX,608-264-XXXX,2
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,608-264-XXXX,1
608-264-XXXX,608-283-XXXX,1
608-283-XXXX,123,1
608-283-XXXX,415-571-XXXX,10
608-283-XXXX,123,2
608-283-XXXX,415-354-XXXX,3
608-283-XXXX,202-501-XXXX,1
608-283-XXXX,952-220-XXXX,17
608-283-XXXX,123,1
608-283-XXXX,415-642-XXXX,2
608-283-XXXX,202-501-XXXX,1
608-264-XXXX,608-283-XXXX,1
608-283-XXXX,202-501-XXXX,1
608-283-XXXX,202-375-XXXX,47
608-264-XXXX,608-283-XXXX,5
608-283-XXXX,202-501-XXXX,33
608-283-XXXX,952-220-XXXX,1
608-283-XXXX,202-501-XXXX,31
952-220-XXXX,608-283-XXXX,9
608-283-XXXX,202-375-XXXX,8
Blocked NBR,608-283-XXXX,64
608-283-XXXX,123,2
608-283-XXXX,202-501-XXXX,11
608-283-XXXX,415-263-XXXX,1
608-283-XXXX,123,2
608-283-XXXX,608-264-XXXX,3
608-283-XXXX,202-375-XXXX,43
608-283-XXXX,202-501-XXXX,1
608-283-XXXX,608-652-XXXX,1
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,608-264-XXXX,16
608-283-XXXX,415-497-XXXX,2
608-283-XXXX,415-497-XXXX,1
608-283-XXXX,202-501-XXXX,6
202-501-XXXX,608-283-XXXX,31
608-283-XXXX,123,2
608-283-XXXX,209-743-XXXX,1
608-283-XXXX,415-263-XXXX,14
202-375-XXXX,608-283-XXXX,3
608-283-XXXX,608-264-XXXX,1
608-264-XXXX,608-283-XXXX,1
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,202-501-XXXX,25
202-375-XXXX,608-283-XXXX,12
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,123,1
608-283-XXXX,202-213-XXXX,4
608-283-XXXX,202-213-XXXX,2
608-283-XXXX,619-817-XXXX,17
608-283-XXXX,415-868-XXXX,2
608-283-XXXX,202-213-XXXX,2
608-283-XXXX,608-264-XXXX,3
202-213-XXXX,608-283-XXXX,11
608-283-XXXX,608-264-XXXX,1
608-264-XXXX,608-283-XXXX,1
608-652-XXXX,608-283-XXXX,66
202-213-XXXX,608-283-XXXX,3
608-264-XXXX,608-283-XXXX,3
202-501-XXXX,608-283-XXXX,30
608-283-XXXX,608-264-XXXX,1
608-283-XXXX,415-566-XXXX,3
608-283-XXXX,781-569-XXXX,4
608-283-XXXX,123,2
608-283-XXXX,202-213-XXXX,2
608-283-XXXX,202-501-XXXX,4
608-283-XXXX,608-264-XXXX,2
608-283-XXXX,123,1
608-283-XXXX,415-441-XXXX,3
608-283-XXXX,202-501-XXXX,30
608-283-XXXX,123,2
608-283-XXXX,202-766-XXXX,28
```

A portion of the telephone call data used for Relationship Visualizer. Each line represents one phone call or connection, including the originating number, the destination number, and the duration of the call in minutes. This simple structure enables the tool to work well with a wide range of data sets.





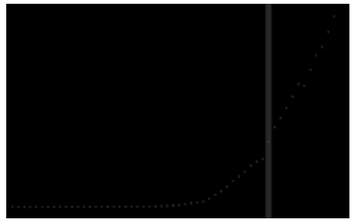
Stills from Audible Particles.

simple sketch with particles that are attracted to and orbit the mouse. A corresponding background hiss follows the particles from left to right, so if there are more on the right, the hiss is louder on the right, and so on.

Next, Colin challenged me to take it a step further and consider how to present data to a blind user. If *visual*ization was not an option, what about an *aural*ization? I created the *Aural Data Plot*, which takes x/y values and plays through them, like musical notes. It's essentially just the audible version of a simple bar chart, with left/right panning indicating the x position, and pitch indicating the y value. Together, these projects started me thinking about expressing data through many output channels, visual and otherwise.

With each of these projects, my design process has evolved. I remember my first review session at DMI, during which a visual mapping I had designed for my "you are here" project was criticized as not meaningful. I took that to heart, and realized that I needed to consider not only the visual elements of my choice, but all the visual elements that others would perceive. Position is important, but so are color, shape, size, relationship, motion — everything. With each project, I get better at forecasting what my users will perceive and adapting the design accordingly.

Aural Data Plot's optional visual representation of the data points being played.



A selection of the web search data from AOL.

AnonID	Query	QueryTime	ItemRank	ClickURL
507	ebay electronics	2006-03-05 11:26:37	104	http://electronics.listi
507	ghost hunting equipment	2006-03-05 16:14:29	1	http://www.paranormality
507	ghost hunting equipment	2006-03-05 16:14:29	3	http://www.prairieghosts
507	ghost hunting equipment	2006-03-05 16:14:29	5	http://ghostinvestigator
507	ebay	2006-03-29 00:10:30	4	http://www.motors.ebay.c
507	vintage auto wiring harness	2006-03-29 10:33:43	2	http://www.car-stuff.com
507	mapquest	2006-05-05 15:56:07		
2708	revenge tactics	2006-03-01 07:33:28	1	http://www.ekran.no
2708	revenge tactics	2006-03-01 07:33:28	2	http://www.linkbase.org
2708	revenge tactics	2006-03-01 07:33:28	3	http://www.totse.com
2708	marblehead massachusetts library	2006-03-01 17:35:48	4	http://www.marblehead.co
2708	the woman's book of revenge	2006-03-01 17:39:48	3	http://www.revengelady.c
2708	how to torment someone	2006-03-01 17:57:05	8	http://www.webmd.com
2708	the woman's book of revenge	2006-03-01 18:00:09	8	http://www.wpr.org
2708	the woman's book of revenge	2006-03-01 18:00:09	10	http://home.earthlink.ne
2708	marblehead massachusetts library	2006-03-01 18:00:16		
2708	alt.revenge	2006-03-01 18:03:20	1	http://www.student.uit.n
2708	alt.revenge	2006-03-01 18:03:20	4	http://www.ekran.no
2708	alt.revenge	2006-03-01 18:03:20	7	http://www.harley.com
2708	alt.revenge	2006-03-01 18:03:20	9	http://www.xent.com
2708	the woman's book of revenge	2006-03-01 18:11:23	1	http://www.amazon.com
2708	dirty tricks for chicks	2006-03-01 18:14:01	7	http://www.shopping.com
2708	out of business	2006-03-01 18:15:06		
2708	out of business by dennis fieriy	2006-03-01 18:15:24		
2708	21st century revenge	2006-03-01 18:20:39	4	http://www.loompanics.co
2708	victor santoro's samples	2006-03-01 18:22:11		
2708	stories or samples of revenge	2006-03-01 18:22:38		
2708	encyclopedia of revenge	2006-03-01 18:25:40	2	http://en.wikipedia.org
2708	encyclopedia of revenge	2006-03-01 18:25:40	3	http://www.paladin-press
2708	voice changer	2006-03-01 18:42:01	4	http://www.spyequipmentg
2708	voice changer	2006-03-01 18:42:01	7	http://www.pimall.com
2708	help for revenge	2006-03-01 20:55:55	1	http://www.armageddongam
2708	help for revenge	2006-03-01 20:55:55	2	http://www.rotteneggs.co
2708	mass lottery	2006-03-02 07:10:15	2	http://www.superpages.co
2708	bravin home services	2006-03-02 18:27:54	1	http://www.servicemagic.
2708	how to humiliate someone	2006-03-03 07:47:13	1	http://www.rateitall.com
2708	cd's to order and buy	2006-03-03 16:47:54	4	http://www.musichristian
2708	12 cd's for the price of one	2006-03-03 16:49:00		
2708	bill me pay later for cd's	2006-03-03 16:50:44	9	http://www.shopping.com
2708	bill me pay later for cd's	2006-03-03 16:52:36	20	http://forums.dealofday.
2708	scams to play on people	2006-03-03 16:53:19	3	http://itim.tamu.edu
2708	playing tricks on someone	2006-03-03 16:55:22		
2708	how humiliate someone	2006-03-03 16:56:05	4	http://www.iron-rose.com
2708	how humiliate someone	2006-03-03 17:03:47		
2708	how humiliate someone	2006-03-03 17:04:11	10	http://www.sysopt.com
2708	how to make someone misreable	2006-03-03 17:05:12	2	http://www.sketchyorigin
2708	how to make someone misreable	2006-03-03 17:05:12	5	http://www.ojar.com
2708	how to make someone misreable	2006-03-03 17:05:12	10	http://www.depnet.ae
2708	how to drive someone crazy	2006-03-03 17:09:27	1	http://www.23nlpeople.co
2708	how to drive someone crazy	2006-03-03 17:09:27	2	http://www.answers.com
2708	how to drive someone crazy	2006-03-03 17:14:46		
2708	how to get revenge on an old lover	2006-03-03 17:15:12	1	http://www.urdumped.co.u
2708	how to get revenge on an old lover	2006-03-03 17:15:12	9	http://www.moviecall.org
2708	how to get revenge on an old lover	2006-03-03 17:19:49	20	http://www.sfgate.com
2708	i hate my ex boyfriend	2006-03-03 17:20:49	2	http://www.angry.net
2708	how to really make someone hurt for the pain they caused to someone else	2006-03-03 17:22:50 6		
2708	names of dogs	2006-03-04 20:24:18	4	http://www.petplace.com
2708	advice on how to get revenge on an old lover	2006-03-04 21:00:497		http://www.def
2708	advice from women who have seeked revenge on old lovers	2006-03-04 21:02:42 5		http
2708	download men's voices	2006-03-05 13:17:24		
2708	funny sounds and voices	2006-03-05 13:32:32	1	http://www.stonewashed.n
2708	makehimsweat.com	2006-03-05 14:22:41		
2708	makehimsuffer.com	2006-03-05 14:22:56		

Search Explorer

In August of 2006, AOL released three months' worth of search queries entered on its website. The data was released to the academic community for research purposes. Although the records were "anonymized" in the sense that email addresses and usernames were omitted, researchers quickly discovered that they could identify the real people behind each randomly generated user ID number, simply by reading the content of the queries themselves. So it turns out that we can be identified (and perhaps defined, to a degree) by what we search for, since our searches expose our values and desires.

Although a handful of websites offer limited access to the AOL search data (aolstalker.com, aolpsycho.com, aolsearchdatabase.com), for the most part it remains inaccessible to the general public, archived in massive text files. I wanted to take this fascinating content and reveal it in an interactive, visual format. I wanted to create a tool that would encourage free-form exploration of these queries and help people visualize meaningful patterns across them. From the beginning, I was trying to create a deliberately voyeuristic experience.

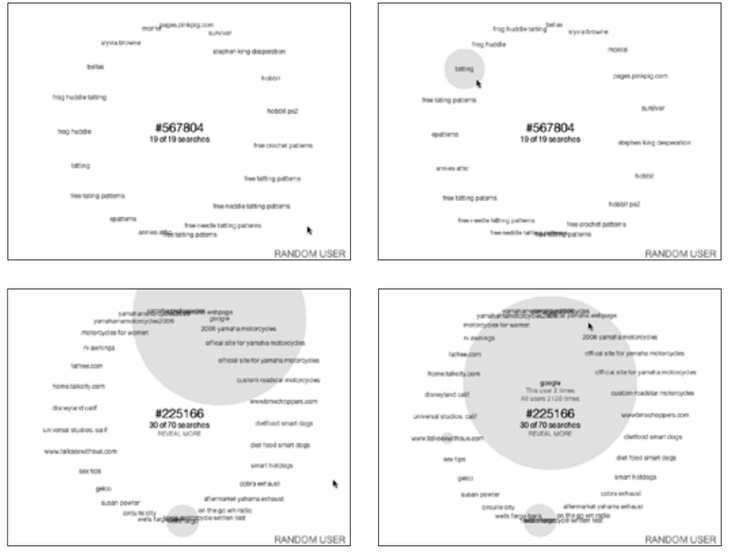
That was the dream, anyway, but the project fell short. Due to all kinds of technical complications I could not have foreseen, I

didn't make as much progress as I had expected. Yet, I learned a great deal in the process of trying, both in terms of designing the interface as well as developing its back-end functionality.

The initial problem was the data set itself. Until this point, the largest data set I had worked with was a dictionary of English words, a few thousand lines long. The AOL data set, though, was 21,011,340 lines long — one line for each search query — for a total of 2.2 gigabytes of data. Plus, each line was not just a word or two, but contained multiple data points: the user ID number and the query text, plus the selected search result (i.e., what URL the user clicked on *after* making the search) and date and time of the query. There were 10,154,742 unique queries, and 19,442,629 user click-throughs. The sheer volume of data presented an issue of scale: How could I store and access so much information?

I set up a MySQL database, and then tried to import the AOL data. And it took *forever.* Or at least it felt like forever. I timed the import process, and the database was taking in about 5 megabytes per minute. At that rate, it would take 7.5 hours to import the whole thing. At that point, I had to question the scale of my ambition. Maybe I could illustrate the concept without using all 21 million queries. So I limited my database to only the first 15 megabytes of data, a mere 0.6% of the whole, with only 251,032 queries from 1,582 different users.

Even with "only" a quarter-million queries, I found that my interface was painfully slow. When clicking the "new user" button, for example, it could take the system about 10 seconds to locate all of that user's records. Then, in order to visualize how popular certain search terms were across *all* users, the system would count how many times each individual query appeared in the database. ("Sex" is popular, of course, and so is "google.") That second round of database



Stills from Search Explorer.

work could take a full minute to complete. In the meantime, my interface was just frozen, apparently crashed, and totally unusable.

It's not a pleasant experience when an interface stops responding for anything longer than about 100 milliseconds. When that happens, it's time to introduce threading, a programming technique that separates chunks of code into different processes. So, for this project, I learned how to compartmentalize all of the database operations into their own "threads," so the UI could continue, smooth as silk. (Think of all the Flash "loading..." indicators you've seen on web pages. The "loading..." animation is one thread, while all the hard work is happening in another thread.) This was my first multi-threaded interface, and, although the technical side wasn't too complex, it required me to think differently about designing the interface. What should be happening in the UI while the background processes complete? How is progress communicated to the user?

Assuming I could get all the technical bits working, how should I visualize this content anyway? The biggest problem with this project was that this important question was nearly left until the end, since I let the technical issues occupy so much of my time. I experimented and then settled on an interface with the user ID number at center, surrounded by the associated queries moving around in a gentle orbit. Although this didn't convey any additional information about the queries (such as date and time of occurrence), I liked what it represented: that the identity of the individual was somehow collectively made up of his or her questions and desires.

I wanted to raise questions about identity and privacy in a time when we routinely enter very personal information into ostensibly impersonal search engines. Although the project wasn't completely successful, I'm still interested in revealing personal information that we expect will always remain hidden. Since then, I've started working on a project that listens to network traffic, extracting any passwords that happen to pass by. It will be interesting to make that invisible information visible for others to see.

In the meantime, we can learn a lot from *Search Explorer* about working with large data sets. In order to keep the interface responsive, I had to limit the scope of the content being used. A friend whose occupation is maximizing efficiency in enormous data sets advised that, while some streamlining could be done, it would require more technical experience than I have, and would not bring that much benefit. At a certain point, technical limitations force design changes. So when embarking on a data-heavy project, we must always keep the structure and size of the data in mind.

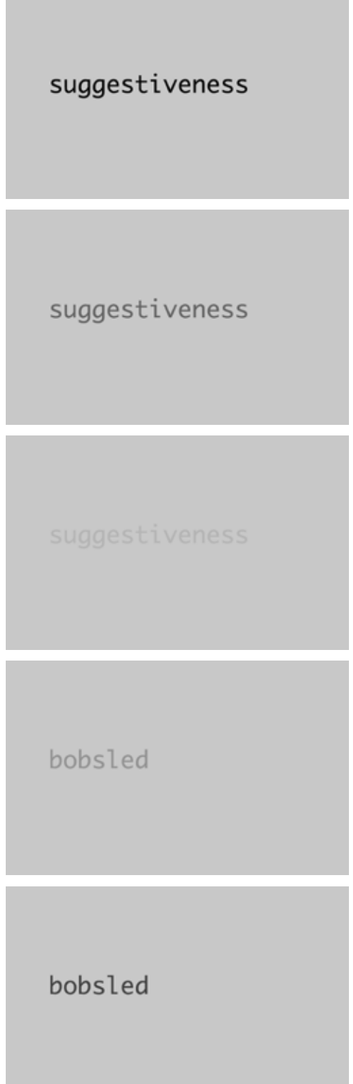
Systems

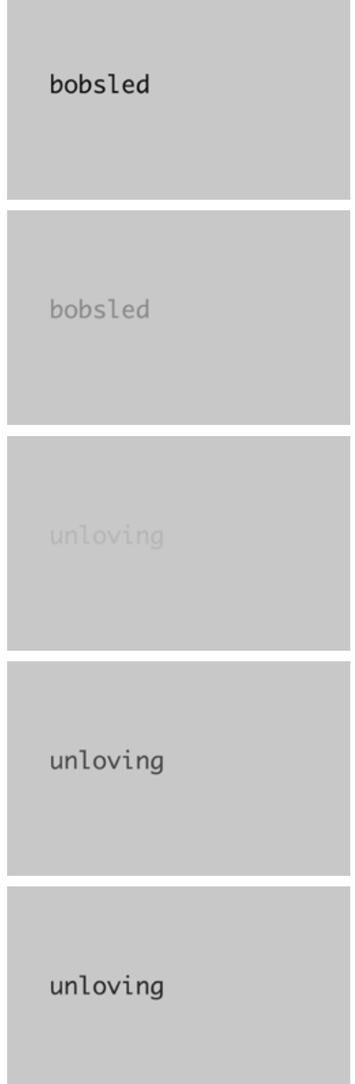
In my art history class at Vassar, we once visited the art museum, where there was an installation by Sol LeWitt. Only, he hadn't done the installation himself — he had "merely" provided instructions to the museum, which Vassar students then executed. The result was an array of crisscrossing strings tied to nails, placed at semi-random points along a wall. At the time, this was presented to us as a great work of conceptual art (which it was), but today I think of it in terms of design systems. LeWitt had composed a set of rules, thereby defining the scope and parameters of the piece. Students had simply carried out his instructions (and, as a result, only LeWitt's name appeared on the placard adjacent to the work). On a different day, another group of students, faculty, or really anyone could have executed LeWitt's instructions, with a similar result. Although the work wasn't visually appealing to me, the approach made perfect sense. The students were just his medium of execution, the computer carrying out the designer's instructions.

I now consider myself a "systems designer." Part of why I was drawn to Processing in the first place was that I wanted to write my own rules, and watch the computer carry them out *in cooperation with the user.* The rules are fixed, but they leave room for user input, so the end expression or output of the system is different every time, within certain constraints, depending on who uses it and how they use it.

Casey Reas, one of the co-authors of Processing, gave a brief talk at the start of a workshop I taught. In it, he discussed his process, and explained how all of his works start as rules written in plain English, like LeWitt's piece. But Reas asks the computer, not undergraduates, to execute his instructions (Reas 2009). The main advantage of this, obviously, is speed. Iterations that would have taken years to execute by hand can finish in seconds. With this speed emerges complexity. Reas's visuals appear very organic and natural despite the simple rules and logic underpinning them.

Both Reas's and LeWitt's works, however do not change in response to their observers. LeWitt's pieces are static. Reas's are, for the most part, dynamic, but not interactive. My challenge is to design interactive systems that are accessible, meaningful, and engaging. Anything programmed is a system, but not every system is interesting or valuable. When designing interactions, we must always consider the ongoing feedback loop between users and system. User input is captured, processed, and responded to with output. That output, in turn, initiates a response from the user, and so on. Of course, these events do not occur sequentially, with each party taking turns; the interactions, hopefully, are much more fluid than that, with each party actively influencing, co-creating, and collaborating with the other, moment to moment.





Stills from the initial version of Dictionary Words.

Narrative

During my fourth semester at DMI, I stopped obsessing about data visualization and immersed myself in narrative. I had suffered too many blank stares when sharing my visualization projects with others — I had learned that visualization can engage people with eye candy, but in order to hold their interest, there has to be a *story*, and that story has to be meaningful to them as individuals. Narrative was my ticket to making engaging projects that would live on in people's imaginations even after the experience had ended.

I studied "non-linear narrative," and realized I'd created such a thing long ago: my early experiments in HyperCard were choose-your-own-adventure-style narratives, albeit simple ones. Clicking on different parts of the screen would perform actions and move you to other parts of the imagined world. It was like the early CD-ROM game Myst, but with hand-drawn, black-and-white graphics.

Dictionary Words

I also remembered my dictionary experiment, from my second semester at DMI. It was a gray screen, with one randomly selected word shown at center. The word would fade out to gray, only to be replaced by a new word, which would fade in. It was simple, yet hypnotic. People responded to it and searched for meaningful associations between the words.

I expanded on that project by adding a second word. Now two randomly selected words would appear and disappear at once (such as "nourish opera" or "telecommuting aphid"). This version was an even greater hit. It was fun to watch the unexpected, often silly word combinations. Of course, the word pairings weren't intrinsically humorous, but we *perceived* them that way. This project gave me a taste of the power of our ability to create meaning out of random data elements.

Around this time, I was introduced to the writings of film theorist Edward Branigan, who posits that narrative is more process than product. Branigan defines narrative as "a fundamental way of organizing data" (1992). Seen in this way, narrative no longer represents a static composition, but

> a perceptual activity that organizes data into a special pattern which represents and explains experience… Thus the word "narrative" may refer to either the product of storytelling/comprehending or to its process of construction. (1992)

This view of narrative as process is essential to understanding how data-driven, non-linear, non-"composed" works like *Dictionary Words* can be understood to have successful narrative elements. Users may or may not perceive patterns when being shown empirically random stimuli. So the ideal data-driven work would be consciously designed to *ensure* that a narrative thread is perceived, *even when the data behind it are essentially random*. The primary challenge of the design process is to avoid presenting what Branigan calls a "catalogue" — a collection of meaningless, unconnected chunks of data.

Stills from the revised Dictionary Words.

This agrees with philosopher John Dewey's view, when he writes that "to perceive, a beholder must *create* his own experience" (1979). The creator may assemble the work with specific narrative goals in mind, but those goals may or may not be fulfilled by the users, whose interpretation of the piece is informed by their prior experiences, background knowledge, cultural contexts, and predispositions. Narrative and meaning, like beauty, are in the eye and mind of the beholder.

Research on visual perception by Fritz Heider and Mary-Ann Simmel helps to validate these claims, illustrating how people interpret ostensibly non-narrative elements (e.g., unnamed circles, triangles, and lines "moving" around each other) as fully narrativized characters, complete with motivations, personalities, triumphs and tragedies (1944). Later research by psychologist Albert Michotte examined visual perception's role in perceived or *attributed* causality, in which people interpret events as being related and causal (or not), based purely on visual attributes of size and motion changing over time (1962).

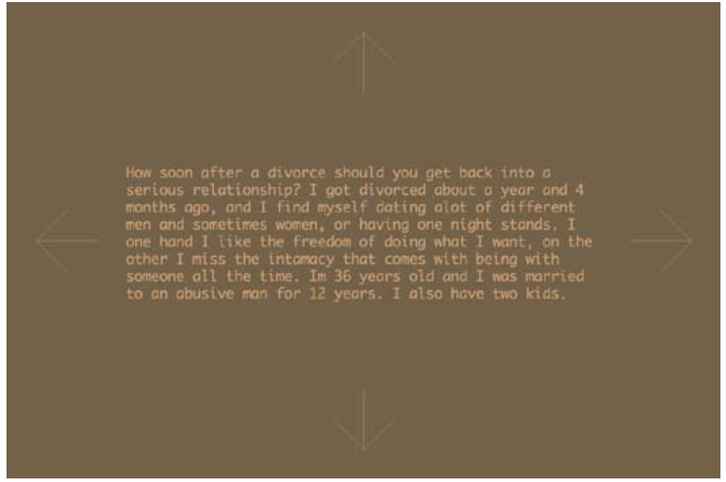
Questions & Answers

I finally had a narrative framework that was applicable to dynamic systems. So I began another project designed to exploit it.

Asking a question reveals information about oneself. In the process of asking, we give clues to our interests, personalities, and concerns. What do we want to know, and why? What can be inferred about our identities from our questions alone? How we respond to others' questions, too, is revealing.

Questions & Answers is designed to encourage exploration of the vast range of questions that people ask (and answer) online. Question and answer data from Yahoo! Answers and colors from ColourLovers are presented in a fluid, minimalist interface, allowing the user to focus on the content and pursue the subjects that interest him or her.

Only the arrow keys are used to navigate (there is no mouse input), and content is arranged in a virtual space with questions above and answers below. By "moving" left, right, up, and down, users can explore the content in any order, whether sequentially — question, answer, question, answer — or not. For example, users could ignore the questions and review only answers (which are often interesting or entertaining without the context of the original question).



By going up, the user accesses the top-level menu, which lists all of Yahoo! Answers' main subject categories, plus two additional options: *Random*, which retrieves content across all categories, and *Search*. Below the *Search* option is a space where text may be entered for a custom query. After text has been entered, moving down again reveals the content found for that query.

The interface is designed to "get out of the way" of the content. Content is shown front and center, and the only other interface elements (just four arrows and one cursor) are presented as unobtrusively as possible. The typeface Monaco was selected for its plain appearance and its strong, even strokes, which make it very legible on-screen. As a monospaced typeface, Monaco connotes early computer terminal interfaces and raw, undecorated data. Very little processing is done to "clean up" the content, no styles (bold, italic, etc.) are used, and the original author's capitalization, punctuation, and phrasing are all preserved.

Color is used to provide texture and offer a more visually engaging experience. The text of each Yahoo! Answers category (e.g., "Business & Finance") is used to retrieve the most relevant colors from ColourLovers, an online community of people who name and rate colors. Each category is then assigned a palette of colors, and navigating between categories cycles the background and foreground colors through the appropriate palette. Although most users won't be aware of this back-end functionality, the palettes help give each topic category a distinct feel when, for example, the "Business" section uses steely blues and "Environment" reveals shades of green. Also, since it would be impossible to predict which color values are returned from ColourLovers, an algorithm is used to ensure enough visual contrast between the background and foreground, so that text is always legible.

The intent with this project was to present inherently engaging content in a way that would encourage exploration and facilitate the narrative-creation process in users. Anecdotally, the reception was mixed. Some people enjoyed it (especially when using the search function), but overall, *Questions & Answers* was nowhere near as consistently engaging as the far simpler, more reductionist *Dictionary Words*. Although Branigan advises against presenting a mere "catalogue" of data, it seems, in this case, that the smaller and more random the chunks of data, the more the piece is open to interpretation.

Stills from Questions & Answers.

Engagement

If a viewer/user/participant is not engaged with a project, then they won't experience it. If a design doesn't draw them in, then the design isn't working. With each of my thesis projects, I was determined to offer participants a true *experience*, in Dewey's sense of the term — a memorable interaction with definitive closure that participants would reflect on long after its conclusion. Dewey distinguishes between experiences generally and the creation of "*an* experience," meaning one that

> has a unity that gives it its name, *that* meal, that storm, that rupture of friendship. The existence of this unity is constituted by a single quality that pervades the entire experience in spite of the variation of its constituent parts. (1979, italics original)

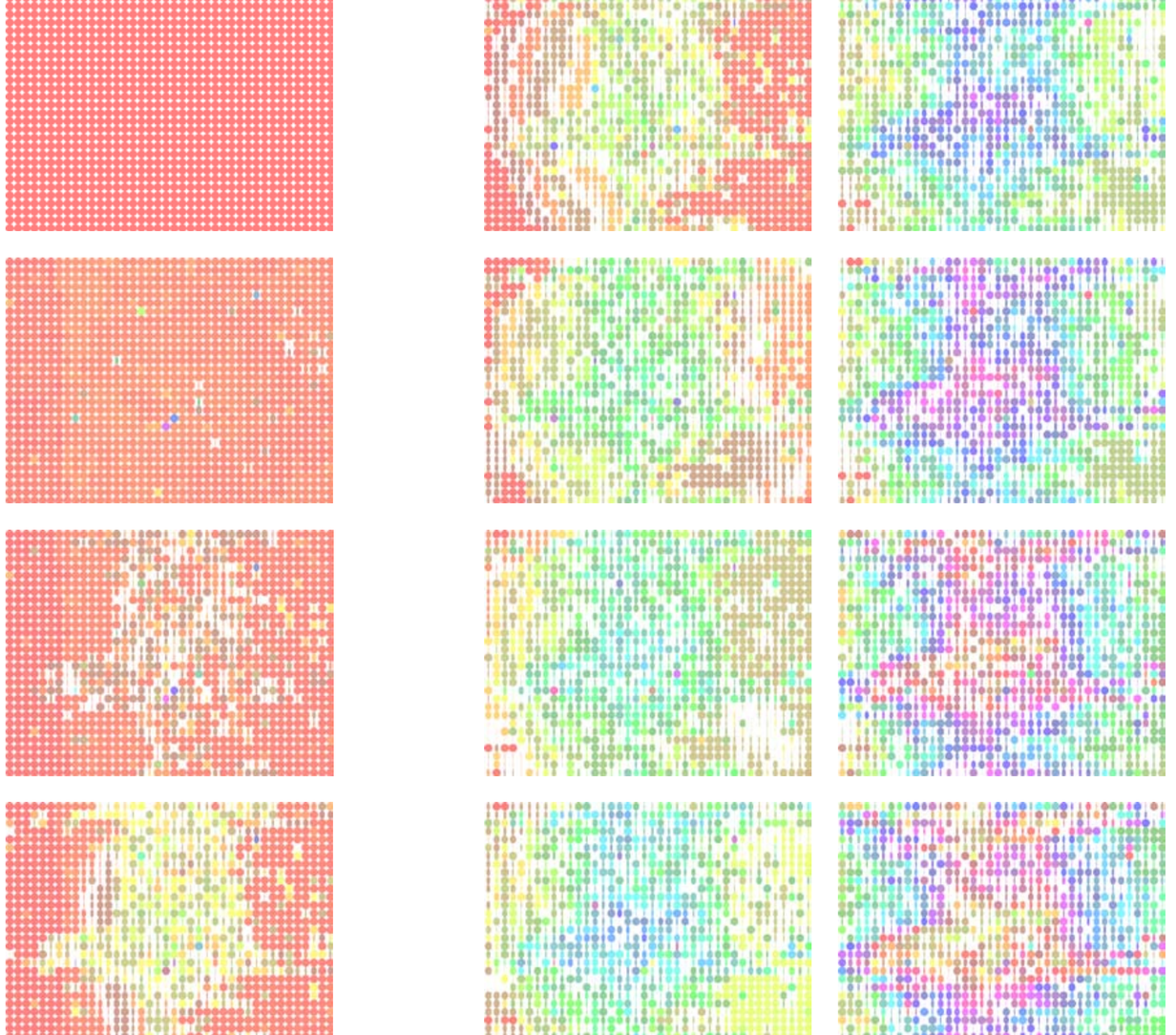
If dynamic work can maintain one meaningfully unifying quality, despite being driven by potentially random, non-meaningful data, it may constitute an experience, in Dewey's sense of the term, meaning that the experiencer found meaning within it. But in order to stand out as an experience, it also needs closure, for

> …we have *an* experience when the material experienced runs its course to fulfillment… A piece of work is finished in a way that is satisfactory; a problem receives its solution; a game is played through; a situation… is so rounded out that its close is a consummation and not a cessation. (1979, italics original)

So, another challenge of meaningful engagement is to design methods for closure of the experience, even when the input data are random and therefore unpredictable.

Many of my past data visualization projects used interesting content, but the presentation wasn't engaging. I had tried to pull people in with a narrative hook, but that had proven extraordinarily difficult to execute well. Reflecting on which projects *had* reeled people in — and that, anecdotally, qualified as memorable experiences — I recalled my *Gesture Project* and *ASCII Photo Booth*. The *Gesture Project* captured the user's motion and mirrored it with spinning, colorful discs, while the *Photo Booth* reflected the users back on themselves, and even printed a hard copy — physical evidence that the experience had occurred and completed. Both of these projects used video as the primary input, and in both cases, I was struck by how entranced people were with watching themselves on-screen. Thus began my interest in mirroring as a means of engagement.

I decided that my thesis projects, too, would use some form of mirroring, since it so quickly establishes a connection between the user and the system. The instant that people see their own reflections, they are engaged, interested, and participating. There is no mouse to click, no interface structure to learn *before* the interaction begins — it starts as soon as the participant enters the space and passes in front of the camera. Holding up a digital mirror is all you need to do to grab someone's attention. (Keeping that attention long enough to sustain engagement, however, is another matter.)

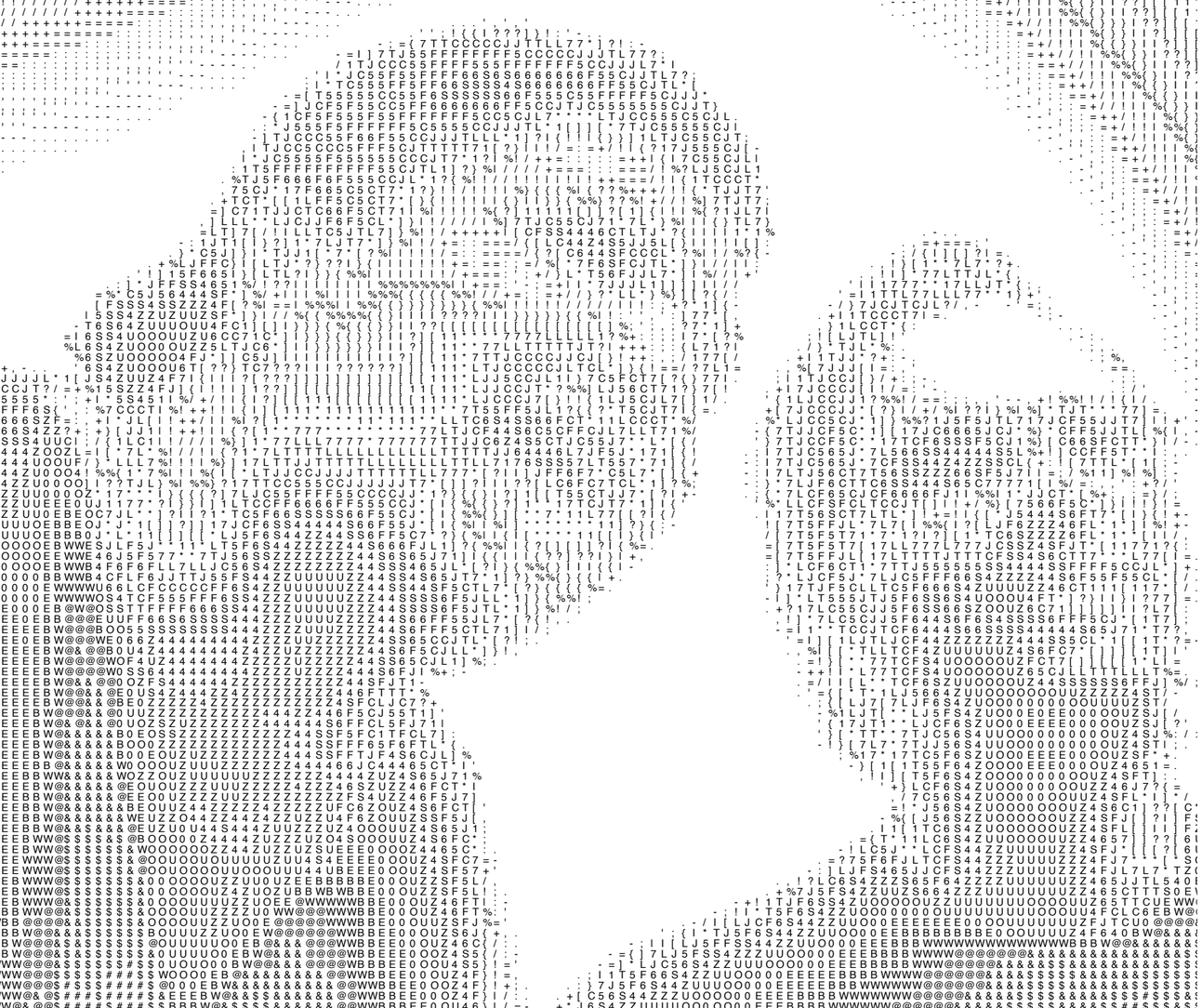


Gesture Project

The *Gesture Project* is a simple concept that produces beautifully complex results. A grid of red circles covers the display, and a video camera points outward, toward participants. As participants move, the system detects their motion, and rotates the discs accordingly. Where there is more motion, the discs spin faster. As they rotate, the discs change hue, gradually cycling through the rainbow. So, over time, visual hot spots emerge, revealing patterns in the history of motion. The entire grid automatically resets every minute, so participants can begin again with a clean slate.

While this project didn't employ literal, visual mirroring of participants, its visualization was coupled tightly enough to their motion that participants felt a direct connection to the system. I remember introducing the project in class, and being surprised by the extremely positive response. People loved it, both as direct participants in front of the camera and as observers. The colors and rotations were visually attractive, and the perceived physical connection was practically addictive. Even now, every time I show this project to someone new, they inevitably call out "Oh!" in excitement, upon realizing that the discs are responding to *them*.

Stills from Gesture Project.



ASCII Photo Booth

The *ASCII Photo Booth* was, far and away, one of my most successful projects at DMI, and it wasn't even done for a class. In December of 2007, at the conclusion of my first semester, Colin Owens announced that he was organizing an ASCII art show, to be held in February. I had just begun teaching myself Processing and knew I had to create something for the show.

Background — ASCII stands for American Standard Code for Information Interchange, the official name of the first widely adopted standard for binary encoding of human-readable text characters. It supports only characters from the Roman alphabet, plus basic punctuation, which is all that the Americans who first created it needed. The small amounts of memory in early computers prevented the inclusion of any characters not used in American English, such as ï, ñ, and ø. (Later, other encoding standards, such as ANSI and Unicode, were developed to offer support for many more characters, so one encoding could support letters from many different alphabets.)

ASCII art — visual work made of only these plain text characters — originated with the earliest terminal systems, before computers had true graphical capabilities. Engineers (the only people who used computers at that point) needed to document and communicate technical designs, and they ingeniously appropriated basic keyboard characters intended only for punctuation (such as: / | \ [ ] { } * # ^ < and >). Using these characters, the only visual elements available to them, they could draw simple flowcharts and schematics with lines, boxes, and labels. Being human, those engineers eventually realized that ASCII could be used to create non-technical images with strictly creative, and often humorous, elements. Today, so-called emoticons like :-) and ;-P are the most popular uses of text characters as visual imagery.

ASCII art now is a creative form in its own right with a semi-underground cult following. Although most people don't know the term ASCII, they will recognize its visual forms right away. With the computing power and graphical abilities of today's machines, we can now create dynamic, generative ASCII art, automating and accelerating a process that was originally done manually.

Concept — One of Processing's built-in example projects takes a video input and converts it, in real-time, to an ASCII art image. Dark pixels in the video are translated into characters with greater visual density (such as 8, W, and E), while lighter pixels are assigned less-dense characters (i, -, and `). A pure white pixel of video, of course, is represented with a space character ( ). By using a monospaced typeface, as found on early terminal systems, it is easy enough to create a grid of characters that corresponds exactly to the input image. So, a video image of 320 by 240 pixels can be translated into an ASCII grid of 320 by 240 characters. Each line of text would be 320 characters across, and the monospaced characters will sit evenly along the grid.

I was taken with Processing's ASCII video sketch, and decided immediately to adapt it to create an ASCII photo booth, an interactive installation in which participants would not only see themselves interpreted in plain text form, but could walk away with a hard copy in-hand. It would be the retro-digital, modern adaptation of the classically analog, coin-operated, Polaroid photo booth. My goals with the project were, first, to create an engaging and fun experience, and, second, to learn how to create an interactive art installation with Processing along the way.

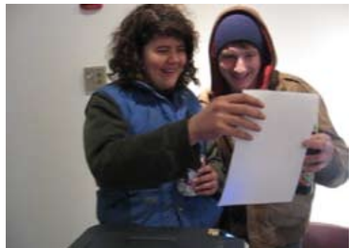
Process — I began by outlining each component that would need to work in the final piece:

1. Capture video image with camera
2. Convert video image to ASCII
3. Output ASCII image to display
4. Take user input from button (to signal when to take the picture)
5. Count down on-screen before "taking" the photograph
6. Generate and save ASCII image
7. Send image to printer

In addition, I needed a physical space designated for participants to pose in front of the camera.

Looking back, it's clear that the project involved a great deal of technical complexity, especially for my very first project in Processing and first-ever interactive installation. But ignorance kept at bay any anxieties that would have arisen otherwise, and I forged ahead, spending many hours over winter break with my Processing book in hand

The installation at the ASCII art show. The participants shown at bottom are captured in the image opposite.



and a clear vision in my head.

Steps 1, 2, and 3, above, were already being handled by the example ASCII video project included with Processing, but I made several modifications to achieve a simpler, cleaner rendering. Also, I wanted the video to occupy the full screen (not just a small window), and programming that change involved learning a great deal about how computers interpret video as arrays of pixels.

For step 4, I wanted a plain-and-simple button: ideally a spring-loaded, red plastic button with some tactile feedback — a nice, clean "click" sound would have been nice. Today, I would feel comfortable acquiring such a button and figuring out how to get it to communicate with Processing (via an Arduino board or some other hardware input), but given the timeframe for this project and my limited experience, I opted for the simplest input solution: a one-button mouse. This compromise ended up being technically successful, although it detracted from the retro aesthetic I was trying to capture. (Analog photo booths don't have mice, at least not electronic ones.) Any computer user today is familiar with mice, and recognizes that they are used for both motion (left, right, up and down) as well as selection (clicking and dragging). So it was not the ideal input mechanism, since it would be underutilized in this context, but it was faster to implement.

Step 5 was simple enough to develop. When the mouse was clicked, a white box masked a strip along the bottom of the display, in which would appear:

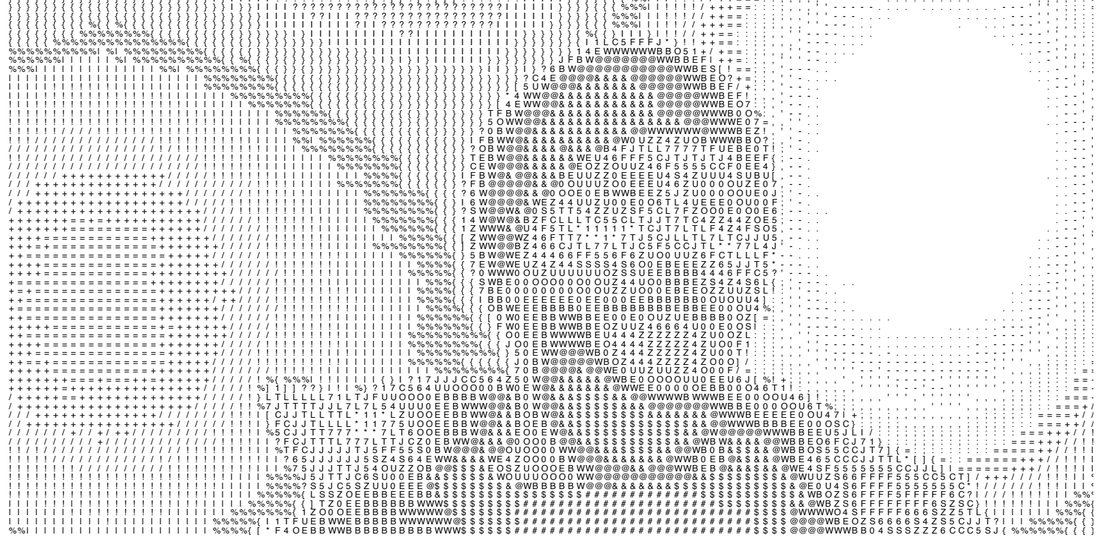
3...
3... 2...
3... 2... 1...
3... 2... 1... Smile!

At that point, the system "froze" the screen and saved the current frame of video, converted into ASCII (step 6). So far, so good. Now how was I going to get the image from the screen into the user's hands?

Printing (step 7) turned out to be the most technically complex piece, as Processing does not have a built-in way to talk directly to a printer. Processing could only export the image, after which point it was up to me to send it to the printer, but how to do that? I ended up writing an AppleScript that would "watch" a designated folder for any new images saved there. As soon as it detected a new file, it would open and print it. So Processing just had to create and save the images, and my script was in charge of printing them out.

At the time, I had an inkjet photo printer, which was far too slow for the experience I wanted to create. I needed people to be able

ASCII Photo Booth image generated on 2008.2.23 18:14. Project developed by Scott Murray with Processing, incorporating code by Ben Fry. alignedleft.com



to pick up their printed artwork immediately after the photo was taken — no waiting necessary. Only a laser printer would be fast enough to output a full page in just a few seconds. After failing to scrounge up something compatible (and free) from MassArt's computer labs, I found a great deal on a brand new laser printer at MicroCenter.

Now I had the speed that I wanted, but I wasn't satisfied with the quality of the image. At first, I tried simply capturing the screen image and sending that to the printer, but the output was blocky and pixelated — just as it appeared on-screen. Again, unaware of the complexity I was about to introduce, I decided to export my ASCII art not as a bitmap file, but as a vector-based PDF. For technical reasons, that turned out to be a non-trivial task, and involved taking the ASCII text that appeared on-screen and generating a completely separate image, using the same text characters, and then saving it to disk. But once I got it to work, it was beautiful: high-resolution ASCII art — possibly a contradiction in terms — output to paper in just seconds at 600 dpi. I loved using new, "hi-fi" technology to process and present an old, "lo-fi" art form. Also, for self-promotion purposes, I included a footer with my name and website address on every print-out.

On the afternoon of February 23, 2008, I hauled my new printer, a Mac Mini, an old 17-inch CRT display, my iSight camera, and a mouse over to the Doran Gallery and began setting up. Without much time, materials, or carpentry skills, the final, physical installation didn't resemble a classic photo booth as much as I would have liked, but I did manage to designate a defined space by positioning white display podiums around the work. Just as with an analog photo booth, it was clear when someone was inside or outside of the space, either being photographed or merely observing the process. The display and video camera were positioned about six feet away from the white wall, where a chair was placed. In order to see the display clearly, then, participants had to sit in the chair between the display and the wall — so the only clear view of the display was from directly in front of it.

Experience — The *ASCII Photo Booth* took 67 photos over the course of the evening, all of which I have archived, and some of which, I hope, made it home with their subjects. (I like to think that at least one of those dynamically generated, original artworks ended up on a refrigerator somewhere.)

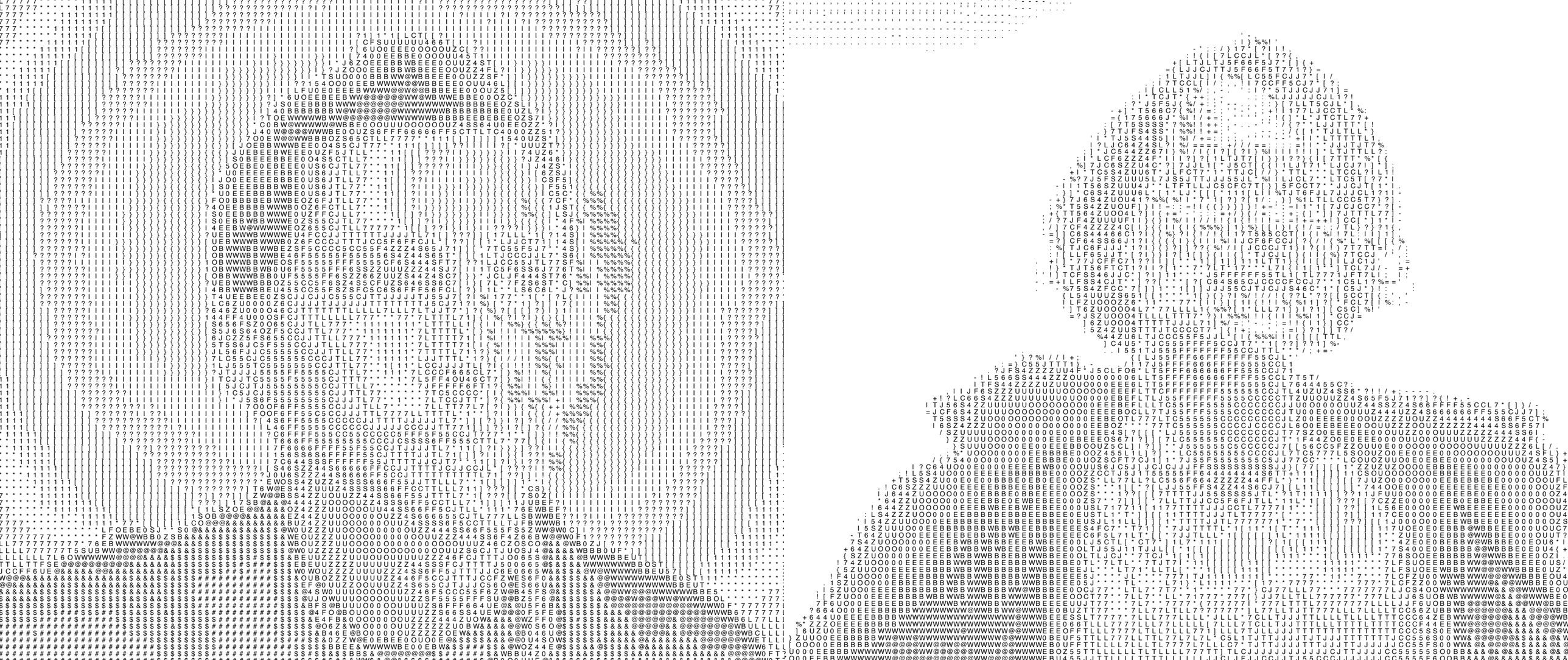
It was lots of fun to watch people use the photo booth, mostly because they were having so much fun watching themselves. Most people understood it right away, and even people who are normally camera-shy were comfortable admiring ASCII renderings of themselves. My first user observation, then, was that most people, like Narcissus, love to look at themselves. The fact that the image was augmented in a unique way kept people engaged and interested. Watching straight video of oneself is not as interesting as exploring how a system reacts to one's motion and recalculates its visual representation on-the-fly. This augmented reflection was highly engaging.

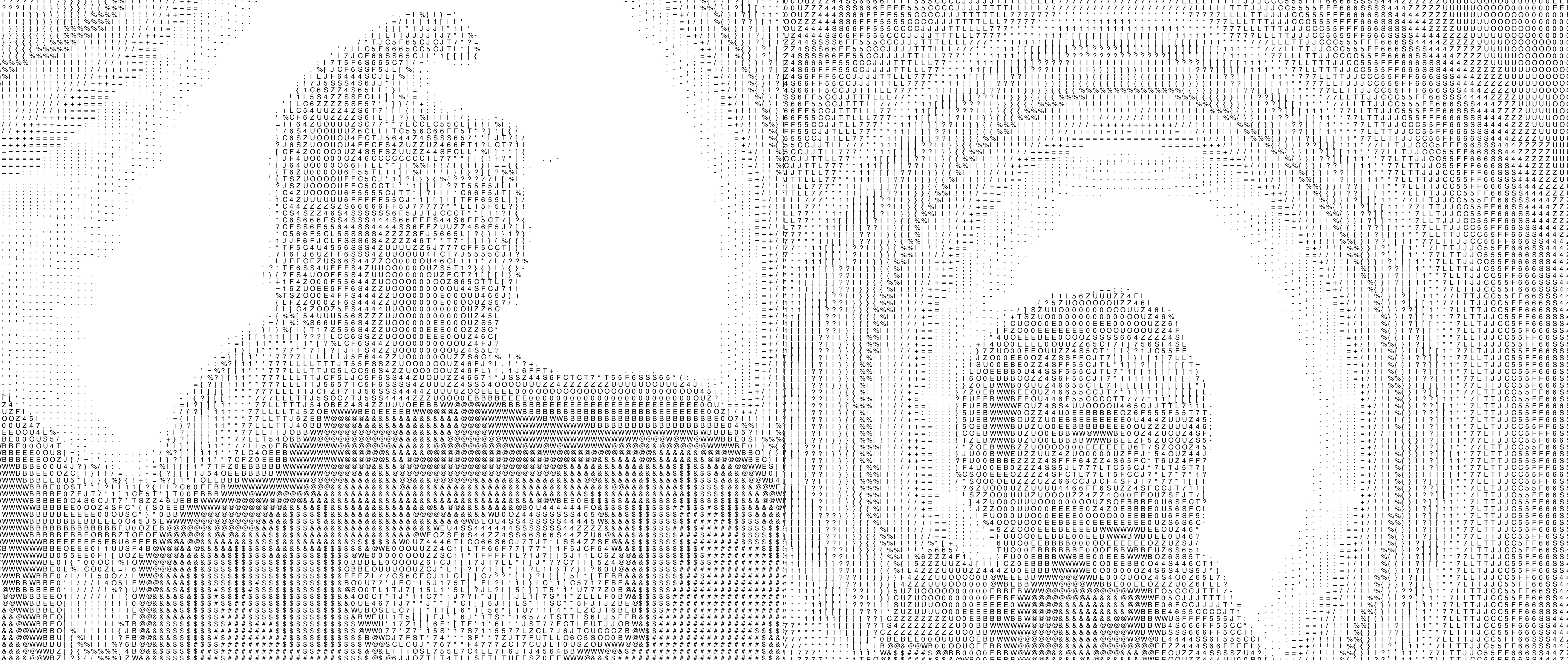
I also observed that the best images were created by people who took time to experiment with the system. They would lean in closer to the camera, then farther back, watching the on-screen text regenerate in response to their motion. A handful of participants spent only a few seconds with the piece, but the people who lingered were rewarded with better images. This was partly due to the camera's automatic exposure, which compensated for changes in light levels as people moved in and out of the frame. Thus, the final images were sharpest when subjects sat completely still before and during the exposure. That felt appropriate, given that ASCII originates from a time when computers were much slower and unable to process images at all. As with early photography, perhaps it *should* take time for a clear ASCII image to develop. In any case, although this quality could be considered a technical flaw, I liked the idea that people who invested more time and energy were rewarded with higher quality output — a theme I have carried forward into later projects.

As an added bonus, although I was half-expecting the application to crash at some point during the show, it never did.

Conclusion — Generally speaking, the project was successful because it met both my goals: I learned a lot about Processing, and the end result was an engaging and fun experience, as evidenced by the smiles on participants' faces. Were I to ever resurrect the *ASCII Photo Booth*, however, I would get some help with the physical construction, and build a space resembling a traditional photo booth, with black walls and a red curtain. I would use what I've learned about hardware integration to replace the mouse with a more analog-like button, and maybe even incorporate a system-controlled flash. (Although a flash is not technically necessary, I consider it an essential component of the analog photo booth experience.) The laser printer, too, could be concealed within the booth, and a system devised so the print-out could be deposited in a slot for retrieval. Also, instead of 8.5 by 11" paper, a non-standard size photo paper could be used, to better match the strip of images output by a traditional photo booth.

In short, I would like to remove all indications of the project's digital core except for the ASCII characters themselves. The ideal installation would take advantage of widespread familiarity with photo booths' physical form to make users completely comfortable within the space. Then, the use of ASCII becomes a clever surprise, the only outstanding element in an otherwise commonplace interactive object.





Mirroring may be a highly effective tool for triggering immediate engagement, but that engagement may be fleeting if not sustained by other means.

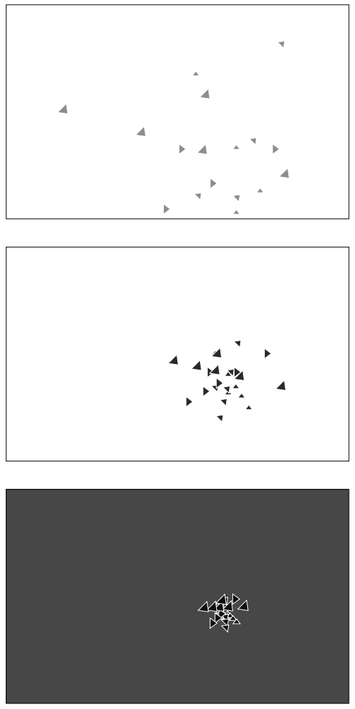
In my experience, sustaining engagement is best accomplished by establishing a sense of *anticipation*. This can be done by providing aural and visual clues of what's about to happen next. This requires systems to be somewhat secretive. To maintain engagement, we must provide hints, but not complete answers.

Anticipation Study

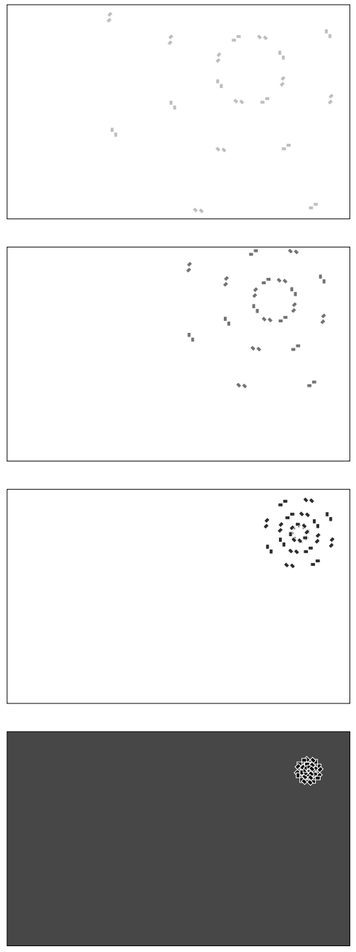
The first studio assignment in my second semester at DMI (spring of 2008) was to produce a response to this theme of *anticipation*. At first, I was stymied. How could one indicate that something was *about to* happen before it actually happened? How could I hint at what was about to occur, while maintaining tension during those moments of anticipation? If the end result was revealed too soon, then the anticipatory moments would fall flat. With a foreseen conclusion, there would be nothing to anticipate.

I decided that this concept was too amorphous to study in the abstract. I needed to create a working, interactive prototype, not just static mockups. I had just begun delving into Processing, and chose it as my tool. Next, I decided to use the mouse as my input device. The fluid, two-dimensional motion of the mouse seemed better suited to a study of anticipation than my alternative: the keyboard, with its discrete, on/off, binary switches. Via the mouse, a user could express position, motion, direction, and velocity — all concepts that connected with my traditional understanding of anticipation. Position: Where am I now? Direction: Where am I headed? Motion: How fast am I moving? Velocity: At this rate, when will the event I am waiting for occur?

Of course, most of the thinking above was developed only in retrospect. At the time, it was really just an intuitive decision to use the mouse. It simply made sense, and connected with the project's theme in ways I couldn't articulate at the time.



In any case, I began my experiment by using Processing to connect on-screen events with the mouse's position. I had the computer draw simple shapes — circles, squares, triangles — that followed the mouse around. Interesting, but there was no goal, no destination: a prerequisite for anticipation. So, still thinking very literally at this point,



Stills from Anticipation Study.

I defined a new x/y coordinate as the target "destination." But, not wanting to spoil the moment of surprise (and thereby deflate those expectant moments), I did nothing to indicate where that destination was. The user had to discover it on his or her own, and it was revealed only upon arrival. To make the experience even more abstract, I hid the mouse pointer, so only the moving shapes were visible.

Since the system's feedback was entirely visual, a dramatic visual change to indicate "arrival" seemed appropriate. Mousing over the target location triggered a full-screen fade to black. The shapes changed (e.g., from circles to squares, or squares to triangles), and then the background returned to normal. The target destination was then invisibly moved to a different, randomly selected point.

This treatment successfully indicated completion of the task, but there was still no tension of anticipation. I needed a way to tell the user — visually, and without words — whether or not they were on the right track. I reflected on these statements by Mihai Nadin:

- Anticipation is an expression of the connectedness of the world, in particular of quantum non-locality.
- Anticipation is an attractor within dynamic systems. (2004)

Perfect — connection and attraction were the answer! I had already linked the position of shapes to follow the mouse, but they needed to also be *attracted* to the target destination. The shapes' motion needed to express not just mouse position but the *relationship* between mouse position and destination. By observing this relationship, the user could deduce where they needed to go.

I modified the rules so the shapes would not just *follow* the mouse, but also *converge* on it when close to the target. This was not clear enough, so I factored in opacity, so when the mouse position was far from the target, the shapes would be spread out and very light gray (or completely invisible). Moving the mouse closer to the target brought the shapes closer together and increased the opacity (to dark gray or black). So, in a way, the user was rewarded first with shapes appearing on-screen, and second, by seeing them cluster together. Then, once the target destination was found, the fade-to-black transition was triggered, and the process began anew. User testing during the in-class critique revealed that this approach worked quite well.

Incredibly, this was only a one-week project, but the experience was invaluable. These issues of non-verbal instruction, challenge, and reward have arisen in many subsequent projects, including my thesis projects, as we'll see in a moment.

An element of anticipation is essential for engagement of any substantial duration. The question is how to elicit that emotional experience of anticipatory tension. Equally important is communicating instruction to the user; if they need to do something to achieve the goal and trigger the anticipated event, then they must be told or given clues on how to do that. The system has to communicate the terms of interaction with nonverbal, visual cues, or the user will remain hanging in suspense — until they give up and walk away.

Practice

Practice is the culmination of my explorations at DMI. But to call it a culmination implies that it is something of a terminus, when it really is just the beginning of a new course of study. Let me explain how the project originated before I delve into that new direction.

I knew that I wanted to complete the year with one substantial project that integrated much of my learning from the program. Yet after two years of immersing myself in interaction design, data visualization, systems design, and narrative studies, I was at a loss when formulating a thesis project. I couldn't imagine what the content would be, but I had a list of elements that it should incorporate:

- data-as-narrative
- interface and visualization
- dynamic, streaming, or live data sources
- visualizing the invisible

These were my favorite bits and pieces from past projects, but this list needed simplification. What was the one, single quality I wanted in my project? *Immediate engagement.* I wanted my thesis project to be engaging. And for that reason, I selected mirrored, interactive video as the primary input/output medium. Processing, of course, would be used as the development and execution medium.

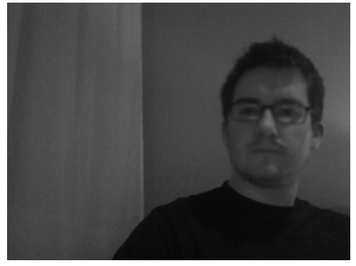
Early Experiments

Using video as an interface input allowed me to explore some fun technologies, like computer vision and live video processing. One of the computer vision libraries available for Processing offers face detection, which lets the system look at each frame of incoming video and analyze whether or not any faces are present in the image.

Although face detection is more popularly associated with the security industry and, more recently, point-and-shoot digital cameras (which can identify faces, and adjust exposure settings accordingly), this technology is ripe for exploration in the context of interactive art. Faces are emotionally loaded entities; they are our primary means of both identifying others and recognizing ourselves. The emotional responses we have to faces cannot be overstated, and this power makes them prime targets for artistic exploration.

More pragmatically, face detection can be used to isolate a person's position against a background of visual noise. Many interactive video projects track "motion" by looking only at which pixels changed from one frame to the next. (*Gesture Project* used this simple definition of motion.) By using computer vision algorithms to look for faces, our systems can disregard all other visual input, such as objects moving in the background. Different face detection "profiles" can be used to identify faces from different angles (head-on, ¾ profile, or full profile, for example), so we can even differentiate between people facing directly toward the camera, and those turned away from it.

Of course, computer vision is nowhere near perfect, so good lighting is critical. The algorithm needs to be able to see two eyes and a mouth in order to identify a face. If only half the face is well-lit, it will not be detected. In my experience, false positive identifications are actually more common — such as when the system "sees" a face that isn't really there, in the folds of a shirt or among shadows cast



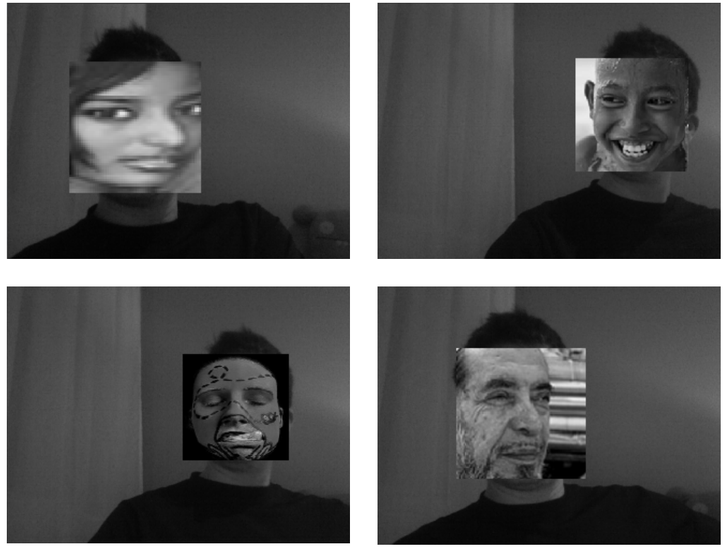
Stills from the "blurred face" experiment.

far in the background. The computer has no sense of physical depth, so a small circle near the camera will be perceived the same as a much larger circle very far away. Both could be interpreted as "eyes" of the same face, even though they are hundreds of feet away in physical space. The camera knows only pixels.

Despite these technical considerations, I was committed to exploring face detection's possibilities. My first experiment was ostensibly very simple. It captured the video image, analyzed it for faces, and then blurred a portion of the video around the user's face. The blur gradually intensified, then diminished, cycling through varying levels of clarity. I incorporated easing, so the blurry box would move smoothly, gliding, not jumping, into place. Technically, it supported identifying only one face at a time, so introducing a second or third user confused the system. (The blurry box would jump quickly from face to face.) But for a quick sketch, it worked well.

Users hated it. I showed the project to a number of friends, and the universal reaction was to move around, trying to dodge the blurry rectangle. A couple people tried to physically push the blurred box away from their faces using their hands. I quickly realized that this blurring effect was operating contrary to the engagement effect of mirroring. Users were seeing themselves, and the system's motion (of the blurry box) was coupled to their motion, but because the mirroring effect was interrupted visually, users were immediately annoyed and put off. People wanted desperately to see their own faces, not just torsos and necks topped by fuzzy squares. I had designed a frustration machine.

I interpreted this user frustration to mean the project was not successful, but Gunta Kaza encouraged me to explore it further. She pointed out that the experience triggered a strong emotional response from users (if not a positive one), and for that reason, at least, was worth exploring further.



Stills from the "replaced face" experiment.

For the next iteration, I wanted to know how people would react if the blurry box was gone and they were shown a face — just not their own.

I tapped into the Flickr API to retrieve the most recent photographs tagged with "face" or "person." Then, the system would run face detection on those photos. If no obvious face was found, the image was discarded. If a face was found, the image was cropped and stored in memory. Then, when someone stepped in front of the camera, one of the Flickr faces was selected at random and mapped over that of the user's. Every 10 seconds or so, a new face was selected and displayed.

Users found this version much less frustrating, but extremely creepy. They wanted to know where the other faces came from, and why they were being placed over their own. Some juxtapositions were more entertaining than others, such as a baby's face, or a face with a shape and hairline that visually matched the user's. In the event of the latter, users would reposition their bodies to best fit the image being shown, like the inverse of a carnival cut-out: a clown's face on top of your body.

Clarified Direction

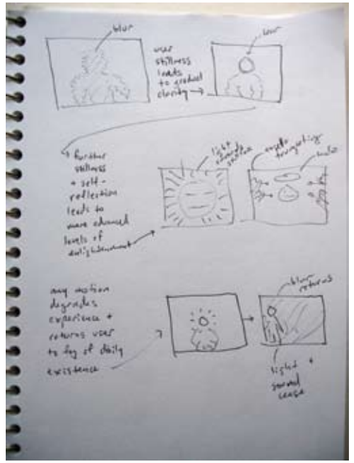
Neither of these quick projects were the engaging, rewarding experiences I wanted to create, but the discomfort they induced provided me with some valuable insights. Gunta encouraged me to consider framing the system-user interactions in terms of challenge and reward. My *Gesture Project*, for example, was 100% reward — there was no challenge. But maybe a reward would be sweeter if users had to work for it by tolerating some amount of intentional, *designed* discomfort.

Yet, how can we challenge someone while keeping them engaged? When a user has essentially no commitment to the project (e.g., they have not paid money to see it or worked to create it), how can we incentivize them to overcome the discomfort of the challenge? As described earlier, establishing a sense of anticipation is essential. By hinting at the rewards to come, without revealing them too soon, we hold out a proverbial carrot for our users to pursue. But such a challenge/reward structure may be too simple. We must remember that we are trying to create a positive, aesthetically unified experience on the whole, not just sequential alternations of bad and good elements.

In *Rules of Play*, game theorists Katie Salen and Eric Zimmerman review the research of psychologist Mihaly Csikszentmihalyi, who has written extensively on cultivating the peaceful mental state of "flow." Csikszentmihalyi's state of flow corresponds roughly to my earlier definition of satori: a quiet, concentrated mental state, in which the subject's focus is entirely on the task at hand in the present moment (1990). Salen and Zimmerman apply this research to game design theory, arguing that any challenges must fairly match the user's skill level (2004). A challenge that is too difficult leads to anxiety and failure. One that is too easy leaves users bored and disengaged. To elicit an engaged state of flow, the challenge must be of appropriate difficulty. As the user gains experience with the system over time, the challenges must escalate at a corresponding rate in order to sustain the same level of engagement.

Game designers, of course, strive to create games with "replay value" and even addictive qualities. Gamers should be engaged not only "in the moment," but over time, as they return for subsequent sessions. Salen and Zimmerman consider games as systems of rules and actions. They write, "If you create a space of possibility that rewards players for exploration, then you are likely to have players that want to see more permutations of how the rules play out" (2004). All the while, the ideal design would elicit what Csikszentmihalyi calls an "autotelic" experience, meaning "a self-contained activity, one that is done not with the expectation of some future benefit, but simply because the doing itself is the reward" (1990). Use of the system, then, becomes a reward in and of itself, inclusive of all the challenges presented within that interaction.

So the primary design challenges for my thesis projects became: to engage participants, construct a sense of anticipation, and then reward them for tolerating the discomfort elicited from sustained engagement. My approach had shifted from an attempt to create a universally engaging and memorable experience to a study of human behavior. By designing increasing levels of challenge and reward, I could gather data on the intensity of participants' engagement, tolerance for discomfort, and patience for reward.



Social Context

So, then, what reward to offer, and what challenges to precede it? To my mind, the ultimate reward is enlightenment: all-knowingness, clear perception, right seeing, oneness with the universe, and the thoughtless awareness of the present moment that comes with satori. Enlightenment cannot be literally bestowed upon users, of course, but I could allude to it figuratively, and at least try to trigger enough physical discomfort and emotional ambiguity to elicit a state of satori.

With enlightenment as the reward, stillness would be the challenge. Stillness can be interpreted both metaphorically, as in stillness of mind, as well as physically, as in a lack of motion. The latter is easily measured by computer, so users could progress toward enlightenment by being physically still. Movement would trigger regression away from the goal.

Nearly all video-based installations reward motion — jumping up and down, waving one's arms, ducking and jumping. *Practice* rewards stillness, which should not be confused with inactivity or a lack of interaction. My early user tests showed that maintaining stillness is, indeed, quite a challenge, and the level of interaction, as reported by users, can be intense. Once participants understand that the key to success is *not moving*, they grow very quiet, still, and focused, watching and listening closely to whatever the system does next.

As described earlier, video-based projects are always fundamentally social projects, since they can be used by multiple people at once. In many installations, participants may either interact independently of each other or work together. In my project, I built functionality that allows me to test either giving one user control of the piece (tracking the motion of only one individual) or giving all users some input (all user's motions are considered). My hope was that, by considering the motion of all participants, the system would encourage social negotiation between its users. This would make the challenge of stillness even more difficult to attain, since interacting with others requires some amount of motion (at least a gentle nudge or mumbled instruction).

Also, we are used to staring silently at screens, but ignoring a fellow human being is impolite. So there is not only tension between the users and the system, but among the users themselves, as they struggle to maintain focus on the project while negotiating with each other. Thus, the participants inadvertently become performers, and the system expands to include not just the screen and sound, but also the people in front of that screen, and the observers who are watching those people. As with a flash mob, anyone present becomes an active participant at some level, whether willingly or unwittingly so. Even pure "observers" are engaged in the dynamic, because their very presence serves to distract participants from the task at hand.

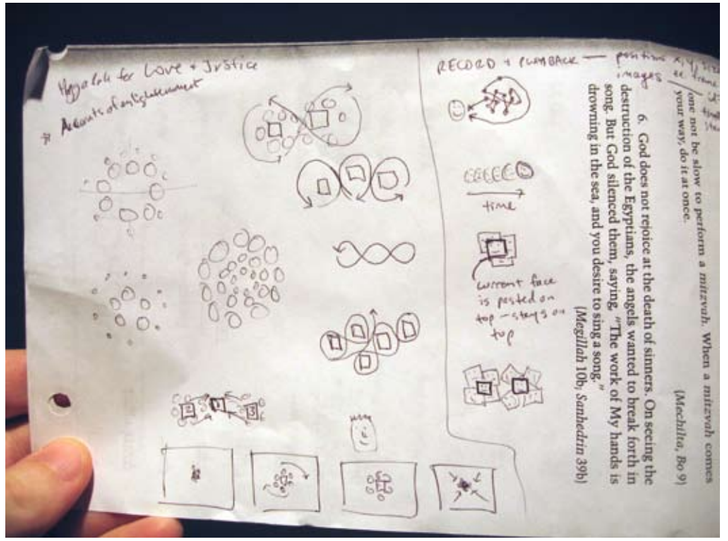
Structuring the System

With a primary metaphor of stillness as a means of progressing toward enlightenment, *Practice*'s initial display had to be grayscale, blurry, and dark. The experience begins with my interpretation of the hazy state of everyday life. We move quickly, going through the motions, without reflection or clarity around why we make the decisions that we do.

But with stillness comes clarity, so when the user stands still, facing the screen, the system gradually removes the blur, and the image comes into focus. Beyond that, color is slowly restored, until the participant finally sees him or herself reflected clearly, just as though looking into a mirror.

Of course, any participant motion disrupts the stillness, in which case the system regresses — color fades away, and the sharp image grows blurry. We return to our default, unclear state.

I knew that the system would do at least this much, but I also envisioned the addition of several more advanced stages during which the mirroring would be augmented with increasingly complex imagery. This sequence of stages would culminate in a final interpretation of enlightenment.



Soon after beginning work on the project, I had to consider the structure of my code. How could I organize these different programmatic elements in such a way that would support the experience I wanted to create, while also ensuring a straightforward process for developing and inserting additional stages that I hadn't yet designed or considered?

I settled on using a single number — a *progress* value — to track the participants' "position" within the sequence of stages, while each stage was assigned a whole number. For example, imagine each stage as a point along a line, starting with 0, then 1, 2, 3, and so on. In the beginning, the progress value is 0.0. At 0.5, we are halfway through the initial stage (stage zero). The scene is grayscale, but some blurriness has been removed. At 1.0, the scene is clear, and at 2.0, we see in full color.

This structure was useful because, for each frame of video, the system had only to reference one number to know where it was in the sequence. So during each frame, two steps occur: First, the face tracking algorithms are applied, and an averaged "stillness value" is calculated. If the stillness value is above a certain threshold (meaning, there was little or no motion), then the progress value is increased by a small amount. If the stillness value is low (meaning, there was a lot of

motion), then the progress value is decreased by a substantial amount. The increment/decrement values are unequal, so progress is lost more easily than it is achieved — another way in which the operating metaphor was expressed in code.

Once the stillness and progress values are calculated, the system simply renders the appropriate stage. So, if the progress value is 4.7, the system executes the code for stage four. In addition, a normalized value is used the determine the position within each stage. That is, 4.7 tells the system both that we are in stage four, and that we are 70% of the way through the stage. That normalized value is then used to drive different events within the stage, such as how much blur to apply or color to restore.

Events that happen between stages, such as triggering sounds or resetting elements' positions, are controlled by comparing the previous frame's progress value to the new one. So, if progress moves from 5.99 to 6.02, the system knows that we've just entered the sixth stage, and it executes the appropriate actions. Similarly, moving from 3.1 to 2.84 means we're regressing, so any audio played during stage three should fade out.

Structuring the project around multiple, self-contained stages tied to one numerical progress value gave me a great deal of control over each stage, and simplified the process of adding new stages.

User Testing

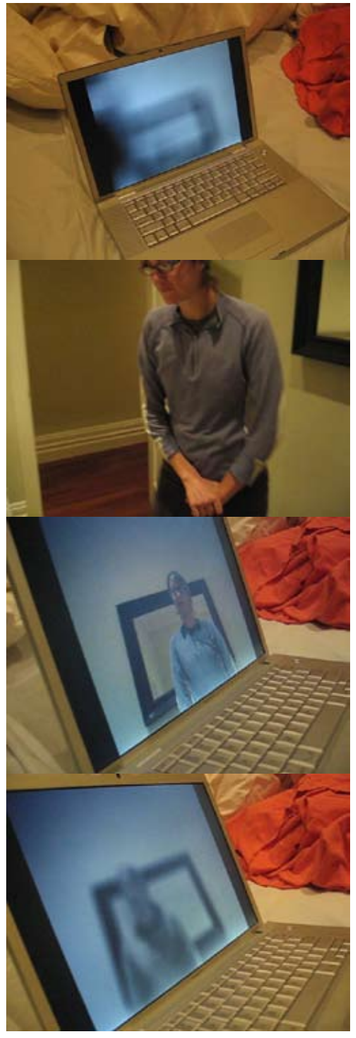
Since my professional background is in user-focused web design, usability research is always an essential part of my design process. I had tested the "blurry box" and "replaced face" experiments only informally. But with *Practice*, I applied more structure to my user testing and conducted tests at least weekly, enlisting around 20-30 different people over the course of a few months.

The earlier user testing can occur in the design process, the better. Ideally, the designer can test early and often, making refinements to the design along the way.

The most difficult design challenge with *Practice* was figuring out how to instruct stillness. In my first tests, users were content to see themselves reflected in blurry grayscale. But they never progressed through the system, because nothing was telling them to be still. My testers would invariably wave their arms and jump around in front of the camera for about half a minute, and then try to act impressed, despite their obvious disappointment that there wasn't more to my big art project.

I thought it might help to communicate that the system was seeing participants' faces, so I tried drawing a primitive, rotating spiral shape around each detected face. That only made testers move more, since they enjoyed watching the spiral spin and change size as they moved closer and further away from the screen.

Thinking the problem was in the rotating motion, I tested a simpler treatment: every time a new face was detected, a brief tone



played and a soft, blue circle appeared, covering the face. The circle quickly faded away, so I thought there would be no incentive for additional motion. But I was proven wrong, and my testers only moved further and faster this time, confusing the face tracking system, triggering many more blue dots and xylophone-like tones.

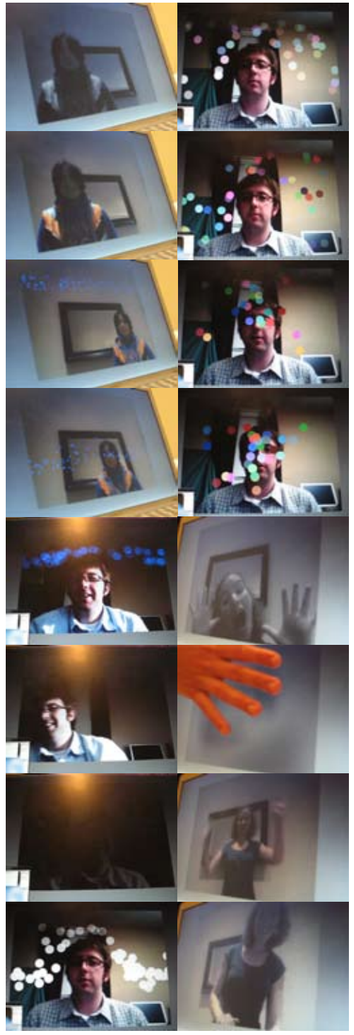
It was time to reevaluate my approach. This initial instruction was absolutely critical to the project's success. I was certain that, once participants understood the interaction model (stillness leads to progress and reward), they would smoothly advance through each stage. But stillness is such a foreign concept in our culture, especially within the sphere of dynamic media, which always incites us to go, go, go and never stop moving. So a completely new treatment was needed.

Since all testers found the mirroring element invariably rewarding (even when blurry), why not "punish" their motion by taking the mirror away? I removed the circles and sounds, and coded the system to quickly fade to black upon detecting too much motion. This worked perfectly — finally, people stopped moving! By removing the visual stimulation altogether, I could provide immediate, negative feedback. With a blank screen in front of them, users have no incentive to keep moving. That, combined with confusion around why the video suddenly "went away," causes them to hold still. They lean in close, squinting and perplexed, asking "What happened?" Then, once they become still, the black fades out and the video returns. It usually takes no more than two or three of these disappearances for participants to understand the interaction model.

At this point, I observe an "aha" moment in participants, when they understand that the system is looking for stillness. This initial stage zero is completely silent; the first audio plays only upon reaching stage one, when a bell chimes. Entering each subsequent stage triggers another discrete bell-like chime (a different chime for each), and begins looping other ambient audio. (In testing, users correctly understood these chimes as indicators that something is "about to happen.")

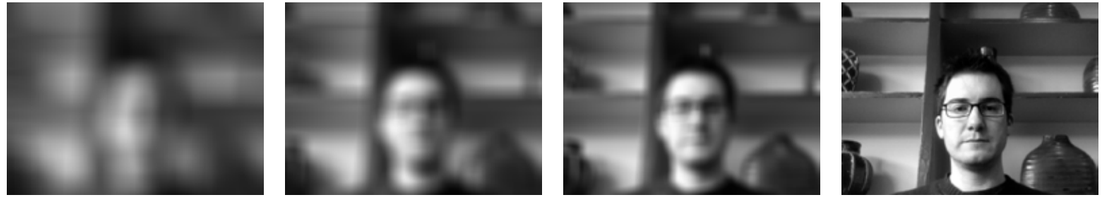
The visual transitions, however, are deliberately very gradual and subtle, which contributes to the sense of anticipation. For example, it is only about halfway through stage two that users realize that the video is being colorized. The subtlety of these early stages primes them for focused observation and keeps participants watching and listening closely as they progress.

Stills from user testing documentation.



The Stages of *Practice*

There's no substitute for experiencing *Practice* in person, but for the purposes of documenting the project here, I will describe each of the eight stages and articulate the intent behind the design of each stage.

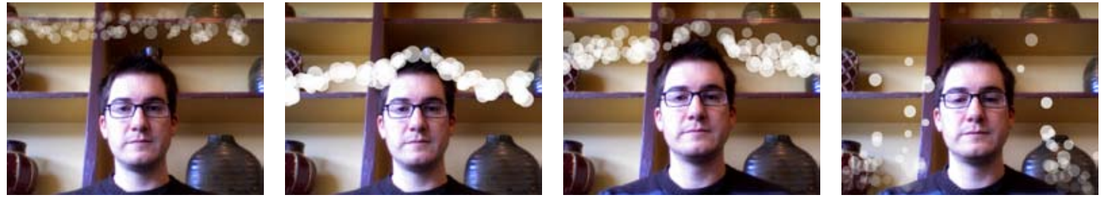


Stage 0 — Clarity

Video is initially grayscale and blurry. The blur is gradually removed, and by the end of this stage, it is in full focus.

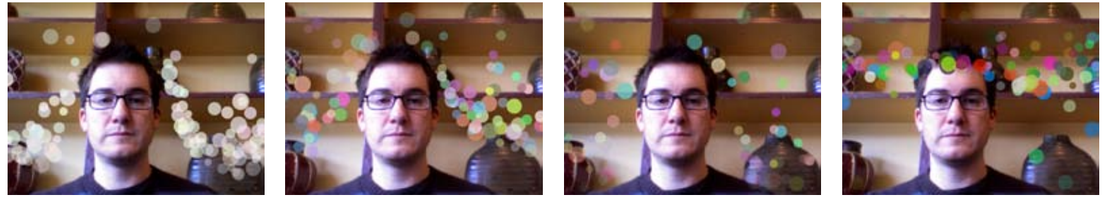
Stage 1 — Color

A chime sounds, and video gradually transitions from grayscale to full color. Too much motion in this stage triggers a blackout of the screen. After this point, motion causes only regression (a decrease in the progress value), but not a fade to black.



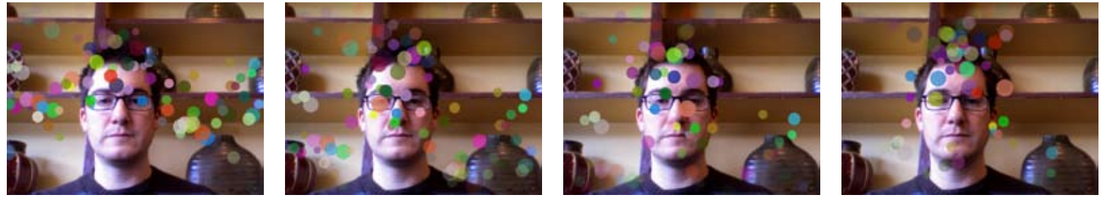
Stage 2 — White Snow

A chime sounds, and ambient drumming sounds fade in. Assorted, semi-transparent white circles begin falling down from above. They seem to be responding to gravity, yet they cascade around participants' heads. This is the first indication to users that the system "sees" them. They discover that the shapes are indifferent to waving hands and arms. Moving their faces (or whole bodies) affects the circles' trajectories, but that motion also triggers regression within the sequence. In testing, some participants look up to see where the circles are falling from. The circles' relative sizes are proportional to how close users are to the screen, so standing up close magnifies the circles' diameters.



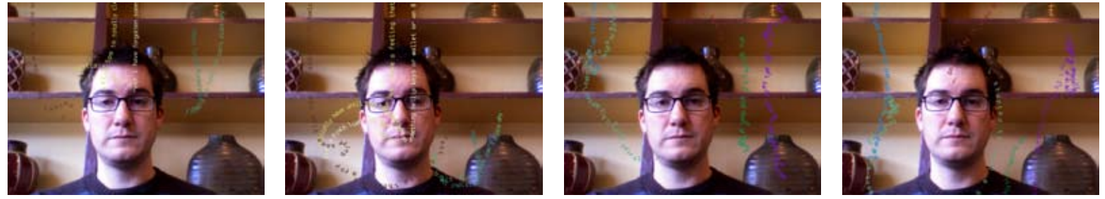
Stage 3 — Colorful Snow

Another chime sounds, and the drumming loop continues. The physics of the cascading circles remains the same, but they gradually transition from white to assorted colors. This transition mirrors the earlier shift from colorlessness to full color.



Stage 4 — Orbiting

A chime sounds, the drum loop stops, and spacey, ambient audio fades in. The circles no longer avoid participants' faces, but are attracted to them. The circles then orbit and obscure users' view of their own faces. This deliberate obstruction is intended to heighten the tension and discomfort of stillness. Just as my early experiments obstructed users' faces, so do the circles here. With the mirroring interrupted, there is greater incentive to move. But moving to the side to reestablish the view of oneself, of course, triggers regression. In testing, participants find this stage somewhat disorienting, and many tried to eat the circles as though they were floating pieces of candy.



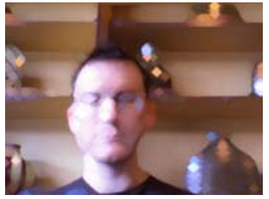
Stage 5 — Emotions

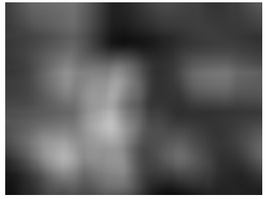
A chime sounds, the space-like audio stops, and a friendlier, yet contemplative audio track begins looping. The circles have disappeared, and now strings of text cascade down from above, pulled down by gravity. Those that pass near a face latch on and slowly orbit the face. The words are feeling statements, taken from the We Feel Fine API. The intent is to enhance the contemplative mood of the piece by forcing reflection on the statements presented, which are recent "I feel…" statements from blogs on the Internet. By visually attaching emotional statements to users, participants must consider whether or not they want to be associated with those statements. "I feel sad because of what I did today." "I feel happy that we were able to spend so much time together." Does the participant relate to these feelings? Do they cause him discomfort, possibly by exposing his own feelings that he would not have otherwise exposed in this public setting? In testing, since the source of these statements is not fully explained, a few users felt the system was somehow reading their minds.




Stage 6 — Personal History

A chime sounds, and the emotion statements continue to drift and orbit. But a new visual element appears — a sort of flaming, colorful line that zigzags about the screen, ultimately coming to rest on each participants' third eye, at the center of the forehead. The line moves differently for everyone, because it is a visualization of individual participants' movements, as recorded by the system throughout their session. This element serves to bring each participants' focus back to his or her own self. As it cycles through a range of colors, it appears to pulsate and flicker, much like the flame of a candle, as observed in meditation.

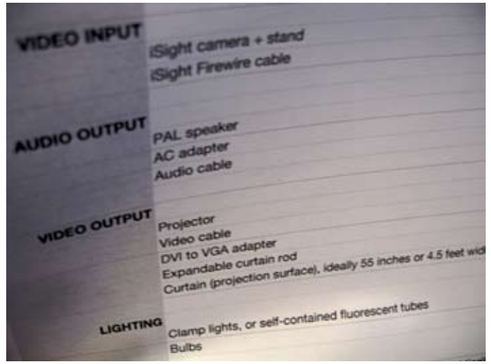





Stage 7 — Enlightenment

Warm, electronic tones gradually build and crescendo, as the video blurs using a method that produces diamond-shaped patterns, like a photographic lens filter. As the diamond blur increases, the whole image grows brighter, until it is solid white. Just after the audio peaks, it becomes suddenly silent, the screen fades to black, and then video is restored once more, blurry and gray. We achieved the clarity of enlightenment, but only for a few fleeting moments, and now we are back where we began.

Process documentation from the setup at Revamp Salon.



Street Test

After months of experimentation and refinement, it was time to test the project with people who I'd never met and who didn't know anything about the project.

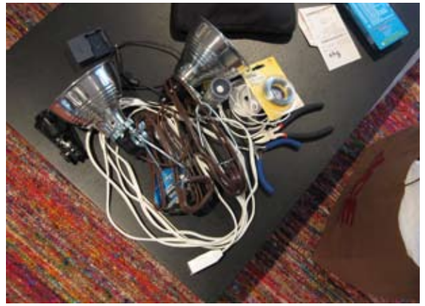
I contacted my friend Huy Le, who owns Revamp Salon here in San Francisco. I had installed the *Gesture Project* at the salon a few months earlier, and he has been a great supporter of my artistic process. Revamp is near 16th St. and Guerrero St., a busy corner in the Mission District with lots of foot traffic, especially on weekend nights. The salon has a window facing Guerrero, and an entryway protected by a lockable gate. It was a great opportunity. I could set up the video camera and projection screen in the window, and place speakers just behind the gate. Random passersby could experience the project, and no equipment could be stolen. I proposed this plan to Huy, and he agreed to let me take over the front window for an evening. I chose Friday, February 12, the start of the Valentine's Day weekend, hoping to reach the crowds of people headed out for dinners and drinks.

I sketched out possible layouts, and then went over to the salon to take measurements. I would need cables long enough to connect the computer to the camera, projector, and speakers without interfering with the projection. And I would need light — lots of strong, even lighting directed out onto the sidewalk — to ensure that the face detection would work at all. Finally, I would need to find a material onto which I could project video from inside, and have it be seen clearly outside, on the street.

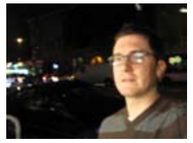
Another friend, interactive artist Mary Franck, offered me the use of her projector for the weekend. On Tuesday the 9th, I picked up the projector, and made a trip to the hardware store to purchase clamp lights, an extendable curtain rod, and wire. Back at the salon, I maneuvered the camera, curtain rod, computer, and projector in place, and a plain white curtain I brought served well as a projection screen. The physical setup was ready, but I still had some changes to make to the software before Friday.

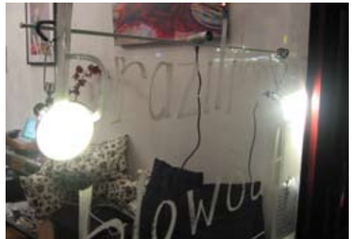
The final features I built into the *Practice* application enabled the system to save much of the data it was already capturing during operation. This included:

- A record of every face detected by the system, including when it first appeared, how long it was present, and how much progress it made through the various stages.
- For each face, a complete history of its size and x/y position for every frame in which it was seen.
- Once each minute, a still image capture of whatever was on-screen at that moment. (I tried saving these screen captures more frequently, or even recording full motion video, but doing so slowed the system to an unacceptable degree.)



I made anecdotal observations during the installation, but all of this data helped generate quantifiable answers to my core questions: How long would it take people to first understand that stillness was the key to success? What was their tolerance for remaining still? And how many people would make it all the way to the final stage?

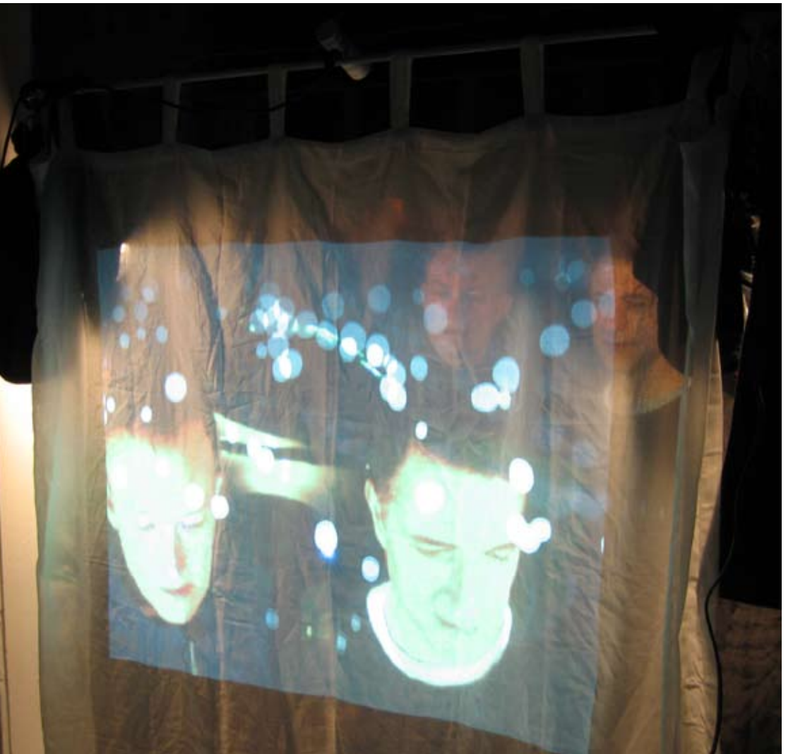




Anecdotal Observations

That Friday, *Practice* ran from 9:00 p.m. until 11:00. A number of friends that I'd invited stopped by, and a handful of other people stopped to observe or engage, most only briefly, but some for sustained periods. During those two hours, I made the following observations and conclusions.

1 — The physical context is critically influential to the overall experience, as it informs participants' expectations and what behaviors they will consider socially acceptable. A busy street is not necessarily the best place for an interactive installation. I had hoped for a lot of foot traffic, and got some, but most of those pedestrians were destination-oriented, not casually strolling. Guerrero is not very pedestrian-friendly, and the volume of car traffic encourages people to walk quickly, until they can turn off on a more welcoming street with wider sidewalks and fewer cars. That said, for my purposes, this element didn't make the installation a failure, but only amplified the discomfort and challenge of both remaining still and simply stopping in the first place. As the importance of context became clear, I understood that just one installation would not be enough to definitively answer any questions about people's tolerance for stillness. If *Practice* were in an art gallery, or even a less hectic public space, I would see very different results.

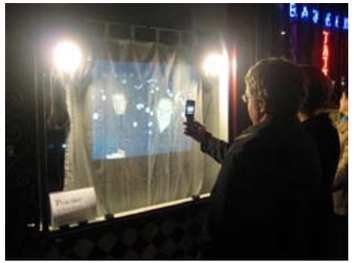


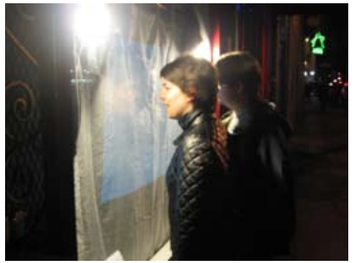
Friday night participants.

2 — Each physical space offers its own technical challenges and opportunities. Although the salon window was a perfect size for my purposes, the street was noisy, and the outdoor setting dictated that I could only show the project at night, so the lighting conditions and aural environment were not ideal.

3 — Over the course of the evening, not one person stopped to observe or interact with the project *unless* there were others already engaged with it. It was critical that I had invited friends, because it's possible no one would have stopped otherwise. With only two or three people present, passersby tended to glance at the projection, but they wouldn't stop walking. But once there was a crowd of 10-15 people, every pedestrian stopped, partly out of interest, but also due the physical necessity of navigating between so many people along the narrow sidewalk. This number of 10-15 functioned as a sort of critical mass, which would draw in new people as others left, and, for a time, was self-sustaining.

4 — To my great satisfaction, the social interactions that I had anticipated and designed for were fully present. While active participants interacted with the system, observers interacted with each other, and coached the participants. A fascinating dynamic evolved between the initiated and the uninitiated. In an ironic twist, the initiated — those who had already progressed to higher stages and understood that stillness was the key to success — quickly grew impatient with newcomers, becoming frustrated when the uninitiated would move too soon. Initiated observers were torn between withholding the "secret" and encouraging others. They seemed to be content watching for a while, but when a newcomer would "give up" or turn away from the screen, the initiated would be quick to offer urgent instruction — "No, no, don't look over here!" or "Stand still!" I anticipated this dynamic, to a degree, but not its intensity. I hope it reflects that the interactive experience itself is so emotionally engaging that, having completed it, participants are motivated to coach newcomers so that they, too, may share in that experience.





Data Analysis

By the end of the evening, the *Practice* application had generated about five megabytes of face tracking data and 168 screen captures. I then dove back into Processing to write a new program that would read the data files and generate some visualizations. But first, some numbers.

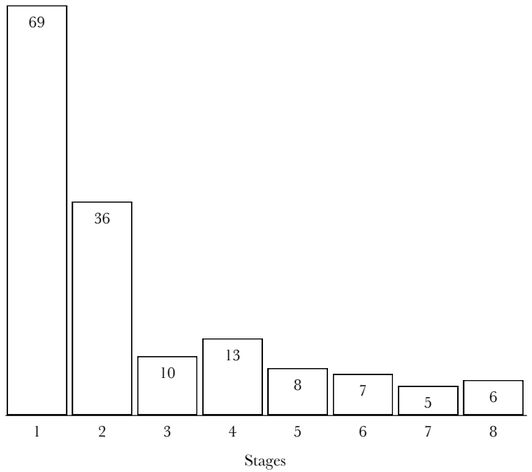
The system detected and tracked 1,198 faces over the course of the evening, but about half of them "existed" for fewer than five seconds. I excluded these faces from my analysis, assuming that they were mostly false positives. Of course, many of the 605 faces that lasted for five or more seconds were also incorrectly "seen" by the face detection algorithms, as we'll see in just a moment. In any case, I proceed with the understanding that the data set is not quantitatively accurate, yet can still be used to derive some valuable insights.

My primary questions of the data were: For how long were people engaged, and how many stages did they complete? Of those 605 faces, the average "lifespan" was 13.8 seconds — not very long. But an average isn't meaningful in this case because only a few participants were engaged for significant periods of time. The longest-lasting face existed for 5 minutes and 4 seconds, a considerable amount of time to stand up straight, look straight ahead, and resist urges to turn your head and acknowledge the people around you. The 50th percentile for time spent was only 7.6 seconds, and the 70th was 10.6 seconds. The top 10% of faces lasted longer than 24 seconds, and only 7 of those lasted more than a minute.

Time spent aside, how far did people progress toward the final stage? The chart below shows all the numbers. Of the 605 total faces, just 69, or 11%, made it to the first stage (as indicated by the first bell sound, and removal of blur), and only 6 experienced the enlightened ending, stage eight.

We would expect the numbers to decrease for each subsequent stage, and for the most part, that's the case. But notice that for stages four and eight, they actually increase. This indicates either face detection inaccuracies (whether false positives or faces "lost," then "found," between stages), or instances of people stepping into the frame, joining others who have already progressed to a higher stage.

Number of Faces Arriving at Each Stage

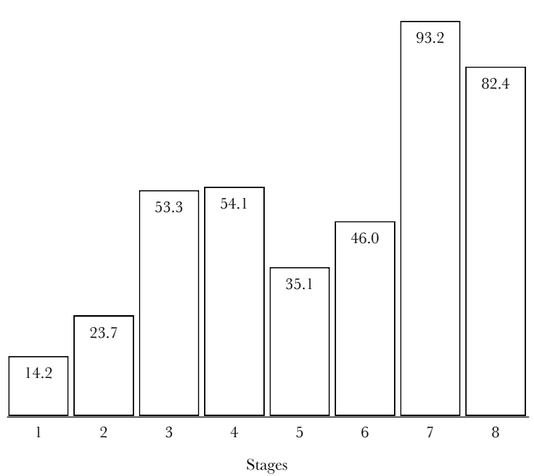


This technical limitation is also expressed in this analysis of the time needed to reach each stage (see chart below).

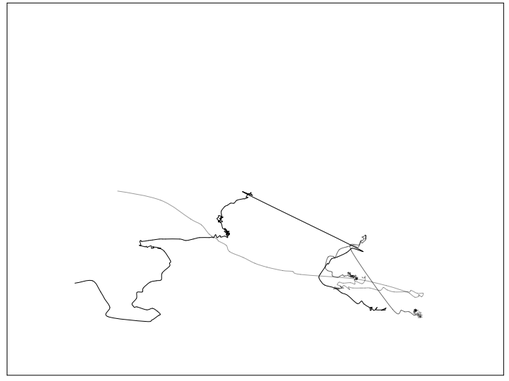
Since the stages are always shown in the same order (five is after four, e.g.), we would expect later stages to always require more time than earlier ones. But, instead of an upward trend, we see some downward dips, representing imperfections in the computer vision capabilities.

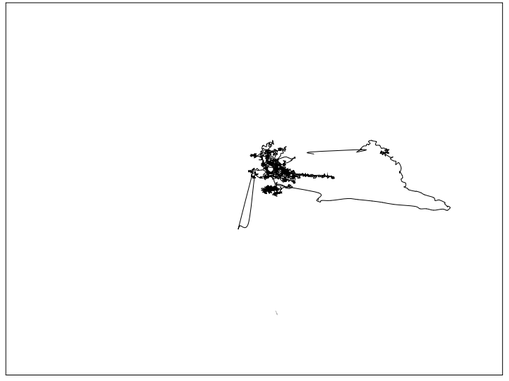
Although these numbers are not wholly accurate, they are still valuable. Participants may have been engaged for several minutes, then turned away for a few seconds, and returned to face the screen — which would have counted as two separate "face sessions," not one. Until computer vision systems can recognize and track people from all angles (and outside of the video frame), it won't be possible to automate this level of data collection around engagement.

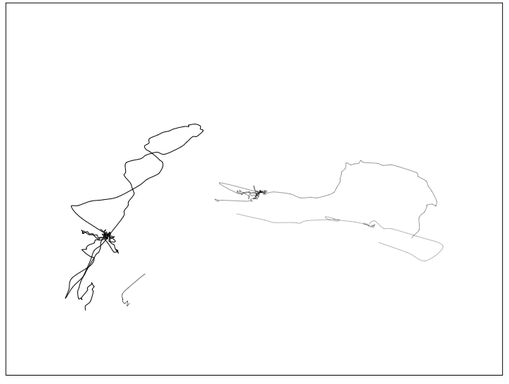
Average Time to Acheive Each Stage (seconds)



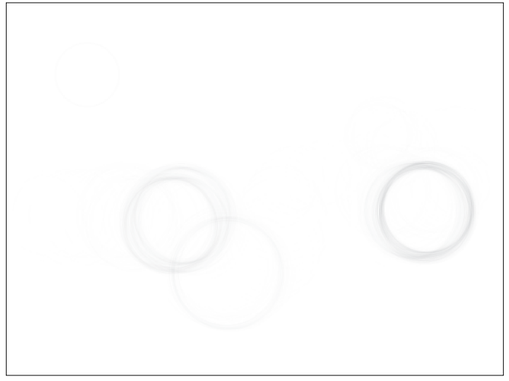
Engagement is not limited to visible presence in any case. In my observations, several people, especially those who came with friends, spent anywhere from 15 minutes to an hour at the installation, engaged either directly with the system or as observers, socializing and communicating with others. This social engagement, that occurred outside the realm of the digital system, could not be tracked by the system, but was an equally important part of the overall experience. The social interactions, along with the physical environment, formed the context in which the interactive system was experienced, and thereby informed participants' tolerance for different forms of interaction. Having to balance internal curiosity (or lack thereof) with external, social stimulation ("Stand still!") and anxieties ("Everyone is watching me up here, and I look stupid!") supplemented the emotional range of an experience that, absent the social dynamic, would be quite different.

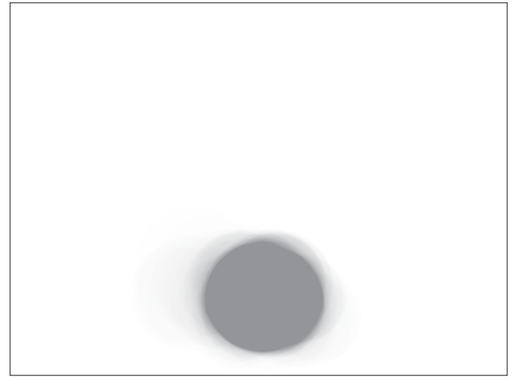


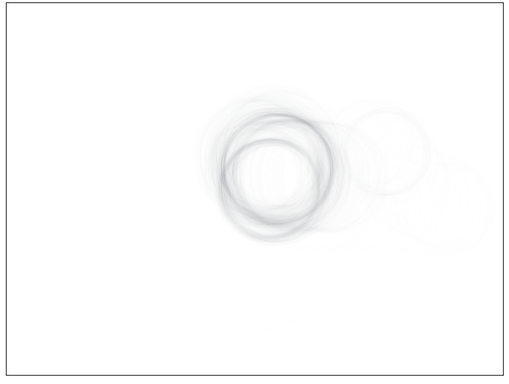


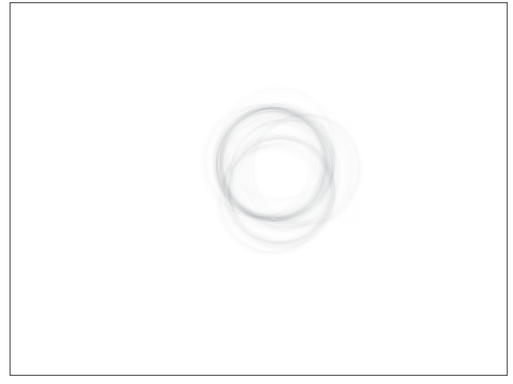


Visualizations of face tracking data. Each line represents one person (or "face") and its position within the video frame.

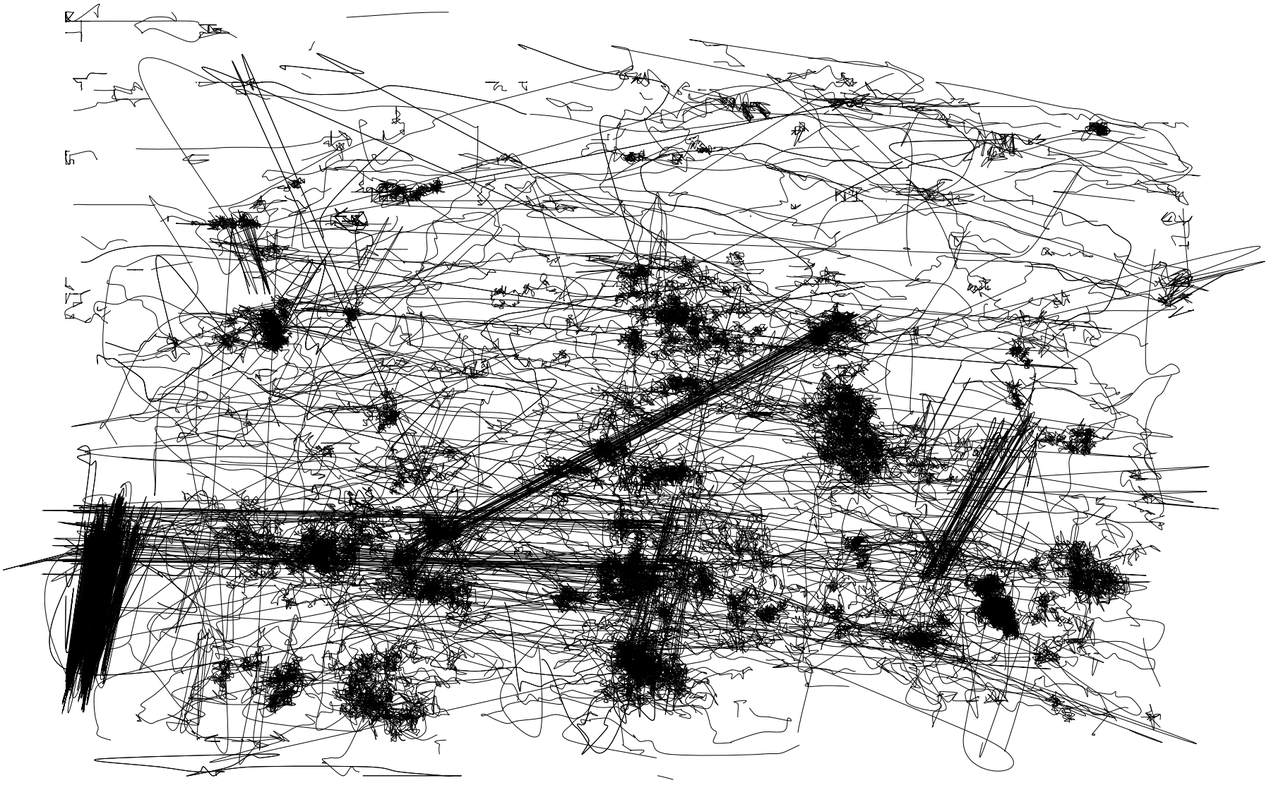


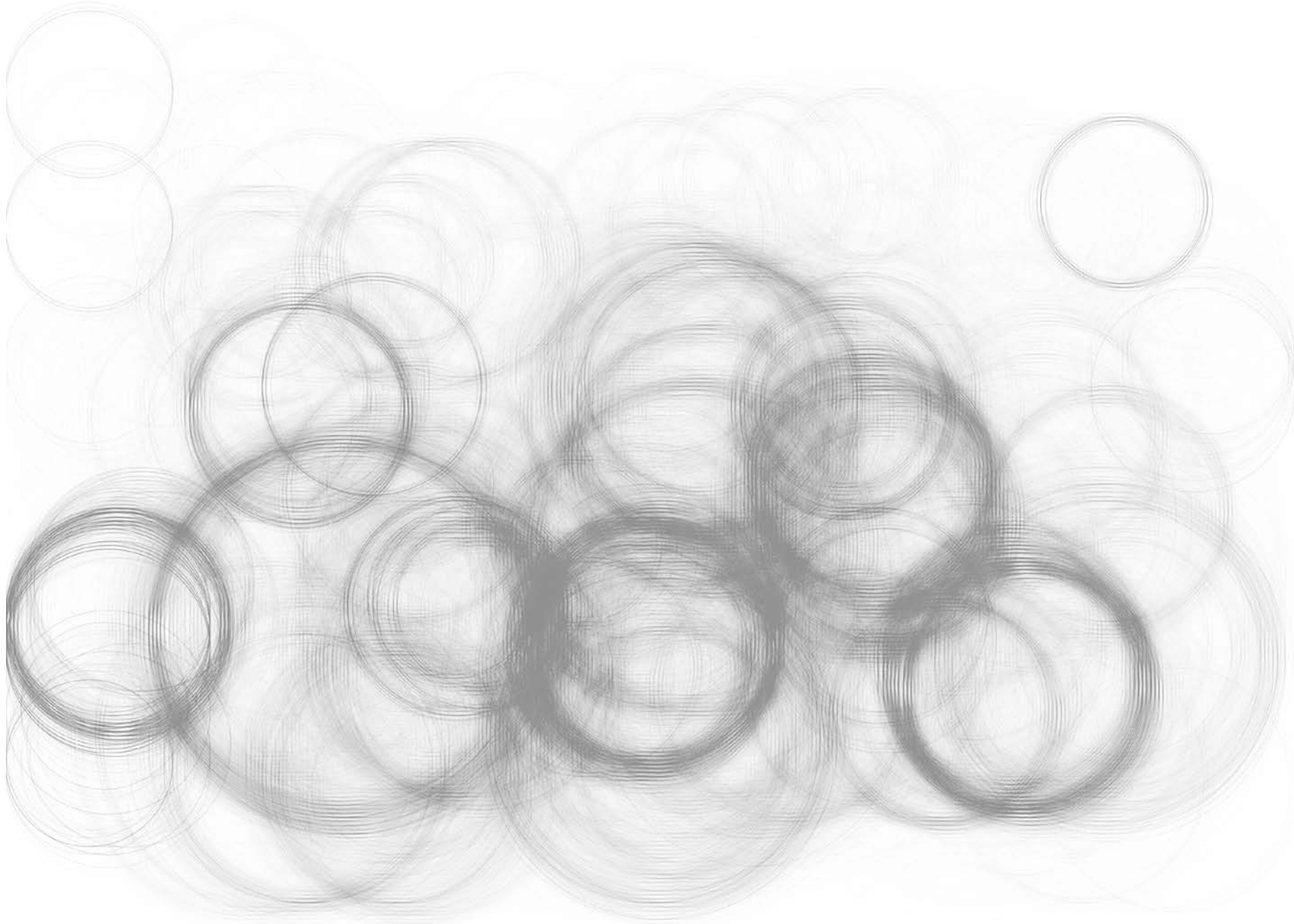






Same as at left, but with the addition of face size, represented by circle size. By looking at size, we can see how near or far participants were to the camera.
Following spread: All face motion paths from the 2-hour installation.





Future

I am satisfied with these findings, and believe that they illustrate how, for the most part, the design successfully engaged participants on many different levels. In a world where video advertisements and motion graphics compete for eyeballs and measure success in duration of participation, I think it's valuable to take a broader view of engagement that accounts for the entire experience, including its physical and social contexts.

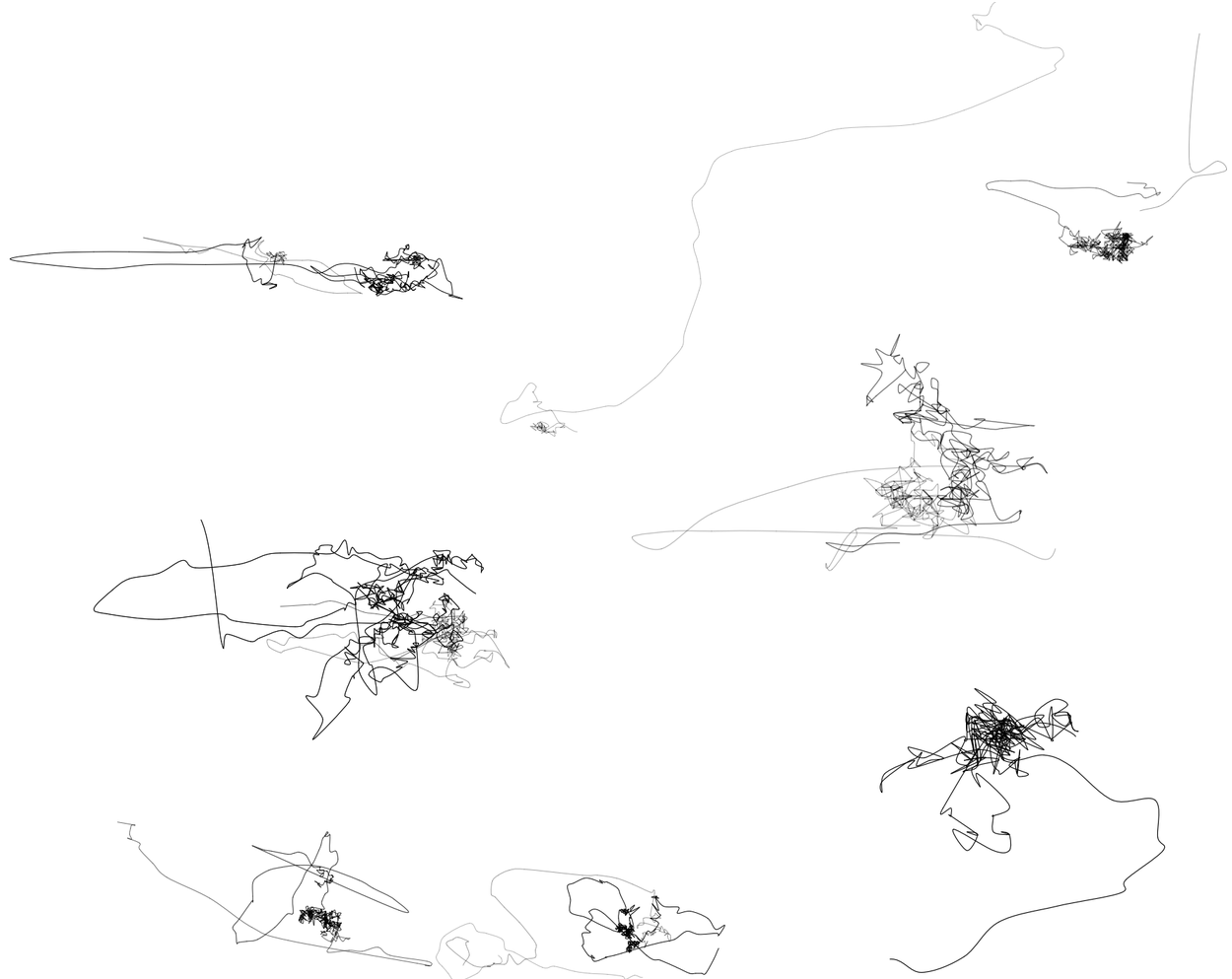
The data can only tell us so much, especially in an uncontrolled environment, such as this busy street in the Mission District. User testing and ongoing refinements to the design are still essential, but future research could focus on expanding the role of automated data collection — perhaps using a second overhead camera to monitor the meta-space around the installation, counting both direct participants and third-party observers, their physical proximity to each other, and correlating noise levels to events in the system. (For example, do cheers erupt upon successful achievements?) An ongoing challenge will be improving the quality of the data collected without dampening the challenge and joy of the overall experience. An highly controlled environment (such as a gallery space) may increase accuracy, but that physical context will trigger very different responses than that of a more chaotic, public space (such as a city street). Both are valuable for study, and each presents its own challenges and opportunities.

Credits

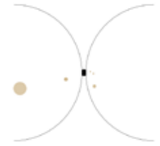
Practice uses data from We Feel Fine (wefeelfine.org), by Jonathan Harris and Sep Kamvar, and incorporates recordings from Freesound (freesound.org) by the following authors: acclivity, chipfork, fauxpress, Freed, Jovica, kerri, suburban grilla, suonho, and zuben.

X	Y	Size	Progress
371.0	310.0	224.0	0.0
367.0	313.0	236.0	0.0
367.0	313.0	236.0	0.0
367.0	310.0	236.0	0.0
365.0	311.0	233.0	0.0
367.0	310.0	230.0	0.0
365.0	311.0	233.0	0.0
364.0	310.0	224.0	0.0050
369.0	311.0	227.0	0.01
366.0	311.0	220.0	0.015
365.0	311.0	233.0	0.02
367.0	313.0	243.0	0.024999999
368.0	311.0	233.0	0.029999997
369.0	311.0	246.0	0.034999996
367.0	316.0	243.0	0.039999995
370.0	316.0	243.0	0.044999994
369.0	315.0	240.0	0.049999993
363.0	312.0	240.0	0.054999992
367.0	313.0	243.0	0.05999999
366.0	308.0	240.0	0.06499999
367.0	312.0	249.0	0.06999999
369.0	311.0	246.0	0.074999996
369.0	311.0	246.0	0.08
369.0	311.0	246.0	0.085
367.0	309.0	243.0	0.09
369.0	311.0	246.0	0.095000006
369.0	311.0	246.0	0.10000001
367.0	309.0	243.0	0.10500001
369.0	312.0	240.0	0.110000014
370.0	303.0	249.0	0.11500002
367.0	300.0	236.0	0.12000002
364.0	306.0	249.0	0.12000002
364.0	303.0	236.0	0.12500001
372.0	299.0	246.0	0.13000001
365.0	295.0	259.0	0.13000001
367.0	296.0	249.0	0.13000001
364.0	290.0	236.0	0.135
374.0	284.0	230.0	0.135
377.0	284.0	236.0	0.135
375.0	285.0	233.0	0.135
367.0	271.0	243.0	0.135
357.0	271.0	217.0	0.135
351.0	274.0	211.0	0.135
353.0	273.0	195.0	0.135
241.0	193.0	182.0	0.16999997
238.0	193.0	182.0	0.16999997
250.0	189.0	169.0	0.16999997
260.0	180.0	169.0	0.16999997
257.0	161.0	131.0	0.16999997
255.0	153.0	166.0	0.16999997
250.0	145.0	131.0	0.16999997
234.0	145.0	131.0	0.16999997
231.0	145.0	131.0	0.16999997
225.0	145.0	131.0	0.16999997
221.0	132.0	131.0	0.16999997
221.0	122.0	131.0	0.16999997
231.0	119.0	131.0	0.16999997
234.0	119.0	131.0	0.16999997
234.0	116.0	131.0	0.16999997
250.0	113.0	131.0	0.16999997
257.0	119.0	131.0	0.16999997
266.0	116.0	131.0	0.16999997

A portion of the face tracking data recorded by Practice. Each row represents the x, y, size, and progress values for one face during each frame that it was detected.



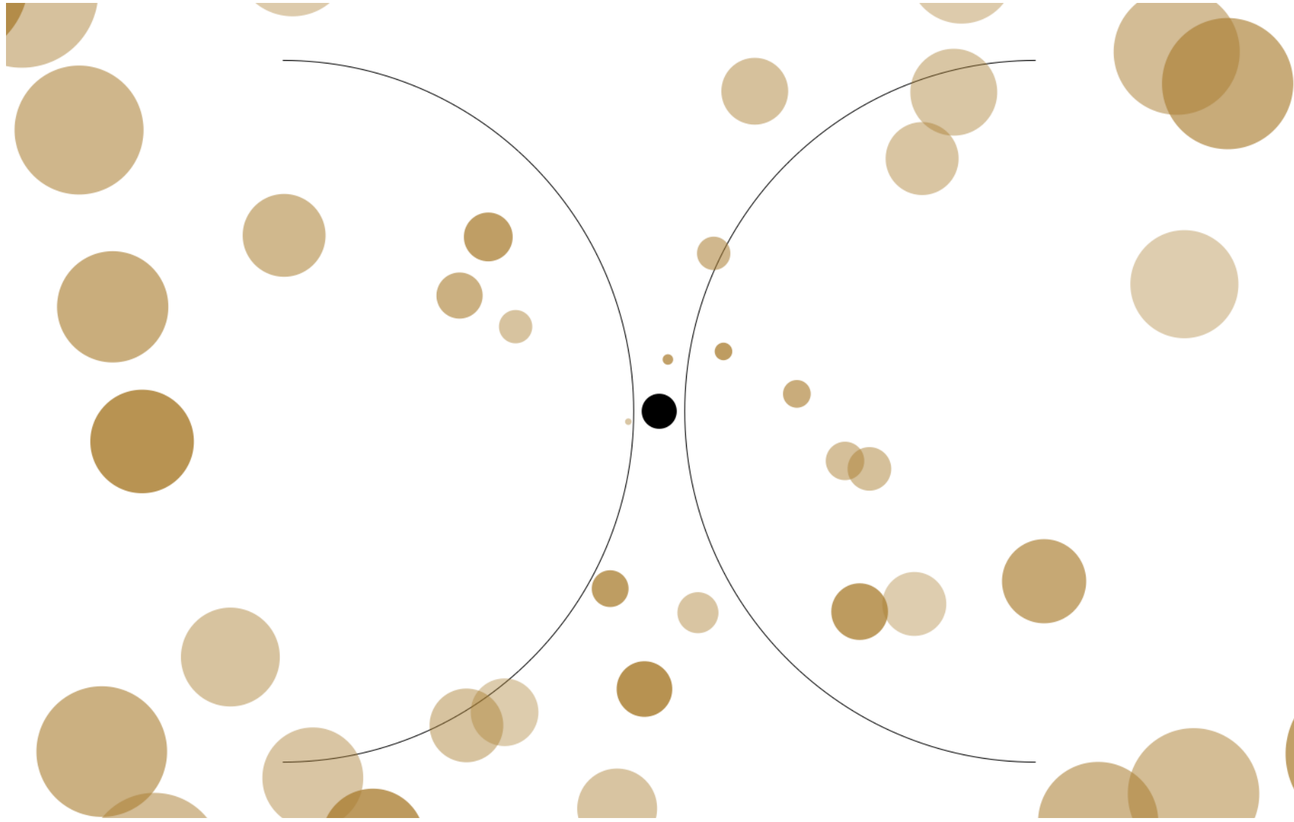
Cheeky



In early January 2010, fellow DMI student Lou Susi announced that he was organizing a gallery show entitled "American Cheese" for the 28th of the month. This was a rare opportunity to apply the conceptual framework and technical skills behind *Practice* toward something less aspirational.

Over the course of two evenings, I came up with a concept, downloaded and edited the necessary audio samples, designed the visual system (which was rudimentary, due to the time constraints), and adapted the face tracking code from *Practice* for use in *Cheeky*. When viewing *Cheeky* from afar, the participant sees a white display with a small double-curve shape at center, sort of like a rounded *X*. As the participant moves in for a closer look, the *X* shape slowly expands, growing roughly in proportion to the user's distance from the display. Face tracking monitors the size of the participant's face, from which it estimates how near or far they are from the video camera and display. A small amount of easing is incorporated, so even if the user jumps up close, the *X* shape will not jump in size, but smoothly scale up over a few seconds.

Upon reaching a target size of about 80% of the display height, the curved *X* separates in two, vibrating rapidly while an offensive, yet instantly recognizable sound is played, and brown, semi-transparent "gas bubbles" are ejected out from center screen and fly away in all directions.

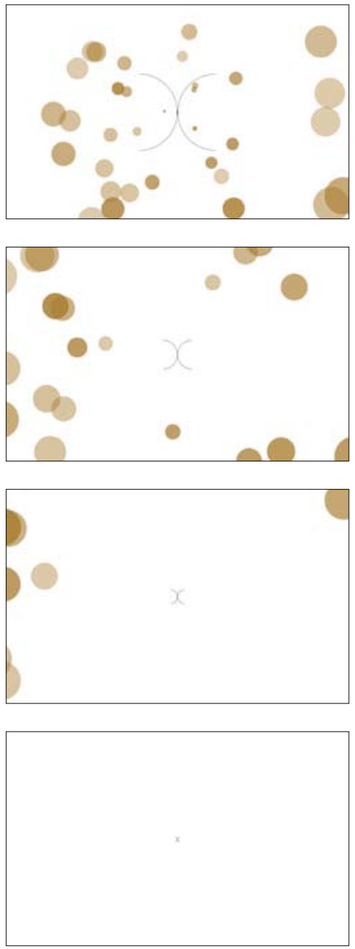


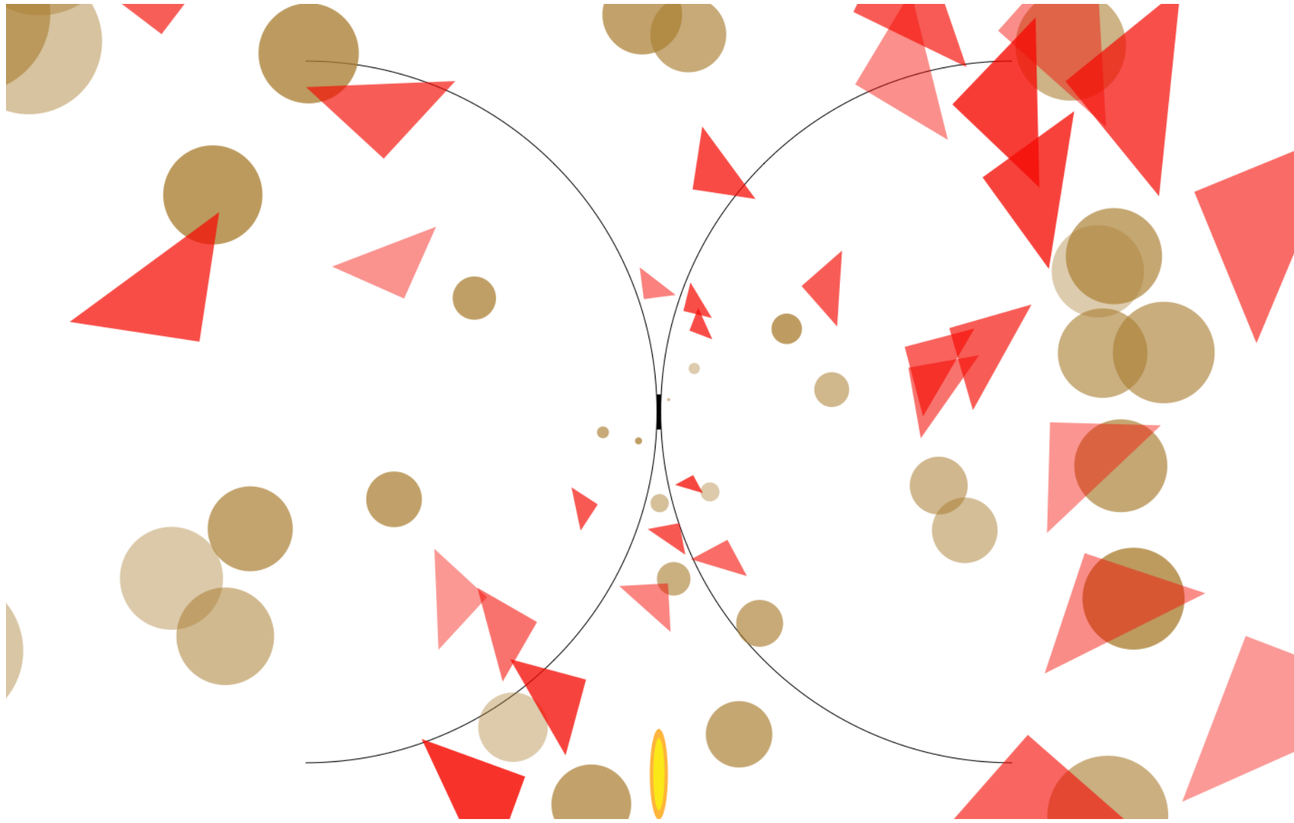
User testing reveals that moment to be a very emotional one, triggering a range of reactions, often all at once, including embarrassment, shame, joy, disgust, love, and giddy excitement upon the realization that a computer has just released a digital fart in the participant's face. The abstract *X* can no longer be seen as anything but two cheeks resting against each other until the next evacuation.

The slow expansion of the *X* builds tension around anticipation of what will happen next. The tension is released in a very visual, highly audible form — a representation of a bodily release that is both very common and socially off-limits (except for eight-year-old boys). So the social context is important, of course, and informs the range of feelings and decisions to be made by the participant in that one airy moment. At a performance art show called "American Cheese," it is probably okay to laugh at a fart joke, and not take yourself too seriously. But what if the work were presented in a more serious context, as though it were high art? Or, what if you are the artist, and your friends who have seen *Cheeky* can't stop recommending it to people you've only just met? How does that recommendation influence the first impression you're in the process of making, and do you position yourself closer toward or further away from the project? Even when *Cheeky* is not on display, it has ways of making me uncomfortable while making me laugh.

Although *Cheeky* is designed for only one active user at a time (which is appropriate, given the intimate nature of the bodily function on display), it employs many of the same design principles as *Practice* to engage third-party observers and become a social experience. In *Cheeky*, the visuals are important, but the audio is critical, since it expands to fill the space around it. At the "American Cheese" show, the display was positioned so that it couldn't be seen right away, but required some navigation within the gallery space to get a good look. So, as I had hoped, the uncomfortable audio caught the attention of uninitiated observers, who looked over and witnessed the active participant, face close to the screen, either shocked and appalled (and, soon thereafter, even further shocked to observe that everyone else was watching him), or laughing uncontrollably, or both. The uninitiated knew, then, that this was some kind of fart machine, but without a clear view of the screen, its interaction model remained a mystery. This curiosity motivated them to experience it for themselves, but in the meantime, the sight and sounds of someone else experiencing the piece was almost funny enough. This is comedy, in dynamic, digital form. And while our social self-consciousness may not be intense enough to induce a satori-like state, it elicits enough discomfort that we just have to laugh. It may not be enlightenment, but it's not bad.

It gets funnier when people come back for repeat performances. In my testing, new users tended to jump back in surprise, or at least look away, trying to connect with the other people in the room. ("Oh my god, can you believe what it just did?" "This is in such poor taste!" "Make eye contact with me so I can communicate to you that I do not approve of this filthy, so-called art!") But those who came





back for more were rewarded with any one of ten pre-selected fart recordings, all acquired from Freesound, the online sound sample archive. In these subsequent interactions, participants may have noticed some of the subtleties of the design, such as how the cheeks' vibrations are synchronized with the audio (so louder noises produce greater cheek separation), as are the speeds of the outgoing bubbles (which move faster for loud ones, and slowly for softer ones). The small, black circle at center is glimpsed only briefly, but its identity and function are unmistakable and cannot be forgotten. The ten sound samples are selected and played in random order. While that randomness helps explain part of what keeps the experience varied, it doesn't tell us why people are willing to (and even excited about) voluntarily approaching a machine that will fart, loudly, in their faces, *even once they know* that it will do so! I guess fart jokes never stop being funny.

Once *Cheeky* was working, the randomized sounds and vibrating cheeks weren't quite enough; I felt it needed one more thing to keep users coming back for more. So I created the "lighter" stage. About ten percent of the time, when a release is triggered, the audio will play, gas bubbles escape, and cheeks vibrate — everything as normal — but with an additional twist: A small "flame" rises from the bottom of the screen, which lights the emerging gases on fire. Corresponding audio is played (*click, click, WHOOOOOOSHHH!*), and red triangles crudely represent the gas that has been set aflame. (Unfortunately, due to a version control issue, the *Cheeky* shown at "American Cheese" triggered the lighter nearly every time, instead of just ten percent of the time.)

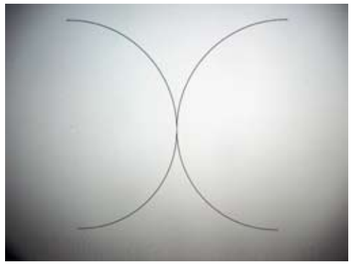
The lighter sequence adds another layer of depth and possibility to the interaction, and provides an incentive for participants to knowingly put themselves in an uncomfortable position, deliberately sustaining repeated blows, until they achieve the reward of the flaming fart. Because the lighter sequence is played at random, users may have to suffer through the humiliation of anywhere from one to twenty or more evacuations. When the reward does come, it is very big and very loud, a release befitting the discomfort and effort required to attain it.

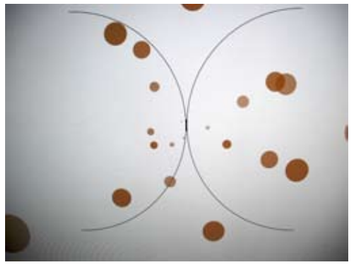
Despite the low-brow content, I hope to have made clear how a number of techniques were consciously used to make *Cheeky* emotionally engaging, especially that of exploiting the tension of anticipation, pairing physical and social discomfort with satisfying rewards, and employing random elements to ensure variety in the experience. In the end, *Cheeky* may actually be more engaging and successful (if less philosophically pure) than *Practice*. It is certainly a shorter experience, since it requires the user to stand still not for minutes, but only moments. *Practice* is slow-moving, deliberate, peaceful, and contemplative, while *Cheeky* is explosive, shocking, humiliating and downright offensive. Yet people love it.

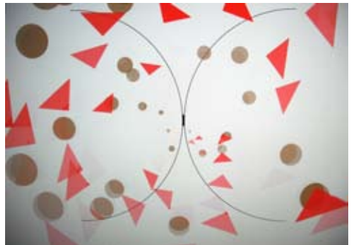
Participants engaging with Cheeky at the "American Cheese" show organized by Lou Susi. Left column photos and top right photo by Lou Susi. Center column photos and bottom right photo by David Tamés.

Credits

Cheeky incorporates recordings from Freesound (freesound.org) by the following authors: elmomo, IFartInUrGeneralDirection, monterey2000, NoiseCollector, scarbelly25, and Walter Odington.







Conclusion

It is difficult to conclude this journey of discovery, when each new insight only leads to more questions. Conclusions ring with finality, yet this process has opened up countless new directions. Nonetheless, this is a good time to consider potential future explorations.

We now know that dynamic systems absolutely *can* elicit experiences that are engaging, meaningful, and memorable. Visual and aural design elements can be used to capture participants' attention and establish anticipation of events to come. Engagement can be sustained by maintaining anticipatory tension while offering periodic rewards, emotional releases won by the discomfort of the challenge. And visual mirroring is perhaps the perfect medium for engagement, given our biologically innate understanding of it: mirroring is the ultimate interface medium, with zero learning curve. By considering the social context of the work, we can design systems that actively engage participants on multiple levels, both with the systems directly and with each other.

Yet more user testing is always valuable. I would like to test both *Practice* and *Cheeky* in a wider range of physical and social contexts. More controlled environments could offer consistent lighting, which, in turn, would elicit more accurate and meaningful data on motion and behavior. As face tracking algorithms improve, this data could become accurate enough to be useful for quantitative studies of human behavior. There may even be applications for psychological research, testing subjects' tolerances for stillness, anticipation, and discomfort.

Practice and *Cheeky*, though, are just two systems born from the same design principles; there is certainly room for many, many more. *Practice* could be expanded with additional stages, and *Cheeky* may be just the first in a series of projects inspired by taboo bodily functions. It may not be high art, but *Cheeky* triggered such powerful reactions that it is worth exploring further. I am fortunate to have had one project that inspires such pure giddiness, and I hope to create many more.

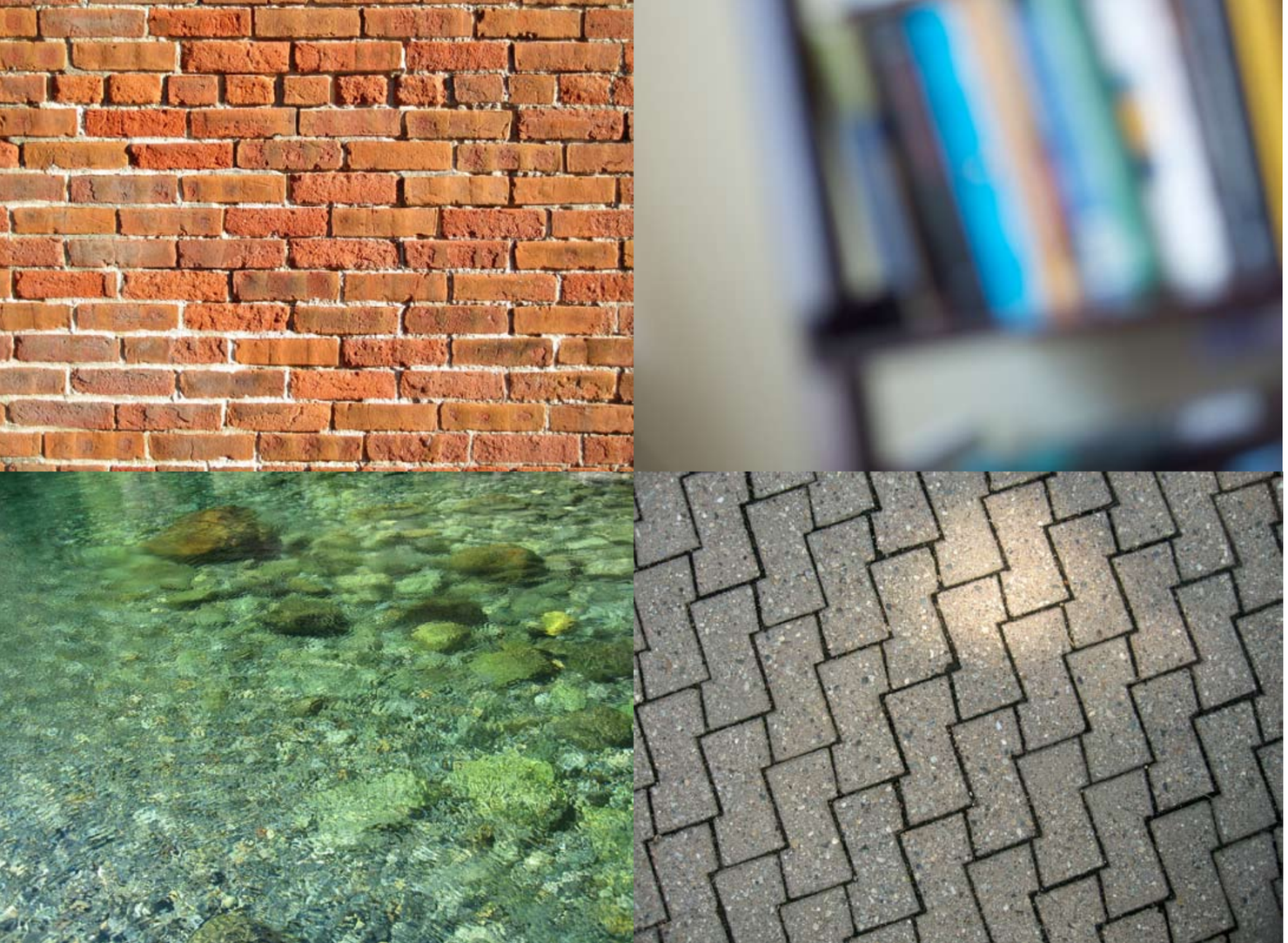
That brings me to another observation: humor is severely underutilized in dynamic media. Both of my thesis advisors loved *Cheeky*, but were surprised that I had created something so outlandishly silly. Apparently, for the past three years I had succumbed to the seriousness of design research and had been treating the entire field of design with the gravity it projected. But humor is too powerful to ignore, especially when we are interested in engagement — hilarity trumps gravity every time. I am grateful to have stumbled upon *Cheeky* before completing my time at DMI, as it's opened the doors to a whole new area of research. The experience has also reminded me of what I so easily forget: that it's acceptable (even preferable) to integrate my full personality into my design work. In short, it's okay to be who I already am — farts and all.

I recently presented *Cheeky* to a class of freshmen design students at the University of San Francisco. It was at the end of a guest lecture, which until that point had been fairly tame. But when a female volunteer first approached the screen and said it looked like "boobs," her visible discomfort amplified my own, and I was suddenly very focused on avoiding a harassment lawsuit. In a way, I was right back at Rainbow Pools, experiencing something unknown, unable to predict how my immediate future would play out. It was another moment of satori, elicited by my uncomfortable predicament.

In the end, I've learned that discomfort should not be avoided but embraced. Every time we feel uncomfortable, awkward, or tense, we have an opportunity to evaluate why and learn something about ourselves. For a designer, these insights are invaluable, offering clues toward how to elicit similar emotions in others. Of course, resolving those moments of tension is essential; we want participants to enjoy our designs, not dread them. But the path from a concept to an engaging design is rarely obvious and straightforward.

For me, a successful design can only emerge from a structured process of describing the problem, designing solutions, observing users, and revising the design in response. My insistence on structure may be a psychological response to my own fear of ambiguity, the uncomfortable state that both *Practice* and *Cheeky* induce in others. Perhaps I find others' experiences of ambiguity so interesting because those experiences are so significant in my own life.

Ambiguity, however, is a particular variety of discomfort that brings with it a special reward: an awareness of the present moment. In those uncomfortable moments, our minds grow quiet, allowing us to focus. When experienced collectively, ambiguity can also bring us together, connecting us with a shared experience. In the end, dynamic systems are not just about engaging users, but establishing human connections and deepening our relationships with each other.



References

Branigan, Edward. *Narrative Comprehension and Film*. London: Routledge, 1992.

Csikszentmihalyi, Mihaly. *Flow: The Psychology of Optimal Experience*. New York: Harper & Row, 1990.

Dewey, John. *Art as Experience*. New York: Paragon, 1979.

Heider, Fritz and Mary-Ann Simmel. "An experimental study of apparent behaviour." *American Journal of Psychology*. 57 (1944): 243–259.

Hutchins, Edwin L., James D. Hollan, and Donald A. Norman. "Direct Manipulation Interfaces." *User Centered System Design: New Perspectives on Human-Computer Interaction*. Eds. D. A. Norman and S. Draper. Hillsdale, NJ: Lawrence Erlbaum, 1986.

Laurel, Brenda. *Computers as Theatre*. Reading, Massachusetts: Addison-Wesley Publishing Company, 1993.

Manovich, Lev. *The Language of New Media*. Cambridge, Massachusetts: MIT Press, 2002.

Michotte, Albert. *The Perception of Causality*. Andover, MA: Methuen, 1962.

Nadin, Mihai. "12 Definitions." anticipation.info. November 22, 2004. Web. Accessed March 31, 2010.

Pines, Malcolm. "Mirroring and Child Development." *Psychoanalytic Inquiry*. 5 (1985): 211–231.

Reas, Casey. Untitled talk during the workshop "Creative Coding: An Introduction to Processing," Gray Area Foundation for the Arts, San Francisco, October 3, 2009. Lecture.

Salen, Katie and Eric Zimmerman. *Rules of Play*. Cambridge, Massachusetts: The MIT Press, 2004.

Sekida, Katsuki. *Two Zen Classics: Mumonkan and Heikiganroku*. New York: Weatherhill, 1996.

Utterback, Camille. "Luscious Complexity: Transcending the Doohickey." The Art, Technology, and Culture Colloquium, UC Berkeley, October 5, 2009. Lecture.

Ware, Colin. *Information Visualization: Perception for Design*. 2nd ed. San Francisco: Morgan Kaufmann, 2004.

Winnicott, Donald W. "Mirror-role of mother and family in child development." *The Predicament of the Family*. Ed. P. Lomas. London: Hogarth Press, 1967.