Music, Mind and Health:

How Community Change, Diagnosis, and Neuro-rehabilitation can be Targeted During Creative Tasks

by

Adam Boulanger

B.A., Berklee College of Music (2004)
M.S., Massachusetts Institute of Technology (2006)

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

Doctor of Philosophy in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September, 2010

Author ....
Program in Media Arts and Sciences
September, 2010

Certified by ....
Tod Machover
Professor of Music and Media
Thesis Supervisor

Accepted by ....
Patti Maes
Chairman, Department Committee on Graduate Students

Music, Mind and Health:
How Community Change, Diagnosis, and Neuro-rehabilitation can be Targeted During Creative Tasks

by Adam Boulanger

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, on September, 2010,
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Media Arts and Sciences

Abstract

As a culture, we have the capacity to lead creative lives. Part of that capacity lies in how something like music can touch on just about every aspect of human thinking and experience. If music is such a pervasive phenomenon, what does it mean for the way we consider our lives in health? There are three problems with connecting the richness of music to scientifically valid clinical interventions. First, it is unclear how to provide access to something as seemingly complex as music to a diverse group of subjects with various cognitive and physical deficits. Second, it is necessary to quantify what takes place in music interactions so that causality can be attributed to what is unique to the music experience compared to motivation or attention. Finally, one must provide the structure to facilitate clinical change without losing the communicative and expressive power of music. This thesis will demonstrate how new music technologies are the ideal interfaces to address the issues of scale, assessment, and structured intervention that plague the ability to introduce creative work into healthcare environments. Additionally, we describe the first neural interface for multisensory-based physical rehabilitation, with implications for new interventions in diverse settings. This thesis demonstrates the design and implementation of devices that structure music interaction from the neural basis of rehabilitation. At the conclusion of this research, it is possible to envision an area where users are empowered during scientifically based creative tasks to compose neurological change.

Thesis supervisor: Tod Machover
Title: Professor of Music and Media

Thesis Committee

Thesis supervisor ....

Tod Machover
Professor of Music and Media
MIT Program in Media Arts and Sciences

Thesis reader ....

Gottfried Schlaug, M.D., Ph.D.
Associate Professor of Neurology
Beth Israel Deaconess Medical Center, Harvard Medical School

Thesis reader ....

Hugh Herr, Ph.D.
Associate Professor
MIT Program in Media Arts and Sciences, Health Sciences and Technology

Acknowledgments

This thesis was made possible by the care and support of many friends, colleagues, and mentors. In particular, I'd like to thank the following people:

My advisor, and friend, Tod Machover, who has supported this work since the beginning, providing tremendous opportunities to entertain ideas, applications, and new directions.

My thesis readers Hugh Herr, and Gottfried Schlaug.

Fellow students of the Hyperinstruments Group, past and present.

Frank Moss, Edith Ackermann, and Hiroshi Ishii,

My collaborators, Dan Ellsey and the Tewksbury Team,

My family,

And most importantly, Cati Vaucelle, my love, and Anaïs, mon coeur.

Table of Contents

List of Figures

CHAPTER ONE

Introduction and Motivation

Healthcare is part of everyday life. The decisions we make regarding daily activities, diet, education and exercise are well studied as to their direct contribution to the prevalence of disease and divergent mortality across cultures[Mokdad et al., 2004, Khaw et al., 2008, Stringhini et al., 2010]. However, beyond lifestyle factors, what about our lifestyles in value? What are the things we care about, the things that we, as a society, are invested in? If healthcare is part of everyday life, there are currently no signs of connecting healthcare to the aspects of daily life that people find personally fulfilling.

I imagine a future in which healthcare is seamlessly integrated into the areas of life that we are the most passionate about. Personally, this means music. We are at our best as a culture when we are being creative, actively pursuing something to enrich our lives or the lives of others. These moments are the opportunities to prevent, to improve, or overcome disease. Individuals can become empowered managers of their own health as in the moments of life they are personally invested in.

The scientific dilemma is how to embed valid, meaningful healthcare practice into elusive creative experiences. The throes of creative experience could not be more antithetical to the process of clinical intervention, especially in the face of research standards including randomized controlled trials, and placebo control studies. The problem is exacerbated the further one defines the tier of healthcare under consideration. Diagnosis, intervention and community-care are vastly different aspects of an overall system. Each requires different handling to discover the inroads by which clinical practice can be connected to creative work.

Creative work, without significant rethinking, is not appropriate for incorporation into the methodologies that substantiate clinical research. Before embedding any kind of health experience into creative activities of daily life, do we have the right tools to be creative?

Any new platform for clinical intervention must satisfy the ability to scale within the population. This is largely a problem of access. Many existing treatments, despite their efficacy, cannot be introduced to the population at large due to cost or lack of infrastructure [Richardson, 2001]. What is our current creative infrastructure? Educator John Holt challenges Suzuki method[Suzuki, 1983] tenets such as the necessity to learn an instrument at the earliest possible age [Holt, 1989]. And yet, if an adult wants to learn an instrument, will they easily find the time and structure to overcome the initial hurdles to produce music?

New tools can simultaneously provide access to creative experiences to diverse users, present an experience where participation targets or is driven by the cognitive or physical regions of interest for diagnosis or rehabilitation, and finally, structure the experience so that as the user evolves in their creative work, they simultaneously evolve in the clinical intervention. This thesis presents the research and technology to address each of these necessary criteria for capturing valid and meaningful healthcare as part of creative work.

1.1 Creativity - an overview

Creativity is the ability to produce work that is both novel and appropriate [Sternberg and Lubart, 1999]. The origins of one's creativity remains a lifelong debate. However, researchers have found a correlation between open-ended play and imagination in writers, poets, and scientists [Singer and Singer, 1990]. David Edwards, writing on the intersection of art and science, cuts a swathe through examples where - evoking Bronowski[Bronowski, 1965] and Kuhn[Kuhn, 2009] - scientific revolutions are engaged by creative aptitude rather than application of the scientific method[Edwards, 2009].

Rather than understanding how to be creative, motivated, and innovative [Amabile, 1996], throughout this research I am investigating design solutions to access creativity.

1.1.1 Constructionism

Developing students' creativity and their potential as creative thinkers is one of the most important goals of education [Baer, 2005]. Researchers investigate design creativity training activities for education [Baer, 1996, Baer, 1997, Baer, 1998] and have found that creative individuals are usually polymaths who think in trans-disciplinary ways [Root-Bernstein, 2001].

To enable the access to creativity, I design interactive applications for anyone to create music. I ground this research on the field of *constructionism* [Harel and Papert, 1991a]. First defined as *learning by making*, constructionism relies on the notion of perspective-taking in learning [Ackermann, 1996] and the interaction between both cognitive and affective processes as a central role in building connections between old and new knowledge. Researchers in constructionism analyze the relationship between designing and learning to provide personally meaningful contexts for learning. They often employ new technologies to provide access to gaming or software design experiences, at the service of learning [Harel and Papert, 1991b].

1.2 Music - the ultimate interface

Music is the kind of environment that people devote themselves to. Music, both within clinical environments and, simply, in the world, has the potential to be highly motivating. We care deeply about music as a culture and as individuals. It is inseparably woven into the fabric of everyday life.

In the past decade, our scientific understanding of music has accelerated. Neuroimaging, in particular, has provided insight into the the way music touches on just about every aspect of thinking imaginable. Leveraging both the science, and the art, a rich field has emerged pertaining to the clinical utility of music. However, both of these domains remain far afield from one another. Efforts to capitalize on targeted neurological interpretations of music often fail to capture the richness of music, in favor of something overtly more definable from a scientific standpoint.

In contrast, clinicians currently lack the tools to introduce music at scale, or with the right assessments to attribute clinical outcomes to the music taking place in complex and highly social interventions.

Finally, music may be the best environment to establish an intervention that has significant value beyond encouragement, or assessment. And yet, a framework to capitalize on the latent scientific and creative opportunities in a reproducible and empowering way does not exist.

1.3 Music Therapy - current clinical uses of music

The American Association of Music Therapy defines music therapy as follows:

> clinical and evidence-based use of music interventions to accomplish individualized goals within a therapeutic relationship by a credentialed professional who has completed an approved music therapy program.

Music therapy is often sculpted from the raw materials of instrument and voice interactions into fledgling therapeutic relationships. Interventions, in the best practices of music therapy, are highly dependent on social dynamics and relationship building such that the American Music Therapy Association defines the profession in terms of the therapeutic relationship. Music excels in this respect. It can be non-threatening. Patients tend to have personal and positive experiences with at least some kind of music. Furthermore, music interactions tend to give a patient more control over their environment than is typically afforded by clinical treatments. The personal relationships that people have with music, and the flexibility of music as it is composed or performed to fit the needs of a particular patient, make for personalized treatments.

Finally, music can be fun. It can also be challenging, deeply serious, and even emotionally straining. The ability to connect to emotion, and to employ it purposefully around music interactions further contributes to relational and social dynamics.

Clinical domain	Goals	Techniques
Special education	Social/emotional, motor skills, communication, pre-vocational skills, preacademic skills.	Behavioral
Psychiatric disorders	Social interaction/social awareness, reality orientation, group facility, social/emotional, awareness of behaviors, behavioral change.	Insight-oriented/process therapy, psychoanalytic, behavioral, activity-oriented
Institutionalized elderly	Strength/range-of-motion, social interaction, stimulate memory, reality orientation, social/emotional, communication.	Behavioral
Autistic spectrum disorders	Language development, social/emotional, pre-academic, sensorimotor.	Developmental/behavioral, neuropsychological
Neurological rehabilitation	Strength/range-of-motion, language, cognitive reasoning, executive function, problem solving, communication, adaptation, activities of daily life.	Neurological techniques (melodic intonation therapy, rhythmic entrainment)
Correctional psychiatry *Sensory disorders* *Physical disability* ...		

Figure 1-1: Therapeutic goals and techniques employed in the clinical use of music.

Although it is difficult to generalize from the multiple models of music therapy, with vastly different treatment styles, treatment techniques, and goals (Figure 1-1), consistently, music therapists use music as a catalyst for constructing therapeutic relationships. As relationships around music and music making develop, therapists leverage the relationship to address clinical goals and objectives. The target clinical goals and objectives are derived from the patient's disease process, and treatment program[Hanser, 1999]. Therapists may target speech symptoms in singing tasks[Wan et al., 2010], motor symptoms in instrument playing[Davis, 1999], or psychiatric symptoms in song writing groups [Cassity, 2006].

Since 1994, Medicare supported partial hospitalization programs have reimbursed music therapy as long as it fits the following definition of *active therapy*:

- Be prescribed by a physician
- Be reasonable and necessary for the treatment of the patient's illness
- Be goal directed and based on a documented treatement plan
- The goal of the treatment cannot be to simply maintain the patient's current level of functioning. The individual must show some level of improvement.

These criteria can be considered best practices for the clinical use of music both within partial hospital programs and in other domains of treatment.

Insurance reimbursement codes (Figure 1-2) have been used to describe music therapy treatment objectives and seek third-party insurance reimbursement for services. However, the majority of music's use in clinical environments is not reimbursed by third-party payment, but rather, paid for as an *out-of-pocket* expense by consumers. Also, music therapists are often employed by recreation departments of healthcare institutions, where, although the music therapist may adhere to their educational backgrounds and expertise to conduct interventions, their role in the larger hospital structure is to provide increased quality-of-life and engaging activities for the patient population.

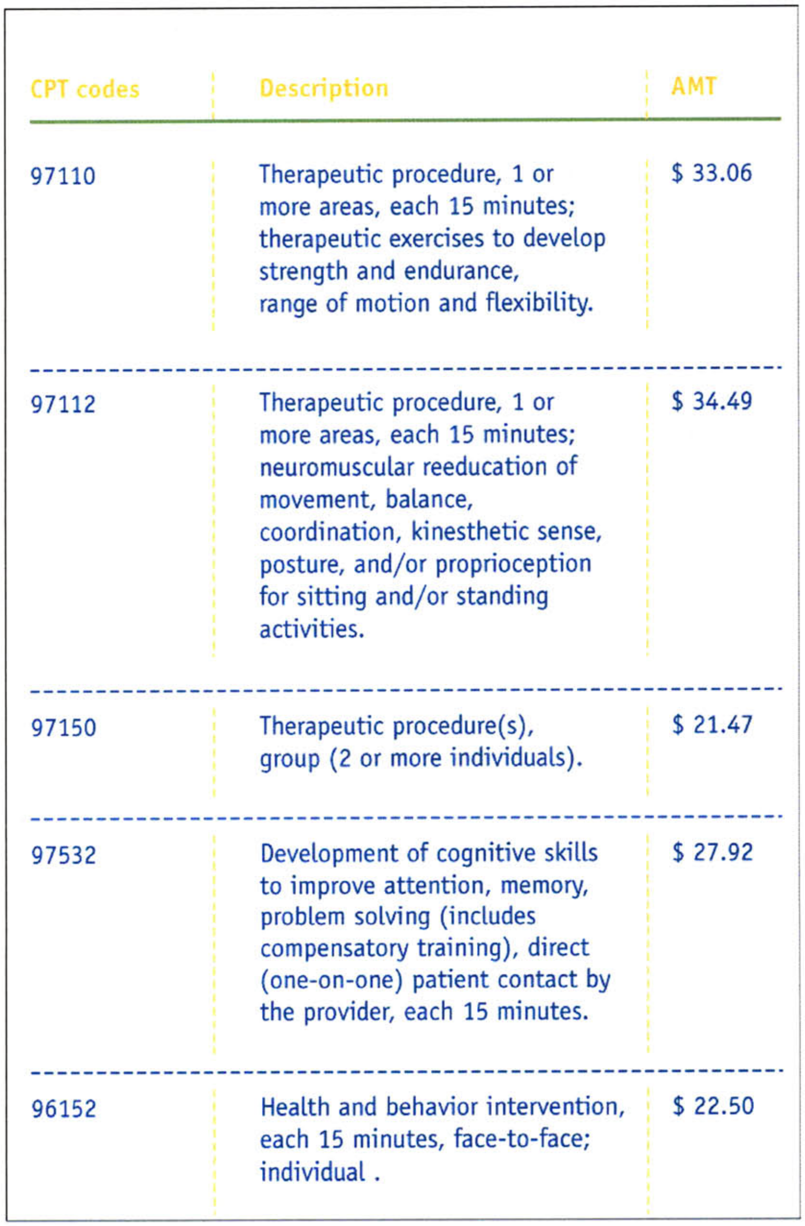
1.3.1 Assessment

Music therapy assessment strategies abound. In best case scenarios, therapists working in healthcare institutions such as hospitals and rehab units are integrated into treatment teams consisting of other practitioners including nurse staff, physical therapists, occupational therapists, and, usually, a physician. The team considers advances over disease displayed in the subject's therapies, updates the patient chart directly, and then recommends new goals and objectives as a group. Music therapists outside of healthcare institutions adopt the methods that best suit their purposes. Many seek insurance reimbursement, which requires assessment to demonstrate progress in standardized goal areas.

Despite these assessment models, it is difficult to determine whether music is responsible for driving any observed cognitive, physical, or behavioral changes, or whether, the relationships built around music interventions are the causal variable. Current research hardly addresses this question. Extracting the social variable in a blinded, controlled study, would require a suitable control for music without a therapeutic relationship. However, this hypothetical control, without an interpersonal dynamic, would restrict the resultant content and creative affordances of the music control to an extent that would make direct comparisons between the two music groups anecdotal, at best. Music therapy is facilitated by practitioners. Relationships are inherent. Establishing research control is always a struggle.

Even if the relational components were not critical to the success of current clinical music practices, assessment would still be the limiting factor of the field. Music is simply too complex. At any given time in a clinical music environment, a patient can be operating on entirely different features of an overall music experience. On some level, music is just a collection of notes, even before that, pressure waves, impinging on the inner ear. Simply listening, we perceive notes, groups of notes, motives, layers of harmonic and melodic structure, sections, movements, and eventually pieces. Our percept is assembled from constituent, and yet, interdependent musical elements. The situation is only more complex as an individual performs, or engages with music. In the moment, she perceives and executes these structures, with the addition of an expressive intention, and ensuing musical result.

When a therapist facilitates the creation of music with an individual, it is impossible to extract the salient feature of the overall music experience that



CPT codes	Description	AMT
97110	Therapeutic procedure, 1 or more areas, each 15 minutes; therapeutic exercises to develop strength and endurance, range of motion and flexibility.	$ 33.06
97112	Therapeutic procedure, 1 or more areas, each 15 minutes; neuromuscular reeducation of movement, balance, coordination, kinesthetic sense, posture, and/or proprioception for sitting and/or standing activities.	$ 34.49
97150	Therapeutic procedure(s), group (2 or more individuals).	$ 21.47
97532	Development of cognitive skills to improve attention, memory, problem solving (includes compensatory training), direct (one-on-one) patient contact by the provider, each 15 minutes.	$ 27.92
96152	Health and behavior intervention, each 15 minutes, face-to-face; individual .	$ 22.50

Figure 1-2: Example CPT codes frequently used by practitioners who incorporate music into clinical practice.

may be responsible for the observed benefit. Again, the controlled experiment would require music without a particular feature. Stockhausen would approve of such environments. However, direct comparisons between music with metric rhythmic structure (or whichever feature is the variable in question), and music without would be inappropriate. The two environments would not be equivocally perceived as music. Music is not simply the additive sum of its parts.

1.3.2 Scale

The clinical use of music is propagated by music therapists. According to the American Music Therapy Association, there are currently several thousand licensed, practicing music therapists in the United States[AMTA, 2009]. Music performance and composition takes place in facilitated environments. Clinicians guide patients through the use of music. They lead groups of performers. They guide patients to introduce motives, melodies, or lyrics into music improvisations. They rely on their own talent as performers to buttress the musical material offered by patients so that it integrates into the context of the improvisation. They also actively maintain a dynamic where patients cannot fail despite lack of musical talent or background.

The ideal music facilitator is uniquely capable of entering into almost any clinical environment, immediately making connections with patients, and engaging in music. Music quickly gets wrapped to the patient's needs, and treatment begins. However, this model is limited by its inability to scale. Personnel are expensive, and music is resource intensive. Clinical uses of music cannot be prescribed because the intervention can only be administered by talented musicians. Very few musicians within the overall musician community have the requisite skills to improvise in clinical scenarios, making great music out of whatever complete novices give you as material.

If a hypothetical comparative analysis of music therapy's efficacy reveals that music therapy is better than current treatment alternatives for something like autism, the disease population cannot be effectively treated by the number of people capable of facilitating music in any meaningful way.

As a limited number of practitioners conduct group sessions with increasingly large numbers of patients, it becomes impossible to adapt the music experience to incorporate the contributions of individual participants with sufficient frequency to provide specific treatment. Clinicians lose the ability to tailor the musical experience to meet targeted patient needs. The interventions deteriorate into mere activities.

Other than distributing treatment to a diseased population, issues of scale bar the integration of clinically useful music into home environments. This is perhaps the most unfortunate aspect of the scale issue. Despite that one can pursue piano, finally, at any age, the connection of the effort to meaningful health goals is restricted by access to specialized practitioners. We want to embed healthcare into creative environments to enable access. The notion that one cannot take advantage of the opportunity for lack of a music therapist, in the home, to guide one through the process is simply unacceptable.

1.3.3 From designing interventions to designing applications

Due to the social aspects of collaborative music making, it is difficult to introduce a music intervention at scale and develop an experimental or assessment paradigm that can answer what is unique about music within an intervention. Arguably, these questions are not paramount in a behavioral model where an intervention is evaluated largely based on its outcomes in the target population[Madson, 1987]. However, the research developed in this thesis enables *embedded healthcare*, where empowered consumers manage their health as part of their daily lives and through the use of applications that have significant secondary value in addition to health benefits. Embedding healthcare into applications with significant secondary value is a shift from interventions to applications.

An application distributed in a population implies a working solution to issues of scale. A central tenet of this thesis is that creativity is a universal opportunity, given the right tools. Music is a platform for creative work that impinges on our health. There is no reason that access to creative experience need be relegated by music professionals, if the right tools can sufficiently empower motivated consumers.

Furthermore, I want to distance music, mind and health from behavioral paradigms to be able to conduct music-based intervention, rehabilitation, or even diagnosis at a scientific level that is competitive with any clinical standard available today.

1.4 Music as a biological science

Beyond the storied creative value of music, the past two decades have garnered a scientific revolution in our understanding of the neural basis of music perception and production [see Peretz Zatorre, 2005, for review]. Music touches on diverse cognitive faculties from memory [Särkämö et al., 2010, Zatorre et al., 1994, Janata, 2009], to movement [Zatorre et al., 2007, Lahav et al., 2007, Janata and Grafton, 2003], and language [Fedorenko et al., 2009, Schellenberg and Peretz, 2008]. Compelling, scientifically based interventions are emerging that rely on music's distribution across the brain to drive the rehabilitation of speech after core language areas have been ablated as a result of stroke or traumatic brain injury[Norton et al., 2009, Schlaug et al., 2009]. Similar interventions are being developed for language acquisition in autism spectrum disorders. The mechanism of these interventions is to draw on intact music processing regions of the brain to initially compensate, and through training, eventually subsume the functionality of dysfunctional areas of processing.

When clinical intervention is centered on music processing it tends to be fairly unmusical. Often, the richness of music is disassembled into constituent elements that are intended to overlap with the target dysfunctional domain.

1.4.1 Right-brained musicians and Broca's aphasia rehabilitation

In language based music rehabilitation, speaking is rehearsed as sequences of simple pitched tones; music is decomposed to an exercise in pitched language [Albert et al., 1973]. These methods are currently being explored for language development in autism [Wan et al., 2010], and have met with significant success in rehabilitation of language in aphasic patients, post-stroke or traumatic brain injury.

Patients with Broca's aphasia have difficulty initiating speech, despite knowing what they want to say. A patient with severe aphasia might perseverate on a single vowel, "o...o...o...," rather than successfully executing a phrase,"Happy

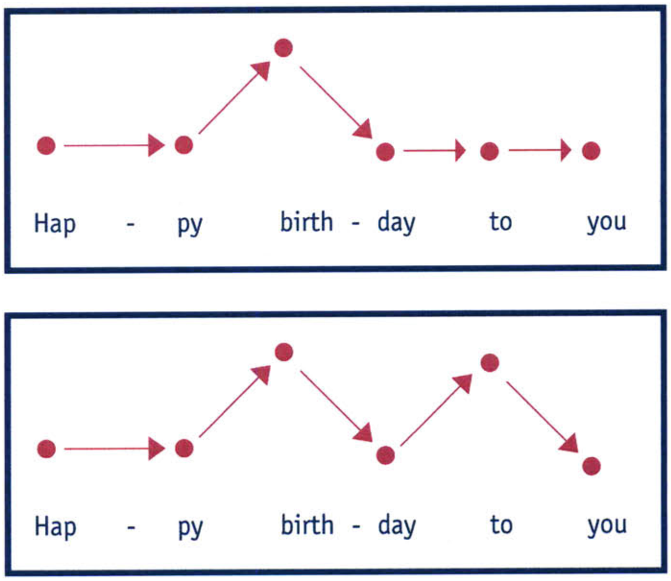


Figure 1-3: Happy *birth*-day, in a typical conversation, might stress the third syllable. The same phrase as a target in melodic intonation therapy alternates between pitches in a "sing-song" manner.

Birthday to you." In moderate cases, the patient belabors over a sentence, only to execute a few nouns, in a disjointed statement that lacks various functional grammatical elements and morphemes, "you...me...go...go...tuesday..."

In melodic intonation therapy, patients with Broca's aphasia rehearse singing pitched target sentences. Singing is not the correct word to attribute to the intervention, which typically selects only two different pitches to be intoned over the course of the target sentence while the patient simultaneosly taps their left hand to the meter of the sentence's syllables (Figure 1-3). After an intensive rehabilitation lasting between forty and seventy-five sessions, patients doubled or tripled the number of information-containing words that they produce in a target phrase, outperforming a control patient that underwent repetition-based language rehabilitation without intonation. Furthermore, performance transferred to various untrained language measures[Schlaug et al., 2008].

The neurological basis of melodic intonation therapy has made it one of the leading examples of the clinical use of music with a basis in biology. The patients do not make wonderful music at the conclusion of treatment, nor are they even thinking of the treatment as a musical endeavor. They are, however, engaging the right, intact homologue of their abated, left hemisphere.

Hemispheric specialization for auditory processing has long driven generalizations regarding *right-brained* musicians. Research has established that spoken language is predominantly perceived, constructed and produced via the engagement of right hemisphere structures, such as Broca's area, and Wernicke's area. Pitched melodic content, phrasing, meter, and sung speech, are predominantly processed in the right hemisphere. These musical features are part of both music perception and expressive language.

Intensive melodic intonation therapy drives structural and functional changes in the homologous right hemisphere of patients with Broca's aphasia. The current theory is that the intact, right hemisphere is bootstrapped, or perhaps, re-specialized, to function similarly to the language centers of the disordered left-hemisphere[Schlaug et al., 2009].

Hemispheric specialization for various auditory content is employed during the perception of music. The same mechanisms are relied upon for the foundation of an innovative intervention. Our experience of music is much more than the perception of intonation differences, but nonetheless, both music and the intervention rely on the same structures.

1.4.2 Harmonic perception and Alzheimer's Disease

A frequently cited, albeit anecdotal, observation from the clinical use of music is that patients with even the most severe cases of Alzheimer's Disease can often remember the music of their early childhood. These patient's may not be able to recognize their environment, or family-members, but if you sing or begin to play the songs buried in their long-term memories, they can orient to the music and interact, recalling words, melodies, and often, the song in its entirety.

The medial prefrontal cortex has been identified as a region of the brain that tracks melodic contour as a melody passes through major and minor keys in Western tonal music [Janata, 2005]. The medial prefrontal cortex is one of the last areas of the brain to atrophy as part of sweeping whole brain changes that take place in the advanced stages of Alzheimer's Disease. Recent studies have further connected the medial prefrontal cortex to *binding* processes between music and memory [Janata, 2009]. The medial prefrontal cortex is a complex center of executive function, additionally supporting emotion regulation, autobiographical memory, mental states, and representations of the self. If such a region binds between these functions, while supporting music structure processing, it may be a single site responsible for the unique status of music in the Alzheimer's patient's life.

This foundational research explores a link between a particular region of music processing and the etiology of Alzheimer's Disease. Further study will be needed to directly correlate the facility of late stage Alzheimer's patients for autobiographically meaningful music with medial prefrontal cortex function. Furthermore, one can imagine the design of an intervention to engage intact music processing in the medial prefrontal cortex to serve *rehabilitation*, or, in the case of a neurodegenerative disease, perhaps, compensation, of diseased memory and reality orientation.

However, neither the study nor intervention currently exist to make the direct connection between these innovative music neuroscience findings and clinical work. As a result, the ongoing music therapeutic work is, at best, informed by the research, without gaining the benefit of direct correlation between research and proposed treatment. Neuroimaging studies have identified a region of interest pertaining to music processing, and the region also relates to the progression of Alzheimer's Disease. The observed symptoms and proposed music function may be related.

Whereas the empowering social environment of music is unwieldy to attribute causal relationships between observed clinical change and elements of music, the neural basis of music is yet to be employed in a way that enriches people's lives while structuring brain plasticity. The concept of healthcare being part of everyday life requires new interventions that can both empower individuals and communities while targeting known underlying mechanisms of disease.

1.5 The opportunity defined exclusively from the standpoint of healthcare

Examining the role of music and healthcare belies a larger opportunity in the healthcare system. Healthcare is changing. Consumers are becoming more empowered. The nature of the healthcare institution, and its reimbursement infrastructure, has been significantly reformed. Healthcare is a key component of national political dialog, and yet, we remain as unhealthy as ever[Docteur and Berenson, 2009].

Staggering costs, inefficiencies, and the inability to accommodate aging populations in its current form, the United States healthcare system is desperately in need of innovation. It is possible to define the opportunity for novel, creative tools for healthcare exclusively from observations regarding the current domains proposing to ameliorate the difficulties in a failing system.

1.5.1 Recent innovations in social health

The next revolution in healthcare is social medicine. Mobile health, home-based assessment, sensor-based patient monitoring are all poised to be disruptive forces. Several technology developments highlight current approaches to capitalize on the emerging social health movement.

First, many top-tier hospitals maintain some kind of innovation-centered, in-house, technology development trying to capitalize on various technology trends from electronic medical records, to hospital systems integration, medical device development, or various application extensions of bioinformatics. As an example, Children Hospital Boston's Informatics Program (CHIP) has developed projects ranging from mobile epidemic tracking[Brownstein et al., 2010], bioterrorism surveillance[Reis et al., 2007], electronic medical records [Mandl et al., 2007], and algorithms for prognosis modeling[Liu et al., 2006]. The projects are all remarkable, vastly different from one another, and developed by small teams working within the hospital environment.

A traditional hospital-based innovation program centers on technology licensing and intellectual property management. A group like CHIP privileges multi-institution relationships to gain patient access throughout the development process, while recruiting developers as graduate students or research fellows. CHIP is an informatics program, and the research questions stem from molecular biology. However, recent CHIP projects have utilized new platforms, such as the web, to translate informatics into the hands of technology-empowered communities. Healthcare institutions have the opportunity to become centers of innovation as they converge patient access, physician expertise, strong applied research standards, and fundability from various sources.

Second, the electronic medical record has been a central focus in recent healthcare re-form efforts, as well as multi-billion dollar government reimbursement initiatives for outfitting existing infrastructure with *meaningfully used* technology[Detmer, 2010]. Database initiatives such as Microsoft's HealthVault have shifted from medical health record implementation towards managing the transfer of health information from various devices and sources into centralized data repositories, where consumers define the parameters of sharing and access both with practitioners, and, ostensibly, other communites. Whether from schools, work, healthcare institutions, or home diagnostic devices, consumers are compiling and managing the transport of their information to various service providers and practitioners, incorporating appointment management, medications, and assessment measures, with *wellbeing* and *life goals* situated amongst the traditional healthcare metrics [Schemas, 2010].

Third, home healthcare assessment and diagnostics have been lacing sensor networks into the home since the beginning of ubiquitous computing. Established groups in firms such as Intel and Philips have major research initiatives in sensor-driven, home-based health strategies[Dishman, 2004, Reiter and Maglaveras, 2009].

Information itself is empowering if consumers utilize the availabilty of information to engage in preventative medicine, and to better manage their treatment process, decreasing the frequency of re-admission. Employers anticipate their healthcare related costs will decrease as consumers manage more of their own healthcare plans and portfolios[Mandl and Kohane, 2008].

Furthermore, healthcare consumers are accessing information in different ways. Physicians are a close second place to the internet as the primary source of information used to make healthcare decisions[Survey, 2010]. In the future, the expectation may simply be that individuals are more responsible to manage their own information in an otherwise burdened system.

Facilitating transparent, direct, and clear patient-doctor communication, applications such as John Moore's CollaboRhythm, propose a multi-platform approach to facilitate doctor-patient interactions where doctors coach patients through simple tasks, such as pushed medication reminders, or encourage patients in more sophisticated decision making endeavors that may affect their health in their daily lives, in between infrequent office visits. Part of new models for information access is redefining the relationships where the subject of care delivers, receives and acts on information in partnership with their providers.

However, for all of the efforts to innovate in emerging platforms of healthcare information technology, consumers often lack the tools to make better decisions about their own health. A compelling example is derived from sub-acute rehab hospitals. In the few months following disease onset, whether its stroke, traumatic brain injury, or cardiac related, a patient often has a critical window to accomplish the majority of their rehabilitation before entering the chronic phase. Biologically, the critical window is where patients will maximize their rehabilitation based on their effort and commitment to treatment. Treatment takes place in intensive, residential hospitals, where patients undergo multiple physical and occupational therapy sessions per day, five days per week.

Patients in residential sub-acute rehab facilities are fully aware of what's at stake. They know this is their limited opportunity to move again. And yet, a large number of patients are noncompliant, moving only when the occupational therapist comes through the rotation to insist that they try. They do not adhere to their treatment program. They are, "deconditioned and predisposed to a sedentary lifestyle," to the extent that stroke survivors can, "benefit from counseling on participation in physical activity and exercise training[Gordon et al., 2004]." As a result of noncompliance, patients are less capable, moving into the chronic phase, than what may have been otherwise achievable[Duncan et al., 2002].

Despite efforts to deliver information to patients, or to embed or acquire information from various sources and on new platforms, there is a dearth of applications to ensure that preventative medicine is empowering. We have the tools to lead healthier lives right now, to eat differently, to live differently, but as a population, we remain unhealthy. We need something more than innovation in healthcare information technology.

Creative tasks can marry significant healthcare treatment to applications that are fundamentally rewarding for individuals. With the existence of new platforms for healthcare, and mainly to capture and integrate disparate health information sources into healthcare provider systems, all vectors point to healthcare as part of our everyday lives. This thesis creates a map of how applications can be developed to integrate healthcare practice with the creative applications that consumers genuinely care about, such as learning an instrument, music listening, and music sharing. Then, the innovative landscape of healthcare information technology will have an empowered and proactive community of consumers as constituents.

1.5.2 Everyday healthcare and the chronic case

The ideal scenario for empowering new interventions is envisioned in chronic care post-stroke. After a major paralyzing neural incident, stroke, traumatic brain injury, or other disease, there are roughly three months constituting a critical window for rehabilitation[Cramer, 2008]. During that time, subjects undergoing extensive physical and occupational therapies are expected to achieve the major extent of their capacity to rehabilitate. Once patients are basically self-sufficient, ambulatory, with some facility to address their daily needs, they tend be discharged and to move into other forms of less frequent rehab, if any. From a clinical perspective they are mostly rehabilitated. Insurance reimbursement is curtailed at this point[Deutsch et al., 2006]. However, these subjects are not cured, and never will be. Often, fine-motor control will never be rehabilitated due to its lack of priority in current treatment systems.

An enormous population of stroke victims in the chronic phase have little to no recourse to further pursue rehabilitation in a home environment. The tools simply do not exist. Finally, it is unrealistic to imagine that any rehabilitation treatment in its current form would be palatable to someone in a chronic condition. Stroke rehabilitation even within the critical window is painful, arduous, and incredibly repetitive. Even when patients are aware of the need to reach their fullest potential within a critical window, they are often noncompliant [Gordon et al., 2004].

The situation is similar across many motor pathologies. Even in neurodegenerative cases, such as Parkinson's Disease, pharmacology provides a temporary respite from the first years of debilitation, but the long-term result tends to be an amalgam of symptoms in addition to the side-effects of treatment [Coelho et al., 2010].

Where technology has been proposed, it is typically expensive, and relegated to research institutions or select hospitals with wholly different services than what are available to the average healthcare consumer. Physical and occupational therapy in sub-acute rehab centers across the country are understaffed, desperately in need of basic new equipment, and may not be in a position to absorb new models of treatment, let alone new tools .

This environment could be revolutionized by new technology to provide access to creative experience, to structure an intervention that would persist beyond the critical window, into the home environment and chronic case. Patients could be empowered around something of value, music, while addressing their fundamental needs to move better no matter what stage of treatment they are in. Finally, with the right low-cost tool in the home, the first steps could be made to pursue research to determine what are the factors and costs associated with long-term rehabilitation interventions on the scale of months or years of involvement.

1.6 Thesis goals

The goal of this thesis is to demonstrate that new technologies can create access to interventions for diverse users, clearly structure the creative task in a way that targets neural regions of interest, and interface directly with symptomatic biology. Furthermore, the power of computation can be utilized to provide assessment and personalize the experience to an individual's ongoing rehabilitative process. These advantages solve the key problems of scale, and scientific credibility that plague the application of creative work to novel health contexts.

The thesis is a model for embedding scalable, valid, healthcare intervention in creative work, or other applications that have significant value in addition to directly impacting clinical change. I demonstrate how the interaction between the creative potential of music, neurobiology, diverse cognitive processes such as memory, auditory, and visualspatial processing, and a disease with a remarkable description, can be utilized to build technologies that have the potential to impact people's lives.

Three tools are introduced with accompanying studies. Each tool is developed to highlight an area in which music has implications for healthcare: within the community, diagnosis, and rehabilitation. The design of the underlying technology, as well as the supporting interventions, are evaluated with respect to observed clinical change over disease, the ability to scale the intervention, the ability to target neural regions of interest, and finally, the ability to afford opportunities for expression and communication.

Observations regarding these tools and their effectiveness within their target populations are incorporated into the final design of a surface electromyogram(sEMG) interface for multisensory-based rehabilitation. The sEMG interface serves as the input to an expressive music environment. Together, the sEMG interface and creative music environment enable the pursuit of a long-term fine-motor rehabilitation task that is inseparable from a long-term creative music opportunity.

Chapter two introduces accessible creative work in communities of individuals with physical and cognitive deficits. The role of new technologies to empower patients in their disease process, and the effect of such interventions on the healthcare institution are discussed.

Chapter three outlines initial research to validate diagnostic features of emerging creative healthcare applications. Alzheimer's Disease is a target for a new generation of auditory-visualspatial diagnostic tools, built from music. The research to establish a baseline in the general population, and test the diagnosticity of the test for a known-group of mild cognitively impaired and mild Alzheimer's Diseased patients is presented.

Chapter four presents the design and implementation of rehabilitation tools to connect with underlying biology, and structure rehabilitation as part of instrumented music performance environments. The rationale for technical design choices is elaborated, and a methodology presented to implement future designs in the space of empowering musical expression for neurorehabilitation.

Chapter five concludes with a summary of salient findings, and a methodology for music, mind and health, integrating across community-scale creative opportunity, diagnostic specificity, and neurorehabilitation tools.

CHAPTER TWO

Creative Applications in Clinical Communities

The Tewksbury Project was a series of composition workshops with residents of Tewksbury Hospital, a long-term chronic care facility. Patients had diseases as diverse as schizophrenia, severe bipolar disorder, Alzheimer's disease, spina bifida, and cerebral palsy. It was a heterogenous clinical group, where patients had marked differences in physical, cognitive and social capabilities. I introduced patients to the Hyperscore composition tool in two different departments of the hospital. The first department was a lock-down psychiatric unit, and the other was the department of physical health. The goal of the project was to observe how patients with diverse cognitive and physical symptoms would take advantage of the creative opportunity provided by a technology that gives anyone access to composition, with little to no requisite background in music.

Throughout the Tewksbury Project it was clear that new technology could provide access to creative opportunities in healthcare environments: Schizophrenic patients were observed exhibiting significantly fewer halucinations over the course of a 45 minute workshop than during any other activity. Patients graduated from being completely uncommunicative to finishing pieces, sharing them with their peers on the unit, opening new lines of discussion with their doctors, and decreasing negative and self-injurious behaviors over the course of the session work so that they would remain eligible for the composition workshops. In the department of physical health (DPH), patients arduously labored through the prohibitive interaction with the traditional desktop computing environment to gain access to composition. Hospital staff eventually converged on the Hyperscore sessions, observing clinical change was drastically different than what had previously been observed with our patients. Clinical changes were then incorporated into the patient's treatment plans.

Hyperscore meant more to the broader hospital community than mere access to music. It was an intervention that could scale across cognitive and physical deficits, operate on a community level, and allow patients to address issues of communication and personal expression as they related to their disease process.

When I first encountered the field of "accessible computing" I was surprised at the emphasis on interfacing to existing computing platforms without attached applications - a world of *adaptive* inputs. Access is more than being able to do something, its about doing something that brings you further. If a technology just lets you taste what it is to be involved in an experience, it's not doing enough. By integrating Hyperscore into the Tewkbury Hospital environment, I established access to music and provided the structure for patients to emerge as composers. As a result, patients with diverse backgrounds, and varying cognitive and physical capabilities, were able to take advantage of the opportunity to move forward in their clinical process.

Despite this research's existence solely in the residential hospital environment, I derive the design principle that devices which embed heatlhcare into home environments must provide access to experiences with significant value for the consumer. The opportunity to lead a healthier life is not significant enough. Healthcare needs to be part of cherished experiences, such as music. Also, regarding accessible experiences, it isn't enough to give users some facsimile of being great. Accessibility means the structure to fulfill their expectations of quality, to exercise creativity, and to grow throughout the process.

2.1 Hyperscore in the Hospital

Music therapists are the sole arbiters of the clinical use of music therapy. Part of the reason is that the therapist as a facilitator is uniquely capable of connecting the music interaction to the ongoing treatment process and goals of the patient. Another, simpler, reason is that music therapists are highly trained in the art of taking naive musical utterances, or the fumbling, awkward instrument attempts of complete music novices, and to provide the improvisatory musical support to morph those raw musical materials into something encouraging. Without the music therapist, the average patient, no matter what their cognitive or physical profile, is not in a position to make great music, nor do they feel confident to try.

This is a technical problem. Anyone can make music. We just need the right tools. Tod Machover's work in the area has shown that technology can provide access to creative opportunity without stifling the ability to make individual creative statements. What is particularly remarkable about Machover's music composition and performance technologies is that they offer access to the creation of completely original material. The clinical use of music tends to wrap the musical contributions of novices in the expert musical contributions of facilitators. Machover's technology begins and ends with the contributions of the participants, and is stronger for it. By bringing accessible composition tools to residential hospital communities, patients emerged as composers, instead of merely participants in music groups.

2.1.1 The Hyperscore Composition Program

Hyperscore is a graphical music composition[Farbood et al., 2004]. Originally part of the Toy Symphony project, Hyperscore gives anyone access to music composition, whether or not they have any formal musical training or music background. During Toy Symphony, Hyperscore was brought to large numbers of children who would compose intensively for two and three-weeks at a time.

Researchers, educators, and various other interested collaborators would work with the children in a mentorship model. At the conclusion of the composition phase, Hyperscore output the children's compositions in a format conducive to orchestration. Toy Symphony culminated in public performances of the children's compositions by leading symphony orchestras such as the Berlin Philharmonic, Scottish BBC Symphony Orchestra, and the Boston Modern Orchestra Project. The accomplishments of Toy Symphony participants are celebrated, not in some kind of children's concert, but rather, for the vibrant, exciting, and original work that they are able to accomplish with composition tools that structure access to creative opportunities.

Composition in Hyperscore first involves creating thematic material, referred to as motives, and then painting motives in a score window. Line, shape, and color abstract aspects of the composition process from the user while maintaining the original thematic intention. Scores are constructed as layers of interweaving lines along a timeline.

To create a melody, a user opens a motive window and enters in a series of notes. The height of the note in the window corresponds to its pitch class, and its horizontal position in the window corresponds with when you will hear the note in time. Users can create motives of varying lengths. Each motive is assigned a corresponding color which labels the melody. Users paint their motives into the score window.

In the score window, the shape of a colored line interjects the corresponding motive into the composition. The pitches of the motive are adjusted along the contour of the line. Layers of lines representing different motives together create harmonies.

The majority of the research behind the Hyperscore program encompassed the development of an algorithm to quantize the harmonies of different motives. A "harmony line" runs the length of the score window. Deformations to the harmony line modify the functional categories of the chords in the position of the score where the deformation is made. The user's motives are modified accordingly. Hyperscore successfully pushes motives into tonic harmonic relationships in addition to enabling deformations of harmony within a tonic system.

Unlike any other composition application, Hyperscore allows the user to focus on the "composition scale" after defining a handful of musical ideas. After completing a series of motives the user is free to focus on the structure of a composition. By emphasizing the relationships of various colored lines to one another, user's spend a large amount of their composition working on the relationships of motives to one another in the score window. Sections emerge. Users can construct a piece as a structure of musical ideas, without having to arduously realize their structure according to the rules of western harmonic tradition. They can paint the hierarchical representation to realize their composition's structure.

At no point in Hyperscore work, either with kids or within the general population, is it necessary to discuss the rules of harmony, key, or counterpoint. Rather, discussions privilege structural ideas in composition: how to organize a piece, how to make a beginning, middle, and end, how to make a climax, melodic parts, or the construction of supporting harmonies. These concepts are almost intuitively realized in the sketching that takes place in the score window. As a result, within two or three weeks of working on Hyperscore compositions, every day, children with little to no music background are able to complete compositions with remarkable structure, and compelling musicality.

2.1.2 Intervention Design

Tewksbury Hosptial and Recruitment

Tewksbury Hospital is a five hundred bed state hospital, with treatment programs that span many different diseases and populations, primarily divided by mental health and acute, chronic care categories. The sprawling historical campus consists of several Queen-Anne style group homes and a main hospital. The constituent departments are fairly autonomous from one another.

Two different patient groups were recruited for Hyperscore in the Hospital, one in the department of mental health, and the other in the department of physical health. Before the intervention began, a presentation was made to both interested patients and practitioners from DMH and DPH. Recruitment was conducted by the unit staff. Interested patients either volunteered based on their initial impressions of Hyperscore, or were recommended Hyperscore sessions by the practitioners.

The DMH group consisted of 8 patients, between the ages of twenty and forty-five, recruited from a lock-down psychiatric unit. The unit tends to have forty to fifty residents at a time. Patients had diseases such as schizophrenia, bipolar disorder, substance abuse, eating disorder, acute-manic depression, and borderline personality disorder. The majority of mental health patients in the Hyperscore sessions had prior serious suicide attempts and were non-compliant in other interventions including, work programs, music therapy, and expressive arts therapy.

None of the patients in Tewksbury's DMH unit come to the hospital directly. They are referred from other institutions that either aren't succeeding in treatment, or that lack the resources to provide adequate care. Many of the patients were on various psychotropic medications. Other symptoms under active management included self mutilation, auditory and visual hallucinations, and violent behavior.

The DPH group consisted of ten to thirteen patients. Patients self-selected or were recruited based on their prior interest in expressive arts and music programs. Patients had disease such as Huntington's Disease, cerebral palsy, mental retardation, multiple sclerosis, Alzheimer's disease, spina bifida, and traumatic brain injury, exhibiting a range of cognitive and physical deficits. Dominant symptoms included gross and fine-motor deficits, memory impairment, developmental delay.

Session Structure

One-hour Hyperscore sessions were conducted once per week. Initially, sessions started with ten to fifteen minute tutorial presentations focusing on particular aspects of the software, such as motives, putting motives into the score window, or use of the harmony line. Eventually, initial presentations were jettisoned from the session plan as patients voiced their desire to dive into composition as quickly as possible.

Patients were left to compose either individually or in collaborative pairs. In DMH, all of the patients worked on their own. In DPH, collaboration both with one another and with staff was more common. Physically limited patients would partner with less cognitively competent patients, or physically limited patients would work with staff to manage the computer interface and express execute their compositional ideas.

Sessions continued for three months, after which there was a celebration of the patient compositions. Hyperscore outputs the compositions in a format that can be easily orchestrated for public performance. The Lowell Symphony Orchestra came to Tewksbury Hospital and gave a performance for the broader hospital community as well as the general public. At no point in the sessions were patients overtly working towards the final performance. In the final weeks of the Tewksbury sessions we mentioned that the hospital wanted to celebrate the compositions of all the patients in the groups, and patients could decide whether or not they wanted their work featured, performed, or identified in any ensuing performances. Part of the success of the program was implicit when the patients, in entirety, were comfortable with showcasing their finished pieces.

Composition sessions were facilitated and mentored by myself, Tod Machover, two senior music therapy students from the Berklee College of Music, and one

MIT undergraduate. Facilitators had the dual roles of helping patients to gain autonomy in Hyperscore by answering questions pertaining to use of the software and, more importantly, encouraging patients to experiment, come up with ideas, implement those ideas, and reflect on the process. The mentorship model favors a reflective process to a directive one. Patients are never shown successful or even model Hyperscore compositions. Instead, they are challenged to generate original material at the outset, and to refine that material over time. The mentorship model has been highly effective in Hyperscore sessions for nonpathological populations, as documented by the Toy Symphony project [Jennings, 2003].

Initially, hospital staff did not participate directly in the sessions. However, after approximately six sessions, when patients began producing complete musical ideas, the staff began to become integrated into the session work. By the eighth session, unit managers, specialists from physical therapy and occupational therapy, and expressive arts therapists started to observe and, in some circumstances, collaborate with patients.

2.2 Patient outcomes

All of the patients that entered the Hyperscore program completed the full, three-month intervention having composed finished pieces. This alone is an outcome for a group where their pathological distinction bars such accomplishments as a function of debilitating low self-esteem and lack of feelings of self-worth, or, physical deficit.

2.2.1 DMH outcomes

Two-cases illustrate disease specific outcomes in the DMH.

When Pam came to Tewksbury Hospital she was clinically depressed and homeless. Shelter programs had referred her for an evaluation of mental illness after several serious suicide attempts. She was twenty-nine years old when we started the Tewksbury program, and had lived at the hospital for eight months. Her treatment team deemed it impossible to consider discharge at the time, as she was injuring herself, "almost constantly," in the words of one staff member involved in her daily care.

Pam excelled in the Hyperscore sessions. She refrained from self-injury because she did not want to have to stay on the unit and miss Hyperscore. It wasn't us, or even her peers that she looked forward to coming back to week-after-week. It was composing. She worked on her own. She did not request or seem interested in too much guidance. When we'd ask to listen to what she was working on, she always obliged and showed both confidence and ownership early in the sessions.

When the final concert was held, she was able to say she was moving into a new apartment and being discharged. No self-injurious episodes were observed during the three-month duration of Hyperscore interventions. Hyperscore wasn't a reward mechanism, but rather, a template for accomplishment that helped Pam to self-regulate her self-injurious tendencies throughout the week leading up to composition. In Pam's words, "there was new hope." Hope is wholly different than something to look forward to.

Daniel was forty-five and had lived at Tewksbury for nine months. In that time he had eight suicide attempts on the unit. He was multiply diagnosed with severe depression and schizophrenia. His negative ideation was so strong that it manifested into voices, dictating that he end his life for he had no possible contribution to society. During his participation in Hyperscore sessions over a three-month period, drastic changes were observed in Daniel's behavior.

He became compliant in rehabilitation sessions, consistently attending where no significant commitment was previously observed. He came off of safety checks, due to the physician's observation that there was a gradual decrease in his suicidal tendencies. He then used his progress in the Hyperscore program to petition for a job within Tewksbury Hospital, managing the computer labs. Finally, at the completion of the Hyperscore sessions, he was discharged, returning to his prior vocation as a peer job developer. Hyperscore enabled him to compensate for severe feelings of worthlessness and eventually construe his involvement with technology to align himself with vocational opportunities. Concretely, Daniel transitioned from roughly one suicide attempt per month to no such attempts over the course of Hyperscore sessions.

2.2.2 DPH outcomes

Outcomes in DPH were observed in partnership with the physical and occupational therapists that participated in the later part of the Hyperscore sessions. They noticed that the majority of patients with fine and gross motor difficulties were moving differently than observed in previous therapies. Whereas, many of these patients were not expected to be able to interact with prohibitive computer interfaces without significant assistance, in the Hyperscore work, many patients exhibited the minimum facility required to execute their musical ideas, and compose.

For instance, Sarah, a middle-aged woman with spina bifida, exhibited symptoms including mild cognitive impairment, gross motor and fine-motor difficulties. She was ambulatory and self-sufficient, but initially could not interact with a computer due to motor deficits. Reaching for mouse, grasping the mouse, and pressing a button were exceedingly difficult. Meaningful computer use was deemed inaccessible to her by her treatment team. She had been a resident at Tewksbury for more than five years and had previously shown some interest in music groups as part of the expressive arts program.

Sarah initially worked with a mentor, who guided her movements, hand-over-hand, using the mouse to enter musical ideas and explore the system. The mentor gradually decreased direct physical support, and by the seventh week, the patient was making musical entries on her own, for the duration of the hour-long session. The entries of the patient were slow, arduous affairs, but the ability to manage any kind of computing environment had not previously been observed. The treatment team then reconsidered the role of computing interfaces in Sarah's treatment plan.

Outcomes in expression were more pronounced than those in motor control. Physical disability is a disorder of expression. The individuals in DPH had a tremendous capacity to communication, albeit at a different time-course than their peers. The quality of expression exhibited by patients during the Hyperscore session was not observed in other interventions, including novel therapies such as expressive arts therapy, and music therapy without accessible technology.

Dan was twenty-eight when we first started working together in the Hyperscore sessions, and had lived at the hospital for over a decade. He has cerebral palsy.

He is paralyzed below the neck and uses an infrared pointer on his head to communicate via a text-to-speech computer. Typically, he spells out simple sentences at a frequency of one per thirty seconds to a minute in conversation. He has control over head and eye movements, however, the movements are symptomatic. He will often go into ataxia as he concentrates on hitting letter targets on his text to speech device. His muscles will involuntarily contract and he will lose what ever button he was trying to press.

Dan had been active in expressive arts groups prior to the Hyperscore sessions, including theater, and music therapy. His physical and occupational therapy in the chronic phase centered on the requisite facility to use his text-to-speech controller. His treatment team worked with him on a semi-regular basis (one-session per week, but not consistent) to increase strength, range-of-motion, control, and decrease spurious ataxic incidents such as spasms and involuntary contractions in the head and neck.

We began working together in a mentorship model. I would ask Dan what he wanted to accomplish with a given motive or section in the piece. I would then proceed down an exhaustive decision tree of questions, trying to decipher what Dan wanted to do with a particular note or phrase. He would answer "yes" and "no" by either rolling his eyes up or to the side. The process was inefficient and somewhat directive. At the third week, we modified the head-pointer that he used with his text-to-speech device so that it would interface to the Hyperscore application as a mouse. The result was that Dan could compose in the Hyperscore without a facilitator trying to unearth his musical ideas.

Dan composed on his own. The musical ideas he was able to execute were a surprise to his treatment team and peers. Furthermore, his physical therapist observed that when Dan was composing in Hyperscore he had a significantly lower frequency of involuntary movement than when he was using his text-to-speech device, despite the disproportional amount of time he had spent in tailored rehab environments, expressly working on control in the text-to-speech head pointing environment.

2.2.3 Patient mentors

Hyperscore use at Tewksbury has grown out of unit specific interventions. Prior to the existence of Hyperscore sessions, the hospital was split between DMH and DPH with little interaction, and a modicum of contention, between the two departments. Now, novel interventions have been created where patients from the DMH mentor patients in the DPH who may have difficulty managing the computer interface.

Consolidation between the two departments not only increases the sustainability of the intervention, granting some autonomy from department specific staffing and budgetary concerns, but also constitutes an intervention of its own. As patients in DMH become experts, peers, or mentors, they are increasingly capable of establishing the coping mechanisms to become healthier. They show the capacity to externalize their process, developed in their personal use of Hyperscore, to the benefit of others. Communication, but also empathy, and perspective, are clinical objectives for the majority of patients in DMH.

2.3 Continued use of Hyperscore

Hyperscore continues to be used as a clinical intervention at Tewksbury Hospital, long after the conclusion of our initial sessions there. Every day, patients on the units have a variety of activities and interventions available to them. In collaboration with their treatment teams, patients elect to go to some activities as part of their active "schedule". Hyperscore is available on the schedule. Some patients are prescribed Hyperscore, especially when they are non-compliant or uninterested in other activities. Hyperscore is first documented in the patient charts when it is initially recommended. It is then continuously assessed pertaining to exhibited patient compliance and engagement throughout the duration of a patient's involvement with the tool.

The following case illustrates a typical success of Hyperscore separate from the resources rallying and motivational aspects of MIT's initial involvement.

2.3.1 Continued use case

Emily was forty-two years old, and had lived in the Tewksbury DMH psychiatric unit for . She was unmanageable in community oriented psychiatric care. Previously she had lived in group homes but had taken to sequestering herself in her room, refusing to get out of bed for even the most basic food or toiletry needs. She was effectively mute, and the fear was that she could not sustain herself. At Tewksbury she received medical treatment to keep her alive.

Emily was uninterested in any form of treatment requiring her to leave her hospital room. It took occupational therapists eight weeks to get her to follow a staff member to Hyperscore. After making the initial steps to leave the room and get in front of the composition interface, Hyperscore structured the access to composition, and the patient started her work as a burgeoning artist. Over several months of coming to Hyperscore twice a week, she composed a series of varied and distinctive compositions. At this time, a pattern of engagement and interactivity measures were reflected in her chart, glaring in the face of many unsuccessful prior attempts in other environments.

The treatment team needed to assess her ability to transfer gains made in the Hyperscore environment to the external world, to see whether her gains were reflective of coping mechanisms or improvement over disease. Instead of asking her to explain her progress, the team attended a presentation, conducted by Emily, of her compositions. Her treatment team interpreted the resultant compositions as evidence of "cognitive organization". Based on her accomplishments, and ability to present those accomplishments to the community, a new treatment schedule was introduced that transitioned her back into community care, and out of Tewksbury.

2.3.2 Limitations to continued use

The primary limitation of Hyperscore's current use at the hospital pertains to the activity model. With a large number of subjects on the unit, and a limited number of staff, the availability of one activity or another is rarely decided based on the efficacy of the programs compared to one another. All programs suffer from the energy and enthusiasm of the staff who need to maintain the activity's availability per patient request. Despite the powerful and often unique gains made with the Hyperscore application, unless staff become advocates for the application, it is equally considered amongst many possible activities as a function of staff competence and interest. The structure to comparatively gauge the value of Hyperscore compared to other interventions for a given patient does not exist.

2.4 Considerations, moving forward

2.4.1 Completing a composition

Hyperscore in the hospital challenged existing models of music in healthcare by emphasizing the creation of finished, high quality compositions, in turn, de-emphasizing the process by which patient's came to create those compositions.

A central tenant of traditional music therapy is that the "process" of being engaged in tailored music experiences, together with a professional music therapist, is a key contributing factor to observed clinical change. Process doubles both as a reference to the evolving therapeutic relationship, as well as continual engagement with music experiences constructed to address a patient's ongoing treatment. The sentiment is that being pre-occupied with the end result of music making, the final product, is counter-productive to therapeutic process. It puts undue pressure on the patient, and unnecessarily introduces the possibility of failure if the product doesn't meet expectations. Expectations are derived from what a patient knows or observes from the music in their environment, further barring them from acceptance of music therapy as something fundamentally different than preconceived definitions of music and music performance. Succeeding or failing by creating either a good piece of music, or a good performance, is not considered a key factor for patients to progress in their clinical outcomes.

The ability of the entire patient group to create excellent compositions was critical to the success of the Hyperscore at the Hospital program. We didn't expect this to be the case at the outset. Patients were mentored at their own pace. There was no pressure to move from focusing on thematic material to transition to creating larger musical structures. Patients didn't make these transitions at the same points over the course of the intervention, but all eventually did. We didn't tell patients that there would be any kind of celebration or performance of their work at the conclusion of the program for fear of damaging the delicate process and fragile feelings of capability. In the end, not only did every patient complete pieces of music, they all participated in the final performances of one another's pieces. They celebrated their work and the work of others.

Despite being product rather than process oriented, patients participated in sophisticated music experiences. The gains in communication and community are attributed to the patient's use of technology to express themselves as artists, rather than derived from insight into their process or therapeutic relationship. Patients were empowered by accomplishing what most people feel they can't do - by composing.

Disease inextricably defines us. To have disease, to be changed by it even if only biologically, assails our definition of self. The case is only more pronounced in residential health communities, where the sense of community, of the patient's society, is constructed around the provision of services to manage disease. Hyperscore in the Hospital provided the structure for patient's to transcend that definition, elevated as composers both in the moments that they participated in the Hyperscore group, and afterwards. They hold onto more than the memory of what they had access to musically. They own the artifact, the finished work that bears their mark of creativity, the final product.

As an accessible technology, the emphasis on empowering patients to produce original work of high caliber is the mark of a technology that provides access without stifling or automating the creative process. Upon description of how Hyperscore works, it is almost impossible to believe that all of the compositions coming from the program don't sound the same. However, it simply is not the case. The greatest contribution of the application is not as a graphic interface for composition, as it is often described, but rather, as a compelling algorithm to adjust harmonic material without homogenizing the music of the entire cohort of Hyperscore composers.

2.4.2 The Next Tewksbury

There are two different ways to generalize Hyperscore at the Hospital to other hospital environments. The most direct method would take into account the essential features of the intervention and reproduce as close of a facsimile to the original as possible. The essential features, as elaborated in this chapter, were the following:

- Open enrollment.
- Mentorship model.
- Pieces are completed, and celebrated

2.4.3 Constructing a patient group

Expected caveats for a one-to-one mapping of the project to new healthcare environments stem from the unique patient composition of residential state hospitals, which, even as a group, differ significantly across state borders and even among one another within a state. Tewksbury has a fairly stable population of patients. Many of the patients we worked with in both the department of mental health (DMH) and department of physical health (DPH) units had been residents for several years. Some of the patients from the DPH unit had been members of the hospital community for decades. At extreme lengths of stay, caretakers often use the term "institutionalized" to refer to the inability of patients to even consider transitioning out of the hospital environment that they have come to know as their home. The reception of a novel intervention is expected to interact with the length of a subject's stay and the number of alternative resources, both recreational and therapeutic, available to a patient.

DPH groups had a motivated contingency of patients who were active in other social and alternative interventions available at the hospital. At the beginning of the program, staff had no question that these subjects would excel by nature of their personalities and how they had previously shown commitment in various groups, such as music therapy, expressive arts therapy, and theatre. There attendance in other programs elevated them to a kind of "regular" status across expressive arts therapy programs. Some were "stars" of multiple novel interventions. This turned out to be advantageous to the group. Despite functional limitations that made simple access to the software application difficult and perhaps even frustrating for many patients, the group had a stable set of members that were simply "going to music." The existence of the intervention was enough for them to be initially motivated. The expectation is that they would have self-selected to participate in other forms of music, music therapies, music making, or other activities, solely based on their novelty relative to a long stay in a residential hospital.

The DMH patients were selected based on the their clinical, and especially behavioral, needs, irrespective of their history of compliance in other groups, or expressed interest. Orthogonally to the DPH group, the majority of the patients in the DMH group were noncompliant in almost any other program. Some had intermittently attended vocational training programs, or music groups, but rarely with any regularity. The DMH patients, due to their shorter time as residents, behavioral difficulties, and the social dimension of their diseases, did not come to the initial session with a perspective that they would try all or any novel intervention for its social implications. In fact, the development of this kind of social thinking would have been grounds to reconsider the subject's treatment.

Both units exemplify extreme ends of the spectrum of consumers. In DPH, there is a group that is involved and motivated because the program exists. In DMH, largely unmotivated subjects try the intervention because it has been recommended to them in terms of potential clinical outcomes. Both groups quickly discover that the creative opportunity encoded in the Hyperscore application is significantly different than whatever their expectations were. It is composition, rather than assisted music improvisation. They are responsible for the creation of their own material. However, the finding that almost all patient's that we worked with moved into new phases of treatment as a result of progression in the Hyperscore sessions is particularly true of the DMH side. DPH patients accomplished a great deal, and in some cases were observed to move differently than in other physical or occupational therapy treatments. However, the sessions were presented in the context of a history of novel programs primarily intended to increase quality of life. This affected the ability to attribute causality between patient change and the session work. The DMH patients, although more likely to be noncompliant based on their history and clinical description, immediately began work in the context of their therapeutic process. The clinical results were stronger in this group as a result.

Healthcare consumers within an institutional setting enter into novel interventions with strong expectations derived from the duration of their residency, the availability of other, similar programs, and whether those programs are undertaken as social or therapeutic endeavors. As creative interventions and new therapies become annexed by a hospital's recreation departments or activity centers, obfuscating the intention of the program at the outset, patient's expectations are changed, affecting the clinical outcomes. Creative interventions, like the Hyperscore work at Tewksbury and future programs, need to be delimited as interventions despite their ability to improve quality of life.

Mentors

Another critical plank of the Hyperscore in the Hospital program was the mentorship model. Having proactive guides during the composition process was instrumental to encourage the patients to move forward despite encountering creative and technical difficulties with the creative process and software.

Mentors also served the function of brokers and spokespeople to garner attention from the hospital community at large as to what patients were accomplishing in the session work, especially during the earliest phases of the program. Initially, Hyperscore in the Hospital was received as an experimental expressive arts program. Our first sessions consisted of the mentors from the MIT Media Lab, Berklee College of Music, and an M.I.T. undergraduate, with a few nurses and a unit manager observing the session but not interacting.

Mentors insisted on being team members during weekly treatment team meetings, creating weekly post-session assessments for each patient, and having hospital staff initially visit and eventually participate in the sessions by becoming mentors themselves. By the end of the three month intervention, treatment teams were coming to the Hyperscore sessions of their own volition, to observe what their patients were uniquely accomplishing in the Hyperscore environment. Physicians were coming in to observe and reconsider the treatment trajectory of their patients.

2.4.4 Mechanism to celebrate patient accomplishment

Finally, the culmination of the Hyperscore in the Hospital process was a celebration of the patient compositions. The Lowell Symphony Philharmonic orchestra, local community orchestra, came to Tewksbury hospital and gave a performance of the patient compositions to the entire patient community, the practitioners, hospital staff, administration and the community at large.

Aside from the obvious value that such a performance carries by nature of its size, it functioned to further many of the clinical outcomes of the patient-composers. Often, residential healthcare environments have peculiar relationships to their surrounding communities. As members of "institutions" in the pejorative sense, the patient community is segregated from the external community. There are no obvious bridges between the two worlds, to the detriment of the segregated patients.

The final performance gave patients the opportunity to redefine their identities and capabilities both within the hospital and to the external community. Patients demand acknowledgement as composers, despite physical limitation or mental disfunction. The community is forced to reconsider their definitions of disability, and what patients can and cannot do. For many patients, the redefinition of identity is the springboard for clinical gains, fostering feelings of self-worth, ownership, responsibility, and capability.

For many of the patients in the DMH, the ability to succeed within a community, both personally, and interpersonally, is grounds to transition from the lock-down unit to new phases of treatment, perhaps into half-way homes or other kinds of integrated healthcare communities. The performance was critical to exhibit the patient's ability to succeed not just within the music intervention, but during the connection to the community at large.

2.5 Summary

The Hyperscore at the Hospital program was a highly effective integration of accessible creativity technologies into a healthcare environment. It became the cornerstone of many discharge strategies, most notably in the DMH. Patients used accomplishment in an accessible composition environment to overcome disease and transition into new phases of their treatment. In the Department of Physical Health, patients managed computing interfaces that were previously thought inaccessible to them, because they were motivated to gain access to composition. Many disease specific outcomes were observed.

Given an assessment of our initial intervention, in addition to an assessment of the successes and challenges of integrating the program into the hospital as a sustainable intervention, it is quite feasible to reproduce the Hyperscore in the hospital program for other institutions and healthcare communities. However, more importantly for this thesis, I have shown that it is possible to use a technology to provide access to creative work at scale, and that creative access can be leveraged to show remarkable gains in health.

Throughout the process of this work, patients demanded that their caregivers reconsider their available access to the community at large based on what they were showing as clinical gains in the composition environment.

The limitations of Hyperscore stem largely from the fact that it is not intended to target specific aspects of disease, or to structure rehabilitation for specific diseases. It is a highly accessible interface. The tradeoff is that it isn't tailored to individuals. It is similar to traditional music interventions by the criticism that it is premature to attribute causality to features of the music environment, or aspects of music perception exercised in Hyperscore.

After the Hyperscore sessions, I am confident that new, accessible creative opportunities can impinge on patient's health. Furthermore, such applications can have impact in clinical communities ranging from groups of patients to healthcare institutions. As patients emerge as composers, the patient community and the hospital overall is elevated beyond patterns of care. The challenge is to create new applications that embed assessment and structure tailored interventions into accessible creative applications.

CHAPTER THREE

Musical Diagnostic Tools

The field of Alzheimer's Disease diagnosis is almost contentious at the moment. Neuroimaging has recently proven to be a compelling measure of disfunction [Fennema-Notestine et al., 2009, Mueller et al., 2005]. Complications with neuroimaging arrise from the invasiveness of MRI, PET, and fMRI and their costs. Furthermore, if Alzheimer's Disease is present years before behavioral and memory symptoms emerge[Jack et al., 2010], it is unclear how to integrate, neuroimaging measures into a general check-up. As a result, neuroimaging is unrealistic as a preventative measure for at-risk individuals, despite the recent attention it has garnered from researchers, the media, the National Institute of Aging.

Having examined the role of technology in health communities, I considered how music could be decomposed to target specific neurological regions-of-interest. Part of the rationale for applying music to new health applications is that music touches on many of the neurological faculties that are regions-of-interest for evolving disease processes. Furthermore, a key limitation of the Tewksbury Project pertained to the inability to target specific areas of cognition where patients had the potential to improve.

3.1 Targeting Alzheimer's Disease

I have adapted visual-spatial associative memory tasks that have been found to be effective for distinguishing between individuals with Alzheimer's disease, questionable dementia, frontal-temporal dementia and depression [Swainson et al., 2001]. Tailored diagnostic tests, by focusing on the cognitive processes where the first biological signs of Alzheimer's appear, may better capture early stages of Alzheimer's disease, such as Mild Cognitive Impairment when compared with the large scale memory batteries that are often used in the field [Smith et al., 2007].

Alzheimer's disease is neurodegenerative. People who contract the disease have little recourse from the fact that their memory will deteriorate as part of a whole brain progression that will affect mood, personality, and orientation in the world around them. It is no surprise that Alzheimer's is a stigmatized disease[Blay and Toledo Pisa Peluso, 2010, Graham et al., 2003], where individuals tend to avoid seeking a diagnosis on average around two-years after the onset of symptoms [Lewis and Irvine, 2006]. The situation is critical considering that emerging pharmacological interventions are best introduced at the earliest stages of disease, and may delay disease progression [Birks, 2006]. Recent findings suggest that memory decline in genetically at-risk individuals occurs long before a clinical diagnosis [Caselli et al., 2009, Bäckman et al., 2004]. For these reasons we need new interfaces that are not only sensitive to the earliest stages of the disease, but can be embedded in new environments. Patients should be empowered in the longitudinal assessment of their own cognitive health.

I conducted several studies to establish a normative population and refine the salience of task performance. I examined subjective measures such as strategy formation and perception of task difficulty. Finally, in collaboration with McLean Geriatric Hospital and Harvard Medical School, I measured the diagnostic validity of the task for a group of elderly with definite diagnoses of either mild cognitive impairment or mild Alzheimer's disease compared to age-matched, healthy controls.

Currently, there are no diagnostic tasks sensitive to Alzheimer's disease that engage auditory processing, despite some interest in speech perception and production[Almor et al., 2009, Vestal et al., 2006]. The advantage of broadening the scope of perceptual domains available for diagnosis is to simultaneously broaden the scope of applications in which you can perform assessment while introducing new platforms for research into areas such as longitudinal or embedded assessment. New applications are implied to motivate individuals to become pro-active managers of their personal health as part of their everyday environments. Cognitive assessment is a critical part of this new generation of applications. I have developed a series of technologies to embed the auditory

Alzheimer's test, including a mobile application, a tangible tabletop application, and a room-scale gesture tracking application.

It is irrelevant to discuss the diagnostic tool in terms of clinical change over disease as it is not an intervention, but rather, an assessment. What it lacks in facilitated change, it makes up for in scale and neural specificity. By validating the assessment tool for the diagnosis of Alzheimer's disease we have created a platform that can be embedded in many new environments and applications to test for the onset of disease. The assessment in itself is not as interesting as the validation of a way of working with auditory information and visual-spatial location. We have substantiated a way of working with auditory and visual-spatial information that we hope can be a platform for many embedded applications for disease diagnosis.

The resultant diagnostic tool is limited in terms of expression and communication. As I began embedding the diagnosis in mobile devices, room-scale sensing environments, and tabletop applications, it became clear that the focused attention required of a concerted effort in a diagnostic environment limits the ability for open-ended creativity and continuous expression in time.

3.2 Development of an Auditory-Visualspatial Tool

3.2.1 Early Prototypes

A Glove-based Sound-design Interface for Computer Music Composition

My first application efforts in the Alzheimer's diagnostic space were not particularly successful. My strategy was to build a compelling music application from the outset. The user would manipulate the interface to design individual sounds. Then, after designing a sound, the user would switch over to a composition program. While in the composition program, users would select from all of their previously designed sounds, and layer them into a sequence of audio tracks. Compositions would be built up sequentially from a palette of previously designed sounds.

Problems began long before a user would reach the composition phase. The sound design process was cumbersome and awkward. The input device consisted of a virtual glove interface, with onscreen sliders to visualize parameter changes. Users would make movements in a two-dimensional plane, in front of them, while glove position and finger presses were recorded using infrared sensing and force-sensitive resistors along the fingers. Hand movements would make musical changes. Finger presses would save parameter changes and move on to new parameters, save the sound as a whole, or in general, manage other interface routines such as 'undo'.

The sound parameter changes available to a user can be described as a decision tree. The user begins at the top of the decision tree. At this tier, their hand movement in the two-dimensional plane merely selects from three available *sound design templates* - starting with additive synthesis, granular synthesis, or formant synthesis. The second tier also specifies meta information for the sound design process. The user hears a repeating *tick* sound at a certain period in between ticks, and a given fundamental frequency. By moving along the y-axis, the user changes the fundamental, while changing the period on the x-axis. When the user is satisfied with a certain fundamental frequency and period, they move a finger to make the selection and proceed to the third tier, where sound design begins in earnest. For all tiers after following a period and fundamental frequency selection, the sound that they are actively modifying continuously loops at the period specified at the second tier.

At this stage the user can move up and down the decision tree with finger presses. At each tier one sound parameter is arbitrarily mapped to the x-axis and another to the y-axis. The user moves through the space of possible states until they finish designing a sound to their approval. A finger press coincides with saving the sound and exporting it to the composition interface.

Diagnostic objects from sound design

The connection to Alzheimer's diagnosis, and working memory, was derived from a visualization created during the sound design step. As the user makes sound parameter changes, an abstract, three-dimensional image is contorted according to how the parameters are being modified (Figure 3-1). Users are meant to attend to the visualization as it changes according to their sound design changes. Users are instructed to remember the final three-dimensional image as they save their completed sound before entering the composition step.

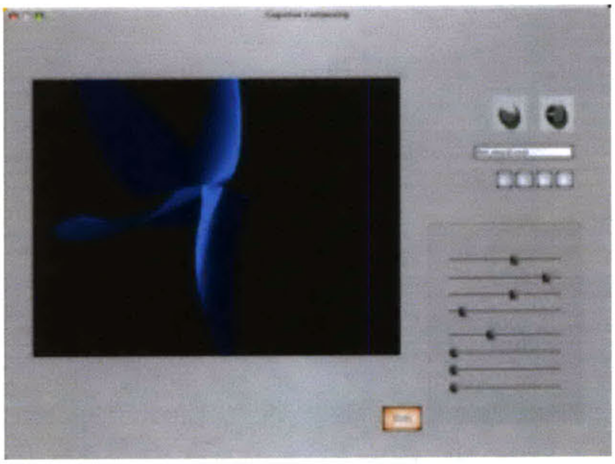


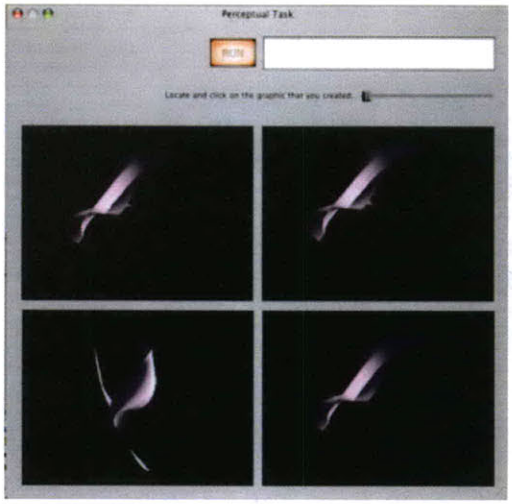
Figure 3-1: Sound design interface. The user does not interact directly with the software application. They use a separate glove interface to modify sounds, and move through the application. As visual feedback of glove input, sliders reflect parameter changes, while the paired image, creating during sound design, takes precedence.

The image is associated with the completed sound. They belong together as a pair.

Randomly, when transitioning between sound design and composition, the user may be interrupted by a diagnostic task. Users are asked to take a *delayed matching to sample* test before proceeding to composition. In the test, the user is presented with four abstract images and one of the sounds that they had previously designed. The user must match the sound with the correct abstract image, the image associated with the sound when the sound design was completed. The three confounding images are generated according to a rule set: some are completely random, some are based on the features of the correct image, simply change the lighting angle or color of the rendered image (3-2).

Prototype evaluation

There were three factors of our application design that needed significant restructuring. Firstly, the application was immutable because of the number of state dependencies of each stage of the sound editing process. Each set of parameters is altered in pairs, in order, as a user descends a tree of parameter edits to create complex sounds. Second, stimuli generation was completely



Locate and click on the graphic that you created.

Figure 3-2: Presentation of delayed matching to sample task.

non-modular, and the user was restricted to moving from one set of parameter changes to another, with no sense of direction, history, rationale behind the ordering of parameters, or conception of how editing parameters early in the chain would affect dependent parameters later in the chain. The application proceeded as a horizontal chain of state changes, with no dynamic relationships to ever form between objects created. The objects may have been interesting for testing stimuli, but were awkwardly formed and of limited creative potential.

Most importantly, because sound parameter changes were distributed over a multitude of states accessible one after another, it is virtually impossible to substitute a different type of sound source into the application, which the user would then manipulate accordingly. For instance, the application is good for changing computerized sounds over many sets of parameter changes, where parameter changes are equally important. What happens when our neuropsychological validation yields that we need to be working with pre-recorded melodies from the subject's past musical experiences? The current arrangement of parameter edits where each parameter receives equal and random distribution in a series of editing steps no longer seems appropriate. Furthermore, the interface suffered from other design flaws, such as a limited set of user inputs being remapped to different functions at each step of the application, making it virtually impossible to quickly make any changes as necessitated by creative demands.

Second, we sought to survey the field as to basic cognitive science examining auditory cueing for visual-spatial tasks. My rationale was that if I could constrain the possible auditory stimuli that users would be working with, say, recorded classical music, or user-generated short melodies, I could then build the basic architecture of the interaction while being flexible to adjust a category of predetermined auditory stimuli. Unfortunately, there is little to no understanding as to the differences between types of auditory stimuli as they relate to visual spatial associative memory, or even memory at all. Music is certainly perceived differently than tone pips and sine waves, but does one of the two auditory stimuli disproportionately convey an advantage to user's that have familiarity with music creation, or a more robust history of listening? The difficulty of the question is only exacerbated when considering the whole-brain pathology of Alzheimer's disease. Research exploring Alzheimer's disease and the perception of music is is still in its early stages[Omar et al., 2010]. To date, foundational studies lack the resolution to garner design strategies and build applications around how an Alzheimer's diseased patient perceives musical stimuli as they operate on the stimuli.

The conclusion is that for an application where the basic perceptual stimuli being operated upon are likely to change based on research, and where the basis of user interaction requires completely different metaphors for assembly and editing based upon the underlying stimuli, it is counter-productive to define an interaction architecture which would incorporate these likely-to-change objects. Furthermore, it is critically important to evaluate interaction throughout the design process, rather than simply certifying an application's adequacy after its creation. Developing a full system was premature from a research standpoint.

The music application should privilege involvement with the music from an individual's everyday life, the preferred music of their environment. After the prototype work, rather than build a music composition or performance interface, I instead focused on pursuing a valid neuropsychological diagnostic tool that would be *about* the music in your everyday life.

3.2.2 Music Grid

After evaluating the prototype composition interface for Alzheimer's diagnosis, the majority of remaining effort consisted of validating a neuropsychological auditory task with the intention of making it the core of an ensuing music application. The first version of the auditory-visualspatial task was used as a pilot to examine differences between musicians and non-musicians. A valid diagnostic task should first establish normative data, to ensure that the general population performs similarly enough on a task such that the performance of individuals can be compared against a group[Spreen, 1998]. I then made several interface changes before testing the application in an elderly population.

3.2.3 Establishing a Normative Population

Our task randomly maps 10 different auditory stimuli onto 10 different on-screen locations. During the first iteration of the test, (i=1), one of the on-screen locations "flips open" (Figure 3-3). The user then hears a 1.5 second audio stimulus associated with that location. When the stimulus is done playing, the location "flips closed". Then, after a short pause, the user hears the audio stimulus as a cue, and needs to click on the screen location where he originally heard the corresponding audio stimulus. If the user is correct in their selection the test proceeds to the second iteration, (i=2), where the user will have to remember two (i) audio stimuli and their corresponding locations. The user continues in this fashion until they incorrectly respond to a cue, after which, the test ends. Audio-location pairs are randomly distributed before a test iteration. Furthermore, during testing, the order of presentation of the audio cues is also randomized, making it unlike *Simon* or other games where you chain one new audio-location pair after another in a single, persistently lengthening sequence. The sequence in the Music Grid is different for each iteration, making the test about novel pairs of associations, rather than memory of a sequence of associations.

Instead of testing one type of auditory-visual spatial memory for the normative population and the at-risk elderly population, a major feature of the application is the ability to dynamically change many of the training and stimuli set parameters. Experimental variables include: (1) type of audio stimuli, (2) time allotted to answer after an audio presentation, (3) musical experience, (4) time between stimuli presentation in training phase, and (5) time between training and testing phases. It is important to stress that with no available information on auditory processing and associative memory for the given population, this testing needs to establish basic principles for stimuli presentation and test construction.

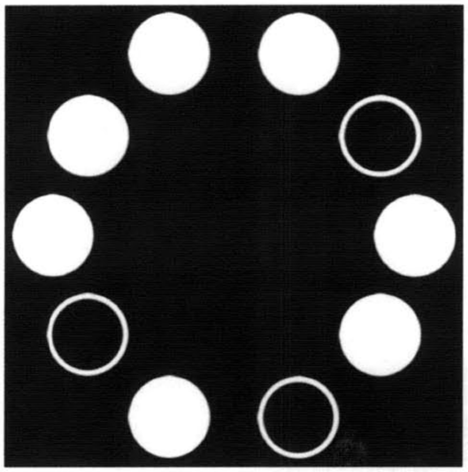


Figure 3-3: Layout of the testing interface. Visual locations reveal their associated audio stimuli.

During preliminary studies, I determined that the following stimuli presentation parameters were the most effective - buttressing the performance of the normative group, while keeping the general time-course of the test in a range that suited short-term associative memory.

Duration of auditory stimuli: 1.5 s
Time between training and testing phases 3 s
Time between stimuli presentations in training phase: 2 s
Time between stimuli presentations in testing phase: Wait for user.
Time between end of test level and start of next training phase: 5 s

If musicians gain an advantage from their expertise, one would expect to see differences in performance compared to the general population in stimuli sets containing features that musicians have scrutinized over the course of their study (as in musical phrasing[Pantev et al., 2003]), or features where musicians discriminate differently on a known neurological basis (as in pitch perception[Bidelman et al., 2009]). The study incorporated seven different stimuli sets (Figure 3-4). From the single note level, to timbre, contour, dynamics, loudness, tonality (harmonic content), phrase, and genre, we attempted to introduce a spectrum of feature differences that encompassed both objective and subjective content.

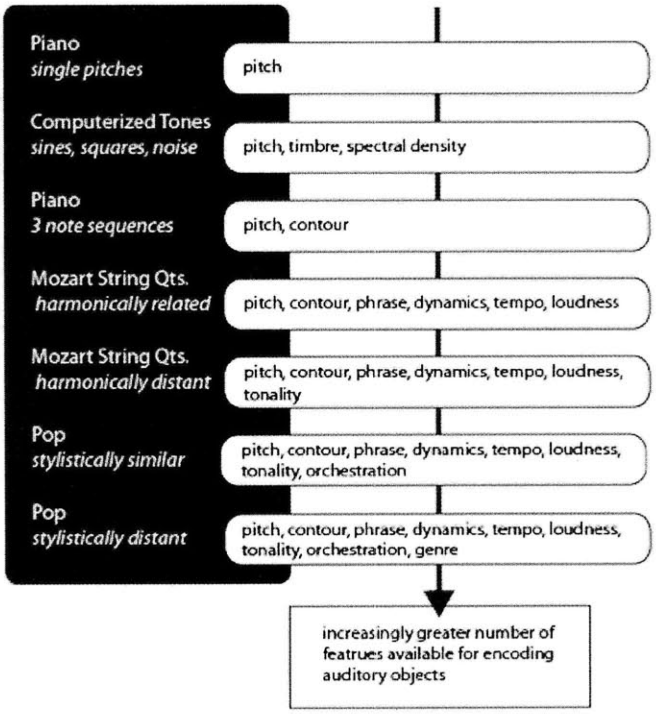


Figure 3-4: Stimuli sets, descending in order of complexity. Each stimuli set builds on the features available in the previous set.

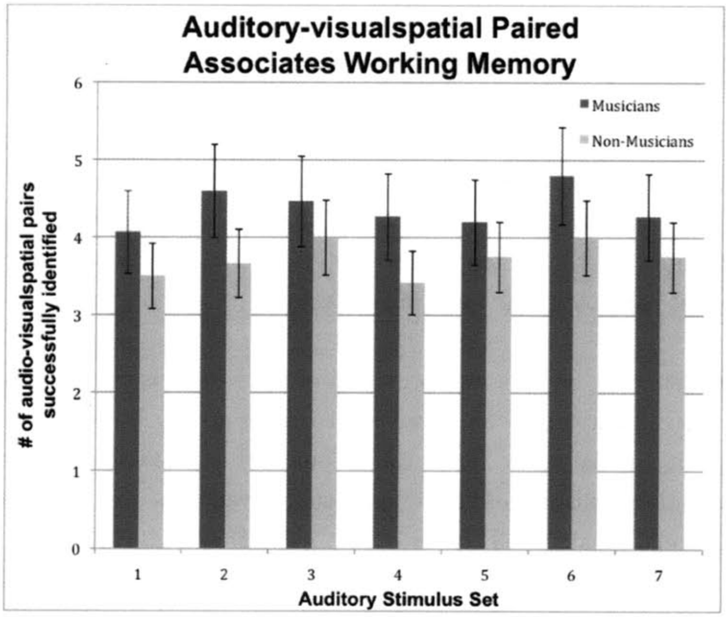


Figure 3-5: Results. Set1: Popular music (same song), Set2: Popular music (different songs), Set3: Piano (3 notes), Set4: Piano (single notes), Set5: Mozart String Quartets (harmonically related), Set6: Mozart String Quartets (harmonically unrelated), Set7: Computerized Tones. Graphed error bars are confidence intervals at the .05 level. Sets 2 (Popular music, different songs), and 4 (piano, single tones) yield statistically significant differences between musician and non-musician groups (set 2: t(25)=2.81, set 4: t(25)=2.46, p¡0.05)).

I tested twenty-seven subjects, (N=27, musicians=15, non-musicians=12). Musicians did perform significantly better than non-musicians in two conditions, for single piano tones (t(25)=2.46, p¡0.05), and for popular music with different songs (t(25)=2.81, p¡0.05). However, overall, there was no significant difference in sets that included music where experts had entire lives of dedication. Some of our subjects were able to name individual pieces of music after 1.5 second stimuli, however, it did not convey a significant advantage (Figure 3-5).

In the case of single piano tones, our musicians found the task to be similar to their formal music training for pitch discrimination in solfege and sight-singing exercises. Also, pitch discrimination is an area where musicians have clearly defined perceptual differences compared with nonmusicains as evidenced in neurological study[Pantev et al., 2003], as well as behavioral examination[Magne et al., 2006]. It may be that musicians perform differently on memory tasks that offer a memory component as a direct analog of a previously learned task, drilled through rehearsed practice as part of music theory or performance training. Most of our musicians had undergone over two years of music theory training (N=11, 73%) in addition to formal instrument study for the majority of their lives. In the interpretation that the single pitch stimuli sets is not a memory test for complex auditory objects, but rather, a pitch discrimination task in a visual space, then it stands to reason that musicians would perform better than non-musicians for the task as a direct analog of their experience. Complex auditory stimuli, mapped to the visualspatial space, do not offer such an immediate analog to the musician's training.

Additional support for this interpretation of the results is that single computerized tones did not convey an advantage to either group. If it were merely that the simplicity of the single piano tones somehow enabled the musician group to access well-established auditory perceptual advantages over the general population[Moreno et al., 2009, Hyde et al., 2009, Dowling, 1986], than perhaps simple computerized tones would also convey an advantage. However, this was not observed in our data. Despite that computerized tones were as elementary as the piano tones, with only a modicum of increased complexity in timbre, musicians performed similarly to non-musicians for single computerized tones. The single piano pitches most likely relate to a directly analogous trained task.

Any advantage conveyed by increased pitch discrimination abilities in musicians was not carried into the remaining stimuli sets, including the 3 note pitch contour set, where the only additional variable to pitch was the contour of the 3 note sequence. This finding also indicates that advantages for one feature are not additive with other features. Pitch as an underlying feature of later stimuli sets did not convey an advantage to musicians. One final interpretation is that non-musicians may need additional features to the elementary pitch and temporal characteristics of single notes to distinguish between auditory stimuli as objects, more than discrete acoustic stimuli. Musicians may outperform non-musicians where the stimuli are the basic elements of auditory perception. However, for complex audio, non-musicians and musicians may have the same capacity to construct associative objects from complex auditory stimuli. non-musicians may need more features to perform similarly to musicians.

With regard to the popular music (different songs) set, it is unclear how genre as the distinguishing feature could disproportionately advantage the musician group. Furthermore, many of the nonmusician subjects considered the popular music set to be the easiest of all the data sets. This seems inconsistent and will need to be borne out in further study. The popular music (different songs) set contained the most familiar stimuli for the nonmusician population, with most non-musicians being familiar with all of the presented genres throughout their extensive daily listening experiences. It may be that familiarity confounds the assessment of degree of difficulty in a memory test constructed from familiar stimuli. This is a positive finding, considering our goal of constructing a test where people are motivated to participate, using familiar stimuli in a difficult task. Our non-musicians were significantly older than our musician group (mean age, NMus=32.5, SD 12.8, Mus=25.7, SD 5.8). One possible interpretation is that familiarity and age interact to advantage a group to better label and distinguish between differences between stimuli in a set. We wouldn't necessarily see this difference where musicians are more familiar with the classical music sets, given our exclusive use of string quartets for a musician group that contains many non-string players. As we develop tools that use a subject's own music media to construct a test, we would expect this advantage to disappear given a roughly equal level of familiarity between individual user's and their personal media.

There was a great deal of variability in the subject's perceived difficulty of one stimuli set to another (Figure 3.2.3). For most stimuli sets, what several subjects perceived to be the easiest set was another subset of subjects perceived to be the most difficult. Without a single set emerging as distinctively more difficult than another, it may imply that subjects are not devising strategies that they can apply consistently across stimuli sets.

Subject's strategies were assessed via questionnaire. We initially asked what type of strategy, if any, was used to assist in memorizing the order of the auditory stimuli. Then, we asked for elaboration, and to distinguish if different strategies were applied for one set rather than another. Questions continued with whether the subject considered their strategies to be effective and whether any of those strategies relate to their formal music training or "musical knowledge". We also asked subjects whether or not they felt there was improvement over the course of the test, to distinguish between strategy and overall improvement from practicing the test over the course of the experiment session. Strategies were quantified on a scale of 0-5: 0, wholesale melody placement or no strategy specified, 1, single pitches and pitch reductions of more complicated feature sets in the visual space, 2, pitches in the visual space with one additional abstracted feature such as a semantic (pneumonic) tag or loudness, 3, contour

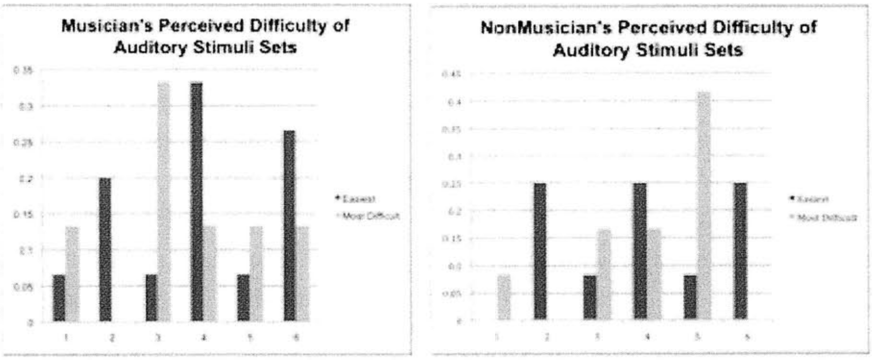


Figure 3-6: Percent of subjects that felt that the given stimuli set was more easy or more difficult than the other sets. Set1: Popular music (same song), Set2: Popular music (different songs), Set3: Piano (3 notes), Set4: Piano (single notes), Set5: Classical Music, Set6: Computerized Tones. Since it was not obvious to subjects what was different between the two classical music sets, we combined them into one super set, "classical" for survey purposes.

or some other geometric abstraction, 4, contour with a complex character abstraction encompassing idiom, motion, phrase, or other high level music features in combination, 5, three or more strategies in parallel. We quantified strategies according to the number of audio features that a subject would necessarily operate on to successfully use the strategy. For instance, if a subject reported that they were trying to remember the individual pitches, they would be at the first tier of strategy complexity. If subjects recounted that they attended to the individual notes and the melodic contour of the musical line, they would be at the second tier of complexity. Other features included timbre, dynamics, and semantic tags.

If a subject is performing operations on a large number of features to encode the stimuli into the visual space, it can be argued that their strategy is more sophisticated than an individual using fewer features to encode the audio in the space. With more features available to the user, one may be lead to conclude that their process would convey an advantage to remember the associations between stimuli and visual location. However, despite the tendency of our musicians to average as slightly more sophisticated than our non-musicians in strategy formation, this finding is not quite statistically significant due to the large amount of variability in strategy formation even among similarly expert musicians (t(25)=1.89, p¡0.05, musicians mean=3.4, SD 1.88, non-musicians mean=2.08, SD 1.68). This has interesting implications for pursuing testing in the everyday environment. Unlike current neuropsychological assessment,

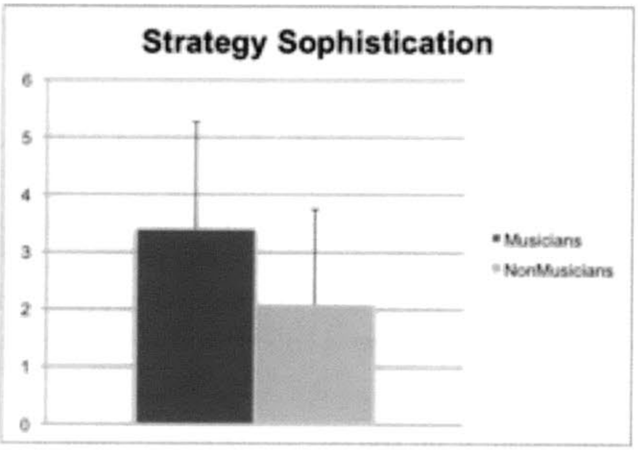


Figure 3-7: Measured strategy sophistication on a scale of 0-5. Higher rating implies larger number of features attended to during encoding audio samples into the visual space, left, Musicians (mean=3.4, SD 1.9), right, non-musicians (mean=2.1, SD 1.7).

testing in the everyday environment implies that a user has many iterations within the same testing environment with which to perfect a strategy. In our study, we have very little consistency in either the musician or nonmusician groups regarding strategy formation and a fair amount of variability within groups on performance for any given stimuli set. Over time, subjects may adopt more consistent strategies and deliver more salient, consistent performance on the memory test (Figure 3-7).

In conclusion, complex, orchestrated, culturally engrained, idiomatic musical stimuli do not convey an advantage in expert populations for this particular type of memory testing. Having established normative data, I began testing in a diseased population.

3.2.4 Known-groups Validity Study

To determine the diagnostic validity of the auditory-visualspatial test, I partnered with McLean Psychiatric Hospital's Geriatric Psychiatry Department, an affiliate of Harvard Medical School. The hypothesis was that auditory-visualspatial tasks are valid for differentiation between mild cognitive impairment, Alzheimer's disease, and healthy controls.

Subjects were recruited from various clinics and outpatient programs connected to McLean Psychiatric Hospital, as well as the current patients of Dr. Ellison.

In the memory impaired group, patients were referred to the study if they were known to have a diagnosis of either mild cognitive impairment or mild Alzheimer's disease from their normal treatment or initial assessments at McLean Hospital (N=6, mean age=75, SD=8.7). Subjects were between the ages of sixty and eighty-nine. They were required to be competent to provide informed consent, and to score within the following criteria on standard memory batteries:

Geriatric Depression Scale (GDS-15) less than or equal to six.
Mini-mental Status Exam (MMSE) greater or equal to twenty-four
Montreal Cognitive assessment (MOCA) greater or equal to twenty.

The cutoff scores are conservative for inclusion in a memory impaired group. Furthermore, they neglect subjects that may exhibit stronger diagnostic features, or those exhibiting Alzheimer's disease in a more progressed form. However, we were committed to examine diagnostic validity for the earliest forms of the disease, or, in the case of mild cognitive impairment, precursors of Alzheimer's disease. Additionally, a depression scale was included because depression has been shown to confound working memory performance[Channon and Green, 1999] and performance on Alzheimer's diagnostic measures[Swainson et al., 2001].

Control subjects were recruited from lists of control subjects kept by McLean Hospital, and from the general population (N= 10, mean age=72, SD=8.6). Subjects in the control group were between the ages of sixty and eighty-nine, but with no diagnosis, and no reported history of memory complaints. Their scores on inclusion tests needed to meet the following criteria:

Geriatric Depression Scale (GDS-15) less than six.
Mini-mental Status Exam (MMSE) greater than twenty-six
Montreal Cognitive assessment (MOCA) greater twenty-six.

Both groups were highly educated, with most having completed college (Control/Experimental = 16.8/18.4 mean number of years in school, SD = 2.5, 1.9). The majority of both groups continued to be active in some form of part-time or volunteer work. On a five-point scale reflecting the frequency with which subjects listen to music, between never and everyday, our mild AD and MCI group listened slightly more frequently than the control group (Control/Experimental mean listening frequency = 4.2/3.5, SD = 1.0/1.1).

Subjects were seated at a large, touchscreen interface, and presented with an identical looking application to the one used in the pilot and normative studies. Stimuli presentation parameters, such as the time between training stimuli, the time between training and testing phases, were kept the same as in the normative study. However, we did modify the application to introduce a *three strikes* rule. In an attempt to increase the salience of the performance of the non-impaired group, and to decrease the standard deviation within a group, we gave study participants three tries on a single tier of the test. For instance, after getting an associated audio and visual pair wrong, the application stops, notifies the user that something was incorrect without identifying which stimuli-location pair was entered incorrectly, and reverts back to the training phase for that tier. The test is only complete when the user fails a total of three times on a single tier.

The rationale behind giving the users multiple attempts to correctly associate stimuli is that the control group should benefit from repeat exposure, while the impaired group should not benefit from re-trying a task that impinges on a domain of cognition where they exhibit functional deficit. Mild cognitively impaired, and mild Alzheimer's diseased subjects should not be able to quickly rebound from difficulties in associating stimuli in a working memory task. One can imagine a scenario where due to a lapse in attention, stress, or fatigue, an unimpaired subject makes a mistake early that does not relate to their capacity to perform well in this cognitive test. On a subsequent retry, the subject has the opportunity to overcome spurious entries. A diseased subject would have a similar opportunity, but would most likely not be able to benefit from repeated attempts. The intended result is to make the performance of the control group more salient relative to the experimental group.

I first tested for equality of variances between the two samples. The expectation is that a diseased population will have different variance than the general population on a neuropsychological test intended to be diagnostic for a domain of deficit. Each of the four sample sets were tested independently for equality of variance with a welch's f-test at a 95% confidence interval. The two pop-

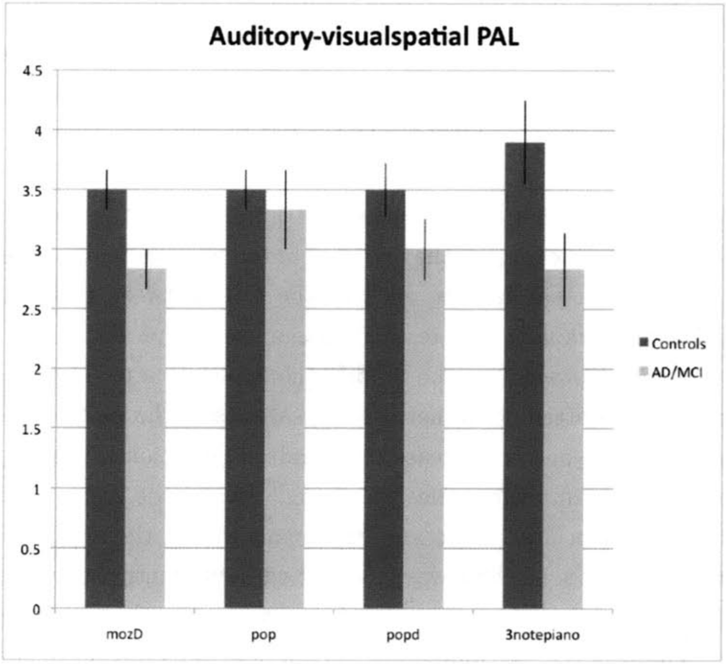


Figure 3-8: Known-groups validity test, of auditory-visualspatial measure for mild AD, and MCI. Set1: Mozart String Quartets, harmonically distant, Set2: Popular music (same song), Set3: Popular music (different songs), Set3: Piano (3 notes), error bars reflect standard error about the mean.

ulations displayed unequal variance for harmonically different string quartet excerpts (F=8.0, p=.001) and for three note piano motifs (F = 5.3, p = .034), but not for popular music excerpts drawn from a single piece (F = .2, p = .67), or popular examples from different pieces (F = 2.1, p = .17). A one-tailed Welch's t-test was chosen for all further tests of significance to adjust for unequal variances.

The null hypothesis was rejected in both the harmonically distant string quartet examples (t(12.9) = -2.93, p = .007), and the three note piano melodies (t(13.6) = -2.30, p = 0.02), but not for either the popular music examples drawn from the same piece of music (t(7.6) = -0.44, p = 0.33), or from the popular music examples drawn from distant vastly different pieces (t(11.7) = -1.46, p = 0.08) (Figure 3-8).

The data indicate that the MCI and mild Alzheimer's Diseased population perform significantly worse than healthy, matched controls for short excerpts of simple groups of notes, and short, orchestrated complex, idiomatic, excerpts of harmonically distant orchestral music during an auditory visualspatial task.

An interpretation of the results hinges on considering our inclusion criteria for the study. Recruited subjects were evaluated with two established tests for Alzheimer's Disease, the MMSE and the MOCA. Many studies have documented the insensitivity of the MMSE to detect mild cognitive impairment and mild dementia[Tombaugh and McIntyre, 1992, Tierney et al., 2000]. The MOCA is a stronger test for our target population. Even so, our diseased population scored on the threshold of the recommended cut-off for consideration of impairment, which is 26 out of a possible 30. The average MCI subject is expected to score closer to 22 on the MOCA, and Alzheimer's Diseased patients, exhibiting memory deficit and functional impairment, are expected to score around 16 out of 30 [Nasreddine et al., 2005]. All of our subjects had a definitive diagnosis of MCI or mild AD, but scored significantly higher than the mean MCI score reported in previous studies of the MOCA when subjected to a one-tailed, one-sample t-test, at a 95% confidence interval ($t(5) = 3.0$, $p = .015$) (Figure 3-9). Outliers within our six subject group had scores of 21.0 and 29.0, and not surprisingly, the highest scoring MCI subject on the MOCA was also our highest scoring subject. Despite the intention to conduct a known-groups validity test, our diseased group was less impaired, as demonstrated by MOCA performance, than the broader population of MCI and mild AD subjects. That any of the stimuli sets would be significantly diagnostic at this level is encouraging and allow us to move forward to the next stages of validating a diagnostic, namely, increasing statistical power, validating the construct, and determining test-retest reliability.

Other difficulties in the current study include low sample sizes (MCI, mild AD group, $N = 6$, Controls, $N = 10$), raising concerns as to whether the study has sufficient effect size to be representative of performance in a population of diseased patients. The study offers initial observations that must be confirmed in larger, multi-center, controlled studies. Use of the term diagnostic may be premature.

The next step for validation of a diagnostic tool is to determine the test-retest reliability, and the construct validity. Test-retest reliability is especially pertinent to the intended goal of distributing diagnostic tools into the general population, embedded in the applications of their user's daily lives. If user's

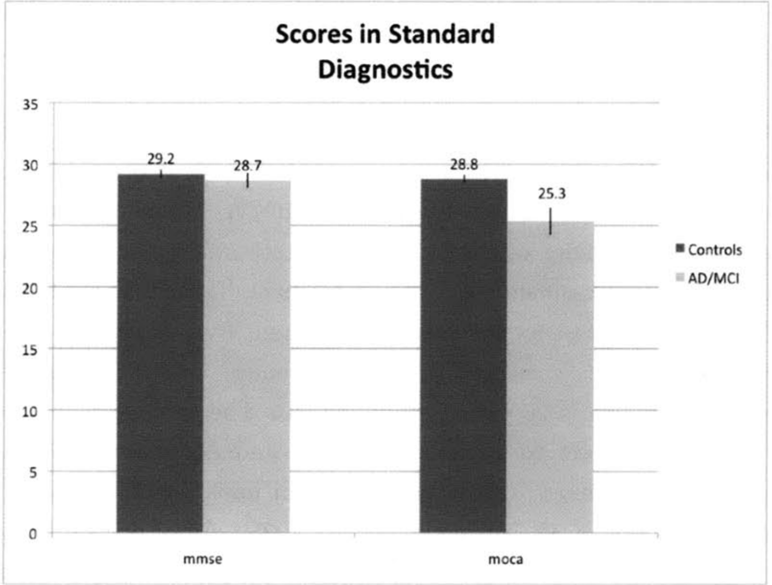


Figure 3-9: Standard Alzheimer's Disease memory batteries, the Mini-Mental Status Exam, and the Montreal Cognitive Assessment .

take a test everyday, or even every month, they would be expected to develop strategies and expertise over the course of long-term interaction with a tool. The effect of experience on high-frequency cognitive assessment is a new area for neuropsychological research, predicated by the possibility of embedded assessments. During a longitudinal study of test-retest reliability, if the recruited population is large enough, one can also examine the sensitivity of the measure to indicate when a healthy subject transitions into disease.

Construct validity is a measure of the specificity of the test to target the intended domain of interest, rather than some ancillary or superficially related domain. The auditory-visualspatial test was constructed after reviewing the etiology of Alzheimer's Disease, and its initial presentation in the entorhinal cortex, a region of the hippocampal formation specialized in associative memory [Van Hoesen et al., 1991, Raji et al., 2009]. It has been hypothesized that the effectiveness of paired associate learning tasks and delayed matching to sample in early Alzheimer's Diagnosis is derived from their relationship to the type of memory related to entorhinal cortex function [Jo and Lee, 2010]. Future studies will have to include known tests, not just for MCI and AD, but also for visual associative memory, or tactile associative memory. By examining associative memory function across several domains, including auditory-visualspatial, one can draw conclusions as to the specificity of the measure for the target cognitive domain.

3.3 A Platform for Embedded Diagnostic Tasks

The goal of this work is not to deliver a single neuropsychological test for Alzheimer's disease, but rather, to enable application development that embeds ways of working with auditory and visualspatial information that is diagnostic for the earliest phases of Alzheimer's disease. The key to rich and rewarding applications, that go far beyond the test itself, is a strong foundation in a validated interaction technique. Otherwise, applications become stained by a test that gets awkwardly wedged into an otherwise rewarding experience. By building upon a layer that incorporates the auditory-visualspatial interaction that we have found to be useful for assessment purposes, we can finally develop applications that leverage the unique aspects of music while simultaneously providing assessment.

The advantage of music has always been that music is an interface that people care about. It is intrisically engaging and rewarding. By developing technologies that provide access to music for the general population, and pairing those applications with rigorous assessment, we can postulate a scenario where people will be highly motivated to proactively engage with their cognitive health at a stage where they are at-risk of developing Alzheimer's disease.

In one instantiation of the diagnostic test, I built a mobile device application to deliver auditory-visualspatial testing in a user's everyday environment. Currently, the test is automatically generated on application launch from whatever audio media exists on the user's mobile device. This new application bridges between our previous experiments and a situated assessment. In particular, key research questions can be addressed including: what is the validity of a cognitive assessment conducted with a frequency unprecedented in the neuropsychology literature, how do users manage their cognitive profiles as they become the first point of feedback for that information in terms of privacy, and community access to the data, and finally, how does the user's perception of healthy and unhealthy cognitive decline in relation to aging change as they are engaged with their cognitive profile on a daily basis.

Beyond enabling future studies, novel embodiments of this research will exist at various scales and in many different contexts. Location based services can pair auditory and visual representations of location to diagnostic ends. Mapping interfaces, guides, planners, and any application that engages the establishment of visual locations and auditory pairs can benefit from this underlying research. Whether it is room-scale sensing environments, tabletop interfaces, or mobile devices, we are often being asked to interact with diverse materials with physical, visual, and auditory components. By validating the elements being used in new interfaces, diagnostics can be embedded in our everyday tools.

Musically, what happens when the sound of the world around us becomes diagnostic, when we can probe how, and where we hear to actively listen in a way that is indicative of our cognitive health?

3.3.1 A Tangible Example

Together with Elena Jessop, we attempted to take the auditory-visualspatial diagnostic tool into a tangible media environment. The tangible design employs several key components to establish the bridge between physical and digital media. Tangible representations are computationally coupled to underlying digital information. The same physical representations embody mechanisms for interactive control. Physical representations are perceptually coupled to actively mediated digital representations. And finally, physical state of the tangible environment embodies key aspects of the digital system[Ullmer and Ishii, 2000]. This conceptual framework effectively couples digital information with controllable physical representations.

Elly had previously developed a gesture controller for vocal performance that uses hand grasping movements to capture and "hold" and "manipulate" recorded sounds[Torpey and Jessop, 2009]. The VAMP glove controller was used to capture audio in the enviornment, and "place" captured sounds in smart containers. Each container is equipped with a proximity sensor and a two-axis accelerometer. When the lid is removed from a container, the proximity sensor can detect the presence of a hand held above that container, so that the system can correctly interpret when a captured auditory object is dropped into a particular container and when an auditory object currently held in a container is again grabbed by the user. The accelerometer is used to detect when the container is inverted and shaken, the user's gesture to remove a currently held note. Additionally, conductive materials on the lid and top are used to detect the removal or replacement of the lid for temporarily silencing the note. Power is sent up one side of the container and across the lid to an input on the other side of the container; when the lid is replaced, this circuit is closed and power flows to the input. All sensor and interaction data is sent to the computer via Xbee enabled Arduino boards.

The Audio Oxide project is intended to provide a series of smart containers where a user stores small audio keepsakes, to be remembered at a later time. By associating individual audio stimuli with separate visualspatial locations, here in the form of containers distributed around a room, the interaction paradigm is consistent with what we validated in this thesis for diagnosticity of Alzheimer's disease. The network of containers could then be probed at the user's discretion as a cognitive "check-up". The user hears a series of auditory cues and has to interact with the containers where the audio was originally stored.

CHAPTER FOUR

Enabling Musical Expression to Target Neurorehabilitation

When providing access to creative opportunity while assessing emergent change in cognitive or physical domains of interest, what remains is to structure an intervention that takes place over *time*. Unfortunately, open-ended creative environments, at scale, are orthogonal to the incremental scaffolding necessary of an intervention. The hallmarks of intervention are the ability to reproduce treatment and attribute causality. New tools are implied that simultaneously provide access to creative experience, establish a vector of assessment seamlessly integrated into the task, and structure an experience from the outset through to the target goal. This goal should be both creative and clinical. Our patients can emerge empowered in their lives, and in their health, as part of the same experience.

In our creative lives, there is no disease, only a particular definition of expression. As Machover states, "we can all be creative, we simply lack the right tools." Human expression similarly lacks the right transport. Part of our concept of identity is our individual expression executed in the world. But in which contexts are we being expressive? How are we using our expressivity? With the right tool, expressive intention can be captured, crystallized, and amplified. Music instruments can wrap to individuals no matter what their musical backgrounds. Personal instruments can amplify the intention to perform, and to express oneself.

The issues are the same for the general population and for profoundly disabled individuals. Technology can capture personally meaningful expressive movement, or the beginning of expressive intention, and connect it to the creative context.

In the Enabling Musical Expression projects, I first built an instrument for one of the composers from Hyperscore in the Hospital. Dan Ellsey is in his early thirties and has cerebral palsy. He cannot move below the neck and he cannot speak. He uses a text-to-speech device which consists of an infrared pointer that he wears on a headband and a computer screen that gets mounted approximately one foot in front of his face. The computer tracks the infrared light as he attempts to navigate a series of on-screen menus and icons to string together words and formulate sentences. It takes him approximately one minute to execute a sentence in a normal conversation.

Dan's instrument allows him to control the expressive performance of one of his Hyperscore compositions, in real-time, with nuance and precision, despite that being his fundamental limitation. The experience has been transformative. Not only does Dan exhibit physical control in the creative performance context that is significantly different than in other physical therapies, but after demonstrating the ability to express himself in time, with control, his treatment team convened to reconsider how the hospital was structuring access to the outside community. Dan has since given performances for thousands.

In examining the design and implementation of Dan's personalized instrument, several principles are elucidated to create new, home-based, physical rehabilitation devices.

4.1 Enabling Musical expression

The first Enabling Musical Expression project consists of a vision-tracking interface that captures limited and symptomatic movement of the quadriplegic user. In software, the system introduces a minimal amount of filtering, while calculating descriptive statistics of the movement from basic kinematics to quantifiable measures of curvature. The analyzed signal is then mapped to parameters relevant for musical control.

The user can simultaneously control both discrete musical events, such as orchestration changes, as well as continuous events such as volume changes for

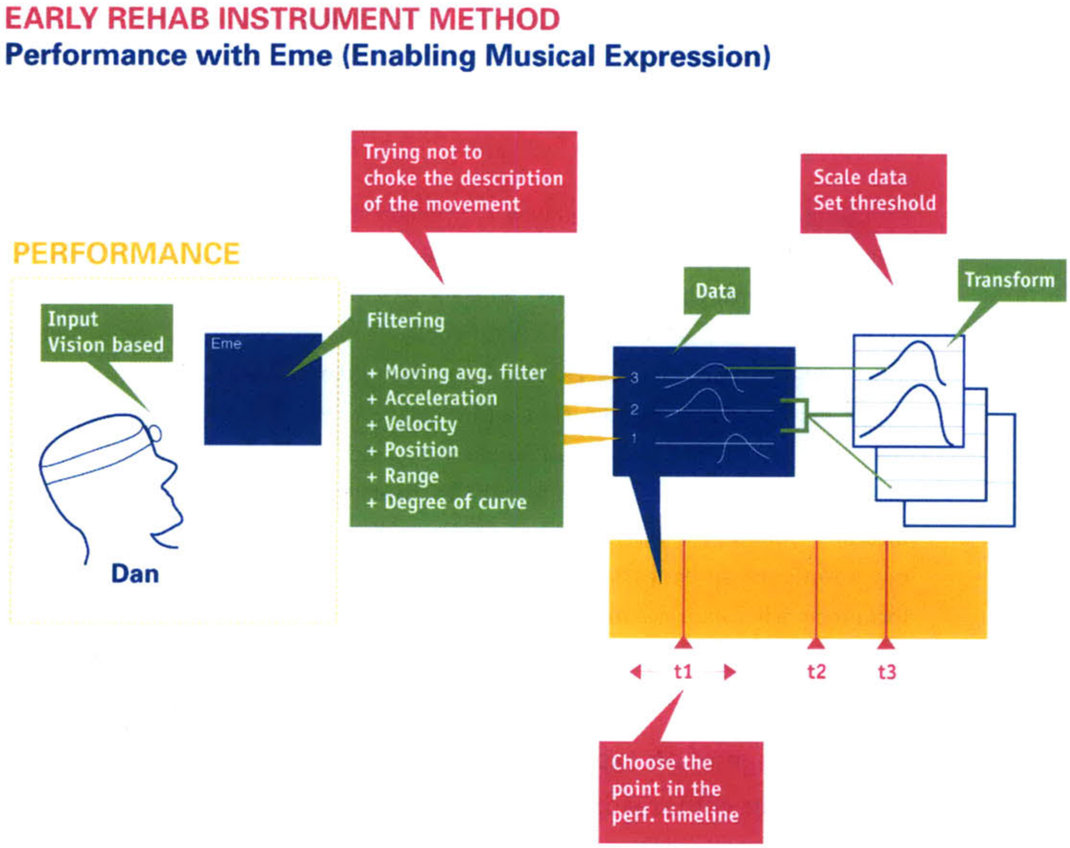


Figure 4-1: Infrared LED input, on the left, in two dimensions, is filtered, as seen on the right. In addition to mechanics we measure qualitive features of the resultant shapes such as degree of curvature versus angular movement.

individual tracks of the composition (Figure 4-1). Rather than presume that one mapping scheme is effective for all music, the parameterization of mappings changes over the course of a piece. Typically, we work together with a subject to determine how to parameterize around key moments or sections in a piece of music, so that the subject may convey the musical expression that they would like to at a given time in a piece of music.

The key finding of the Enabling Musical Expression project is that an expressive controller can drive functional improvement. After several months of intensive rehearsal with the instrument, the subject's primary physical therapist was invited to observe any outcomes and offer an analysis relative to ongoing treatment goals from traditional physical therapy. As reported by the subject's primary physical therapist, the subject moves with more sustained accuracy and control when using the expressive music controller than what has been observed in physical therapy over the past twelve years of treatment. Additionally, he exhibits decreased involuntary tongue movements, and fewer episodes of ataxia.

Furthermore, the ability to convey expression in time, with nuance, and control, despite being severely paralyzed, has allowed the subject to redefine his relationship with the community. The subject has now given a series of performances for audiences of up to 13,000 people. He is employed as a spokesperson for a text-to-speech device. He also frequently gives lectures in the community about his process of composition and creative work. The technology provides the structure for the individual to show marked improvement in a domain of deficit, while engaging the community in creative and empowering work. This interface enables an individual, who takes approaximatley one minute to spell out a sentence on his text-to-speach device, the ability to shape a musical performance with expression, control, and nuance. The subject does not have similar opportunities for communication in other adaptive environments.

4.1.1 Design of an expressive music controller for a Hyperscore composer with quadriplegia

The following design stages outline the process by which an expressive music controller was developed for one of the composers from Tewksbury Hospital's Hyperscore sessions.

Development begins with fairly unstructured observations of the subject's expressive movement during selected music excerpts. Then, input strategies are explored and decided upon. A metaphor is chosen to fit the observed movements within a music performance context. For example, when choosing a metaphor, designers may select a conducting system [Boulanger and Mathews, 1997], performance system [Weinberg et al., 2002], hyperinstrument [Young et al., 2006], or generative composition system [Xenakis, 2001]. Within a system, ideas as to whether the resulting experience should be *similar to* string instrument performance, or piano instrument performance, dictate which variables best describe the movement, and how to map the movement to particular kinds of musical events. After choosing a metaphor, features of the movement are quantified, and assigned to musical events that fit the performance metaphor for a given piece of music. Finally, expressive control is rehearsed for the given piece. Scalars

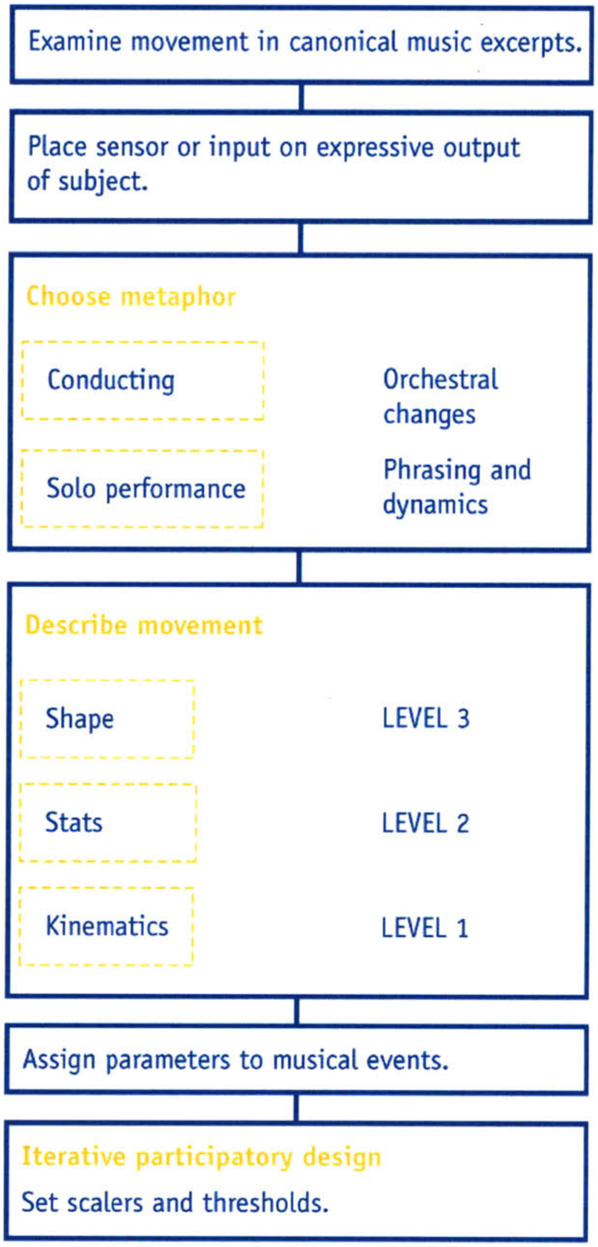


Figure 4-2: Design stages to develop expressive music controller for an individual with quadriplegia.

and thresholds are parameterized based on whether the subject can successfully achieve their intended expression within a section of the piece (Figure 4-2).

Initial assessment during movement to canonical expressive music excerpts

Dan is quadraplegic. He only has control of his head movement, and that control is imperfect and limited. It is difficult for him to sustain deliberate and continuous control. Often, his body will go into ataxia, in which involuntary muscle contraction will interrupt his attempt to point with his head or maintain a position with his head. Furthermore, his movement is more pathological in one hemisphere than another.

At the beginning of the music performance work with Dan, I recorded his movement both during and after listening to short musical excerpts. The goal was to assess to what extent Dan was able to make concerted, reproducible movements that he felt accurately represented the types of expressive elements that were indicated in the music. Sessions were filmed and reviewed before beginning any kind of technology development.

We would first listen to a musical excerpt several times without movement. I would then ask Dan if he had an idea regarding what made the excerpt "interesting" and "expressive". If he did, then we would listen to the excerpt again, and Dan would try to move in a way to represent the expressive features of the excerpt. He was told to move as if he was playing the music, emphasizing the parts that made the excerpt distinct. We never identified the salient expressive features of an example directly, through discussion or as part of the observation format.

Musical excerpts were chosen from various genres and styles, with different numbers of instruments and with a range of salient expressive features. For instance, in one particular cello example, the performer starts softly, slowly, before launching into a frenzied series of notes in the upper register of the instrument. This contrasted less salient exceprts such as the middle of a popular tune, where the expressive characteristics were fairly consistent with little to no departure from a set pattern.

As a result of movement with recordings, I confirmed that Dan is comfortable with head movements to abstract expressive features in the context of musical excerpts. The resultant head movements were consistent for expressive features such as dynamics. When given a musical example that highlighted dynamic changes, Dan would change his range and speed of head movement in persistent circular movements. In excerpts where dynamics remain constant, Dan would resort to head movement corresponding to note events, such as melodic contour or global section changes. For instance, he would move his head with a series of concerted down-strokes to correspond with a series of single notes, clearly demarcated one after another.

Dan's temporal resolution was poor. He could not successfully produce simple, head nodding movements in a consistent meter either on his own, with a metronome cue, or along with a musical excerpt. However, when an orchestral change would come at a distinct moment in time, Dan would characteristically make a large cue-ing movement to signify the event. The cueing movements tended to fall slightly after the the event occurred, even during repeated rehearsal with the same musical excerpt and learned anticipation of the event. Despite poor temporal resolution, Dan consistently attempted to represent discrete events, sparsely distributed in an excerpt.

Dan communicated to me that he was not confident that the types of movements he was making for the presented musical excerpts were "correct". However, given his lack of a performance background, the observed consistency of movement to musical dynamics and orchestration was encouraging. An input technology and musical performance metaphor were chosen to reinforce movement according to musical dynamics and orchestral changes.

Choosing the input

A head-mounted infrared (IR) pointer was chosen as the input mechanism for Dan's expressive controller. A firewire camera is mounted the same distance from Dan's head as his text-to-speech device, approximately two feet in front of him. The firewire camera has an attached filter that registers the IR spectrum. The IR point is tracked by its color in the streaming camera feed.

Head-mounted IR was chosen as an input method because Dan has more control and accuracy over head movement than elsewhere in his body. This is due to the nature of his paralysis, and the attention that head movement has received in both physical and occupational therapy throughout the course of Dan's life.

Dan also relies on head-mounted IR as an input to his commercial text-to-speech device. The text-to-speech interface provides visual feedback via a monitor mounted approximately two feet in front of Dan's face, suspended on an armature attached to his chair. The IR input is a pointer for a series of on-screen targets corresponding to word categories, phrases, individual words, and letters. Selecting on-screen targets causes words and sentences to be delivered by a speaker system.

By using an IR sensor for input we couple to movement where Dan has maximal physical ability, and we overlap with an application context where he has prior knowledge. Dan is familiar with the use of head-mounted movement for control. Furthermore, choosing a head-mounted interface allowed post-intervention assessment by Dan's physical therapist who could make direct comparisons between head-movment in the expressive music environment, and the text-to-speech environment as a baseline.

The disadvantages of using head-mounted IR for input pertains to interpolating from prior knowledge for discrete menu selections in the text-to-speech environment to continuos control in a music performance environment. The key difference between the proposed music performance environment and the text-to-speech environment is that the music environment requires the user to make continuous movements in time rather than a series of discrete menu selections. It was unclear whether Dan could adapt to the demands of continuous movement if he had *overlearned* the use of head-movement for discrete selection [Grol et al., 2009]. However, concerns regarding adaptation to continuous head control were dispelled during initial observations of Dan's movement to music excerpts.

Other input devices were considered, including Dan's distribution of weight in his wheelchair [Kunzler, 2007], breath (as in Rolf Gehlhaar's music performance device, Headspace), and head position in multiple axes of movement [Ford and Sheredos, 1995]. Whereas, Dan can shift his weight in his chair, he does so by contorting his neck and upper-body. Shifts in weight are derived from movement in his head and neck, therefore lacking direct control. Although, Dan does have control over his breath, he has little to no control over the musculature of his tongue and mouth. Various researchers have explored breath control in assistive technologies for individuals with palsy [music, text-to-speech]. However, through consultation with Dan's physical therapist, it was deemed that mouth or breath controllers would require several months of training to gain basic facility.

Finally, Dan has the most facility for head movement in the coronal plane. Sagittal and transverse movement were pathological. Dan could not repeat movements, initiate movements consistently, or guide movements with gradations of control along the sagittal or transverse plane. As an example, when given a target head movement from a forward leaning, ventral head position, across the coronal plane, to a dorsal position, Dan would attempt to initiate movement unsuccessfully, a variable amount of time would pass from half-a-second second to three seconds, and he would abruptly throw his head back and arrest his movement in the dorsal position. There were no signs of gradual movement or gradations of controlled movement between the dorsal and ventral positions.

Control parameters and signal processing

An infrared tracked point, streaming from a live camera feed, allows for several levels of control.

The first level of control consists of basic kinematics: velocity, acceleration, and their precedent, position in an x-y plane. At the second level of control we have summary statistics that further describe the mechanics: averages for position, velocity, and acceleration, minimum and maximum range. Summary statistics necessarily restrict the type of expressive movement being analyzed as the designer is forced to choose a bounded period of time to calculate summary statistics from the unbounded, continuous movement exhibited at the first level of control. At the third, and most sophisticated level, there are the models of classification, and categorization that assume the previous analysis while describing shape and pattern. I compare the ratio of arc length to euclidean distance for a tracked line segment to gain a representation of curved versus linear line (Figure 4-3).

The firewire camera has a variable transmission rate curtailed by processor load. Under the full load of the complete expressive music controller, firewire input varies between 125 and 133 frames-per-second (between 8 and 7.52 msec-per-sample) on a 2.33GHz Intel Core 2 Duo processor, with 3GB of 667 MHz ram, running Mac OS X 10.5.8.

Streaming IR position is first downsampled to 15 msec-per-sample. Velocity and acceleration are calculated at 200msec intervals. Minimum and maximum

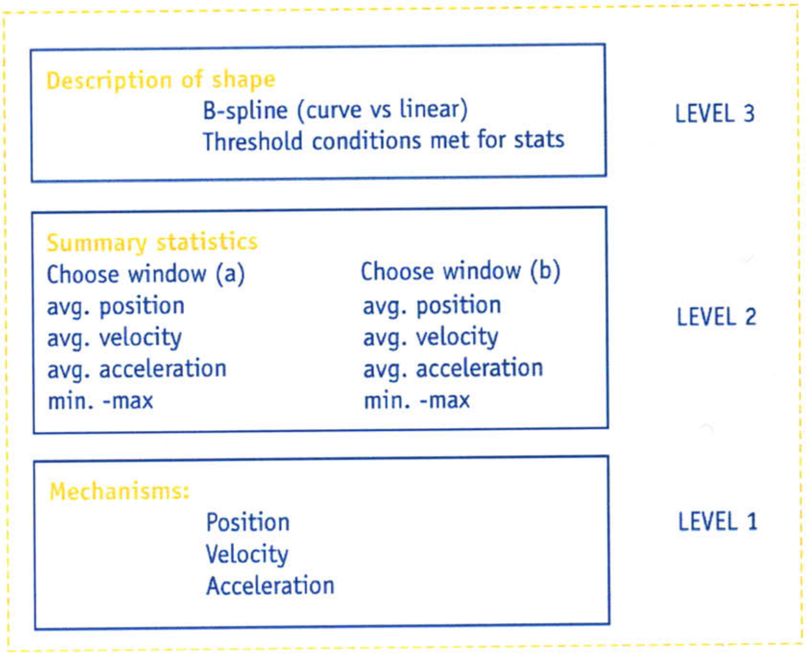


Figure 4-3: Levels of control derived from IR input.

position as well as range are calculated from a buffer of 25 position values, sampled at the frequency of the unfiltered, incoming IR signal (130 fps, 7.7 msec-per-sample). The 25 sample buffer corresponds to 192.5 msec between range calculations.

Downsampling factor, buffer-sizes, and the frequency of calculation were chosen by experimentation. A series of on-screen shapes were presented, such as circles, squares and triangles, in different sizes and locations on the screen. In a timed task, Dan was required to trace the on-screen shapes using his head pointer. He received continuous feedback as to the measured position of the head-pointer throughout the tracing task. I incrementally modified downsampling factor and window parameters through iterations of the task until the shape made by Dan conformed to the target shape with minimal jitter, and the kinematics measures appeared representative of the line segments that Dan would make for various discrete shapes.

To describe the degree of curvature of the incoming IR signal, the unfiltered IR input is first buffered in a 50 sample window, corresponding to 385 msecs. After the buffer is full, the cumulative distance traversed by the line is divided by the euclidean distance between the start and ending position. The ratio between the two values approaches one as the line becomes straight. The larger the ratio between the two values, the further the segment departs from a straight line.

The final ratio of curve to euclidean distance is a weighted sum between the 50 sample buffer and a buffer at 5 times the scale, corresponding to 250 samples, and 1925 msecs. By assessing curvature at both 385 milliseconds and roughly two-seconds, gestures made to correspond to single notes and short phrases of notes were both accommodated. The shorter time-scale was the predominant focus of Dan's movement, and the weights were kept constant at 80% for 50 samples, and 20% for the 250 sample resolution.

Expressive performance is not intrinsic in any subset of these three levels of control. After developing the input system, one has all of the information available for control and must determine the appropriate mapping between movement and musical control that will be considered expressive for a given piece of music. I decided to use metaphor to guide the selection of control parameters, their mapping to music characteristics, and the parameterization that would give the mapping variation over time. To find an appropriate metaphor, I examined Dan's movement.

Choosing the metaphor - assigning control parameters to musical events

Considering Dan's movement to music examples, and the physiological limits of his pathological movement, I decided on a metaphor to map movement data to a performance context based in part on the idea of conducting, and also, partially, on aspects of solo instrumental performance. Metaphor refers to the similarity of the performance context to existing models of activity. By engaging with an environment built from established performance models, such as conducting, or solo instrumental performance, Dan is able to utilize prior knowledge to exercise appropriate movement towards musical goals. Dan had not directly engaged in expressive musical control before we began the collaboration to design his instrument. However, the assumption was that Dan would have some concept of the way that conductors move, or the way that solo performers looked when playing music.

From conducting, I mapped the velocity and shape of his curved input line to control orchestration selections. Fast, angular movement changes instruments. Orchestration selections change between sets of instruments during the play-back. This selection process is constrained. I avoid note selection, in addition to harmony selection or any type of decision process that would significantly restructure the content of the performed work. In freeing the performer from composing in real-time, the performer aligns as closely as possible with their preconceived notion of the types of control indicative of music performance. A conductor will not traverse a decision tree of possible final musical pieces. In contrast, a conductor shapes a performance by modifying a discernible set of parameters for an already finalized composition. To rely on Dan's movement as a subset of his learned associations with traditional performance control, orchestration seemed a sensible mapping target.

From solo instrumental performance, I sought to control dynamics over time. Range of motion and shape of the curved input control the amplitude, in time, of the performance. Slower, curved movements allow the user to shape dynamics and create concerted phrases in the performance. The two predominant parameters for expressive performance in solo instrumental performance are amplitude and time. Because of Dan's pathological movement, which has poor temporal resolution, I decided to filter the movement in time, and to avoid a mapping forcing Dan to be a timekeeper, pushing and pulling the tempo of his performance. Musical dynamics are conducive to holistic, large-scale, shaping movements.

Assigning weights to control parameters based on location in a piece of music

After the input technology and the basic mapping were established, I began to work with Dan as I would with any artist. Repeated performance using the system lead to an understanding of where and how the mapping needed to be tuned relative to the piece, and performance that Dan wanted to give.

A *score follower* is started simultaneously with the beginning of a performance. The score follower yields an integer index for each note in the main theme of the composition, allowing the assignment of parameters based on location in the piece.

The three minute composition was divided into five different thematic sections. At each section, the threshold conditions required to make orchestral changes, and the weighted transform of kinematics values to the amplitude of musical tracks, were changed based on the goals of the section. Section markers and goals were determined over three months of rehearsal, and in discussion with user. Discussion rarely broached the technical aspects of mapping control to music parameters. Dan was unaware of the parameters being modified, and simply discussed whether the software was stifling or enabling him to make the changes that he intended to for a given section of the piece.

As an example, in moment of the piece where the first string accompaniment appears, Dan's system stops targeting orchestration changes and allows for a moment of highly sensitive dynamics changes assigned to the incoming string accompaniment. The result is that Dan is able to bring in the accompaniment and taper its dynamics with a high degree of specificity, before the sensitivity on the orchestration switching increases. When sensitivity of orchestral changes re-emerges in parallel with dynamic changes, Dan, in effect, turns away from the string section which was disproportionately important for that one moment in the piece.

It is not enough to define an input system and a set of mapping targets born out of a performance metaphor. Despite the fact that a concert pianist has his input system in place, namely, the piano, and the wealth of his mapping opportunities available continuously throughout a performance, expressive parameters are not handled equally at all times in a performance. At some points, based on the piece of music, and the plan of the performer, certain parameters dominate the performance. The dynamics and shape of some melodic lines may come to the forefront of the artist's attention and delivery. Dynamics may as a whole remain flat, in a moment, to play with timbre or the quality of sound. All of these decisions are made in parallel. For our technologies to be truly expressive, they need to be contextualized to the pieces and moments in which they afford an individual to be expressive, and exert their will as an artist.

4.1.2 Motivation

Motivation had a significant role in the observed physical effect. The prognosis of an individual with cerebral palsy is that they will not rehabilitate. Observed differences in movement are unlikely to correlate with substantial neurological change regarding motor function. Furthermore, the controller technology lies on the periphery of the subject's symptoms, reading pathological movement and interpreting it in a meaningful way for expressive control. I have successfully targeted the symptomatic behavior of interest, but it remains unclear as to whether I have impacted underlying disease process or merely motivated the individual to move differently. Other limitations of the Enabling Musical Expression project include the necessity for extensive development to parameterize the controller to one particular individual, challenging the ability to scale the interface to other users.

4.1.3 Generalizing the design method of the EME system to other individuals with palsy

The expressive controller built for Dan Ellsey was a collaborative effort. As a methodology, I took a participatory design approach to discover effective control strategies, a musical performance metaphor that would empower Dan to confidently shape a performance of one of his pieces, and the technology to link between control and the expressive context. The connection between control and music performance was born of discussion. The system was manually parameterized through successive iterations and rehearsal. Adherence to the design method required development to take place incrementally, in the hospital, over several months. As a result, the culminating expressive music controller is tailored to a single individual.

Future expressive performance devices can rely on a similar design process, but gain scalability as adaptive systems. Innovations in the areas of free-hand sketch recognition, computer vision, and machine learning each offer techniques to automate different stages of the development process, ensuring that similar devices can wrap to disparate pathological users, and generalize to healthy populations as personalized instruments.

The manual components of the current design process are the following: extracting observed movement features that correspond to expressive music features, parameterizing buffer sizes for kinematics and curvature data, and assigning thresholds throughout a target piece of music. When working with Dan, it was first necessary to determine whether he could move consistently to express features of musical performance. Then, it was necessary to choose a performance metaphor that supported the observed movement. Reproducing the current system with adaptive components avoids initial, open-ended observations as to movement and musical expression.

The design process for an adaptive version of the current system would begin with calculating appropriate buffer sizes for shape representations of a user's input. Interestingly, the field of free-hand sketch detection is similarly concerned with partitioning a drawing, made in real-time, into distinct line segments and curves despite motor noise and error from sampling the incoming signal [Brieler and Minas, 2010, Ye et al., 2009]. To combat the issue of multiple users representing similar shapes at varying scales, Sezgin and Davis [Sezgin and Davis, 2006] propose an implementation of a scale-space representation framework [Witkin, 1983] to derive multiple representations of shape features, and choose a filtering scale that best satisfies their conditions for selecting control points.

The scale-space framework yields increasingly higher level descriptions of a signal by convolving it with a filter containing a smoothing parameter that varies the filter's granularity. The constructed scale-space is then used to compare how the variables in question behave over different filtering scales, and to select a scale that meets predetermined criteria. In our case, we would want to maximize the ratio of arc length to euclidean distance to determine the buffer-size for calculating curvature. In a second scale-space representation, we would want to select control points between the start and completion of a line segment by observing how the maxima of curvature change at various scales[Lindeberg, 1994]. Determining the start of curved shapes versus linear segments would allow a binary parameter to replace a user-determined threshold.

After conditioning the input signal, machine learning methods could be employed to extract expressive movement features for given musical excerpts and to interpolate parameters for dynamic and orchestral changes for a piece that has never been performed. The use of machine learning methods to model expressive music performance is an active area of research [see Widmer and Goebl, 2004, for review]. Often, timing and loudness variations of performed notes are utilized to generate performance signatures for a given artist. Performers have been classified and identified by their signatures [Saunders et al., 2004]. For our application, kinematic and shape data, time-stamped according to its occurrence with regard to the musical excerpt, would be substituted for timing, loudness, timbre, and note attack The resultant performance signatures would be in terms of movement parameters. Given a new piece of music, the user's expressive signature for similar expressive examples could be interpolated to establish baseline measures to map between movements and dynamics, while allowing for variation during actual performance.

In addition to extracting relevant features of movement, the context of movement must also be represented. Score-based methods posit the musical score as the central context for the resulting expressive performance. Different levels of feature extraction have decomposed the score either by individual notes[Todd, 1995], phrases [Friberg, 1995], or global structures[Windsor et al., 2006], and demonstrated the correlation to expressive performance parameters.

Each of these approaches would require significant research to adapt the methods to expressive control in circumstances where the users exhibit pathological gestures and have little to no familiarity with expressive movement. However, adaptive filtering, feature extraction of human performance data, and interpolated mapping based on a learned user-model, are feasible extensions of the existing application. Supporting research for an adaptive version of the current system exists from multiple disciplines, including music.

4.1.4 Key Limitations on the periphery of biology

Technology can do more than provide access. The Enabling Musical Expression project is tailored. It is an expressive interface that wraps around only Dan's movement, what he can do well, and what he cannot do. But, it is limited, existing on the periphery of underlying pathology.

Not only is the underlying biology the genesis of pathological movement, but it is also the dimension in which you are trying to facilitate change. The "rehabilitative" design of the Enabling Musical Expression application is lacking. Despite the observation significant behavioral change, the mechanism of change is lost. It is possible to go back, revisit the project and assess various pre- and post- factors that may correlate to the observed change, but, it is much more proactive to build a better tool that integrates underlying pathology into the interface. In contrast, interpreting symptomatic behavior can only, at best, provide a useful abstraction of the rehabilitative process.

The cost of developing rehabilitation interfaces on the periphery of biology is realized in the areas of scalability to multiple users and generalizability beyond a single domain of deficit. By interpreting summative gestures such as head movement, limb movement, or hand gestures, the technology is mapping inputs to a desired outcome. The complex, multi-storied input is being abstracted and attached to useful output. One user's best metaphors of performance and interaction are unlikely to be appropriate for another user. Furthermore, one set of abstractions from global gesture is unlikely to be appropriate for any domain other than the one being designed for.

The result is an interface that necessitates design from metaphors. The designer is forced to make assumptions regarding the suitability of summary movements and global gestures, as she connects abstractions of movement to target functions. Metaphor and mapping issues characteristically plague the field of music and expressive controllers[Fels et al., 2002]. The emphasis is usually wrong. Rather than interfacing to conducting, developers implement something "like" conducting. The Enabling Musical Expression project also occupies this difficult design space. The user "sculpts" an interpretation of performance. Gestures that seem appropriate for conductor-like functions provide conductor-like functions (they change orchestration). In this particular interface, conducting gestures begin and end at discrete points in time, that have definable shape, and they do not take too long to execute. Gestures that are similar to phrasing gestures are assigned to the control of dynamics. Phrasing gestures permeate larger musical constructs, they are less definable in time, and they are more amorphous in shape.

The best metaphors are well researched. We conducted extensive participatory design sessions with our intended user. The metaphors were largely determined by him, both in terms of his available functionality, his desired range of interpretation for a contextualizing single piece of music, and the gestures that he felt best captured the dynamic changes he wanted to assert in his piece of music. Other researchers similarly explore the metaphor space before engaging in abstraction at the interface. We can do better than abstraction. We can provide access to authentic composition, authentic performance, or instrument learning, using technology to guide the user towards something that is immediately rewarding, with the opportunity to grow indefinitely.

4.1.5 The rationale for connecting to underlying biology

There is something deeply unsatisfying about systems that interpret your interaction and enable you to do something that is just a facsimile of what one really wants to accomplish. The balance between access, scaffolding personal growth and gaining personal investment in learning new creative skills is difficult to achieve. Wildly successful abstractions of performance, such as the Rock Band video game system, fail to transition users to the meaningful acquisition of skills; Although, games to address this market are surely soon to come. Dan's investment in his technology is most likely derived from the extreme circumstances in which he is barred from alternative means of personal expression. He has an interpretive controller with which he can provide immediate and powerful expression, despite that being his fundamental physical limitation. In the creative environment, he is functionally different. However, the interface does not provide the opportunity for him to grow as a performer.

The disconnection stems from the interface existing on the periphery of his underlying biology. The criticism extends beyond expressive technologies for healthcare and into expressive technologies overall. Interfaces need to connect to the most subtle, the most elementary beginnings of the creative act. Rather than create accessible systems that are like guitar playing, use technology to look at fledgling guitar playing, buttress attempted performance with rewarding musical experience from the beginning, and feed-back to push the user further along the vector in which they are trying to perform.

Only by connecting to underlying pathology can we develop a rehabilitative tool that provides access to significant creative experience without the burden of metaphors and abstraction. We avoid mapping, and the prohibitive decisions that one invariably has to make to proffer a rough approximation of appropriate input to a rewarding experience.

Conveniently, in physical rehabilitation and instrumented environments we have access to the peripheral nervous system, and a direct neural transport of the disordered underlying signal. For rehabilitation to be successful, movement must change, and the peripheral nervous signal must be the cause of change. By interfacing with an underlying biological signal to drive the interface, we simultaneously have insight into the signal that is likely to show signs of change throughout successful rehabilitation. If the system can adapt to rehabilitative changes, then the underlying biological signal is both an interface and assessment opportunity.

Finally, the analysis of a disordered but consistent pathological signal in the context where an individual attempts to execute a discrete motor program is the access point to both drive a rehabilitation and to create an interface that generalizes to a diseased population. Beyond creating better applications, there is a strong rationale for using the underlying pathological signal to reinforce the unique multisensory implications of creative interfaces. Novel rehabilitations stand to benefit from accessing intact domains of processing as part of the rehabilitative task [Schlaug et al., 2009]. By partnering with underlying biology within the creative interface, we gain access to neurobiologically significant multisensory rehabilitation both in terms of research and the feedback that is most likely to rehabilitate our intended population.

4.2 EMG interfaces to target underlying biology

To escape the limitations of interpretive mapping in systems that connect users to a facsimile of creative and human performance, and to connect directly to underlying biology where the neural signal in question can be handled both as an input, and an assessment measure, I have implemented a piano-learning and performance interface driven by an electromyogram. The interface is intended to amplify the intention to perform - to model a hemiparetic subject to access piano learning and performing environments, with control, expression, and precision, with little to no finger movement.

The basic structure of the second project in the Enabling Musical Expression area consists of an 8 channel electromyogram (EMG) interface that a user wears on their forearm. The output of the sensors are streamed continuously. The computer will have previously trained on the subject's attempted finger presses. As long as the underlying signal is somewhat consistent, a model is developed that classifies the attempted finger press based on the disordered pattern of EMG signal. The user is presented a score, which defines the rehabilitation context and music performance task. As the user attempts to play notes in the sequence, the attempted finger presses are recognized, and the corresponding notes are played and heard automatically, whether or not the user could successfully produce the movement. The continuous attempts at movement, and the multisensory reinforcement (Figure 4-4).

The electromyogram is a biosignal frequently employed in the clinical assessment of neuropathy associated with movement. The EMG signal is a summary of muscle unit action potentials in the vicinity of a given electrode. Depending on the specific EMG technique utilized, it is possible to mathematically derive local characteristics of muscle units and global characteristics of muscle groups. The most informative EMG statistics are root-mean-squared ampli-

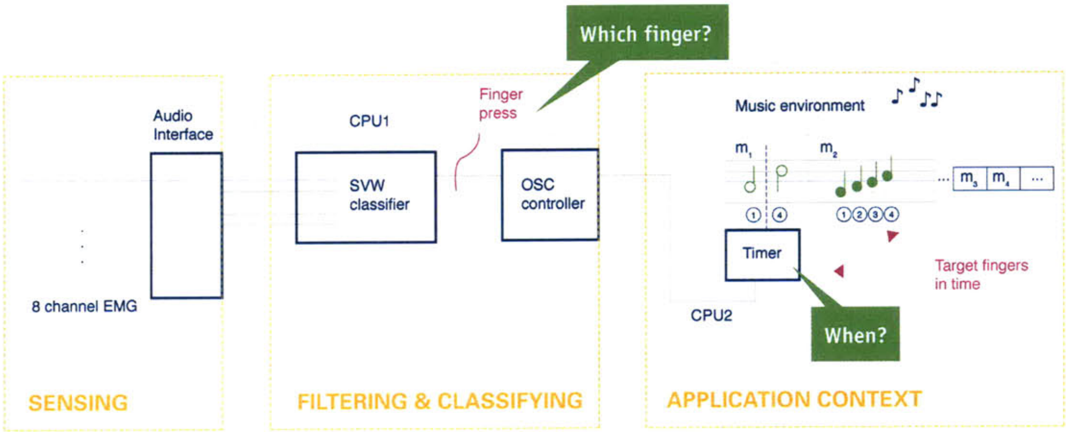


Figure 4-4: Sensing, filtering and application stages built to access underlying biology (EMG), assess as part of the interaction (SVM Classification), and structure context with significant value beyond the rehabilitative task (Piano Learning).

tude and mean frequency, which correlate with muscle atrophy, re-innervation, spasticity, fatigue, conduction velocity and coordination of antagonistic muscle groups by task. The limitations of EMG for assessment include significant inter-subject variability and requisite knowledge of underlying physiology as the EMG signal is affected by variables such as skin conditions, underlying fatty tissue and proximity of the electrode to the innervation site along the muscle fiber. Placement of EMG electrodes is exceedingly difficult.

4.2.1 EMG

In addition to clinical assessment, EMG has been employed in prosthetic control [Chan and Englehart, 2005], rehabilitative robotics [Dipietro et al., 2005, Marchal-Crespo and Reinkensmeyer, 2009, Hu et al., 2009], and human-computer interfaces [Saponas et al., 2008]. In these fields, reconstruction of muscle unit activity is not as important as the measurement of a consistent pattern of muscle unit signal, in a single session, that can be classified with attempted behavior and utilized to control prosthetic limbs or software applications. Machine learning and artificial intelligence are convenient tools to decompose noisy, inconsistent, and non-parametric continuous signals. Many novel control systems are now using these techniques to model a user's EMG response and produce actions coupled to muscle firing in prosthetic devices or software. Rather than definitively associate often unreliable and imperfect pathological signals to individual muscle units, classified patterns of EMG signals are used to represent attempted muscle coordination which then drive the desired output in a prosthesis or software application. Several devices employing this approach include whole-arm flexion-extension devices [Stein et al., 2007], ankle prosthesis control [Au et al., 2005] and arm prostheses[Kuiken et al., 2009].

The choice to use EMG was made in relative to potentially simpler, more cost-effective, and more practical sensors. EMG systems are expensive, and they require fitting and unfitting by an individual who can best place the electrodes with respect to muscle activity that changes from session-to-session. Low-cost EMG sensing solutions are an active area of research, yielding various methods suitable for embedding in home devices[Cheney et al., 1998, Saponas et al., 2010]. Also, machine learning approaches to EMG filtering and classification may be able to account for imperfect signals derived from sub-optimal positioning via attribute selection [Yan et al., 2008], although, such methods have not been necessary in pilot development with healthy subjects.

4.2.2 SVM Classification

This device leverages recent findings in healthy individuals, in which support vector machines were used to develop a user model of finger-press behaviors from 8 EMG channels recorded from the forearm with little to no attention to sensor placement relative to innervation site, or targeting effector muscles directly associated with finger movements [Saponas et al., 2008]. Rather, this quick-and-dirty approach successfully developed user-specific models from whatever noisy and secondary muscles were coordinated with finger presses. After approximately 50 trials, user's finger presses were classified with 95% accuracy.

Elaborating on this technique will eventually allow our platform to be extended into real-time classification of stroke victim's attempted finger presses. The advantage of a machine learning approach is that whatever signal the subject exhibits is meaningful, as long as it is somewhat consistent in the context of an intended movement.

Dipietro and colleagues, investigating the feasibility of EMG methods for robot-assisted therapies, reported that their most impaired subjects exhibited EMG modulation in targeted, impaired muscle groups even when speed of movement in the measured limb was effectively zero. Despite a lack of discernible movement, severely impaired subjects were able to engage emg-trigerred robot assistance. The researchers conclude that EMG triggering might be particularly useful for the most severely impaired subjects [Dipietro et al., 2005].

4.2.3 Training

To train the classifier, the application instructs subjects to attempt to make a finger press when prompted by the computer. Subjects are not expected to be able to successfully complete finger presses. Rather, the imperfections of their muscle response, if consistent within some reasonable margin of error, will be sufficient to model attempted finger presses.

EMG is recorded at that moment of an attempted finger press. The classifier model is updated with each finger press attempt. After sufficient recording, the model is able to identify the most likely attempted finger press from continuously recorded EMG data outside of training tasks in real-time environments. Classified finger presses serve as the input into the multimodal feedback environment.

Classification is conducted with a support vector machine, implemented in Java using the Weka machine learning API [Hall et al., 2009], which utilizes Platt's sequential minimal optimization algorithm for SVM training[Platt, 1999]. The feature vector used to train the model consists of root-mean-squared amplitude, mean frequency, spectral density, and autocorrelation between the channels. Eight channels are recorded simultaneously and band-pass filtered between twenty and four-hundred Hertz. EMG data is sampled at 44.1kHz via a 16 bit MOTU Traveler Audio Interface, set to sample at 12 bits. The Delsys, Bagnoli 8-channel Desktop EMG System delivers 8 channels of amplified analog output.

4.2.4 Feedback

After measuring a user's attempted finger press, this system provides auditory, visual, and tactile feedback in the form of piano notes. The neurological goal is to systematically engage multimodal integration during a task that binds between the domains pertaining to action representation. We want to create the feedback of an anticipated, consistent, and reproducible multimodal event coupled to the intended movement.

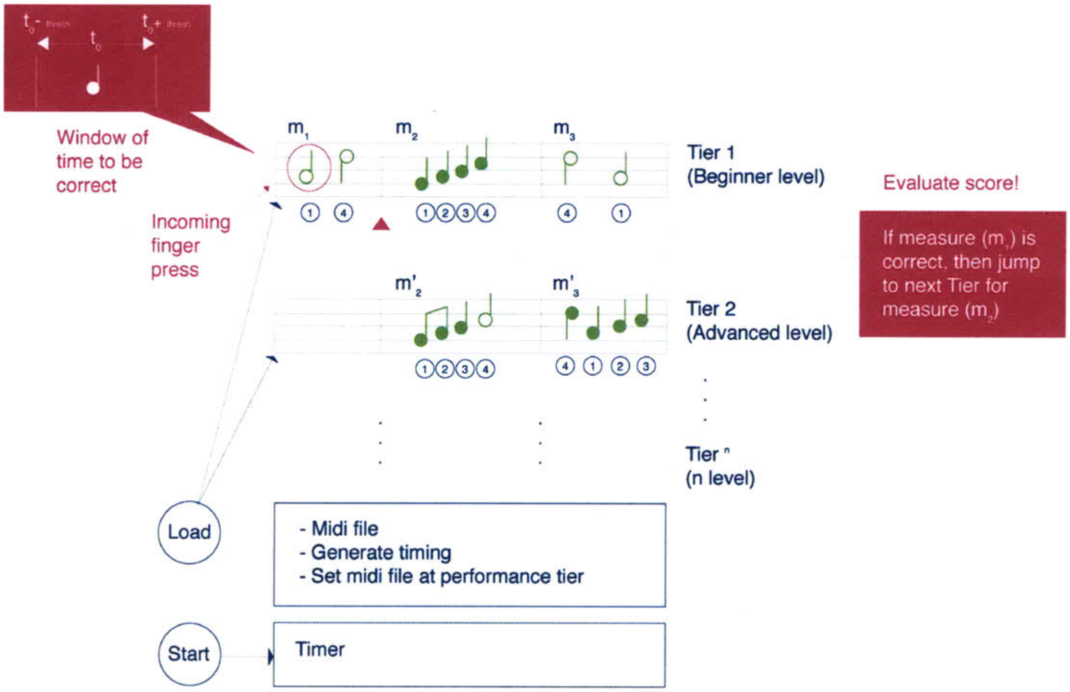


Figure 4-5: Performers are scored at the end of individual measures as to their ability to enter the correct fingers at the correct timings, indicated by the musical score for that measure. If successful, performers move up the tier system to a more difficult measure.

4.2.5 Intervention

Finally, the system structures an intervention. I have developed a creative music learning environment to contextualize the multimodal feedback, so that it becomes subsumed by piano learning (Figure 4-5).

Users have a single piece of music that begins with a two-note sequence. The two-note sequence corresponds to two different finger presses. When the system detects that a user has successfully delivered a muscle unit signal classified as the attempted finger presses corresponding to the two instructed notes, the piece expands on those two notes to a slightly longer sequence (Figure 4-6). In addition to increasing the sequence available to the user, the system also expands in the way it accompanies a user, offering a stronger sense of phrasing and dynamic control based on how efficient a user's measured movement signal is with respect to the target movement. The more that a user changes, functionally, the more the system will respond to their improvement to incorporate musical and expressive control. As the piece grows based on successfully at-

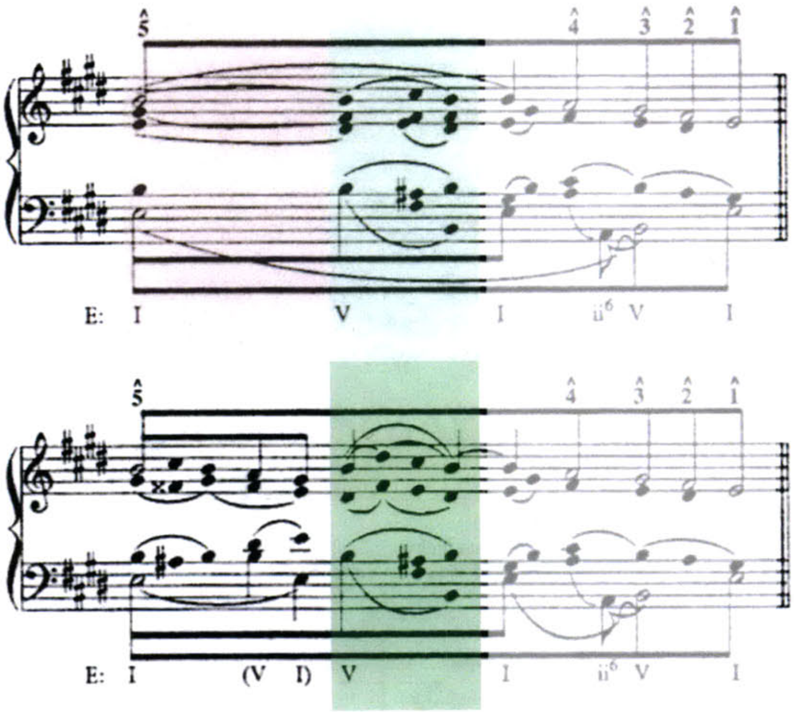


Figure 4-6: As users enter in correct notes, measure-by-measure, performance proceeds up a hierarchy of increasingly difficult note groups.

tempted finger presses the user will be supported in gradual motor learning in an expressive multimodal environment.

4.3 Multisensory Neuroscience

The seamless construction of a movement with a goal, and the successful production of that movement, requires action representation. Action representation is thought to be a neurological embodiment of an intended movement across sensory domains. From the translation of a motor unit potential to the coordinated firing of a motor unit action potential train, to the patterns of motor region neural activity during imagined movement, some coordinating network of neurological motor response is implied to"represent" and establish the intended motion. Furthermore, we know from research into multisensory perception and integration that action representation is a phenomenon that incorporates auditory, visual, and tactile domains of processing with respect to movement. However, no interface currently exists to consistently vary multisensory feedback during the rehabilitation of pathological movement. As physical rehabilitation tends to focus on rehabilitation through use, what would rehabilitation through representation mean? Can the brain, and in particular, intact regions of multisensory processing, be used to bootstrap the rehabilitation of motor representation, and subsequently, movement.

The EMG system is a platform to study multisensory integration during rehabilitation tasks. Human experience is multisensory. We act and observe action in the world amidst parallel sensory inputs. In many circumstances, we rely on information from simultaneously presented sensory domains to act faster and more accurately than in unisensory environments[Hecht et al., 2008, Teder-Sälejärvi et al., 2005, Molholm et al., 2004]. The building blocks of multisensory perception are receptive fields of neurons that respond to auditory, tactile or visual stimuli and often respond greater to multiple sensory domains presented with similar spatial and temporal proximity than presentation of unisensory stimuli.

Whereas multisensory receptive fields may perform perceptual integration, evidence is emerging in support of cortical regions that encode abstract representations between sensory domains and observed actions. Differently than integration, the actions themselves and the sensory stimuli that represent the actions, are thought to be encoded across a distribution of brain areas [Galati et al., 2008].

Two theories, not necessarily mutually exclusive of one another, expressly deal with action observation. First, embodied interpretations of perception, also referred to as motor theories of perception, claim that the perception of action activates motor processing areas that represent the underlying systems responsible for producing the perceived action [Rizzolatti and Craighero, 2004]. Secondly, specialized neurons have been identified in the macaque area F5, thought to be a homolgue to Broca's area in humans[Grèzes et al., 2003], that are responsive to the site of actions, performance of actions, or the sounds of actions [Kohler et al., 2002]. These "mirror neurons" have been proposed as an underlying mechanism that, if established in humans, could be a key neural module for the development of speech, social learning, and integration of high-level abstractions across sensory domains including in action observation [Rizzolatti, et al., 1999]. It is fairly common to discuss the human "mirror-neuron network" as an extension of mirror neuron selectivity into various brain areas that relate to action observation, including multisensory perception, semantic interactions in perception, motor goal perception, motor imagery, and the process of associating motor actions and sensory stimuli.

Direct analogues between single-unit mirror neurons and networks of human brain regions involved in action observation and perception may be premature [Dinstein, 2008] as well as their extension into uniquely human phenomena such as speech [Turella et al., 2009]. It is fairly well established that humans have diverse neural resources available to share representations between motor actions and the sensory environments that are associated with those actions. However, the conceptual distance between single neurons that are responsive to actions as well as unisensory phenomenon and an emergent "representation" of those phenomenon is yet to be established.

This paper examines the extant literature pertaining to the cortical areas proposed as an action observation network, focusing on auditory-motor representations. After outlining the basic auditory-motor representation system, we present findings from motor expertise literature that extends principles of sensorimotor representation into adaptive, dynamic processes. Finally, several critical questions remaining to be addressed in the action representation literature are raised, along with a proposal to examine paretic rehabilitation to address these questions.

4.3.1 Functions of a human action observation network

Neurophysiological and neuroimaging studies support a network of regions that that process various aspects of action observation [see Rizzolatti and Craighero, 2004, for review]. Matching representations to motor templates, motor goals, abstraction of action representation beyond individual effectors, multisensory integration, and encoding semantic or conceptual categories of different types of motor actions are overlapping functions of a shared network of regions. These regions tend to be focused in frontal parietal areas, posterior parietal cortex, dorsal premotor cortex, and superior temporal sulcus.

The goal of an observed or intended action is different than the means to achieve the action often specified as the "what" and "how" of actions. In humans, attention to the completion of action goals has been shown to activate bilateral superior frontal, angular gyrus, left precuneus, and middle temporal gyrus, whereas, attention to their means of completion in the same action context activated ventral PMC and inferior parietal cortex [Hesse et al., 2009]. In scenarios where subjects reconstruct means when given a goal, they incorporate left PMC and right dorsolateral PMC. Conversely, when reconstructing the goal when given the means to an action scenario, subjects activate medial PFC [Chaminade et al., 2002]. Comparing imitation[Koski et al., 2002] and observation[Iacoboni, 2005] that have no obvious goals to those that do, yields activation in both dorsal premotor cortex and Broca's area.

One open question regarding auditory-motor representations pertains to the effector specificity of sensorimotor representation. To what extent does multimodal representation incorporate the concept of the subject's body and the ability to reproduce the action? In primates, mirror neurons have been shown to activate in reaching tasks whether the action was made with the hand, mouth, or with a tool, suggesting abstraction beyond the effector in the representation of an action's goal[Ferrari et al., 2005]. However, studies in humans have found that (transitive and intransitive) actions, when compared with static images of those actions, led to differential activation of PMC and Broca's area based on which effector was used, mapping to somatotopy of known effector representations [Binkofski and Buccino, 2006, Buccino et al., 2001]. Whereas primate mirror neurons may demonstrate abstraction beyond effector specificity, the human action observation network has not yet demonstrated such an abstraction for the means of accomplishing a goal-directed action.

In addition to motor representations being cued by the unisensory stimuli that share their representation, as in the sounds of motor actions [Kohler et al., 2002], human responses to unisensory cues sharing representation for motor actions have been explored on the basis of their semantic relevance within conceptual categories. For instance, auditory cues differentially activate regions pertaining to motor representation for actions compared to non-action sounds [Pizzamiglio et al., 2005]. Schubotz contended that auditory cues could be divided among goal-directed, effector specific, and somatotopically organized activations in premotor cortex depending on which type of representation the auditory cue is accessing temporal, object-related, or spatial[Schubotz et al., 2003]. Additionally, auditory cues corresponding to distinct categories such as tool sounds, and animal sounds, have been shown to activate premotor and medial superior temporal gyrus, similarly to the imitation of the actions represented by those sounds [Lewis et al., 2005].

Multisensory receptive fields are known to respond to auditory, visual, and tactile stimuli, and have long been associated with speech processing [Calvert et al., 1999, Stevenson and James, 2009] and potentially higher-order percepts such as actions and tool use[Doehrmann and Naumer, 2008]. Investigations of multisensory percepts at the representation level have distinguished between cross-modal integrative functions in known multisensory areas, such as superior temporal sulcus, and action representation areas, such as inferior parietal lobule [Kaplan and Iacoboni, 2007]. Whereas STS responds to auditory alone, visual alone, auditory-visual stimuli, abstract images, action representations and semantically categorized percepts, other perceptual areas show selectivity for sensorimotor representations.

An important distinction is inherent to the methodologies in which multisensory function is measured compared to motor representation networks. The majority of multisensory receptive field research exists on the unit recording scale, or via electrophysiology measures. The former characterizes neuron response properties to multiple sensory inputs, and has elucidated the features of disparate sensory domains that multisensory receptive fields are selective to: such as superior colliculus and the superior temporal sulcus. The electrophysiology research, due to its unique temporal resolution compared to imaging techniques, is responsible for an understanding of modulation between unisensory areas in multisensory environments, in which areas previously thought to be unisensory, such as auditory cortex, show modulation by visual stimuli [Kayser et al., 2008, Giard and Peronnet, 1999].

The scope of these findings uniquely situates multisensory processing as functionally responsible for sensory "integration" which is different than abstracted "representation" of higher-level action observation networks. Multisensory integration is understood as an area that integrates between sensory domains as a function of superadditivity [Miller, 1982]. Superadditivity is increased activation beyond the sum of responsivity to the separate unisensory domains (audiovisual $>$ audio $+$ visual). Currently, there is little evidence to support that multisensory integration sites are modulated by representations of actions. However, it has been shown that multisensory sites are responsive to the site of actions, although these actions tend to be devoid of goals, effector specificity, or semantic relevance [Barraclough et al., 2005]. Rather, sight and sound of actions modulate STS by their ability to predict the resulting unisensory activity [Stekelenburg and Vroomen, 2007, Noppeney et al., 2008], which is different than sensitivity to abstracted representation for the multisensory information.

Because of the different level between sensory information that multisensory areas are selective for compared to representation networks, multisensory areas, such as STS, are often identified as key early modules of paired feed-forward and inverse models of motor planning and action[Haruno et al., 2001, Iacoboni et al., 2005]. In these models, multisensory areas produce an efferent copy of desired action, which is fed to higher-level representation and planning cortices. Inversely, frontal parietal areas would receive STS integrated sensory information and incorporate goals.

Inverse and feed-forward models of human action observation networks are unlikely to account for all of the components of action understanding and production. For instance, in a recent review of action comprehension in non-human primates, authors Wood and Hauser argue that goals and intentionality, evaluation of rationality of action, goal prediction, and action reasoning in situations where the agent cannot move, all require inferential mechanisms beyond a motor representation solely based on observed actions and the motor capacities of the subject, as in template matching models[Wood and Hauser, 2008]. The missing piece is consideration of action representation as a process.

By studying expertise, and the neurological changes that accompany developing representations across modalities, we can decompose sensory environments into the components that contribute to representation and those that do not. In the next section, we will examine existing literature in expertise and developing representation through learning paradigms, mainly from auditory-motor research. Then, in conclusion we will propose several unique characteristics of human action representation networks that can be addressed by new rehabilitation devices.

4.3.2 Auditory-motor representation and expertise

Motor representation reflects a process in which multisensory environments mutually reinforce actions perceived in the world. As a result, a significant component of motor representation research examines acquiring representations, and the structural and functional changes that exist after representations have been acquired.

Piano playing is a particular well-suited paradigm to examine motor-auditory representation as a function of expertise. Pianists often report a tightly coupled conceptual link between the act of playing and the perception of music. Various studies have documented involuntary motor activity during passive listening to piano music [Haueisen and Knösche, 2001, Halpern and Zatorre, 1999], or when imagining learned pieces on the musician's instrument [Langheim et al., 2002]. Conversely, when musicians passively observe movement execution without sound on a piano keyboard auditory cortices are activated [Haslinger et al., 2005, Halpern et al., 2004]. These imagery tasks all involve cortical representations that are more substantial than simply cueing one sensory domain from another. Patterns of motor behavior and the context in which they are repeatedly paired with auditory sensory information, develop perceptual relationships that persist despite the absence of one of the contributing domains.

The behaviors of experts should access sensorimotor representations during tasks that incorporate those representations. Although imagery research points to learned associations between different perceptual domains that may relate to representation, they typically fail to demonstrate that the experimental group has built a representation beyond the association of sensory domains implicit in the task. Bangert and colleagues performed an fMRI conjunction analysis of a passive listening and passive performance paradigm in expert pianists and nonmusicians [Bangert et al., 2006]. Their results indicate that experts engage a bilateral network of areas including inferior parietal regions, superior temporal gyrus, supplementary and premotor areas. The passive playing condition is meant to support the inclusion of a control group engaged in a similar tasks to the musicians, thereby demonstrating the sensorimotor representation of the experimental group separately from potentially confounding effects due to the complexity of their learned task [Bangert and Altenmüller, 2003].

The process of acquiring sensorimotor representations has been examined in nonmusicians, as they learn piano sequences over the course of several days[Lahav et al., 2007]. When nonmusicians listen to recently acquired motor sequences, compared to unlearned piano sequences consisting of different notes, Broca's region, inferior parietal regions, intraparietal sulcus and premotor regions are activated. Interestingly, sequences consisting of the same notes but presented in a different order activate some of this network, namely, premotor areas and Broca's region. This serves to characterize the difference between sensorimotor association and representation. Auditory-motor recognition takes place unconsciously and consists of premotor and Broca's area activations, most likely encoding effector-pitch relationships. Auditory-motor representation incorporates regions that have been previously identified with respect to motor representation across modalities, including premotor regions, Broca's area, and inferior parietal regions. Furthermore, if the proposed mirror neuron network is supposed to represent information at an abstract level, it seems odd that a different pitched piano note condition is sufficiently unique from the learned representation to only incorporate part of the proposed mirror neuron network. Abstract sensorimotor representation should be robust enough to generalize beyond individual effectors and pitch associations.

Despite efforts to incorporate na´ive control subjects in auditory-motor representation experiments, or to train nonmusicians to acquire auditory-motor-representations, nonmusicians have deep, culturally engrained expectations regarding the relationships between music instruments, their auditory-motor actions and resultant audio. McNamara and colleagues sought to establish the role of inferior frontal gyrus, and Broca's area, in learning arbitrarily constructed associations between 1.5 second nonsense hand gestures and synthetic meaningless sounds[McNamara et al., 2008]. Over the course of learning the correct associative match, repetitive suppression of activation occurred in the primary regions of interest, BA44, and left-inferior parietal lobule, indicating more efficient processing in those areas through the elimination of redundancy. However, other regions including bilateral inferior temporal gyrus and right hippocampus showed similarly significant suppression of activation as a function of learning. These areas are not typically thought of with respect to auditory-motor representation and may constitute a broader learning network, involving a role for associative memory structures in the early stages of establishing sensorimotor representations. Subjects in this study learned the associations between modalities while in the scanner, over a very short number of trials.

Other auditory-motor learning paradigms emphasize the integrative nature of areas that are also proposed as part of representation systems. In one particular method, subjects are required to keep a constantly changing foreground stimulus in synchrony with a static background[Blum et al., 2007]. The foreground drifts from synchrony with the background in the same way each trial. As the subject learns how the foreground drifts, they can more quickly compensate for the drift by turning the wheel, aligning the foreground and background. In the visual condition, the subjects need to keep tone of a foreground square the same as a background color. In the auditory trial, subjects attended to 8ms tone pips that varied in frequency. By turning the wheel, they attempted to compensate for the variation and keep the varying tones in synchrony with a constant background sine tone that did not change in frequency. This abstract task requires simultaneous mapping of auditory and motor, or visual and motor domains. As subjects increase their performance in the auditory-motor tracking task, the authors demonstrate intra-hemispheric alpha-band EEG phase locking between inferior parietal regions and additional coherence between superior parietal and motor areas. Superior parietal coherence and motor area coherence are thought to relate the interface manipulation with parietal spatial perception. It is difficult to imagine that the multi-domain tracking paradigm would constitute an abstract representation" for auditory-motor content. Rather, these findings highlight the difficulty in distinguishing between integration and representation as a mechanism supported by inferior parietal regions.

We have differentiated between auditory-motor interaction for newly learned auditory-motor associations, and the perception of highly over-learned, ecological stimuli from various semantically meaningful categories, such as environmental sounds, or sounds of tool use. This distinction has illuminated various shared processing centers between auditory-motor representation, and action observation, namely, ventral premotor cortex, including Broca's area, frontal parietal regions, and superior temporal sulcus.

Research into motor representation and its relationship to disparate sensory modalities needs a set of underlying principles that help to define the borders of what is meant by representation. Whereas, definitive mechanisms of representation are lacking, the field of multisensory integration is quite different. Meredith and Stein summarized findings regarding superior colliculus receptive fields properties to establish several defining principles of multisensory processing: the relationship between temporal and spatial properties, the magnitude of multisensory response as a function of unisensory stimulus strength, and the preservation of unisensory stimulus selectivity properties despite multisensory modulation[Meredith and Stein, 1986]. These principles have been a cornerstone for researchers moving forward with a clear vocabulary of integration mechanisms. As a result, exemplary papers now demonstrate the organization of characteristic integration properties at various levels of the cortical hierarchy, giving shape to the informational value of different stages of multisensory integration[Werner and Noppeney, 2010a]. Without a foundation of principles that outline the properties of sensory representations, it is difficult to structure research to uncover the contribution of sensory features to representation.

In rehabilitation environments, it is unclear how representation across multiple domains can be accessed to compensate for pathology in scenarios that may constitute a disruption of representation as part of, or in contrast, exclusively of functional deficit. As an example, in motor apraxias, recent findings suggest a link between motor production deficits and perception of the sounds of those movements [Pazzaglia et al., 2008]. However, it is difficult to untangled conflicting findings from pathological cases where motor function and perception seem to be dissociated in some reports, and not in others[Mahon, 2008]. By establishing the behavioral effect and contribution from intact domains of representation to perception and subsequent rehabilitation in pathological settings, new conclusions may be drawn regarding the function of shared representations.

In addition to behavioral investigation, in the future, it will be important to examine the contribution of known shared representation networks across multiple modalities, throughout the process of rehabilitation. What are the functional and structural changes that take place as an intact sensory modality drives rehabilitation in a motor representation context where the driving sensory modality and disordered motor modality used to share representation? In one recent study, observed actions facilitated post-stroke rehabilitation in moderately impaired individuals (medial cerebral artery infarction) as part of a 4-week treatment compared to a control group that examined geometric image sequences [Ertelt et al., 2007]. Both groups were instructed by non-blinded clinicians who reminded the patients to attend to the images. It is difficult to fathom that the physicians and patients alike could maintain similar attention for geometric image sequences compared to the experimental action observation sequences. Despite this potential caveat, significant functional improvement and motor system reorganization was reported for the experimental group, including increased activation in bilateral ventral PMC, bilateral STG, SMA and contralateral supramarginal gyrus, which are frequently implicated in motor representation research. Multisensory environments that incorporate advances in action representation research have the potential to develop new, non-invasive treatments for rehabilitation of chronic motor deficit.

In many of the aforementioned studies, auditory-motor representation is characterized as a bilateral phenomenon, which departs from classical views of hemispheric specialization for auditory information. Previously, auditory processing was thought to be distributed between right hemisphere functions, such as pitch processing and other bilateral phenomena, such as temporal processing [Peretz and Zatorre, 2005]. However, it is now clear that auditory information processing in addition to hierarchically building from features of the auditory percept, extends significantly beyond auditory features alone as part of a system that maps auditory and motor representations onto a network of regions that are specialized for action observation and production. New models of the distribution of auditory functions will need to include the role of auditory information in multisensory representation contexts.

How is a motor representation created in a human? The mechanism by which effector generalized abstraction from the integration of various domains - tactile, visual, and auditory - create a representation that is behaviorally meaningful, and accessed during the attempted execution of a behavior or the presence of its representing sensory information is yet to be established. In rehabilitation, there exists the unique opportunity to examine the development of elementary motor function as it is paired with elementary sensory stimuli throughout the process. Regions of interest include the relationship between Broca's area, SMA and M1 as function returns, and separately, what are the changes in left IPL as the external sensory environment dominates attempted movement, and later, when the external sensory environment supports movement as a shared representation.

The difficulty in researching the contribution of integration or neural representation networks to the rehabilitation of pathological function is the lack of a tool that can structure a long-term intervention while simultaneously incorporating multiple modalities of presentation and feedback. For instance, in hemiparetic limb rehabilitation, several robotic and wearable devices are becoming platforms for new methodologies to examine long-term rehabilitation [Stein, 2009, Krebs et al., 2008, Krebs et al., 2007]. However, to consistently associate sensory percepts at the moment of attempted motor actions, new tools will need to be constructed to pair stimuli across domains with a degree of coupling that is likely to create shared representations.

The EMG interface, implemented according to the design criteria researched throughout this thesis, is a platform to study the emergence of shared representation in contexts of chronic impairment. Multisensory percepts are linked directly, in the moment of attempted control.

CHAPTER FIVE

Conclusion and Future Work

5.1 Music, mind and health

Our health is engrained in our everyday lives, the decisions we make, the things we care about, and the activities we pursue. In contrast, healthcare is far removed from everyday experience. Despite positive efforts to introduce preventative care, or personalized medicine, we currently lack the transport to partner healthcare with the activities of everyday life.

The emerging *embedded healthcare* area leverages the network and scalable opportunities of ubiquitous technologies to provide healthcare services, to equip patients with the tools to make decisions, or to become more responsible over the management of their own health.

This thesis outlines a different perspective than what exists in the current scope of embedded healthcare. In establishing a framework for music, mind and health, I design and implement the tools to demonstrate that healthcare can be conducted during creative tasks, as part of everyday environments. The research represents a significant expansion of previous uses of music in healthcare by incorporating technology to address issues of scale, specificity and biological interface. These issues are the critical barriers to developing creative applications that are both scientifically valid and specific to disease intervention.

The critical design principles are as follows:

Access - Life is about growing as a person. Access is more than being able to almost do something, its about going further. Technologies that provide access need to provide the value and the structure to grow as a person. Rather than facilitate a facsimile of creative experience, the partnership between healthcare and everyday creative opportunity can enable achievement in the facilitated domain.

Diagnosis - Our interactions with the world around us are displays of ability, capability, competence, and mastery. The role of intervention design is to examine these platforms and kern the information that will help us stave off disease, or validate our tools.

Expression -There is no disorder compared to a particular definition of expression. Amplifying expression is the gateway to everyday creativity. This is equally true of everyone, irrespective of pathology.

5.1.1 Contributions

Community access to creative applications such as Hyperscore contributed to individual and hospital-wide change. The gains that patients made throughout the creative process became the cornerstone of discharge for many of our subjects with diverse mental illnesses. Patients that were previously considered incapable, and noncompliant, emerged from Hypersore sessions having accomplished something as composers. During the Hyperscore sessions, patients' treatment teams frequently documented differences in motor and cognitive function compared to patients' involvement in other activities including music without technology. Furthermore, patients became proactive within their healthcare community, mentoring one another across departments, and creating new avenues for patient-driven communities to emerge in a traditional healthcare institution. New, accessible creative opportunities can impinge on patients' health.

Despite the evidence of clinical change, the adoption of creative tasks in clinical environments is stymied by the difficulty of separating the motivational aspects of music from the scientific basis of its perception. We know that music processing is both distributed and organized in the brain. Preliminary research results contributed to the validation of an auditory test for early Alzheimer's Disease diagnosis. Short musical excerpts are employed in a visualspatial memory task. The results indicate that music can identify individuals in the early stages of Alzheimer's Disease and mild cognitive impairment. The validation of musical material as diagnostic becomes a platform for new applications. As clinical and behavioral markers change in creative interventions, embedded diagnostic components can reflect cognitive change in parallel.

Finally, a personalized instrument, designed and implemented in collaboration with an individual with severe motor pathology, enabled the patient to express compelling interpretations of his previously composed Hyperscore work while targeting motor pathology. When interacting with the instrument, the patient's treatment team reports significant differences compared with ongoing physical and occupational therapies to target similar movements. An adaptive system is proposed to scale the expressive performance controller to individuals with other, equally severe, physical pathology, as well as for other compositions.

Over the course of this research I've made design observations that conclude with the development of an EMG system for physical rehabilitation and music performance. The EMG system consolidates access to a creative environment, embedded diagnosis, and structured intervention. With this tool, we can envision a future of empowered individuals in the chronic-phase of their disease, striving of their own volition, in their homes, to achieve something creative. The rehabilitation process becomes subsumed by the creative process. The technology allows us to consider interventions at scale, and yet, personalized to the naive signal of would-be performers. The interface can grow with users, adapting as the signals driving the system change over time. During the process, we can sculpt the intervention to direct movement in the concerted and careful ways that will produce therapeutic effect.

Future work will consist of the clinical trials to validate the system as a neurorehabilitation for fine-motor control, post-stroke. Additionally, future investigations will elucidate the contribution of multiple sensory domains to emergent motor representations. As the system detects intended finger movement and immediately provides auditory, visual, and haptic feedback to the individual, one can design studies that vary the respective contribution of these domains during the intervention, and determine their correlation with any observed motor processing changes.

5.2 Future Work

5.2.1 Clinical trials for fine-motor rehabilitation

The next step for the EMG system is to implement the technology in clinical trials to realize the rehabilitative potential of the current design. The earliest trials will focus on stroke rehabilitation.

The general strategy is to recruit a small sample of hemiparetic stroke victims in the chronic phase of disease. Ideal subjects will exhibit some gross movement and grasping functionality in the paretic limb, but little to no observable finger movement. After preliminary piloting within a stroke population, subjects will undergo intervention for seventy-five sessions. A modified within-subjects design will stagger an initial period of time whereby each subject will receive no experimental treatment. For instance, for six recruited subjects, all subjects would begin the study at the same time, and only the first subject would immediately start treatment with the currently implemented EMG system for music learning. All other subjects would begin with the control intervention consisting of the use of the fine-motor rehabilitation device, but without contributing auditory and visual reinforcement. After two-weeks, the second subject would discontinue the control intervention, and would begin the intervention including all multisensory aspects of the system. Two weeks later, the third subject would do the same, and so on until all subjects partake in the experimental intervention. At the seventy-five week mark, all subjects would discontinue treatment. This strategy allows the maximally efficient use of a small group of subjects, by relying on them as their own controls and simultaneously testing the role of the duration of treatment.

Several technical considerations are unique to utilizing the current system with a hemiparetic population. First, modeling the EMG signal associated with attempted movement in a paralyzed limb is a very different problem than in healthy controls. Encouraging evidence from recent research suggests the presence of sufficient EMG signal for device control despite little to no observable movement in subjects with severe limb paresis[Dipietro et al., 2005]. Nonetheless, it will be a challenge to determine the most appropriate feature vector for training the support vector machine considering expected heterogenous pathology, including denervation, reinnervation, fatigue, spasticity, atrophy and non-coincident contraction. One would expect gross inter-subject differences pertaining to which features are the most useful. Furthermore, the suitability of introducing feature selection [Weston et al., 2001] on a per-subject basis is not guaranteed to outperform an SVM without any kind of feature selection, despite the cost-benefit of having a smaller feature vector [wei Chen, 2005].

Second, the concept of embedded assessment in the current design stems from the use of EMG as the source of control as well as the measure of pathology. The SVM classifier is responsible for translating a feature vector describing the EMG signal into classified, attempted finger presses. The high-dimensional geometric feature space that represents the multi-channel EMG signal, and the hyperplane that divides the space in the model, are not directly useful as diagnostic measures. However, the change in the model over time is potentially diagnostic. If functional ability improves, one would assume that the underlying signal responsible for the pathological movement has also changed. By examining the change in the model over the course of the treatment, a description of functional change in terms of efficiency, or coordination may be achieved.

5.2.2 Capturing emerging representation

A neuroscience programme, to be conducted in parallel with the initial clinical trials of the system, will examine emerging motor representations and the contribution of distal sensory percepts. A hallmark of the current design is the immediate reinforcement of attempted finger presses with the auditory, visual and haptic stimuli of the intended musical note, played, seen and heard.

The subject, reading from a musical score, anticipates the upcoming note. At the correct time, they try to press the finger corresponding to that note. The system classifies the attempted finger press, whether or not the subject is actually capable of manifesting the movement. The system generates the intended note as long as the subject's composite EMG signal is consistent enough to be classified as that user's signal pattern corresponding to the given note.

The receptor field properties of neuronal populations in "multisensory cortex", such as the superior temporal sulcus (STS), and superior colliculus (SC), have been known for decades [Meredith and Stein, 1986]. STS, as an example, receives terminating neuronal inputs from divergent unisensory domains. The resulting regions are responsive to the site, sound, and touch of actions[Beauchamp et al., 2008]. Beyond the simultaneous presentation of coincident senses, STS neurons have been shown to respond to the quality of information content afforded by contributing unisensory percepts[Werner and Noppeney, 2010b].

There is a difference in scope between cellular properties of multisensory regions and the networks of brain regions that contribute to action-observation and execution. Multisensory receptive fields are considered early modules in higher-order action-observation and "representation" networks[Haruno et al., 2001, Iacoboni et al., 2005]. However, even as action-observation networks inference unified concepts such as action goals [Schubotz and von Cramon, 2009], there is little understood regarding what is representative in a shared representation.

The design of the EMG-driven environment for rehabilitation during music tasks will contribute to a neuroimaging study to examine the formation of shared representations between emergent motor actions and contributing sensory domains. As a stroke victim participates in a music intervention, consistently pairing attempted movement with semantically congruent auditory, and visual feedback the moment of attempted movement, functional and structural neurological changes can be observed in the proposed action-observation system, including Broca's area, supplementary motor cortex, M1 and inferior parietal lobule.

5.2.3 Other essentials of diagnostic validity, statistical power and culture

The Alzheimer's studies were an initial foray into embedded diagnosis, and a model of cognitive testing for Alzheimer's disease detection that departs from the wholesale memory testing batteries that remain the predominant neuropsychological measure for inpatient Alzheimer's disease evaluation. Several follow-up studies will increase the validity of the existing auditory-visualspatial test.

First, having succeeded in a known-groups validity test it is now necessary to increase statistical power, and to validate across geographic regions and cultures. It is important to note that due to the musical underpinnings of this particular diagnostic, the debate regarding the diagnosticity of cognitive measures as a function of culture and demographic are particularly apt discussions that need special attention.

Second, our testing has largely taken place in the clinical environments where the MCI and mild AD subjects have prior experience with evaluation for exactly the type of working memory category that we investigate, albeit, with no prior experience involving auditory stimuli of any kind. These tests need to also demonstrate external validity to be effective outside of the geriatric psychiatry unit and in the world.

Third, test-retest reliability challenges the ability to learn a test of this sort, and whether or not it is effective as a diagnostic at various intervals of follow-up. If expert users can compensate for impending disease based on their familiarity with the tool, then the tool is not useful upon multiple evaluations. To an extent, all tests are expected to be learnable to a certain degree. The open question is whether or not the learning defeats the onset of deficit as exhibited by the measure. The degree to which these tests are learned by the healthy population, at-risk individuals, and those definitively diagnosed are somewhat different validity questions, requiring separate study.

Fourth, construct validity is the sensitivity of the measure to target the domain in question. Because mildly impaired individuals perform badly on our measure, and because the scientific understanding of underlying etiology for emerging Alzheimer's Disease is so strong, our results can be interpreted as relating to the areas of initial deficit for Alzheimer's Disease. However, it is necessary in future work to make this connection directly. By coupling our measure to the pre and post-testing of exploratory MRI, PET, or fMRI studies also pursing Alzheimer's diagnosis, we will be able to substantiate the construct validity of the measure for changes in hippocampal and entorhinal cortex biology, the sites of earliest incident in Alzheimer's Disease[Pihlajamäki et al., 2009].

Fifth, because the prevailing diagnostic measures for Alzheimer's Disease are either neuro-imaging based, or based on infrequent assessments taking place under the guidance of a physician, the validity of testing outside of those environments has not been directly tested against testing within the typical neuroimaging and doctor's visit models. What does a diagnosis mean in the hands of a user? Combined with test-retest reliability, what does it mean for a user to take a test repeatedly and be the sole recipient of their performance feedback?

This also begs various technical questions. Technology can support the transfer of potentially diagnostic information from mobile or at-home environments to anywhere. How does the technology manage privacy in distributed diagnosis? Who should have access to the data? What do we observe as subjects incorporate the diagnostic opportunity into daily life?

The Cognitive Health of Communities

The next step in the diagnostic application is to answer all of the above questions in a mobile environment, with embedded location-based services and a supporting social network component shared by both practitioners and other application users. App-store platforms become research platforms as users sign on *en masse.* The application allows access to itself either daily, weekly, or monthly. Data is collected over the course of a year. The social network serves as both a collaborative environment for proactive users to form groups and mobilize around their efforts to stave off disease. Data visualizations encode a user's performance, protecting privacy while exhibiting global trends, becoming available as displays and thumbnails in the online space. Users maintain global privacy and access rules for their diagnostic information. The application also compiles data relevant to activities outside of the cognitive task as they relate to the propensity to transfer to Alzheimer's Disease, namely, activities of daily life, physician consultations and traditional evaluations.

And finally, by correlating the cognitive health, activity, and application usage of large numbers of users from various residential communities distributed globally, we can leverage social networks around healthcare applications to elevate the discussion of cognitive health away from the individual and into the social space. We can access the cognitive health of communities with the right data acquisition framework. Probing, and prescribing communities as part of the activities of their daily lives. Music, again, is the value add-on of the entire experience. Auditory-visualspatial tasks probe the music of a user's environment, or get constructed from the media shared within the online community. It is the transport for entry into the social network, a key part of my online identity, as well as the mechanism for diagnosis.

5.2.4 Social music medicine

The proposed methodology for pursuing healthcare as part of creative tasks is currently lacking in the social dimension. This is one area where traditional music therapy is put in value. Because of the heavy reliance on interpersonal dynamic and relationship building, music therapists tend to be incredibly good at walking into a room, making real social connections with just about anyone, and using those connections to sculpt music making whether or not the subject has any background, proclivity, talent, or sometimes, even interest. Whether in groups, or on a one-to-one basis, facilitating the social dynamic of music making is a music therapist's gift. If only a technology could facilitate that kind of social connection!

After the Hyperscore in the Hospital research, the diagnostic, performance and wearable systems never provided the structure to integrate neophyte composers and performers into a larger community in quite the way that Hyperscore at Tewksbury did. There isn't anything inherently social about composing. Hyperscore can be isolating as one person settles into realize their masterpiece. Paradoxically, the performance systems, except in the moment where patients were giving performances, where much more isolating than the composition work.

The observation that explains social composition and personal performance or diagnosis is that as researchers we were constructing significant infrastructure around each of these projects. With Hyperscore at the Hospital, we introduced the composition tool to a residential community. The patients were converging on shared resources to come and be a part of the composition sessions using the only available computers at the hospital. The choices we made regarding when and where the technology was available to interested patients, as well as how proactive we were to get hospital staff from various departments involved, helped create at least the initial sense of community. How Hyperscore has continued to be used, its role within the hospital, and its implications for sustainable creative tools in long-term care environments, grew without our active directing.

In performance or diagnosis, we set out to work with individuals either to design tailored performance systems, or to measure executive function from person-to-person. I could have worked with groups of Hyperscore composers-turned performers or communities of elders at risk of transferring to Alzheimer's Disease, but focusing on the individual seemed the most efficient way of doing the foundational research to show a diagnostic and rehabilitation component of the framework, which eventually led to the design of the EMG system.

In future work, much more consideration will need to be given to the way in which these technologies afford either personal exploration, social opportunity or both. The EMG system addresses the need to provide access, diagnosis and structured rehabilitation as part of a single home-based system. However, I have not sufficiently developed the application side to determine whether the opportunities exist solely for the individual pursing their creative work, practicing their instrument, elevating that work to a performance for an intended audience, or becoming part of a community throughout the process.

Communities are key. The lack of a community is part of what is so alienating about the current healthcare system. Unfortunately, a great deal of music as we know it - lessons, practicing, composing - are just as alientating. This is an area where existing social media technology is poised to make an almost immediate difference. Imagine connecting to a network of individuals in the moment they are practicing, striving to achieve mastery on the same material you are.

The same infrastructure would benefit healthcare. A community of proactive healthcare consumers, striving against disease, working together from across the globe, feels very different than what is currently available. At the center of the application, moving forward from the interface technologies that come together in this thesis, reconsidering the role of communities, and how technologies can support communities, both in creativity and health, will further substantiate the advantage of bringing the two fields together. The benefit need not be additive. Rather, I am confident that the richness of creative communities will make wading into a proactive health community all the more feasible despite what it asks us to make public of ourselves and our limitations.

Bibliography

[Ackermann, 1996] Ackermann, E. (1996). Perspective taking and object construction. In Kafai, Y. and Resnick, M., editors, *Constructionism in Practice*, pages 25–36. Lawrence Erlbaum Associates, Hillsdale.

[Albert et al., 1973] Albert, M. L., Sparks, R. W., and Helm, N. A. (1973). Melodic intonation therapy for aphasia. *Arch Neurol*, 29(2):130–131.

[Almor et al., 2009] Almor, A., Aronoff, J. M., MacDonald, M. C., Gonnerman, L. M., Kempler, D., Hintiryan, H., Hayes, U. L., Arunachalam, S., and Andersen, E. S. (2009). A common mechanism in verb and noun naming deficits in alzheimer's patients. *Brain Lang*, 111(1):8–19.

[Amabile, 1996] Amabile, T. (1996). *Creativity in Context.* Westview Press, Boulder.

[AMTA, 2009] AMTA (2009). American music therapy association: Year in review, 2009. accessed from http://www.musictherapy.org/events/09AnnualRpt.pdf, on August 26th.

[Au et al., 2005] Au, S., Bonato, P., and Herr, H. (2005). An emg-position controlled system for an active ankle-foot prosthesis: an initial experimental study. In *IEEE Int. Conf. On Rehabilitation Robotics*, pages 375–379.

[Bäckman et al., 2004] Bäckman, L., Jones, S., Berger, A. K., Laukka, E. J., and Small, B. J. (2004). Multiple cognitive deficits during the transition to alzheimer's disease. *J Intern Med*, 256(3):195–204.

[Baer, 1996] Baer, J. (1996). The effects of task-specific divergent-thinking training. In *Journal of Creative Behavior.*, volume 30., pages 183–187.

[Baer, 1997] Baer, J. (1997). *Creative Teachers, Creative Students.* Allyn and Bacon, Boston.

[Baer, 1998] Baer, J. (1998). The case for domain specificity in creativity. In *Creativity Research Journal.*, volume 11, pages 173–177.

[Baer, 2005] Baer, J Kaufman, J. (2005). Bridging generality and specificity: The amusement park theoretical (apt) model of creativity. In *Roeper Review.*, volume 27., pages 158–163.

[Bangert and Altenmüller, 2003] Bangert, M. and Altenmüller, E. O. (2003). Mapping perception to action in piano practice: a longitudinal dc-eeg study. *BMC Neurosci*, 4:26–26.

[Bangert et al., 2006] Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H. J., and Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fmri conjunction. *Neuroimage*, 30(3):917–926.

[Barraclough et al., 2005] Barraclough, N. E., Xiao, D., Baker, C. I., Oram, M. W., and Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J Cogn Neurosci*, 17(3):377–391.

[Beauchamp et al., 2008] Beauchamp, M. S., Yasar, N. E., Frye, R. E., and Ro, T. (2008). Touch, sound and vision in human superior temporal sulcus. *Neuroimage*, 41(3):1011–1020.

[Belgrave, 2009] Belgrave, M. (2009). The effect of expressive and instrumental touch on the behavior states of older adults with late-stage dementia of the alzheimer's type and on music therapist's perceived rapport. *J Music Ther*, 46(2):132–146.

[Bidelman et al., 2009] Bidelman, G. M., Gandour, J. T., and Krishnan, A. (2009). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci.*

[Binkofski and Buccino, 2006] Binkofski, F. and Buccino, G. (2006). The role of ventral premotor cortex in action execution and action understanding. *J Physiol Paris*, 99(4-6):396–405.

[Birks, 2006] Birks, J. (2006). Cholinesterase inhibitors for alzheimer's disease. *Cochrane Database Syst Rev*, 1(1).

[Blay and Toledo Pisa Peluso, 2010] Blay, S. L. and Toledo Pisa Peluso, E. (2010). Public stigma: the community's tolerance of alzheimer disease. *Am J Geriatr Psychiatry*, 18(2):163–171.

[Blum et al., 2007] Blum, J., Lutz, K., and Jäncke, L. (2007). Coherence and phase locking of intracerebral activation during visuo- and audio-motor learning of continuous tracking movements. *Exp Brain Res*, 182(1):59–69.

[Boulanger and Mathews, 1997] Boulanger, R. and Mathews, M. (1997). The 1997 mathews radio-baton and improvisation modes. In *Proceedings of the International Computer Music Conference*, volume 4, pages 8–21.

[Brieler and Minas, 2010] Brieler, F. and Minas, M. (2010). A model-based recognition engine for sketched diagrams. *J. Vis. Lang. Comput.*, 21(2):81–97.

[Bronowski, 1965] Bronowski, J. (1965). *Science and Human Values.* Perennial Library, San Francisco.

[Brownstein et al., 2010] Brownstein, J. S., Freifeld, C. C., Chan, E. H., Keller, M., Sonricker, A. L., Mekaru, S. R., and Buckeridge, D. L. (2010). Information technology and global surveillance of cases of 2009 h1n1 influenza. *N Engl J Med*, 362(18):1731–1735.

[Buccino et al., 2001] Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fmri study. *Eur J Neurosci*, 13(2):400–404.

[Calvert et al., 1999] Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R., Iversen, S. D., and David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, 10(12):2619–2623.

[Caselli et al., 2009] Caselli, R. J., Dueck, A. C., Osborne, D., Sabbagh, M. N., Connor, D. J., Ahern, G. L., Baxter, L. C., Rapcsak, S. Z., Shi, J., Woodruff, B. K., Locke, D. E., Snyder, C. H., Alexander, G. E., Rademakers, R., and Reiman, E. M. (2009). Longitudinal modeling of age-related memory decline and the apoe epsilon4 effect. *N Engl J Med*, 361(3):255–263.

[Cassity, 2006] Cassity, M. (2006). *Multimodal Psychiatric Music Therapy for Adults, Adolescents, and Children.* Jessica Kingsley Publishers, London.

[Chaminade et al., 2002] Chaminade, T., Meltzoff, A. N., and Decety, J. (2002). Does the end justify the means? a pet exploration of the mechanisms involved in human imitation. *Neuroimage*, 15(2):318–328.

[Chan and Englehart, 2005] Chan, A. D. and Englehart, K. B. (2005). Continuous myoelectric control for powered prostheses using hidden markov models. *IEEE Trans Biomed Eng*, 52(1):121–124.

[Channon and Green, 1999] Channon, S. and Green, P. S. (1999). Executive function in depression: the role of performance strategies in aiding depressed and non-depressed participants. *J Neurol Neurosurg Psychiatry*, 66(2):162–171.

[Cheney et al., 1998] Cheney, P. D., Kenton, J. D., Thompson, R. W., McKiernan, B. J., Lininger, R. E., and Trank, J. W. (1998). A low-cost, multichannel, emg signal processing amplifier. *Journal of Neuroscience Methods*, 79(1):123 - 127.

[Coelho et al., 2010] Coelho, M., Marti, M. J., Tolosa, E., Ferreira, J. J., Valldeoriola, F., Rosa, M., and Sampaio, C. (2010). Late-stage parkinson's disease: the barcelona and lisbon cohort. *J Neurol*, 257(9):1524–1532.

[Cramer, 2008] Cramer, S. C. (2008). Repairing the human brain after stroke. ii. restorative therapies. *Ann Neurol*, 63(5):549–560.

[Cuddy and Duffin, 2005] Cuddy, L. L. and Duffin, J. (2005). Music, memory, and alzheimer's disease: is music recognition spared in dementia, and how can it be assessed? *Med Hypotheses*, 64(2):229–235.

[Davis, 1999] Davis, W. (1999). Music therapy for mentally retarded children and adults. In Davis, W., Gfeller, K., and Thaut, M., editors, *An Introduction to Music Therapy Theory and Practice*, chapter 4, pages 61–80. McGraw-Hill, 2nd edition.

[Detmer, 2010] Detmer, D. E. (2010). Engineering information technology for actionable information and better health - balancing social values through desired outcomes, complementary standards and decision-support. *Stud Health Technol Inform*, 153:107–118.

[Deutsch et al., 2006] Deutsch, A., Granger, C. V., Heinemann, A. W., Fiedler, R. C., DeJong, G., Kane, R. L., Ottenbacher, K. J., Naughton, J. P., and Trevisan, M. (2006). Poststroke rehabilitation: outcomes and reimbursement of inpatient rehabilitation facilities and subacute rehabilitation programs. *Stroke*, 37(6):1477–1482.

[Dinstein, 2008] Dinstein, I. (2008). Human cortex: reflections of mirror neurons. *Curr Biol*, 18(20):956–959.

[Dipietro et al., 2005] Dipietro, L., Ferraro, M., Palazzolo, J. J., Krebs, H. I., Volpe, B. T., and Hogan, N. (2005). Customized interactive robotic treatment for stroke: Emg-triggered therapy. *IEEE Trans Neural Syst Rehabil Eng*, 13(3):325–334.

[Dishman, 2004] Dishman, E. (2004). Inventing wellness systems for aging in place. *Computer*, 37:34–41.

[Docteur and Berenson, 2009] Docteur, E. and Berenson, R. (2009). Timely analysis of immediate health policy issues: How does the quality of u.s. health care compare internationally. The Robert Wood Johnson Foundation.

[Doehrmann and Naumer, 2008] Doehrmann, O. and Naumer, M. J. (2008). Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. *Brain Res*, 1242:136–150.

[Dowling, 1986] Dowling, W. (1986). *Music Cognition.* Academic Press, Boston.

[Duncan et al., 2002] Duncan, P. W., Horner, R. D., Reker, D. M., Samsa, G. P., Hoenig, H., Hamilton, B., LaClair, B. J., and Dudley, T. K. (2002). Adherence to postacute rehabilitation guidelines is associated with functional recovery in stroke. *Stroke*, 33(1):167–177.

[Edwards, 2009] Edwards, D. (2009). *Artscience.* Harvard University Press, Cambridge.

[Epstein and Axtell, 1996] Epstein, J. and Axtell, R. L. (1996). *Growing Artificial Societies.* Brookings Institution Press, Washington.

[Ertelt et al., 2007] Ertelt, D., Small, S., Solodkin, A., Dettmers, C., McNamara, A., Binkofski, F., and Buccino, G. (2007). Action observation has a positive impact on rehabilitation of motor deficits after stroke. *Neuroimage*, 36 Suppl 2:164–173.

[Eslick and Moss, 2010] Eslick, I. and Moss, F. (2010). Collective discovery. Available at https://www.lamsight.org/.

[Farbood et al., 2004] Farbood, M. M., Pasztor, E., and Jennings, K. (2004). Hyperscore: A graphical sketchpad for novice composers. *IEEE Comput. Graph. Appl.*, 24(1):50–54.

[Fedorenko et al., 2009] Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., and Gibson, E. (2009). Structural integration in language and music: evidence for a shared system. *Mem Cognit*, 37(1):1–9.

[Fels et al., 2002] Fels, S., Gadd, A., and Mulder, A. (2002). Mapping transparency through metaphor: towards more expressive musical instruments. *Org. Sound*, 7(2):109–126.

[Fennema-Notestine et al., 2009] Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Jacobson, M. W., Dale, A. M., and Initiative, T. A. D. N. (2009). Structural neuroimaging in the detection and prognosis of pre-clinical and early ad. *Behav Neurol*, 21(1):3–12.

[Ferrari et al., 2005] Ferrari, P. F., Rozzi, S., and Fogassi, L. (2005). Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *J Cogn Neurosci*, 17(2):212–226.

[Ford and Sheredos, 1995] Ford, J. M. and Sheredos, S. J. (1995). Ultrasonic head controller for powered wheelchairs. *J Rehabil Res Dev*, 32(3):280–284.

[Freifeld and Moss, 2010] Freifeld, C. and Moss, F. (2010). Healthmap: Global health, local information. Available at http://www.healthmap.org/en/.

[Friberg, 1995] Friberg (1995). Matching the rule paramters of phrase arch to perfomances of ‘tr´aumerei’: a preliminary study. In *Proceedings of the KTH Symposium on Grammars for Music Performance*, pages 37–44, Stockholm. KTH, KTH.

[Frost and Massagli, 2008] Frost, J. H. and Massagli, M. P. (2008). Social uses of personal health information within patientslikeme, an online patient community: what can happen when patients have access to one another's data. *J Med Internet Res*, 10(3).

[Galati et al., 2008] Galati, G., Committeri, G., Spitoni, G., Aprile, T., Di Russo, F., Pitzalis, S., and Pizzamiglio, L. (2008). A selective representation of the meaning of actions in the auditory mirror system. *Neuroimage*, 40(3):1274–1286.

[Giard and Peronnet, 1999] Giard, M. H. and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci*, 11(5):473–490.

[Gilden, 1971] Gilden, R. (1971). Sigcaph newsletter: Computers and the physically handicapped. ACM.

[Gordon et al., 2004] Gordon, N. F., Gulanick, M., Costa, F., Fletcher, G., Franklin, B. A., Roth, E. J., Shephard, T., and American Heart Association

Council on Clinical Cardiology, Subcommittee on Exercise, Cardiac Rehabilitation, and Prevention; the Council on Cardiovascular Nursing; the Council on Nutrition, Physical Activity, and Metabolism; and the Stroke Council (2004). Physical activity and exercise recommendations for stroke survivors: an american heart association scientific statement from the council on clinical cardiology, subcommittee on exercise, cardiac rehabilitation, and prevention; the council on cardiovascular nursing; the council on nutrition, physical activity, and metabolism; and the stroke council. *Stroke*, 35(5):1230–1240.

[Graham et al., 2003] Graham, N., Lindesay, J., Katona, C., Bertolote, J. M., Camus, V., Copeland, J. R., de Mendonça Lima, C. A., Gaillard, M., Gély Nargeot, M. C., Gray, J., Jacobsson, L., Kingma, M., Kühne, N., O'Loughlin, A., Rutz, W., Saraceno, B., Taintor, Z., Wancata, J., Old Age Psychiatry section, World Psychiatric Association, and World Health Organization (2003). Reducing stigma and discrimination against older people with mental disorders: a technical consensus statement. *Int J Geriatr Psychiatry*, 18(8):670–678.

[Grèzes et al., 2003] Grèzes, J., Armony, J. L., Rowe, J., and Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: an fmri study. *Neuroimage*, 18(4):928–937.

[Grol et al., 2009] Grol, M. J., Toni, I., Lock, M., and Verstraten, F. A. (2009). Spatial representation of overlearned arbitrary visuomotor associations. *Exp Brain Res*, 192(4):751–759.

[Hall et al., 2009] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. (2009). The weka data mining software: An update. *SIGKDD Explorations*, 11(1).

[Halpern and Zatorre, 1999] Halpern, A. R. and Zatorre, R. J. (1999). When that tune runs through your head: a pet investigation of auditory imagery for familiar melodies. *Cereb Cortex*, 9(7):697–704.

[Halpern et al., 2004] Halpern, A. R., Zatorre, R. J., Bouffard, M., and Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, 42(9):1281–1292.

[Hanser, 1999] Hanser, S. (1999). *The New Music Therapist's Handbook.* Berklee Press, Boston.

[Harel and Papert, 1991a] Harel, I. and Papert, S. (1991a). *Constructionism.* Ablex Pub. Corp, Norwood.

[Harel and Papert, 1991b] Harel, I. and Papert, S. (1991b). Software design as a learning environment. In *Constructionism*, pages 41–84. Ablex Pub. Corp, Norwood.

[Haruno et al., 2001] Haruno, M., Wolpert, D. M., and Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural Comput*, 13(10):2201–2220.

[Haslinger et al., 2005] Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., and Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *J Cogn Neurosci*, 17(2):282–293.

[Haueisen and Knösche, 2001] Haueisen, J. and Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *J Cogn Neurosci*, 13(6):786–792.

[Hecht et al., 2008] Hecht, D., Reiner, M., and Karni, A. (2008). Multisensory enhancement: gains in choice and in simple response times. *Exp Brain Res*, 189(2):133–143.

[Hesse et al., 2009] Hesse, M. D., Sparing, R., and Fink, G. R. (2009). End or means–the "what" and "how" of observed intentional actions. *J Cogn Neurosci*, 21(4):776–790.

[Holt, 1989] Holt, J. (1989). *Learning All the Time*. Addison-Wesley, Boston.

[Hu et al., 2009] Hu, X. L., Tong, K. Y., Song, R., Zheng, X. J., and Leung, W. W. (2009). A comparison between electromyography-driven robot and passive motion device on wrist rehabilitation for chronic stroke. *Neurorehabil Neural Repair*, 23(8):837–846.

[Hubert, 2006] Hubert, R. (2006). Accessibility and usability guidelines for mobile devices in home health monitoring. *SIGACCESS Access. Comput.*, 84:26–29.

[Hulme et al., 2010] Hulme, C., Wright, J., Crocker, T., Oluboyede, Y., and House, A. (2010). Non-pharmacological approaches for dementia that informal carers might try or access: a systematic review. *Int J Geriatr Psychiatry*, 25(7):756–763.

[Hyde et al., 2009] Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., and Schlaug, G. (2009). Musical training shapes structural brain development. *J Neurosci*, 29(10):3019–3025.

[Iacoboni, 2005] Iacoboni, M. (2005). Neural mechanisms of imitation. *Curr Opin Neurobiol*, 15(6):632–637.

[Iacoboni et al., 2005] Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol*, 3(3).

[Jack et al., 2010] Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., and Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the alzheimer's pathological cascade. *Lancet Neurol*, 9(1):119–128.

[Janata, 2005] Janata, P. (2005). Brain networks that track musical structure. *Ann N Y Acad Sci*, 1060:111–124.

[Janata, 2009] Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cereb Cortex*, 19(11):2579–2594.

[Janata and Grafton, 2003] Janata, P. and Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat Neurosci*, 6(7):682–687.

[Jennings, 2003] Jennings, K. (2003). Toy symphony: An international music technology for children. *Music Education International*, 2:3–21.

[Jo and Lee, 2010] Jo, Y. S. and Lee, I. (2010). Disconnection of the hippocampal-perirhinal cortical circuits severely disrupts object-place paired associative memory. *J Neurosci*, 30(29):9850–9858.

[Jones et al., 2010] Jones, V., Gay, V., and Leijdekkers, P. (2010). Body sensor networks for mobile health monitoring: Experience in europe and australia. In *ICDS '10: Proceedings of the 2010 Fourth International Conference on Digital Society*, pages 204–209, Washington, DC, USA. IEEE Computer Society.

[Jurik et al., 2009] Jurik, A. D., Bolus, J. F., Weaver, A. C., Calhoun, B. H., and Blalock, T. N. (2009). Mobile health monitoring through biotelemetry. In *BodyNets '09: Proceedings of the Fourth International Conference on Body Area Networks*, pages 1–8, ICST, Brussels, Belgium, Belgium. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[Kaplan and Iacoboni, 2007] Kaplan, J. T. and Iacoboni, M. (2007). Multimodal action representation in human left ventral premotor cortex. *Cogn Process*, 8(2):103–113.

[Kauffman, 1995] Kauffman, S. (1995). *At Home in the Universe.* Oxford University Press, Oxford Oxfordshire.

[Kauffman, 2000] Kauffman, S. (2000). *Investigations.* Oxford University Press, Oxford Oxfordshire.

[Kayser et al., 2008] Kayser, C., Petkov, C. I., and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cereb Cortex*, 18(7):1560–1574.

[Khambati et al., 2008] Khambati, A., Grundy, J., Warren, J., and Hosking, J. (2008). Model-driven development of mobile personal health care applications. In *ASE '08: Proceedings of the 2008 23rd IEEE/ACM International Conference on Automated Software Engineering*, pages 467–470, Washington, DC, USA. IEEE Computer Society.

[Khaw et al., 2008] Khaw, K. T., Wareham, N., Bingham, S., Welch, A., Luben, R., and Day, N. (2008). Combined impact of health behaviours and mortality in men and women: the epic-norfolk prospective population study. *PLoS Med*, 5(1).

[Kohler et al., 2002] Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582):846–848.

[Koski et al., 2002] Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C., and Iacoboni, M. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cereb Cortex*, 12(8):847–855.

[Krebs et al., 2008] Krebs, H. I., Mernoff, S., Fasoli, S. E., Hughes, R., Stein, J., and Hogan, N. (2008). A comparison of functional and impairment-based robotic training in severe to moderate chronic stroke: a pilot study. *NeuroRehabilitation*, 23(1):81–87.

[Krebs et al., 2007] Krebs, H. I., Volpe, B. T., Williams, D., Celestino, J., Charles, S. K., Lynch, D., and Hogan, N. (2007). Robot-aided neurorehabilitation: a robot for wrist rehabilitation. *IEEE Trans Neural Syst Rehabil Eng*, 15(3):327–335.

[Kuhn, 2009] Kuhn, T. (2009). *The Structure of Scientific Revolutions.* Books LLC, City.

[Kuiken et al., 2009] Kuiken, T. A., Li, G., Lock, B. A., Lipschutz, R. D., Miller, L. A., Stubblefield, K. A., and Englehart, K. B. (2009). Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms. *JAMA*, 301(6):619–628.

[Kunzler, 2007] Kunzler, P. (2007). The athlete seat. Master's thesis, Massachusetts Institute of Technology.

[Lahav et al., 2007] Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J Neurosci*, 27(2):308–314.

[Langheim et al., 2002] Langheim, F. J., Callicott, J. H., Mattay, V. S., Duyn, J. H., and Weinberger, D. R. (2002). Cortical systems associated with covert music rehearsal. *Neuroimage*, 16(4):901–908.

[Lehmann et al., 2008] Lehmann, H., Prasad, M., and Scornavacca, E. (2008). Adapting the is success model for mobile technology in health: a new zealand example. In *ICEC '08: Proceedings of the 10th international conference on Electronic commerce*, pages 1–4, New York, NY, USA. ACM.

[Lewis and Irvine, 2006] Lewis, F. and Irvine, J. (2006). I can: Investigating caregivers' attitudes and needs. Harris Interactive Inc., conducted on behalf of the Alzheimer's Foundation of America.

[Lewis et al., 2005] Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., and DeYoe, E. A. (2005). Distinct cortical pathways for processing tool versus animal sounds. *J Neurosci*, 25(21):5148–5158.

[Lindeberg, 1994] Lindeberg, T. (1994). *Scale-Space Theory in Computer Vision.* Kluwer Academic Publishers, Norwell, MA, USA.

[Liu et al., 2006] Liu, H., Kho, A. T., Kohane, I. S., and Sun, Y. (2006). Predicting survival within the lung cancer histopathological hierarchy using a multi-scale genomic model of development. *PLoS Med*, 3(7).

[Lorenz et al., 2007] Lorenz, A., Mielke, D., Oppermann, R., and Zahl, L. (2007). Personalized mobile health monitoring for elderly. In *MobileHCI '07: Proceedings of the 9th international conference on Human computer interaction with mobile devices and services*, pages 297–304, New York, NY, USA. ACM.

[Machover, 2004] Machover, T. (2004). Shaping minds musically. *BT Technology Journal*, 22(4):171–179.

[Madson, 1987] Madson, C. (1987). *Applications of Research in Music Behavior*. University of Alabama Press, University.

[Magne et al., 2006] Magne, C., Schön, D., and Besson, M. (2006). Musician children detect pitch violations in both music and language better than non-musician children: behavioral and electrophysiological approaches. *J Cogn Neurosci*, 18(2):199–211.

[Mahon, 2008] Mahon, B. Z. (2008). Action recognition: is it a motor process? *Curr Biol*, 18(22):1068–1069.

[Mandl and Kohane, 2008] Mandl, K. D. and Kohane, I. S. (2008). Tectonic shifts in the health information economy. *N Engl J Med*, 358(16):1732–1737.

[Mandl et al., 2007] Mandl, K. D., Simons, W. W., Crawford, W. C., and Abbett, J. M. (2007). Indivo: a personally controlled health record for health information exchange and communication. *BMC Med Inform Decis Mak*, 7:25–25.

[Marchal-Crespo and Reinkensmeyer, 2009] Marchal-Crespo, L. and Reinkensmeyer, D. J. (2009). Review of control strategies for robotic movement training after neurologic injury. *J Neuroeng Rehabil*, 6:20–20.

[Massey, 2009] Massey, T. (2009). *Data driven and optimization techniques for mobile health systems*. PhD thesis, Columbia University, Los Angeles, CA, USA. Adviser-Sarrafzadeh, Majid.

[McNamara et al., 2008] McNamara, A., Buccino, G., Menz, M. M., Gläscher, J., Wolbers, T., Baumgärtner, A., and Binkofski, F. (2008). Neural dynamics of learning sound-action associations. *PLoS One*, 3(12).

[Meredith and Stein, 1986] Meredith, M. A. and Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol*, 56(3):640–662.

[Miller, 1982] Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cogn Psychol*, 14(2):247–279.

[Mokdad et al., 2004] Mokdad, A. H., Marks, J. S., Stroup, D. F., and Gerberding, J. L. (2004). Actual causes of death in the united states, 2000. *JAMA*, 291(10):1238–1245.

[Molholm et al., 2004] Molholm, S., Ritter, W., Javitt, D. C., and Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cereb Cortex*, 14(4):452–465.

[Moreno et al., 2009] Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., and Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb Cortex*, 19(3):712–723.

[Mueller et al., 2005] Mueller, S., Weiner, M., Thal, L., Petersen, R., Jack, C., Jagust, W., Trokanowski, J., Toga, A., and Beckett, L. (2005). The alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics of North America*, pages 869–877.

[Nasreddine et al., 2005] Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., and Chertkow, H. (2005). The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*, 53(4):695–699.

[Neyer et al., 2007] Neyer, J., Greenlund, K., Denny, C., Keenan, N., Labarthe, D., and Croft, J. (2007). Prevalence of heart disease — united states, 2005. accessed from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5606a2.htm, on August 26th.

[Noppeney et al., 2008] Noppeney, U., Josephs, O., Hocking, J., Price, C. J., and Friston, K. J. (2008). The effect of prior visual information on recognition of speech and sounds. *Cereb Cortex*, 18(3):598–609.

[Norton et al., 2009] Norton, A., Zipse, L., Marchina, S., and Schlaug, G. (2009). Melodic intonation therapy: shared insights on how it is done and why it might help. *Ann N Y Acad Sci*, 1169:431–436.

[Omar et al., 2010] Omar, R., Hailstone, J. C., Warren, J. E., Crutch, S. J., and Warren, J. D. (2010). The cognitive organization of music knowledge: a clinical analysis. *Brain*, 133(Pt 4):1200–1213.

[Pantev et al., 2003] Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., and Schulz, M. (2003). Music and learning-induced cortical plasticity. *Ann N Y Acad Sci*, 999:438–450.

[Pazzaglia et al., 2008] Pazzaglia, M., Pizzamiglio, L., Pes, E., and Aglioti, S. M. (2008). The sound of actions in apraxia. *Curr Biol*, 18(22):1766–1772.

[Peretz and Zatorre, 2005] Peretz, I. and Zatorre, R. J. (2005). Brain organization for music processing. *Annu Rev Psychol*, 56:89–114.

[Pihlajamäki et al., 2009] Pihlajamäki, M., Jauhiainen, A. M., and Soininen, H. (2009). Structural and functional mri in mild cognitive impairment. *Curr Alzheimer Res*, 6(2):179–185.

[Pizzamiglio et al., 2005] Pizzamiglio, L., Aprile, T., Spitoni, G., Pitzalis, S., Bates, E., D'Amico, S., and Di Russo, F. (2005). Separate neural systems for processing action- or non-action-related sounds. *Neuroimage*, 24(3):852–861.

[Platt, 1999] Platt, J. C. (1999). Fast training of support vector machines using sequential minimal optimization. In *Advances in kernel methods: support vector learning*, pages 185–208. MIT Press, Cambridge, MA, USA.

[Preuveneers and Berbers, 2008] Preuveneers, D. and Berbers, Y. (2008). Mobile phones assisting with health self-care: a diabetes case study. In *MobileHCI '08: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 177–186, New York, NY, USA. ACM.

[Raji et al., 2009] Raji, C. A., Lopez, O. L., Kuller, L. H., Carmichael, O. T., and Becker, J. T. (2009). Age, alzheimer disease, and brain structure. *Neurology*, 73(22):1899–1905.

[Reis et al., 2007] Reis, B. Y., Kirby, C., Hadden, L. E., Olson, K., McMurry, A. J., Daniel, J. B., and Mandl, K. D. (2007). Aegis: a robust and scalable real-time public health surveillance system. *J Am Med Inform Assoc*, 14(5):581–588.

[Reiter and Maglaveras, 2009] Reiter, H. and Maglaveras, N. (2009). Heartcycle: Compliance and effectiveness in hf and cad closed-loop management. In *EMBS '09: Proceedings of the 31st Annual International Conference of the IEEE EMBS*.

[Richardson, 2001] Richardson, W. (2001). Crossing the quality chasm: A new health system for the 21st century. Technical report, Institute of Medicine of the National Academy of Sciences.

[Rizzolatti and Craighero, 2004] Rizzolatti, G. and Craighero, L. (2004). The mirror-neuron system. *Annu Rev Neurosci*, 27:169–192.

[Root-Bernstein, 2001] Root-Bernstein, R. S. (2001). Music, creativity and scientific thinking. In *Leonardo*, volume 34, pages 63–68.

[Sadasivam et al., 2008] Sadasivam, R. S., Tanik, M. M., and Willig, J. (2008). A conceptual framework for personalized and mobile health care. *J. Integr. Des. Process Sci.*, 12(2):31–41.

[Saponas et al., 2008] Saponas, T. S., Tan, D. S., Morris, D., and Balakrishnan, R. (2008). Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 515–524, New York, NY, USA. ACM.

[Saponas et al., 2010] Saponas, T. S., Tan, D. S., Morris, D., Turner, J., and Landay, J. A. (2010). Making muscle-computer interfaces more practical. In *CHI '10: Proceedings of the 28th international conference on Human factors in computing systems*, pages 851–854, New York, NY, USA. ACM.

[Särkämö et al., 2010] Särkämö, T., Pihko, E., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., Autti, T., Silvennoinen, H. M., Erkkilä, J., Laine, M., Peretz, I., Hietanen, M., and Tervaniemi, M. (2010). Music and speech listening enhance the recovery of early sensory processing after stroke. *J Cogn Neurosci*, 22(12):2716–2727.

[Saunders et al., 2004] Saunders, C., Hardoon, D. R., Shawe-Taylor, J., and Widmer, G. (2004). Using string kernels to identify famous performers from their playing style. In Boulicaut, J.-F., Esposito, F., Giannotti, F., and Pedreschi, D., editors, *Machine Learning: ECML 2004*, volume 3201 of *Lecture Notes in Computer Science*, pages 384–395. Springer Berlin / Heidelberg. 10.1007/978-3-540-30115-8$_3$6.

[Schellenberg and Peretz, 2008] Schellenberg, E. G. and Peretz, I. (2008). Music, language and cognition: unresolved issues. *Trends Cogn Sci*, 12(2):45–46.

[Schemas, 2010] Schemas, M. H. (2010). http://developer.healthvault.com/sdk/docs/index.html. Retreived from URL on August 25, 2010.

[Schlaug et al., 2008] Schlaug, G., Marchina, S., and Norton, A. (2008). From singing to speaking: Why singing may lead to recovery of expressive language function in patients with broca's aphasia. *Music Perception*, 25(1):315–323.

[Schlaug et al., 2009] Schlaug, G., Marchina, S., and Norton, A. (2009). Evidence for plasticity in white-matter tracts of patients with chronic broca's aphasia undergoing intense intonation-based speech therapy. *Ann N Y Acad Sci*, 1169:385–394.

[Schubotz and von Cramon, 2009] Schubotz, R. I. and von Cramon, D. Y. (2009). The case of pretense: observing actions and inferring goals. *J Cogn Neurosci*, 21(4):642–653.

[Schubotz et al., 2003] Schubotz, R. I., von Cramon, D. Y., and Lohmann, G. (2003). Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. *Neuroimage*, 20(1):173–185.

[Sezgin and Davis, 2006] Sezgin, T. M. and Davis, R. (2006). Scale-space based feature point detection for digital ink. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Courses*, page 29, New York, NY, USA. ACM.

[Sherry et al., 2007] Sherry, B., Blanck, H., Galuska, D., Pan, L., Dietz, W., and Balluz, L. (2007). Vital signs: State-specific obseity prevalence among adults — united states, 2009. accessed from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm59e0803a1.htm?s$_cid = mm59e0803a1_e$

[Singer and Singer, 1990] Singer, D. and Singer, J. (1990). *The House of Make-Believe.* Harvard University Press, Cambridge.

[Smith et al., 2007] Smith, T., Gildeh, N., and Holmes, C. (2007). The montreal cognitive assessment: validity and utility in a memory clinic setting. *Can J Psychiatry*, 52(5):329–332.

[Spreen, 1998] Spreen, O. (1998). *A Compendium of Neuropsychological Tests.* Oxford University Press, Oxford, Oxfordshire.

[Stein, 2009] Stein, J. (2009). e100 neurorobotic system. *Expert Rev Med Devices*, 6(1):15–19.

[Stein et al., 2007] Stein, J., Narendran, K., McBean, J., Krebs, K., and Hughes, R. (2007). Electromyography-controlled exoskeletal upper-limb-powered orthosis for exercise training after stroke. *Am J Phys Med Rehabil*, 86(4):255–261.

[Stekelenburg and Vroomen, 2007] Stekelenburg, J. J. and Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *J Cogn Neurosci*, 19(12):1964–1973.

[Sternberg, 1999] Sternberg, R. (1999). *Handbook of Creativity.* Cambridge University Press, Cambridge.

[Sternberg and Lubart, 1999] Sternberg, R. and Lubart, T. I. (1999). The concept of creativity: prospects and paradigms. In *Handbook of Creativity.* Cambridge University Press, Cambridge.

[Stevenson and James, 2009] Stevenson, R. A. and James, T. W. (2009). Audio-visual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *Neuroimage*, 44(3):1210–1223.

[Stringhini et al., 2010] Stringhini, S., Sabia, S., Shipley, M., Brunner, E., Nabi, H., Kivimaki, M., and Singh-Manoux, A. (2010). Association of socioeconomic position with health behaviors and mortality. *JAMA*, 303(12):1159–1166.

[Survey, 2010] Survey, P. H. R. I. G. C. (2010). Health reform: Propsering in a post-reform world. accessed from http://www.pwc.com/us/en/health-industries/publications/prospering-in-a-post-reform-world.jhtml, August 26th.

[Susilo and Win, 2007] Susilo, W. and Win, K. T. (2007). Securing personal health information access in mobile healthcare environment through short signature schemes. *Int. J. Mob. Commun.*, 5(2):215–224.

[Suzuki, 1983] Suzuki, S. (1983). *Nurtured by Love.* Suzuki Method International, City.

[Swainson et al., 2001] Swainson, R., Hodges, J. R., Galton, C. J., Semple, J., Michael, A., Dunn, B. D., Iddon, J. L., Robbins, T. W., and Sahakian, B. J. (2001). Early detection and differential diagnosis of alzheimer's disease and depression with neuropsychological tasks. *Dement Geriatr Cogn Disord*, 12(4):265–280.

[Teder-Sälejärvi et al., 2005] Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., and Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *J Cogn Neurosci*, 17(9):1396–1409.

[Tierney et al., 2000] Tierney, M. C., Szalai, J. P., Dunn, E., Geslani, D., and McDowell, I. (2000). Prediction of probable alzheimer disease in patients with symptoms suggestive of memory impairment. value of the mini-mental state examination. *Arch Fam Med*, 9(6):527–532.

[Todd, 1995] Todd, N. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America*, 97(3):1940–1949.

[Tombaugh and McIntyre, 1992] Tombaugh, T. N. and McIntyre, N. J. (1992). The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc*, 40(9):922–935.

[Torpey and Jessop, 2009] Torpey, P. A. and Jessop, E. N. (2009). Disembodied performance. In *CHI '09: Proceedings of the 27th international conference* *extended abstracts on Human factors in computing systems*, pages 3685–3690, New York, NY, USA. ACM.

[Turella et al., 2009] Turella, L., Pierno, A. C., Tubaldi, F., and Castiello, U. (2009). Mirror neurons in humans: consisting or confounding evidence? *Brain Lang*, 108(1):10–21.

[Ullmer and Ishii, 2000] Ullmer, B. and Ishii, H. (2000). Emerging frameworks for tangible user interfaces. *IBM Syst. J.*, 39(3, 4):915–931.

[Van Hoesen et al., 1991] Van Hoesen, G. W., Hyman, B. T., and Damasio, A. R. (1991). Entorhinal cortex pathology in alzheimer's disease. *Hippocampus*, 1(1):1–8.

[Varshney, 2007] Varshney, U. (2007). Pervasive healthcare and wireless health monitoring. *Mob. Netw. Appl.*, 12(2-3):113–127.

[Vestal et al., 2006] Vestal, L., Smith-Olinde, L., Hicks, G., Hutton, T., and Hart, J. (2006). Efficacy of language assessment in alzheimer's disease: comparing in-person examination and telemedicine. *Clin Interv Aging*, 1(4):467–471.

[Wan et al., 2010] Wan, C. Y., Demaine, K., Zipse, L., Norton, A., and Schlaug, G. (2010). From music making to speaking: engaging the mirror neuron system in autism. *Brain Res Bull*, 82(3-4):161–168.

[wei Chen, 2005] wei Chen, Y. (2005). Combining svms with various feature selection strategies. In *Taiwan University*. Springer-Verlag.

[Weinberg et al., 2002] Weinberg, G., Aimi, R., and Jennings, K. (2002). The beatbug network: a rhythmic system for interdependent group collaboration. In *NIME '02: Proceedings of the 2002 conference on New interfaces for musical expression*, pages 1–6, Singapore, Singapore. National University of Singapore.

[Werner and Noppeney, 2010a] Werner, S. and Noppeney, U. (2010a). Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *J Neurosci*, 30(7):2662–2675.

[Werner and Noppeney, 2010b] Werner, S. and Noppeney, U. (2010b). Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cereb Cortex*, 20(8):1829–1842.

[Weston et al., 2001] Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T., and Vapnik, V. (2001). Feature selection for svms. In *Advances in Neural Information Processing Systems 13*, pages 668–674. MIT Press.

[Widmer and Goebl, 2004] Widmer, G. and Goebl, W. (2004). Computational models of expressive music performance: The state of the art. *Journal of New Music Research*, 33:203–216.

[Windsor et al., 2006] Windsor, W. L., Desain, P., Penel, A., and Borkent, M. (2006). A structurally guided method for the decomposition of expression in music performance. *J Acoust Soc Am*, 119(2):1182–1193.

[Witkin, 1983] Witkin, A. (1983). Scale space filtering. In *Proceedings of the 1983 International Joint Conference on Artificial Intelligence*, Palo Alto, California. Morgan-Kaufmann.

[Wohlfarth, 2008] Wohlfarth, J. (2008). Designers tell how music influences creativity. Available at http://www.howdesign.com/article/musiccreativity/.

[Wolfram, 2002] Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media, Champaign.

[Wood and Hauser, 2008] Wood, J. N. and Hauser, M. D. (2008). Action comprehension in non-human primates: motor simulation or inferential reasoning? *Trends Cogn Sci*, 12(12):461–465.

[Xenakis, 2001] Xenakis, I. (2001). *Formalized Music: Thought and Mathematics in Composition*. Pendragon Pr, City.

[Yan et al., 2008] Yan, Z., Wang, Z., and Xie, H. (2008). The application of mutual information-based feature selection and fuzzy ls-svm-based classifier in motion classification. *Comput. Methods Prog. Biomed.*, 90(3):275–284.

[Ye et al., 2009] Ye, J., Zhu, Y., and Li, Z. (2009). Feature point detection for early processing of freehand sketch recognition. In *CINC '09: Proceedings of the 2009 International Conference on Computational Intelligence and Natural Computing*, pages 166–169, Washington, DC, USA. IEEE Computer Society.

[Young et al., 2006] Young, D., Nunn, P., and Vassiliev, A. (2006). Composing for hyperbow: A collaboration between mit and the royal academy of music. In *Proc. Intl. Conf. on New Interfaces for Musical Expression*.

[Zatorre et al., 2007] Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci*, 8(7):547–558.

[Zatorre et al., 1994] Zatorre, R. J., Evans, A. C., and Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci*, 14(4):1908–1919.