Biometrics in Interaction and Interface Design

A thesis submitted to the College of Communication and Information of Kent State University in partial fulfillment of the requirements for the degree of Master of Fine Arts

By
Joshua Kruszynski

August, 2016
# Table of Contents

LIST OF FIGURES .......................................................... iv  
ACKNOWLEDGMENTS .................................................. vi  
CHAPTER I: Introduction .............................................. 1  
CHAPTER II: Secondary Research ................................. 3  
  Interaction and Interface .............................................. 4  
  Interfaces: Past and Future ........................................ 10  
  Biometrics .................................................................. 20  
  EEG ............................................................................ 25  
  Suitable Hardware ....................................................... 29  
  Usability and Cognition ............................................. 31  
  Information Overload .................................................. 37  
  Adaptive User Interfaces .......................................... 39  
CHAPTER III: Primary Research ................................. 43  
  Methodology .......................................................... 45  
  Data and Visualizations ............................................ 46  
  Analysis Overview .................................................. 57  
  Line Graphs .......................................................... 58  
  Scatter Plots and Cluster Analysis ............................. 61  
  Time-lapse Scatter Plot Videos ................................. 66  
  Summary .............................................................. 68  
  Challenges, Issues and Considerations ..................... 70  
CHAPTER IV: Conclusion ............................................. 72  
REFERENCES ............................................................ 75
List of Figures

Figure 1: Garrett’s Elements of User Experience ........................................................ 4
Figure 2: HfG Ulm School’s Man-Machine System ...................................................... 5
Figure 3: Norman’s “Gulf” Model .............................................................................. 6
Figure 4: Author’s Model of Closed-Loop Interaction ............................................... 9
Figure 5: Author’s Interaction Loop with Biometric System ....................................... 10
Figure 6: Evolution of Macintosh Interface ............................................................... 11
Figure 7: GlowCap and Nest ...................................................................................... 17
Figure 8: IFTTT ......................................................................................................... 19
Figure 9: Jain’s Fingerprint Recognition Resolution Levels ...................................... 23
Figure 10: Jain’s Points of Significance .................................................................... 23
Figure 11: Jain’s Fingerprint Binarization .................................................................. 23
Figure 12: Obtrusive EEG ......................................................................................... 26
Figure 13: NeuroSky MindWave ............................................................................... 27
Figure 14: Mattel MindFlex ....................................................................................... 28
Figure 15: Norman’s Mental/Conceptual Model ....................................................... 32
Figure 16: Wicken’s Model of Attentional Resources ............................................... 36
Figure 17: MindWave/Arduino Headset .................................................................. 43
Figure 18: Linegraphs Highlighting Task-Switch ....................................................... 47
Figure 19: Reading Task Linegraphs with Trendline ............................................... 48
Figure 20: Game Task Linegraphs with Trendline .................................................... 49
Figure 21: Combined Task Linegraphs with Trendline ............................................ 50
Figure 22: Reading Task Scatterplots ........................................................................ 51
Figure 23: Game Task Scatterplots .......................................................................... 52
Figure 24: Combined Task Scatterplots.................................................................53
Figure 25: Combined Task Scatterplots with Gaussian Blur Filter.......................54
Figure 26: Combined Task Cluster Analysis ..........................................................55
Figure 27: Time-Lapse Scatterplot Frames .............................................................67
Acknowledgments

Thanks Mom and Dad.
Thanks Debbie.
Thanks Jason.
Thanks Waldo.
Thanks Ken.

It’s been real.
CHAPTER 1

Introduction

In his 1986 title “The Cult of Information”, author and historian Theodore Roczak suggested that a single daily edition of the The New York Times contained more information than the average individual during the 17th century was likely to come across in their entire lifetime (Roczak, 1986). While Roczak put forward this lament during the earliest days of the Internet, the idea rings true to 21st century ears, despite having little evidence to support it. Roczak’s statement is a token; a representation of the feeling that there is an ever-increasing amount of information vying for a finite amount of human attention.

With a limited volume of attentional resources, and an increasing number of applications and agents vying for our attention, we are living with continuous conditions for cognitive strain: too much information for a given task, or simply too many tasks, for a given amount of mental resources. George Miller provided a startlingly low-ball figure for the limit of our mental faculties. Miller posited that humans can only effectively manage seven “chunks” of information (plus or minus two) at a given time (Miller, 1956). But research and intuition suggest there is more to the problem than a simple lack of capacity.

The limits of working memory capacity are compounded by the transient nature of our minds. Continuous effort is required to keep our minds focused, to control the locus—or point of fixation—of our attention (Raskin, 2000). While our working memories can manipulate seven-ish chunks, the locus of our attention can only be occupied by a
single idea or thought at any time, and this position is continuously shifting based on the stimuli we receive. So, management of the information presented by our digital devices—both sought and unbidden—competes both with internal subconscious information and continuous unconscious perceptions of the surrounding environment (Wickens, 1992).

A central tenet behind much interaction design is the creation of more “usable”, or more effective interfaces. Reductions in cognitive strain are achieved by demanding fewer attentional resources, or increasing the speed for the completion of a given task. These successes are a result of better predictions of a user’s needs or preexisting understanding of the function or behavior of some device or software. This has yielded much success in interactions which are discrete and task-completion is the objective. However, this model alone may not serve to support the growing prospect of multi-task, multi-device interactions.

This research explores the creation of interfaces which seek implicit instruction through unconscious biometric signals in lieu of explicit instructions from keyboards and pointers, as a means of redress for the imbalance between a person’s attentional capacities and the persistent nature of digital communication. Research is already underway in both the medical and technical communities to employ biometrics for purposes other than identification. This paper will evaluate biometric systems—specifically EEG—in the context of interface design, for the purposes of everyday interaction, rather than specialized application.

The research covered in this paper includes: Historical survey of interaction models and interface development; Non-technical review of biometrics, with a focus on fingerprint recognition and EEG analysis; Overview of usability and cognition; Primary evaluation of performance for a consumer-grade EEG headset in capturing changes in participant’s mental states between static and dynamic information processing.
CHAPTER II

Secondary Research

Supporting research for this investigation was divided among three main areas: interaction and interface design; biometric and EEG technologies and methodologies; and the relationship between information and cognitive strain. Having some prior knowledge of interaction and interface design stemming from academic training and professional experience, research in this domain was predominantly an endeavor to make better sense—to formalize vocabulary and deconstruct the components—of ideas which were familiar but not intimately understood.

Having very little experience in the technical domain of biometrics and EEG, this portion of research was an entrance into a fundamentally different world. The complexity of computer science does not allow one opportunities for casual pick-up or participation, however a sort of ethnography or role-play was achieved. The methods, approaches and customs of biometrics, as well as the fundamental issues, concerns and operations of working with EEG were introduced through both literature as well as a first-hand classroom experience. These experiences and core concepts will be reviewed.

Finally, what began as a simple pursuit to define the phrase “information overload” eventually gave way to a much broader survey of the relationship between society and information. The painful irony here is that there is simply too much material—both qualitative and quantitative—to adequately explore in one paper. This section will touch briefly on historical context, and then shift into the physiological principles of information overload and how today’s devices differ from those of the past.
Interaction and Interface

Patrick Whitney, dean of the Institute of Design at the Illinois Institute of Technology, explains the value of design as a process for abstracting problems. Rather than seeking to directly improve, design seeks to first understand. Whitney poses the example of developing a better mp3 player against a process of understanding the context of a rapidly-evolving music industry. The former produces a better mp3 player, while the latter produces a rich system of products and services to surround and support a better mp3 player (Whitney, 2009).

FIG. 1: The Elements of User Experience by Jesse James Garrett. 2001.

Setting aside the specifics, the notion of design-as-abstraction suggests the path toward fundamental improvement is not necessarily analysis of the thing itself but of the systems and context which frame a given thing. In the case of this research: if a digital interface is the thing, what is the context?

Looking to the Elements of User Experience (Figure 1), a well-known framework put forth by Jesse James Garrett, principal of user experience firm Adaptive Path, interface design is directly preceded on a scale of abstraction by interaction design.
Garrett’s diagram can be unwieldy in ways—interface and interaction are elements contained within a section labeled “software interface”—but let it suffice for providing a context for interface design in the more abstract setting of interaction.

So, what is interaction? Dubberly, Pangaro and Haque provide a thorough answer in their 2009 article from the journal Interactions. “What is Interaction? Are there different types?” is a tour de force of interaction models from various times and perspectives. Meredith Davis, Richard Buchanan, Donald Norman and more are invoked as the authors seek to analyze the components and dimensions of various interactive systems. The central tenet of the paper is that designers’ models of interaction are too dominated by the use of computers and digital devices as they exist today. The authors assert that ‘click; something happens’ is more accurately described as reaction.

While the perspective of Dubberly et al. could be considered a matter of scope with consideration to the behavior of both humans and devices, the case is sound that human-computer interaction suffers if understood only as a single, closed loop system—or any other singular model. The authors offer that interactive systems consist of components affording varying levels of effect upon one another, creating stratifications and classifications of interdependency and complexity.

In the words of the authors: “in ‘interaction’ the precise way that ‘input affects output’ can itself change” (Dubberly et al., 2009, p.3). Contrast this with the relatively static way that clicking something with a pointer (mouse or finger) makes something else appear or disappear on a monitor. Dubberly et al. conclude and counter this static model with a list which outlines a variety of interaction models, and question what HCI might become if a gamut of interaction models were represented. This list includes:

- Reacting to another system.
- Regulating a simple process
- Learning how actions affect the environment
- Balancing competing systems
• Managing automatic systems

• Entertaining (maintaining engagement in a learning system)

• Conversing

Taking this more robust consideration of interaction as a concept is valuable in that it helps one to consider approaches outside of the traditional feedback loop model. If interactions are being designed with the understanding of a closed loop as the total possible scope, there is also a preconception of what tools and methods are available or possible. If a reactionary system of buttons and outcomes is all there is, then a better button would seem a logical pursuit.

It should be noted—and will be detailed further into this paper—that at least a few of these models are beginning to see representation. But generally, the narrow view of interaction-as-closed-loop seems to be echoed in trade discourse—consider articles, comments and exchanges around skeuomorphic or flat visual styles, or various navigation schemes. Predominant objectives are the incremental improvement of a given thing. There seems to be little abstraction of the problem to drive fundamentally different approaches. This really becomes a feedback loop of its own sort. The input is (ironically) a treatment of a closed feedback loop, which predetermines that the output will be a closed feedback loop, albeit a slightly different version. This is the translation of one idea into a different version of the same idea.

The problem with this is not merely a matter of missing out on what is possible. This cycle misses is the value of abstraction, outlined by Whitney and provided in spades by Dubberly et al. New interfaces would be better guided by fundamentally reconsidering the underlying interaction model and the nature of the parties operating in those models.
The following quote excerpted from Dubberly et al. frames the approach to interaction taken by this research:

In the feedback loop model of interaction, a person is closely coupled with a dynamic system. The nature of the system is unspecified. (The nature of the human is unspecified, too!) The feedback-loop model of interaction raises three questions: What is the nature of the dynamic system? What is the nature of the human? Do different types of dynamic systems enable different types of interaction? (p. 3)

Of primary interest is the nature of the human, and what types of interaction are enabled by a different types of dynamic systems. Biometric input can serve to provide some descriptions regarding the nature of the human. This is the opportunity biometrics offers. With a numeric description of a human, the computer system is provided an on-going input on which self-correcting and learning interaction models can be built.

The model presented in Figure 4 is the author’s attempt at depicting a traditional interaction model while highlighting two nested loops. First is a loop which operates externally, while affecting the sub-conscious user experience. These are environmental influences which reach the subconscious without being consciously sought: notifications, advertisements, a crowded office, traffic etc. Second is an internal loop of unconscious processes and tasks (exchanged between conscious and subconscious). This is comprised of directed and undirected mental content: remembering a password, weighing options, or a busy schedule, a recent rift with a spouse, a general proclivity for mind-wandering.
Of course both of these loops represent stimuli and input affecting the person using a device which are not necessarily generated or controlled by that device or the person's conscious mind. Additionally, it is obvious that these factors cannot be controlled or modified by an information system, or by the person—at least not without great effort. However, through biometrics, the effects of these external and internal sources on an individual can begin to be described, and these descriptions can act as controls and influences on the familiar feedback loop which has characterized much of HCI.

These are some of the subtleties which I believe Dubberly was alluding to. There are variables and sub-processes involved in control and feedback which make a single closed-loop inadequate for describing 'interaction'. HCI has left the domain of discrete, tasks, and has moved into something closer to a way of living.

Our interfaces and information systems have evolved, and now our interaction model should as well. Figure 5 represents such a model, balancing the conscious and subconscious influence between a person and a device. A person can now influence the output of the device through explicit (traditional) and implicit (biometric) means.
Interfaces: Past and Future

For the past 80 years computers—and digital information technology at large—have transformed the way we live. The earliest digital devices began as the love-children of academia and war. After a century of various developments surrounding the mechanization of logic and mathematics, Colossus, the first digital computer, was built by the English to decipher cryptography during the second World War.

After wartime, the progress made on digitizing logic and computation entered the private sector where it was soon married to the newly-invented transistor at Bell Laboratories in the late 1940s, after which sprang the integrated circuit at Texas Instruments in the late 1950s, and later the microprocessor from Intel in the late 1960s. This is admittedly a patchy and grossly over simplified history of computer science, but from this string of invention and intellect, the computer as we understand it today would emerge. For a robust yet still digestible consideration of these developments, see Computing: A Concise History by Paul E. Ceruzzi.

Yet the silicon and circuitry which provide the abilities of digital technology does not ensure accessibility. If a computer is built in the woods but no one can use it, does it really compute? This is the role of the interface. Interfaces provide a means through
which a person can better or more easily manipulate computing. Being able to more easily manipulate computing allows more people to gain access to the abilities of computing, which drives demand for more computing. More computing ability allows for the creation of better—easier, more accessible—interfaces. And the cycle continues.

A similar notion was touched upon by New York Times writer David Pogue in a slide (Figure 6) from a TED talk he gave in 2006 on the conflict between humans and digital devices. Interestingly, Pogue points out a sort of modern inverse of the cycle outlined above. He contends that as devices become more accessible, we have more opportunities to find difficulty with them. This could be construed as an interface failure: interface design has failed to keep pace with growing access.
However, there have been some changes since Pogue’s talk. It isn’t that people no longer struggle with using digital devices, but that there is a new interface paradigm (although the fundamental interaction models have remained relatively static). The smartphone and touchscreen interface have provided access to wider age ranges, and more persistent access for more typical users.

But how did we get to such mass adoption? What is it about an interface paradigm that allows it to succeed? Considering how interfaces have evolved historically suggests something about how they may need to change going forward.

Going back to the earliest digital computers, the interface for the machine was the machine itself. Eventually, cards with holes punched in them—a concept borrowed from the Jacquard loom—would rise as a format for feeding information and instruction into the machine. But beyond the specific mechanisms, the crucial point is that for these earliest computers the interface was highly literal to the machine, and highly abstract to the person operating it. The operator was tasked with understanding every intricacy and detail with regard to how the machine was physically connected and constructed. The machine performed like a machine; if it didn’t receive the proper input, it would not work.

The next interface paradigm would revolve instead around human-readable language. Indeed, we are still in part living in this paradigm. Although pointing-devices known as a light-pen was in use as early as World War II, the new interface for computing would be the keyboard and monitor terminal as controls for a monolithic mainframe computer. Rather than a central processing device which needed to receive manual sets of physical instruction, mainframe computers could be manipulated through a language of instruction, which would be decompiled—decoded—by the computer into a form more native to the hardware.
When comparing keyed instruction to the physical manipulation of circuits, the difference is quite dramatic. By relying on a tool as familiar and manipulable as a keyboard, many more people were able to access the abilities of a computer. People no longer were required to have intricate knowledge of the objects and electronic gates which made computation possible (though it couldn’t hurt). This transference of burden from human to machine seems to be at the heart of interface evolution and success. Access to computing expands as responsibility shifts onto the machine (and its designers) to infer or accommodate the operator’s intentions.

This idea is alluded to throughout “Designing Interactions”, the interaction opus by Bill Moggridge (2006) which traces the development of the modern graphic user interface (GUI) and many of the hardware advances which accompanied it. Moggridge’s phrasing of “kind to chips but cruel to people” suggests many poor interfaces are based on over-accommodating the comfort and performance of a device rather than the comfort and performance of the person operating it.

Successful interfaces need not be cruel to circuits, but they must be kind to (and considerate of) people. Again, this “kindness” can be understood as the allowance of behavior which is natural or literal for a person, while “cruelty” can be construed as behavior which requires unusual or abstract behaviors. For humans, the severe and error-intolerant logic of circuitry and software tends to be unnatural and abstract (unless you are an electrical or software engineer, in which case: good for you).

Moggridge’s tome is an outstanding reference for the trajectory of modern HCI, both software and hardware. For discussions of more arcane devices and their interfaces, please enjoy “The Information” by James Gleick. A more abridged summary of device and interface milestones, and their significance, is as follows:
• **Digital Computation / Machine Languages**
  Transistors, tubes, punch-cards and other components allow humans to harness mechanized logic

• **Terminal Computers / Keyboards and Human-Readable Languages**
  Languages allow communication with computers in a method legible to humans. Sophisticated knowledge of hardware and construction is no longer necessary.

• **Personal Computers / Mouse-pointers and GUIs**
  Mice and GUIs free people from the abstraction of command-line text interfaces. Files and folders can now be represented and manipulated like their physical counterparts

• **Portable Computers / Compact Form-factors and Trackpads**
  Computing power can now be taken out of the office or off the desk. Trackpads allow for greater use of human dexterity through the fingers and create a more direct connection between the objects on screen and a person's body.

• **Mobile Computers / Styli, Gestures and Touchscreens**
  Computing can now be omni-present. Styli and gestures allow humans to use expressive shorthand techniques for manipulating data, and write directly onto/into a device. Touchscreens make control literal.

Today, we seem to be living on the outer rim of the mobile and touch interface paradigm. Artificial intelligence and voice control are offering new levels of control burden transference, though without breakthrough levels of success. Personal assistant software such as Apple Computer's Siri, Google's GoogleNow, and the Echo device from Amazon allow people to simply speak at their device to interact. Asking, “What time is
my lunch meeting tomorrow?” or perhaps “What day is my mom’s birthday?” prompts the device to analyze the recording of your voice, extract relevant phrases and estimate your intentions. Although current shortcomings often draw the notion of ‘intelligence’ into question, when successful the effect is compelling. In some ways, the use of voice as a control could be considered a biometric interface, one that is non-visual, and almost entirely software-based for the person interacting with it.

In the late 1980s and early 1990s Mark Weiser, a researcher at Xerox PARC, lead a vision for a type of computing he called Ubiquitous Computing, which was laid out in the Scientific American article “The Computer for the 21st Century” (Weiser, 1991). Weiser describes three waves of computing based on the quotient of users-to-computers. Weiser’s list is as follows:

- **Mainframes:** one computer, many users
- **PCs:** one computer, one user
- **Ubiquitous Computing:** many computers, one user

On his site (maintained posthumously on Weiser’s behalf by PARC), Weiser explains that Ubiquitous Computing is “roughly the opposite of virtual reality. Where virtual reality puts people inside a computer-generated world, ubiquitous computing forces the computer to live out here in the world with people.” (Weiser, 1996). The original vision for Ubiquitous Computing revolved around “tabs, pads and boards”—computers of a wide range of sizes and capacities. In this third wave, technology is supposed to recede into the background, and we will no longer be sitting with computers “staring uneasily at each other across the desktop” as Weiser put it.
In some ways, Apple’s Siri, GoogleNow and Amazon’s Echo seem to be a mild step in a Ubiquitous Computing direction. Additionally, smartphones—and now smartwatches—tablet computers, and wall-sized touch screens bear something of a resemblance to the idea of tabs, pads and boards. But Weiser also had a more atmospheric conceptualization, best embodied by “dangling string display” created by Natalie Jeremijenko (Weiser, Brown, 1995). As described by the authors, the dangling string is:

“an 8 foot piece of plastic spaghetti that hangs from a small electric motor mounted in the ceiling. The motor is electrically connected to a nearby Ethernet cable, so that each bit of information that goes past causes a tiny twitch of the motor. A very busy network causes a madly whirling string with a characteristic noise; a quiet network causes only a small twitch every few seconds. Placed in an unused corner of a hallway, the long string is visible and audible from many offices without being obtrusive. It is fun and useful.” (Weiser, 1995, para 2).

These sorts of devices are what author, inventor and MIT professor David Rose refers to as “enchanted objects” (Rose, 2014), but are more popularly categorized as the “Internet of Things” (IoT), a term coined in 1999 by British Entrepreneur Kevin Ashton (Ashton, 2009). The idea behind this multi-named movement is the introduction of computation—and the ever-important corollary of network connectivity—to non-computer devices.

Rose provides similar philosophical and quality of life arguments as Weiser: that technology should, and will, fade into the background of our lives. Ashton’s interest are decidedly more practical: that there are huge economic benefits to the masses of data
which can be collected and analyzed from all of the physical objects humans work and live—or interact—with on a daily basis.

An example of one of these devices from Rose’s book—appropriately titled Enchanted Objects—is the GlowCap prescription bottle. Produced by Rose’s own startup company, Vitality, the GlowCap is a network-connected prescription bottle-cap which uses lights and sounds to notify someone that it is time to take their prescription. If the prescription is not taken (or at least, the bottle not opened) a message is sent to a family member or guardian who can then intercede.

FIG.7: The Vitality Glow-cap (left) and Nest thermostat (right).

A more familiar element of the Internet of Things might be the Nest thermostat. This is a thermostat—again, network connected—which keeps track of a wide variety of data such as time, ambient room temperature, weather etc., and over time endeavors to predict necessary heating and cooling adjustments before a home-dweller makes them.

The important point from these examples is something of a paradox. In the future, we may hardly notice computer interfaces, though they might also be in everything
and everywhere around us. Whether we take notice of them seems moot, though the human brain’s predisposition toward sight would seem to suggest that we will at least see some part of them. However, the idea that there will be an explosion in the number of devices we regularly interact with seems inevitable. The rise of smartphones has made information technology a persistent presence, from the bathroom to the bedroom and everywhere in between. In a 2011 report, Dave Evans of Cisco Systems estimated that the number of connected devices surpassed the number of people on the planet between 2008 and 2009. By 2020, Evans projected there will be 50 billion—that’s 50,000,000,000—connected devices (Evans, 2011).

Two compelling examples which straddle, or perhaps defy, specific categorization are Google’s web-based email client Inbox, and IFTTT—an initialism representing the phrase “If this, then that”—which may be best described as a service-control service. Both of these interface models rely upon existing hardware and control mechanisms while pushing into something Dubberly et al. might have considered “learning how actions affect the environment” and perhaps “managing automatic systems” (2009).

Inbox attempts to use automatic analysis to classify and cluster emails together which fit predefined, or user-defined group definitions. To some extent, this capability is already accessible in the existing Gmail interface, and many web clients, as spam filtering. However, Inbox extends this concept by clustering labels like “social” or “promotional” into individual content objects. The value proposition is that this allows better message segmentation and prioritization (an idea which can be appreciated by anyone with an active email flow).

IFTTT is an unusual case, as it is a service which allows for fully or semi-automated control of other services, and the integration of these services as controls
for one another. The IFTTT website suggests pairings such as “If: Twitter; Then: save to Google Sheets”, or “If: Instagram; Then: Dropbox”. These represent “recipes” (IFTTT 2016) which allow different web applications to interact with one another through their own Application Programming Interfaces (API). An API is a library of predefined code segments which allow two pieces of software to interchange data seamlessly, harkening back to text-based programming controls, only without the human!

FIG.8: IMAGE FROM IFTTT MARKETING MATERIAL DEMONSTRATING HOW DEVICES AND SERVICES INTERACT.

In the first example: a service like Google Sheets might download any Twitter message posted with a particular hashtag. In the second example, if an image is posted to an Instagram account, download a backup copy to a particular Dropbox folder. What is interesting about this case, is that the interaction paradigm uses a traditional closed feedback loop to allow users to establish autonomous loops. These are not interfaces we are intended to attend to. IFTTT is succeeding when you are not controlling it directly, a notion which Mark Weiser may have enjoyed.

Whether it is Ubiquitous Computing, the Internet of Things—whatever title you prefer—the trend does indeed seem to be toward something we will be living among, rather than “staring uneasily at each other across the desktop”, as Mark Weiser
(1996) phrased it. Following the trajectory of past computing waves, and patterns in
the relationship between interface and device adoption as reviewed, if there is going
to be a dramatic increase in devices and computational ability, there will need to be
a correspondingly dramatic interface and interaction model. Though the issue may
not be the extent humans are able to access computers, but the extent computers are
able to access people. Explicit communication and direct control will not suffice for 50
billion devices. We will need an expanded range of controls. How do we interface with
computers that are integrated into our lives?

**Biometrics**

Journals, papers and online tutorials can be valuable resources for building and
exercising existing knowledge, but existing knowledge at the outset of this research in
regard to programming and computer science was minimal. For introductory purposes,
it remains difficult to best the experience of interacting in person with an expert. In the
Spring of 2015 I took a class, Biometrics and Multimedia Computing, with Dr. Arvind
Bansal of the Department of Computer Science at Kent State University. This class was an
invaluable experience, and served as a primary source of understanding for the concepts,
vocabulary and methods surrounding biometrics and computer science.

The idea of biometrics is fairly straightforward: to capture the physical
characteristics of biological organisms and translate them into a record of data. This can
be as simple as the information stored about you on your driver’s license. Your height,
your weight, and the color of your hair are simple examples of biometrics, referred to
as “soft” biometrics. Soft biometrics can be easily understood as how you would describe
yourself or another person over the phone. Tall-ish, button-nose, dark skin, large beard.
These traits can change over time, they are not fixed nor are they distinctive; many people have brown hair. The opposite of soft biometrics are hard biometrics. These traits are distinctive and do not change over time (at least not naturally). One’s fingerprint is an example of a hard biometric, as is the pattern in the iris of a person’s eyeball.

The distinction between hard and soft biometrics is important in the application of computational identification. Use in identification introduces another fundamental distinction: the difference between verification and identification. In verification, the system is checking an individual’s claimed identity against a finite pool of potential identities. In identification the system is determining which identity from a pool of potential identities best matches a given set of data. This is a subtle distinction but an important one. If the system is based on verifying against a claim, this dramatically reduces the amount of processing required because the system can simply check the provided data against the data which is known to belong to a particularly identity. To make this more concrete: consider the difference between verifying whether your friend’s face matches the face on the driver’s license in their wallet, or whether your friends face matches the image on one of 50 driver’s licenses.

Biometric data can be gathered from a wide array of sources. Fingerprint and iris have already been mentioned, other sources include: hand-shape, palm print, voice and speech patterns, facial structure, gait, vein patterns, DNA, and even odor. For non identification purposes, biometric information can be gathered on heart rate, breath rate, posture, eye movement, writing and keystroke pattern, language analysis, electromyography (muscle movements) and more. Each of these sources can be captured using audio/visual or other electrical sensors and translated into a set of
numerical features which the computational system uses as benchmarks to compare future inputs against.

The depth and breadth of biometric techniques boggles the mind, but Dr. Bansal’s class in combination with Anil Jain’s Introduction to Biometrics provided an excellent overview of the underlying functions of many biometric identification techniques. What follows is a summary of the basics of fingerprint analysis adapted from Jain (Jain, 2011) which will be valuable in understanding the methods and concepts employed in primary research to follow.

Within each fingerprint are a variety of quantifiable numerical relationships, referred to as ‘features’. These features are based upon abstraction and analysis of the ridges which make up a fingerprint. Jain categorizes these features into three levels (Figure 9).

The first level is based on constructing an orientation field out of the print, which breaks the print into a grid of cells, and measures the angle of the ridge relative to the baseline. This produces an array of representative dash lines as shown in Image B of Figure 9. This array is then analyzed in 9 cell units, where the center cell acts as a reference and the surrounding cells help to identify points of significance (Figure 10). Based upon changes in angle (or lack thereof), the location of specific points known as ‘singular points’ can be identified. These are created by either loop or delta ridgeline patterns. In a loop pattern ridge angles form a ‘core’, where the smallest radii of curving ridges converge with a single ridge in the center. A delta is formed by three distinct angles converging upon a central point. These point types and their locations relative to one another, and the relationship of angles thereby formed, are stored as part of a feature set.

Second-level features are extracted from binarized filters of the fingerprint image. Binarization is a filtering technique which forces each pixel of a grayscale
FIG. 9: Resolution levels of fingerprint recognition. From Jain, 2011.

FIG. 10: Isolating points of significance through the analysis of ridge angles. From Jain, 2011.

FIG. 11: Binarization of fingerprint ridges. From Jain, 2011.
image to a value of either black or white, as opposed to the typical grayscale range of 0 to 255. Binarization is typically requires some sophistication in order to remove noise—unwanted artifacts and image blemishes—and accurately represent the original print. When binarization is complete, a thinning filter is applied and each ridge in the print becomes a one-pixel wide path (Figure 11). From this path, ridge frequency and points called “minutiae” are identified. Minutiae consist of ridge endings, where a ridge stops or there is a gap, and ridge bifurcation is caused by one ridge splitting into two distinct ridges. There are many of these points across a print, and the type and location of each of these is stored as part of a feature set.

Third-level features require significantly richer images. These features are based on identifying and classifying the non-linear shapes of ridges and the gaps formed by the pores of the skin as shown in Image D of Figure 9. Needless to say, the processing and feature-extraction at this level is significantly complex. Both Jain and Dr. Bansal touched only briefly due to the complexity.

So, how do fingerprints relate to integrating biometrics into interfaces? The central idea is that from an informationally-dense piece of data, points and elements of significance can be extracted, stored and recognized by a computer program. A machine can identify an individual from thousands of others based upon a seemingly random set of lines. Today, this can be done so effectively that it has been integrated as a 1-second process used to unlock iPhones and other consumer-level devices. We can already tell our machines something about our individual biologies as a method for identifying ourselves, we do it everyday.

If we could extract the significant features of other biometrics, could we provide some insight beyond identity, into a person’s state of mind? There are practical
considerations, however: a fingerprint is a static trait which produces a still image. Analyzing the brain is a different matter.

**EEG**

Electroencephalography (EEG) is the monitoring and measurement of the brain’s electrical activity. Medical applications of EEG use an array of electrical sensors placed on the scalp to correspond with various regions of the brain (Jasper, 1958). Variations in neural activity in different regions of the brain are detected by these electrical sensors and digitally recorded. EEG is used in medical applications related to seizures and epilepsy, as well as sleep and brain-damage or coma events. Quantitative analysis of EEG (QEEG) has more recently been used in Cognitive Behavioral Therapy for patients suffering from post-traumatic stress disorder (PTSD) as well as clinical anxiety and depression. For a startlingly detailed and technical treatise on these uses see Budzynski, Budzynski and Evans, Chapters 10, 12 and 18.

Typically, EEG has not been used for identification, like fingerprint or facial analysis, due to the the excessive amounts of noise (extraneous information) which must be filtered and processed out, as well as the highly-variable, dynamic nature of the brain. Much of EEG analysis is based on measuring long and short-term changes in EEG levels, referred to respectively as evoked potentials (EP) and event-related potentials (ERP). Measuring EPs and ERPs on different frequency band ranges can describe various mental conditions or states. A basic outline of these band ranges is as follows (Abhang, Gawali, Mehrotra, 2016):

- **Gamma Band:** >35hz, Concentration
- **Beta Band:** 12–35hz, Anxiety-dominant, active, external attention, relaxed
- **Alpha Band**: 8–12hz, Very relaxed, passive attention
- **Theta Band**: 4–8hz, Deeply relaxed, inwardly focused
- **Delta Band**: 0.5–4hz, Sleep

Recently, a system using ERP measurement and analysis was used by researchers at State University of New York to identify individuals with an astounding 100% accuracy (Ruiz-Blondet, Jin, Laszlo, 2016). This system worked by exposing people to a series of images in succession and measuring their various ERPs.

While the progress made in clinical EEG use is impressive, both the hardware and processing involved in such application is extremely complex, requiring expertise in a range of fields—electrical engineering, neuroscience, computer science, etc. Additionally, medical or highly-technical applications involve hardware which is obtrusive due to the number of electrodes (Figure 12), if not highly invasive due to sub-dermal application. More recently new single-node EEG devices have been introduced to the consumer and research markets.
These devices are dramatically less invasive and more affordable, making them accessible to less clinical, laboratory-intensive, research. The company Emotiv produces one such device, the Epoch, at a cost of a few hundred dollars, which uses multiple sensors moisturized with saline solution to improve conductivity. NeuroSky is another manufacturer, producing devices like the MindWave™. The MindWave is not much different than wearing a set of headphones (Figure 13), uses a dry sensor and costs in the low hundreds of dollars—as opposed to the thousands involved in clinical EEG hardware. In fact, MindWave runs on a chipset so affordable it has been integrated into children’s toys such as the MindFlex from Mattel (Figure 14) where your EEG activity controls the motion of a foam ball.

Of course, there are drawbacks. Having only a single sensor, there is far less information available for processing, and the data supplied by the headset is not the raw electrical signal of a typical EEG setup. Instead this hardware puts out pre-filtered values which have undergone significant proprietary processing, with labels such as “attention” and “meditation”. This aspect of consumer-level EEG has drawn skepticism; critics
speculate that NeuroSky sensors may not generate valid data. However, both informal and academic investigations have shown that the hardware retains some research validity.

Crowley et al. used the Stroop Color Word Interference test—a task widely used to induce minor mental stress—to gauge cognitive strain in participants. In part of the same study participants also completed The Towers of Hanoi, a simple puzzle. The results of both tasks affirmed that the NeuroSky headset readings correlated with participants reports of stress. (Crowley, Sliney, Pitt, Murphy, 2013)

FIG. 14: Mattel MindFlex headset. Image from Mattel.

In a more practical setting, Patsis et al. showed comparable hardware was able to accurately gauge the cognitive strain of participants playing games of Tetris (Patsis, Sahli, Verhelst, De Troyer, 2013). Similarly, Bao H.T. showed a NeuroSky single-node EEG capable of discerning when participants were having difficulty with reading and comprehending complex vocabulary. Both videogames and learning are popular areas of research for non-medical EEG applications, with a general aim toward measuring
cognitive experiences in order to adjust task parameters according to each individual’s needs.

While clinical and technical applications of EEG are impressive, they also seem focused on the most challenging aspects of EEG: deriving fine, precise, explicit control inputs from a noisy data source through invasive means in controlled settings. Brain waves and ERPs may be predictable with high-performance hardware on a set corpus of images, as in Ruiz-Blondet et al., but the use of lower-resolution, non-invasive hardware for gauging a broader measure of cognitive strain seems more relevant to interaction and interface design for digital information at large. Enabling an individual to manipulate a screen pointer with their mind might be possible, but it would heinously difficult, and moreover: technical reinvention of the mouse. Enabling a program to gauge when a person is stressed or anxious however, represents a different sort of input altogether which could support a unique interaction model and a different interface paradigm.

**Suitable Hardware**

Due to a total lack of experience with fields applicable to traditional, multi-node EEG analysis, consumer-level hardware seemed the only viable path forward for any primary investigation. Fortunately, past experience with the Arduino prototyping platform provided access to less traditional routes to NeuroSky hardware. How to Hack Toy EEGs (Mika, Vidich, and Yuditskaya, 2010) proved an exceptionally valuable resource for outlining various consumer-level EEG options, as well as guidance on manipulating the Mattel MindFlex hardware in order to access the NeuroSky chipset contained within, and prepare both devices for use with the Arduino and Processing programming platforms. Sparing the technical details, Mika et al. provide a method for circumventing
aspects of the MindFlex’s internal circuitry in order to route that data into the Arduino and then to the computer.

Reading through the work of Mika et al., as well as the NeuroSky documentation provided a deeper understanding of both the criticisms and value of consumer EEG. Within the article are source documents for the Arduino Brain Library (Brain Library, 2010), which serializes (orders) the data coming from the NeuroSky chip and routes it into a computer via USB. Mika et al. also provided a visualization suite (Brain Grapher, 2010) for viewing the EEG sample data in real-time. The Brain Library files “print” the ordered data samples coming from the hardware to a window on screen.

Each sample of data coming from the MindFlex contains 12 values. The first value is a duration timestamp, in seconds. The second value is “signal”, which describes (somewhat idiosyncratically) the quality of the signal from 0 to 200, where 0 is ideal and 200 is unacceptable quality. The third and fourth values are labelled “Attention” and “Meditation” respectively. These values are pre-filtered by the chip inside the MindFlex headset. According to NeuroSky documentation, the Attention value indicates “the intensity of mental ‘focus’ or ‘attention’... increases when a user focuses on a single thought or an external object, and decreases when distracted.” (NeuroSky, 2011) While the Meditation values indicate “the level of mental ‘calmness’ or ‘relaxation.’ The value ranges from 0 to 100, and increases when users relax the mind and decreases when they are uneasy or stressed.” (NeuroSky, 2011)

Delightfully, in addition to these values, the method outlined by Mika et al. also allows for 8 bands of raw EEG data to be captured when using the MindFlex. However, the details regarding these values in the NeuroSky support documentation were somewhat opaque (NeuroSky Support, 2014), and these values were ultimately not utilized, in favor
of the simpler pre-filtered options utilized by Patsis et al. and others. However, with fingerprint and other biometrics as a point-of-reference, cutting edge hardware and analysis are not prerequisites, considering that two of the three levels of fingerprint analysis discussed earlier can be accomplished by a sliver-sized sensor inside a cell phone.

**Usability and Cognition**

As previously discussed with regard to computing power: the means by which we interact, and the enabling effects of those interactions (or lack thereof) help to determine the growth and adoption of digital technology. The pure ability of a machine does not ensure any benefit to a person. The same can be said with regard to a biometric interface. Simply because measuring cognitive strain is possible does not ensure it will provide benefit. So, what benefits might it provide? How will a biometric interface help users become better capable of wielding future computational ability?

This is a difficult question to answer through literature review. One part of that difficulty is that such a question can only be fully answered from practice and experimentation. But a simpler challenge stems from the fact that so much writing and thought around interface and interaction design—especially in professional practice—concentrates on prescriptive sets of rules. ‘Do this; don’t do that’.

It is an understandable approach for a profession driven largely by commercial outcomes. Predictability is a large part of the foundation that interaction and interface stand on. Jakob Nielsen provides this phrasing: “What users believe they know about a UI strongly impacts how they use it. Mismatched mental models are common, especially with designs that try something new.” (Nielsen, 2010). People have expectations, and the interface designer’s job is largely an attempt at creating something which confirms those expectations.
Nielsen’s partner, psychologist Donald Norman outlined this idea in his seminal book “The Design of Everyday Things”. Norman diagrammed (Figure 15) an interface as an attempt by the designer to create a “system image” which resolves their own conceptual model (the “design model”) of some thing with the user’s “mental model” of some thing (Norman, 1990).

FIG.15: NORMAN’S MENTAL (CONCEPTUAL) MODEL.

In closed feedback loop interaction models, the objective has primarily been a best-guess at what a user is going to need to complete a task. Given this objective, it makes sense that rules, best practices and ‘heuristics’ (in the industry jargon) would be a go-to resource. It is indeed the course of knowledge to build upon the work of others. But as discussed, the role of computation is shifting. We are moving away from a strictly task-completion model, away from a closed feedback loop. Or perhaps it can be understood that task-completion is becoming so prevalent, and digital interaction so prevalent, that the way in which we complete tasks is subject to change.

If we are moving toward a more ubiquitous style of access for computational abilities—living with computers rather than sitting at desks and using them as a tool—then it seems logical to employ a more ambient, less task-driven approach. This is not to say that traditional interaction models or interface design are in anyway obsolete, rather
that they can be expanded, just as the role of computing is set to expand. As Dubberly et al. suggested, there are a variety of types of interaction to be embraced. What then is to drive the approach if not task completion?

In a testament to his sagacity, 20 years after laying out many fundamental principles of traditional HCI in “The Design of Everyday Things”, Donald Norman suggested another path forward for computers in “Living with Complexity”, one where machines are fit to not only accomplish tasks, but to live with:

What is needed? Sociable machines. Basic lessons in communication skills. Rules of machine etiquette. Machines need to show consideration for the people with whom they interact, understand their point of view, and above all communicate so that everyone understands what is happening. (Norman, 2011, p. 117).

While Norman’s concerns remain focused on resolving the system-image between designer and user, the statement raises the question of what defines a successful interface outside the scope of task-completion. Sociability? Etiquette? How does a machine demonstrate communication skills? Indeed the purpose of most computers is to communicate as much information as possible. Is this not their skill? Returning to Mark Weiser’s “Calm Technology”, he explains:

“Designs that encalm and inform meet two human needs not usually met together. Information technology is more often the enemy of calm. Pagers, cellphones, news-services, the World-Wide-Web, email, TV, and radio bombard us frenetically. Can we really look to technology itself for a solution? (Weiser 1996)
Overlapping Norman and Weiser suggests that sociable, ubiquitous computing—“calm technology”—may be about not communicating, respecting boundaries or finding the appropriate channel or moment for communication. Are these the communication skills that Norman referred to? Not the ability to communicate plain and simple, but to do so with something resembling grace or tact.

If we are looking for such qualities, we will need a broader bandwidth of interaction models. Consider the spectrum of subtleties that comprise interpersonal communication: body language, non-verbal cues, eye contact—even involuntary emotion. These are the realm of the subconscious, of communication without thought. Without these available to machines, a person must manually translate these types of implicit information into explicit commands for a multitude of digital devices. Although a source of confirmation could not be identified for this notion, it seems an obvious mental tax; unnecessary if not unbearable.

How do we address mental tax? We have guides for alleviating it within task-completion—rule-sets for interface design and interaction guidelines laid down by Cooper et al. and numerous others. Guides for achieving ‘calm’ information technology are only beginning to emerge, but ‘calm’ interfaces or interactions can be better understood by considering a few of the fundamental cognitive devices which underpin human thought.

The first is working memory. Working memory is tied closely to the work of George Miller, author of “The Magical Number Seven, Plus or Minus Two” (Miller, 1956). The peculiar title alludes to Miller’s research findings that participants could hold approximately seven “chunks” of new information in their mind at a given time. As you
may have guessed, some participants managed more, and some managed less. While research has endeavored to build and elaborate on working memory, the fundamental idea has remained: limited, short-term cognitive resources with immediate availability.

These features of memory can be contrasted with long-term memory. Where working memory has a finite size and is time-sensitive, long-term memory has a vast (if not unlimited) size, with the potential for lifelong retention. Long-term memory can be fed into working memory for active use.

A second important device is the locus of attention, a concept discussed in Jef Raskin’s “The Humane Interface”. The locus of attention can be understood as “a feature or an object in the physical world or an idea about which you are intently and actively thinking” (Raskin 2000, p.12). This could be phrased more colloquially as: that which is ‘on your mind’ at any given moment. Raskin continues to draw a distinction between ‘focus’ and ‘locus’, where focusing is a conscious act which is directed to bring something onto or into the locus. The locus is a position which may or may not be maintained through the act of focusing. In his words:

“You can see the distinction when you contemplate this phrase: ‘We can deliberately focus our attention on a particular locus.’ Whereas to focus implies volition, we cannot completely control what our locus of attention will be.”

The salient point is that the ‘mind’ can only be occupied by one thing at a time, and that point or thing is not always entirely a voluntary choice. Raskin uses the example of a firecracker going off. The mental filter which selects which perceptions (everything we experience through our senses and thoughts) are made the locus of our attention—or drawn into working memory—is not infallible. Despite trying to focus
on writing a thesis, the locus may be drawn to a nearby conversation. Or, when trying to focus on completing a task, the locus may be drawn toward extraneous information, within the task domain (screen, object, etc.) or without.

The relationship between these concepts and a mental tax can be seen in Figure 16 from Wickens (1992). Here we can see the connection between the use of attentional resources on our perceptions, our cognitive processes—encompassing our attentional locus, as well as other competing stimuli—and both our working and long-term memories.

**FIG.16: Wickens’ diagram of attentional resources.**

Matthews, Davies, Westerman and Stammers elaborate on this—drawing on the work of Wickens (1986), O’Donnell and Eggemeir (1986) as well as Mulder (1986). The authors position the concept of mental workload as an analogy to physical workload (Matthews et al., 2000, p.87) and explain that “performance deficit may ensue when workload exceeds available resources. Workload also refers to people’s experiences of
cognitive task performance as effortful and fatiguing.” Additionally, the authors note that while resources may become overloaded during single task completion, attempting to complete multiple tasks (in their case, two at once) represents a more attentionally-intensive burden.

Put plainly, one’s attention (locus) and cognitive capacity (working memory and processing capability) are limited. These limitations become more pronounced when they are divided among many tasks. On top of this, the sheer perception of laborious and fatiguing tasks have effect—so our attentional capacity may be affected by a pre-existing mental or emotional state. The negative outcomes of these ideas become amplified when the involuntary nature of the attentional locus are paired with the bevy of modern signals competing for our attention as outlined above by Weiser.

**Information Overload**

Multitasking is one of the central tenets of contemporary interfaces. Email messages, instant messages, text messages, phone calls, social media, web browser tabs and windows, not to mention the physical interactions of the people and environment around us. These all flood in through a single channel to compete for and be addressed by our limited attentional resources. A central tenet of a future ubiquitous computing model would go beyond multitasking, but an interface for ‘ubiquitous tasking’ cannot require continuous conscious attentional resources without producing significant “performance deficit”, to borrow the clinical phrasing of Matthews et al.

These “performance deficits” are most commonly understood as the experience of ‘information overload’ a rather nebulous term popularized by sociologist Alvin Toffler in his book “Future Shock” (1970). As reviewed, the core issue is simply being provided so much
information one becomes less effective at task completion. Or taking a broader view, being provided so much information through so many tasks that one becomes less calm and more anxious. These experiences of information overload can be traced to phenomenon such as map shock characterized by “a negative emotional and motivational reaction to visual complexity” (Danserau, 1998). Additionally, in their book “Positive Computing”, Calvo and Peters highlight that the ability to focus (lack of attentional resources) is a dimension of both CES-D and the DSM-IV evaluations for depression (Calvo & Peters, 2014 p 94-95).

“Positive Computing” is a brief but considerable look forward at technological approaches to improving overall wellbeing. This is a growing research area, one with the attention of MIT professor Rosalind Picard and her Affective Computing Media Lab group. One such project supports automatic stress recognition in call center employees using the biometric of skin conductance—measuring tiny changes in the electrical resistance of the skin due to sweat (Hernandez, Morris, Picard). Calvo and Peters provide a simpler, more behavioral (measuring conscious activity, rather than subconscious biometrics) approach in 750words.com by Buster Benson. This site is a blogging platform which tracks keyboard activity patterns to classify moments of distraction while writing (Calvo & Peters, p. 165).

Perhaps the most passive example of “affective” or “positive” computing, is f.lux (f.lux, 2016), a simple application which does nothing more than adjust the color temperature of a computer monitor in accordance with the location of the sun. The objective is in part to reduce eyestrain due to blue light, but more to make users more aware of their own circadian rhythms, and gently remind them that they might consider ending their work for the evening.

These types of interactions are what Richard Thaler and Cass Sunstein refer to as “nudging” in support of libertarian paternalism, or the notion that one can be supported while remaining independent (Thaler & Sunstein, 2008). Measuring cognitive strain
through EEG could provide passive, continuous input for a system which could reduce information density, or help maintain focus. Attentional strain, or a lack of calm are the sorts of information on which communication ‘skills’ could be developed for our machines. Poor times to interrupt, hierarchy scaffoldings to distinguish essential information from non-essential, even measures for observing one’s own predilection for distraction or attempts at multi-tasking. These are the considerations given by a careful companion—not a task-master. Often offered without request, accepted when appreciated, and declined otherwise. These two models are not mutually exclusive interaction models and interfaces could co-exist; the keyboard remains functional, explicit—directed conscious—input just as it has for 80 years. Meanwhile, EEG and other biometrics can become the implicit—undirected, unconscious—corollaries.

Traditional interaction models and interface guidelines and have helped direct attentional resources within tasks. These are effective; they succeed in producing an intended result. Future interaction models and interface guidelines—in part through utilizing biometrics—can help conserve attentional resources among tasks. These can be affective; pertaining to a general emotional or attitudinal result. That is Weiser’s calm technology.

**Adaptive User Interfaces**

The notion of an interface which changes is in large part counter to much usability and interface best-practice (see Nielsen). The past decades are littered with examples of software trying to be helpful when it is utterly ignorant. The much maligned “Clippy” from Microsoft Word stands out. And while some of the previously mentioned research sets a tone of radical transformation, a worthy example of machine-consideration can be
found today in many websites. Not necessarily in the content, navigation or visual design, but at the structural or skeletal approach (see Garrett) applied through the practice of responsive web design (RWD).

Responsive web design has roots stretching back to “A Dao of Web Design”, an article on the web-based journal A List Apart. In “A Dao of Web Design.”, author John Alsopp built a case against standard design practices of control and pixel-perfectionism by drawing a parallel to the practice of Daoism and the ancient text “Tao Te Ching”. But it would be another 10 years before the technical machinery of responsive web design as it is experienced today was assembled by author and developer Ethan Marcotte in another piece from A List Apart appropriately titled “Responsive Web Design”. RWD began by focusing on the combination of fluid layouts (non-static sizing of text and other elements), media queries (automatic detection of device characteristics) and fluid images (you guess it: non-static treatment of images) to present different layouts and presentations of information based upon the hardware being used to view the content. The content is changed to create a better experience based upon a given individual’s device hardware.

This is another example of interface development to help accommodate users to a new technological ability. As more and more people began accessing the World Wide Web through a growing number and variety of devices (phone, tablet, television, watch), the device took on more of the burden of interaction by handling the presentation of information to better fit the hardware being used by a given individual. Referring again to Garrett’s Elements of User Experience(Figure 1) we can understand how RWD effects the interface. The “visual” layer and the “skeleton” layer become softer suggestions of presentation rather than executions. The design and implementation of
these layers is determined in part on the user's end of the interaction. In this way RWD can be seen as a form of parametric design.

Parametric design is a process involving the definition of elements within a system and the relationships of effect among those elements. Parametric design began being articulated in the mid-twentieth century with architect Christopher Alexander’s “Notes on the Synthesis of Form” (1964) and Karl Gerstner’s “Designing Programmes” (1964), although the underlying ideas, outlined by Murdock (2015), can be viewed as extensions of modularity and systems thinking which can traced back to Gutenberg.

In responsive web design, the parameters can be divided into control and display types. Display parameters are CSS (cascading style sheet) declarations used to tell web browsers how various elements should be displayed. The size of typography, the width of a given column or container, the number of columns or containers in the page grid, or the color of a given element are just a few rudimentary declarations. Control parameters are comprised of CSS media-queries, pieces of code which enable or disable specified display parameters based upon various characteristics of the hardware that is running the web browser application, as well as user-settings within the browser software itself. System characteristics including screen size, pixel density, touch capabilities, even printers and screen-readers can be specifically accommodated.

Through responsive web design, users on different devices are served different interfaces based upon the restraints or opportunities offered by various hardware. This sorts of control/display parameter relationship could be applied to interface design with EEG or other biometrics providing control parameters to describe the cognitive strain or sense of anxiety in a user.
Considering the negative attentional implications of information density and multitasking, a simple f.lux-esque application might “nudge” a user to take a break when attention levels are persistently low, or perhaps reorient the position of elements to simplify the layout of a web-page like the reader-mode option offered by many browsers. These are fairly limited examples attached to the interfaces of today. The very notion of a “display” implies some visual, when that may not necessarily be the case. The “display” parameters could be employed for more atmospheric uses. Notifications, messages even environmental factors (lights, temperature etc.) could be controled based until a user has relaxed, or completed a focus-demanding task.

The most meaningful opportunities for applying biometric control parameters may be difficult to foresee from a contemporary perspective, but will come into focus as computing is increasingly integrated into each aspect of our lives. In preparation for this, some processes for defining these control parameters are explored in the next section.
CHAPTER III

Primary Research

With a better understanding of how biometrics could change interaction models, and a point of reference for interfaces which can adapt to control parameters, the objective of the primary research was to explore how those control parameters—or features—could be extracted from consumer-level EEG information.

As discussed earlier, the work of Patsis et al. detected levels of cognitive strain in accordance with reading tasks, while Bao identified a correlation between game-difficulty and attention levels for participants playing games of Tetris. Both papers suggested there was validity in using low-cost, single-node EEG headsets similar to the hardware used in this research. Both activities also provide rather practical advantages by being general.

FIG. 17: EXPERIMENTAL MINDFLEX + ARDUINO HEADSET
enough in nature to afford participation in the trial from just about anyone (making it simple to find candidates).

In addition, a reading task and a game task approximate passive and active forms of interaction—reading requires only that information be processed, while Tetris requires information be processed and some action be taken in response. Referring back to Matthews (2000), Tetris could also serve as a multi-task attentional burden, where larger performance deficits could be expected. Accordingly, the methodology in this evaluation was established as: A.) have participants read, B.) then have them play a game of Tetris, C.) evaluate the results.

Despite a formal classroom experience and having conducted extensive secondary research, a paucity of first-hand experience in the application of statistical analysis and more sophisticated algorithms (Support Vector Machines, Principle Component Analysis, etc.) remained a limitation. The lack of third-party support from or collaboration with a computer scientist or a neuropsychologist all but insisted on a formative evaluation than a strict experiment.

An initial assumption, due in part to internal and external skepticism regarding the quality of signal and viability of the data being produced by the MindFlex-Arduino hardware combination, was that the collected values would be polluted with noise. A competing assumption was that if the data coming from the headset proved usable, there would a discernable difference across individuals between the reading task and the playing task. Finally, a third assumption was that if the data proved usable, a comprehensible display/control relationship could be identified. With secondary and frankly, common, knowledge suggesting the implementation of such a system would be well outside the scope and expertise of the project, the main goal then shifted to simply
identifying and describing how EEG data could produce “breakpoints” for a parametric or automated interface system.

In consideration of these factors, the knowledge-needs—the questions which would address those assumptions—were as follows:

• Is the data coming from Mindflex+Arduino viable, or too questionably noisy/random?
• What defining characteristics be identified in the data (if any)?
• How might these characteristics be implemented as control parameters?

Methodology

Wikipedia’s Random Article feature was utilized to generate content long enough to necessitate three minutes of reading (estimated using wordcounter.net). Random Article allowed for quick and repeated reloading to find material which was neither overtly formal/specialized nor overtly controversial/disturbing. Pages outlining the lives and history of Renaissance or Victorian personalities were utilized often. Freetetris.org served as the game-task, beginning with the easiest level of difficulty. Three minutes was selected as an average length of time which an ordinary person could sustain a game of Tetris, based on prior informal observation.

As seen in Figure 17, the EEG headset hardware has something of an unusual form-factor. In an effort to normalize the overall strangeness of the hardware, the researcher donned the headset prior to each trial while participants were being briefed on the nature of the tasks they were to complete. For this same reason, data was collected outside of a formal lab setting, in classrooms or spaces which may still have contained other individuals
or activities, but were offered more comfortable to the participant overall.

After sufficient time had passed to allow participants to become accustomed to the headset hardware, they were instructed to begin reading and the headset recording process was initiated. At the three minute mark, participants were instructed to close the window containing Wikipedia and begin playing Tetris. Each participant played Tetris as long as they were able, and when the game ended the EEG recording was stopped. Samples were generated by the headset every second and printed to the serial monitor of the Arduino development environment. These values were then copy and pasted out of Arduino and filtered for error messages, which had occasionally been appended to the Gamma band of EEG data coming from the headset. This rendered the Gamma band of EEG data unusable, though this proved to be inconsequential (see ‘Secondary Research, Suitable Hardware’).

Each sample of EEG data returned a row containing 12 values. The 8 bands of raw EEG data—a low and high channel for Alpha, Beta, Theta and Gamma bands, respectively—proved too unwieldy to be manipulated. The extreme range within each band, the presence of inexplicably large or small values, and the chip manufacturer’s opaque descriptions of the band values made visualizing the data challenging, if not infeasible. Ultimately, only the pre-filtered values of “Attention” and “Meditation” were used for this analysis.

**Data and Visualizations**

In total 8 trials were successfully recorded. From each set of trial data, three types of visualizations were generated: line graphs, scatter plots and time-lapsed scatter plots. Figure 18 shows the first line graph visuals for each participant. These
FIG. 18: Linegraphs for each participant with reading-task on white and game task with orange highlight.

Participant 1

Participant 5

Participant 2

Participant 6

Participant 3

Participant 7

Participant 4

Participant 8
FIG. 19: Reading task values with trendlines, for each participant

Participant 1

Participant 5

Participant 2

Participant 6

Participant 3

Participant 7

Participant 4

Participant 8
FIG. 20: Game task values with trendlines, for each participant

Participant 1

Participant 2

Participant 3

Participant 4

Participant 5

Participant 6

Participant 7

Participant 8
FIG. 21: Combined task values with trendlines, for each participant

Participant 1

Participant 5

Participant 2

Participant 6

Participant 3

Participant 7

Participant 4

Participant 8
FIG. 22: Reading task scatterplots – Attention X-axis, Meditation Y-axis

Participant 1

Participant 2

Participant 3

Participant 4

Participant 5

Participant 6

Participant 7

Participant 8
FIG. 23: Game task scatterplots – Attention X-axis, Meditation Y-axis

Participant 1

Participant 2

Participant 3

Participant 4

Participant 5

Participant 6

Participant 7

Participant 8
FIG. 24: Combined task scatterplots – Attention X-axis, Meditation Y-axis. Reading task shown in cyan, game task shown in magenta.
FIG. 25: Combined task scatterplots – Attention X-axis, Meditation Y-axis. Reading task shown in cyan, game task shown in magenta. With Gaussian blur filtering.
FIG. 26: Combined task scatterplot cluster analysis

Participant 1

Participant 2

Participant 3

Participant 4

Participant 5

Participant 6

Participant 7

Participant 8
graphs shows the full domain of time—however long a participant’s trial lasted, reading-task and Tetris-task in total—along the X axis. Attention and Meditation values are plotted on the Y axis, in red and blue respectively. From the 180th sample mark onward an orange overlay was used to highlight the change from reading to playing Tetris (one sample per second, three minutes of reading). This orange overlay appears at different points and sizes in each visual due to the variable lengths of the Tetris task.

Figures 19, 20, and 21 show a second set of line graphs with linear regression trendlines. These figures show data from the reading task, tetris task, and combined tasks, respectively. The trendlines take the erratic peaks and valleys of the data and find the best-fitting overall trajectory for both Attention and Meditation values, again shown in red and blue respectively.

Figures 22, 23 and 24 follow a similar convention to show scatter plots. These plots created by inputting the CSV files from each participant into a simple program written in Processing, a visualization software application. Attention values are plotted on the X-axis, Meditation values are plotted on Y-axis. In the third column with the tasks overlayed, data points from reading Wikipedia are shown in cyan and Tetris is shown in magenta. Time is not represented. The signal value for each point was used to reduce the opacity of points which had sub-optimal signal, thus reducing their visibility, though there were relatively few of these points overall.

Figure 25 shows the combined-task scatter plots, but with Gaussian filtering and level compression applied. These steps help to reduce the significance of outlying points and generalize the shape and location of each cluster. Figure 26 shows these same plots, but with an overlay representing a primitive cluster analysis.

Finally, in order to combine the time-domain of the line graphs with the clustering
of the scatter plots a video clip was created for each participant to trace their Attention and Meditation values over time. These clips were accomplished through a variation of the Processing program used to create the static scatter plots. For this variation, every sample for each participant was plotted and saved as an individual frame. These frames were then compiled in the video editing application After Effects. Each frame was given a semi-transparent background so that previous data points would be visible as progressively fainter “traces” of the point most recently plotted. See Figure 27.

**Analysis Overview**

At the simplest level, these visualizations reinforce that the values coming from the Arduino/Mindflex headset are not random noise. This hardware may not be highly accurate, or capable of medical use in BCI-controlled prostheses. But the distribution of values coming from the headset are not uniformly random for either individuals or for tasks. That is to say, different participants values look different, and different tasks for each of those participants look different. If the values had come out with no visible patterns, clustering, or variation this would have negated the remainder of the evaluation.

Fortunately, this does not appear to be the case. There are visible trends of inclination and declination over time, and for at least some users, a clear demarcation at the point of switching tasks or clear connections between levels of Attention and Meditation values. In the scatter plots the data points are not scattered haphazardly, they are clustered. And the time-lapse clips show similarly promising trends. That being said, it is plain to see that there is no typical pattern in how participants reacted to either task. Though, this is not entirely surprising. Tetris is not universally delightful, nor
is it universally anxiety inducing. Reading Wikipedia is not universally exciting
nor is it universally boring. This should be obvious and these plots were not intended to
help describe how participants as a group would react to each task. Rather, the intent of this
evaluation was to determine what the different traits of each individual’s data set were, and
to better understand how these commonalities might be used to describe an individual’s
state of mind.

Considering this, I will review each set of visualizations in greater detail and
attempt to highlight the features which could be extracted and modeled for the purposes of
recognizing and reacting to the mental-state of a given individual.

**Line Graphs**

The first set of line graphs—Figure 18—show the point at which each participant
switched from reading to playing Tetris. Reactions to the task-switch appear to be split:
four participants seem relatively unchanged by the switch, the remaining four each
show some definitive changes in the characteristics of their plot. These are as follows:

- **P1:** An overall lowering of both Attention and Meditation values which
  persists, intensifying slightly, for the duration of the Tetris session. There is
  also an overall reduction in the range between the highs and lows of each value.
- **P4:** A dramatic loss in Attention values, sustained for approximately 30
  seconds before they resume a slight ascent with great variation. Generally,
  proportions of Attention and Meditation become inverted at the switch.
- **P5:** A decline in Attention and Meditation values of approximately 80 seconds.
- **P8:** A dramatic dip (≈50% loss) in in both Attention and Meditation values
lasting approximately 50 seconds before both values return to reading levels, albeit briefly.

This data is not robust enough to suggest any explanation why an individual would react more or less to switching tasks, yet it remains encouraging that a reaction was detected in some participants. The propensity for task switching, multi-tasking or task notification—where an alert for another task is injected into the workspace of a current task—is a prevalent negative component of digital devices (see ‘Secondary Research’). Being able to understand how each individual person reacts to various task switching scenarios would be valuable in preventing, or persuading people away from task-switching patterns.

Investigating these reactions further helps to highlight characteristics of interest nested within the basic Attention and Meditation values. First, the proportion created between Attention and Meditation. This can be described in a variety of ways, but the simplest may be as a difference, by subtracting Attention from Meditation. Beginning around sample 90 for Participant 2, Attention levels suddenly drop below Meditation levels; the difference value shifts from being positive to negative, and then back to positive around sample 130. These proportions are more easily understood through the scatter plots in the next section.

A second area of interest visible in the line graphs is the variation from one sample to the next within each value—the rate of change over time. Comparing Participant 3 and Participant 4 shows significant differences in rate of change. Participant 3’s values have relatively mild variation when compared to Participant 4’s wild spiking and plummeting. Participant 5 can be seen to exhibit both P4-like spiking, especially early
on, while later becoming more P3-like, especially around the 400 sample.

The second set of line graphs shown in Figure 19 presents the same data with linear regression analysis trend lines overlaid onto the reading session, Tetris session, and the entire session—in Columns 1, 2 and 3 respectively. Regression analysis can be helpful in understanding what is happening to values more generally. The benefits are very clear when considering Rows 4 or 5, which, as you may have surmised show Participant 4 and Participant 5’s data respectively. The trend lines of Column 3 help to underscore the inversion of Attention and Meditation levels as the session switches from reading to playing.

However, looking at Column 2 for P4 we can see that the trend line shows an overall increase in Attention levels for the playing session, rather than the overall decrease shown in the entire session. This same issue is visible for Participant 1. During the reading session, Meditation values were experiencing an overall increase. But the entire session trendline shows an overall decrease. This is an issue of fitting, where particular values or groups of values stick out from the trend created by the rest of the data. Trend algorithms attempt to find a line which passes through, or as near as possible to, as many data points as possible. The more data points the line passes through, or the more points the line is nearer to, the better the fit. So, naturally, the fit quality depends on whether it is analyzing the entire session, or just a reading session, or just a play session.

None of these analyses are wrong, they simply help to underscore the importance of time and task domains. Determining the scope of time to be analyzed for linear regression is important, and a system utilizing regression analysis might benefit from tracking the proportional-inversion and rate-of-change dimensions discussed above.
Scatter Plots and Cluster Analysis

Line graphs were helpful for visualizing characteristics such as proportional shifts and rate-of-change. In order to identify other characteristics, scatter plots were created by removing the domain of time from the X-axis so that Attention values could be plotted against Meditation, so that each sample becomes an individual point in a two dimensional space. This representation helps to better visualize the proportional shifts outlined previously.

By plotting these dimensions against one another, the ranges on each axes from 0 to 49 and 50 to 100 divide the space into four classifiable states:

- **A1M1**: Attention greater than 50, Meditation greater than 50
- **A1M0**: Attention greater than 50, Meditation less than 50
- **A0M1**: Attention less than 50, Meditation greater than 50
- **A0M0**: Attention less than 50, Meditation less than 50

*1 represents a higher likelihood—or presence—of each state. 0 represents a lower likelihood—or absence—of each state.

A1M1 correlates to a higher probability that a person is both focused and calm. Having a high degree of focus while maintaining a high degree of calm seems aligned with the idea of being “in the zone”, perhaps the feeling of playing a videogame or a musical instrument at a very high level—like Csikszentmihalyi’s “flow” state. A1M0 correlates to a higher probability that a person is focused, but a lower probability that they are calm. This might be a state like nervousness or excitement, perhaps watching a scary movie or reading a suspenseful book. A0M1 would be correlate to a lower probability that a person is focused,
but a higher probability that they are calm. This might be daydreaming, or perhaps boredom. Finally, AoMo would correlate to lower probability that a person is focused and a lower probability that they are calm. This might be a condition like anxiety or fretting.

There is no formal secondary sources to confirm or support this model, it is based on the descriptions given in the NeuroSky documentation and general human experience. However, the exact description or classification of a state may not be strictly necessary. Each individual will have different characteristics and tendencies. The descriptions and thresholds for these areas can and will change in a system designed to adapt to different individuals. This will be discussed further in Section 3.6.

Figures 22, 23 and 24 shows the original scatter plots for each participant. For each plot, the reading task is represented in cyan and the tetris task is represented in magenta. Note, these graphs do not show data points which repeatedly landed at the exact same X-Y coordinates, which may have affected the density of smaller clusters, but a review of the CSV files seems to suggest the effect was marginal. This presentation of data allows for certain characteristics to be seen more clearly than the line graph presentations, though it is interesting to recall that this is the same data, only a different presentation.

The clearest characteristic is the position of each cluster into one of the four classifications outline above. Following this is the relationship created between a participant’s task clusters. The separation of clusters, a distinct translation of position from the reading-task to the playing-task, highlights how each individual reacted to switching from a passive task to an active task.

For some participants there is a clear separation of clusters, as seen in Participant 1 and 4. Participant 1 fell predominantly into the A1M1 quadrant (top-left) for reading, but translated predominantly into the AoMo quadrant (bottom-left) for playing tetris.
During reading, Participant 4 was similarly focused, but significantly less calm, falling into the A1M0 quadrant (lower-left). During the playing task, P4 made considerable gains in Meditation levels, although Attention levels fell substantially. P4’s playing-task plot demonstrates the need for a “non-determinant” (ND) classifier, for when a cluster falls indeterminately—halfway—along either axis. In the case of P4, the playing task session might be best classified as AoND.

Other participants exhibit very little cluster separation, such as P3 and P6. These low-separation results are in part reflected by the corresponding line graphs; no clear significant event at the 180 sample mark. It seems logical that high cluster separation would suggest markedly different experiences between tasks, while low separation would suggest little to no difference in experience. However, this determination may be affected by the density of each cluster, a third characteristic.

This evaluation did not establish a quantitative approach to describing density, only qualitative visual assessments were made. Some participants values fell more frequently into distinct ranges or areas, such as P1 or P3. This creates more dense, and therefore more well-defined clustering. Others, such as P4 or P7 have much larger, less dense clusters.

Additionally, while each of these plots share an equal number of data points for reading, they have a varying number of data points for playing Tetris. This is simply because more skilled players were able to keep playing for longer. So, although more skilled players may have larger clusters for tetris, those clusters were able to remain visually dense.

The exact significance of different densities is unclear, but a few suppositions can be drawn. A narrow range for a fixed number of values—a smaller, more dense cluster—could suggest persistence of a particular state over time. A broader range of values over
the same period of time—a larger, less dense cluster—could suggest either no clear state, because the values are scattered randomly across a given period. A larger, less dense cluster might also suggest a sort of transience between distinct states, if the larger cluster is actually representing one cluster drifting over time.

Another explanation might be that different types of people have different subconscious expressions of their mental states. The relatively narrow range of values in P3—reflected in a lower rate-of-change on the line graphs—or the broad values of P4—similarly reflected in a high rate-of-change on the line graphs—could say something about personality or general disposition. One person daydreaming might look similar to another person studying fervently. However, analyzing these plots in such a way would require a consideration of the time window for tetris, as mentioned above.

The interaction and relationship of density and cluster separation—and therefore with experiential classification—is highlighted in P6. While the cluster separation appears low, with a large number of purple points indicating overlapping data from each task, a considerable number of points for each cluster can be seen to the left and the right of the overlapping area, corresponding to reading and playing tasks, respectively. Returning to P6’s line graph, this pattern is echoed in the rate-of-change, as mentioned above.

The data points from P6’s line graph falling within the approximate range of sample 100 to sample 260—spanning the 180-sample task-switching mark—show a lower rate of change (more consistently adhering to the overall trendline). While the ranges from 0–100 and 261–Finish show more dramatic rate-of-change, more dramatic deviation from the trendline, and correspondingly account for the more erratic scatter plot coordinates. This underscores the importance of time in considering the
translation of clusters from one location to the next, or a perceived lack of cluster separation. P6 had a shift in state-of-mind, it simply didn't correspond with the change in task.

Another shift can be seen in the scatter plots for P2. Two clusters can be identified within the data for P2's reading task. One in the A1Mo and the other in A0Mo. This is another shift which was not brought on by the change of task, but can be clearly seen in the corresponding line graph and was also reiterated in feedback provided by the participant after the evaluation: “I was reading but I wasn't really reading … at some point, I realized I was completely on a different planet mentally.”

While quantitative descriptions of these features and characteristics were outside the scope of this evaluation, Figure 24 is an attempt to demonstrate how one would begin to take put numbers on some of the qualitative descriptors such as “big”, “dense”, “less dense”, “erratic” etc., used up to this point.

Figure 25 shows the original scatter plot visuals with gaussian blurring and histogram compression, applied through a photoshop action. These were employed manually due to the relatively low number of applications, but there is no reason these methods could not be employed programmatically. By applying these filters, the significance of individual outlying points is reduced, while groups of points in similar locations are emphasized. From this, we can begin to visualize how some of these points can be discounted.

Following this, the most dense point within each cluster becomes a reference point against which all other points are measured. Similar to the fitting of the trend lines onto the line graphs, the objective of these measures would be to determine the smallest radius which can contain the greatest number of points. This would in effect create a perfectly circular cluster, however a better fit can be achieved by adding a
second radius, perpendicular to the first, which would provide an elliptically-fit cluster. These clusters are shown in Figure 26. Note that these are not actual mathematical expressions, but rather approximations intended to serve as examples.

What this processing allows is a more precise, numeric description of the cluster characteristics. Of course the center points of each cluster can be stored as a single set of XY coordinates and compared to others, even within the same A/M quadrant classification. The surface area of the cluster can now be quantitatively stated, after which it can be compared to the number of points fit within that cluster. This comparison would give us a number to describe the density. The angles of the radii provide axes with which to further compare tasks.

Additionally, each of these methods could be applied to the overall evaluation session. Considering the nature and use-case of most computer interaction, task switching is common, so much so that tasks may blend together. Taking all of these tasks together as a session could be understood as a meta-cluster. The data from this could be applied longitudinally for comparison of many sessions over time.

**Time-Lapse Scatter Plot Videos**

Removing the domain of time from the scatterplots was a helpful way of developing a classification system and considering the metrics which could be created by such a system. However the significance of time was underscored when seeking correlations between patterns in the the plots and line graphs. Following this, a set of time-lapse video clips were created by breaking each scatter plot point into a frame, which combined the more easily-classified visual of the scatter plot with the time offered by the line graphs.
As in the initial scatter-plots, for each sample Attention values were plotted to the X axis and Meditation values plotted to the Y axis. Each sample was saved out as an individual frame through a simple program written in Processing. These frames were then compiled in AfterEffects at a framerate of 30 frames per second in order to compress the domain of time. Compressing time in such a way—30x its normal speed—helps to highlight the periods where values clustered together. This can be considered comparable to the gaussian filtering and tonal compression applied to the static scatter-plots.

With 1 sample being taken per second and a frame rate of 30 frames per second, the 180th sample, representing the switch from reading Wikipedia to playing Tetris occurs at the 6 second mark of each video clip. Frames from a participant are shown in Figure 27, but being that this is a dynamic/time-based format, these are best viewed on a digital device.

As discussed in the previous section, determining the length of time for a cluster window plays a significant role in the analysis of experience. The window for this evaluation was initially determined to be the point of a known change in task. Each data segment before and after this task-change comprised a “cluster”. However, through the visualizations it was shown that some participants, notably P2 and P6 showed significant grouping across the task-transition. These types of clusters can be described

FIG.27: Frames from time-lapse scatterplots
as “macro-clusters”, since they occur outside the bounds of the task-domain clusters.

In addition to the macro clusters, micro-clusters can also be seen forming within a given task-domain cluster. It is my speculation that tracking these clusters quantitatively, using a finer, but equally straightforward system such as the four quadrant A/M classifier, would be a critical component in developing a more sensitive and accurate biometric input.

The simplest approach could be simply tracking the positional change using an 8-way directional classifier like a compass: Up, Up/Right, Right, Down/Right, Down, Down/Left, Left, and Up/Left. See Figure 24. By pairing this with a distance measurement—which direction AND how far—Every sample could produce a data point which describes how that individual point sits in relation to the point which preceded it. Analyzing the frequency and relationship of these micro-clusters could help predict the larger trends which they precede at the cluster/meta-cluster and macro cluster levels. This is in some ways analogous to the gridded Poincaré index fingerprint classification technique shown in Jain.

Building a tool for analyzing these other sorts of clusters fell outside the scope of this evaluation, however stepping through the frames of any of the time-lapse scatterplots can provide an analogous effect. Combining analysis of this data at such a micro-level was outside the scope of this evaluation, but is an area deserving of it’s own evaluation and study.

**Summary**

After taking sample data and reviewing the visualizations each initial assumption was addressed, if not resolved. To begin, the hardware does not appear to have produced
random noise. There was enough inter-task distinction and other reaction events identifiable that the Mindflex+Arduino hardware, and comparable single-node headsets, seem suitable to provide input for a biometric interface.

Although the line graphs and scatter plots actually depict the same data, the improved comprehensibility of the scatter plots and time-lapse plots made them the focus for defining the important characteristics. Key characteristics include:

**Cluster Position:** Describes overall condition or change in condition

**Cluster Proportions:** angle, height and width: Describes variation within a particular value or proportional variation between values.

**Cluster Density:** Describes the significance or degree of certainty.

And each of these can be observed at four scales:

**Cluster:** points from a single task taken together

**Macro-cluster:** points which group together significantly across tasks

**Micro-cluster:** points which group together significantly within tasks

**Meta-cluster:** two (or more) task clusters taken together

Movement, proportion and density could also be applied to the evaluation of the “routes”, or the linear movements that often preceded micro-clustering. That being said, sample-to-sample data alone would not serve as a valid source for an interface control parameter. Deriving meaning from a single-node headset is garnered from segments or collections of samples rather than any individual sample. Smaller segments could be
used as points of prediction for what the overall trend may be, though this would require significant further work to classify finer trends and relationships, should they exist at all.

Additionally, using this data as an input would require significant programmatic analysis on an individual basis, with some other source of user feedback to gauge and inform the analysis. Although similar, the same values do not necessarily reflect the same effects for two different individuals. A biometric input system would need to be capable of observing baselines and behaviors for each individual user and determining what negative and positive attentional and calming pressures look like. Only after this would they be suitable for use as parametric controls. This issue will be discussed further in the following section.

With these characteristics and considerations in hand, a more experimental analysis needs to be undertaken. Can the AM classification be reliably exhibited and redressed? Can a pattern or relationship be discerned in the traits of larger participant groups? Type A Type B baselines?

**Challenges, Issues and Considerations**

A central challenge with this research was the lack of experience in two of the three most relevant fields. Much time and effort was spent in developing basic understandings; it was difficult to look without knowing what to look for. Technical limitations also set frustrating thresholds. The MindFlex+Arduino headset as well as the manual visualization and analysis were sufficient only for the most primitive of processes, which this was, relatively speaking. The opaque quality of the Attention and Meditation data labels remains a point of trepidation, but more concerning was a lack of more incisive or strategic evaluation techniques for the data that was captured.
Additionally, the task-change as the sole point of differentiation—with the exception of a few pieces of unprompted user feedback—felt limiting. Some second source of data, follow-up questions, screen-captures, eye-tracking etc., may have brought more meaning and opportunities for identifying moments of significance. These techniques went largely unemployed due to a lack of familiarity with designing experiments and trials.

One of the largest areas of consideration is the amount of variation there is across participants. The role of different dispositions, prior experience, or even simply the mood a person is in, created a sort of paranoid solipsism when reviewing the data. A similar challenge faced Hernandez et al. and seems to be the nature of biometric analysis when searching for conditions rather than identities. However, based upon secondary research and primary observation, it seems the answer does not necessarily lie in finding particular patterns that suit all people, but rather in identifying individualized tendencies and EEG characteristics which correlate to other metrics or events.

The final consideration which has been raised repeatedly in responses to this research is: who is going to wear an EEG headset like this and is it really worth it? In part, the latter question helps answer the former. The benefits to be gained from engaging in a biometric interaction model may not be relevant to all people. Long-term conservation of attentional resources would most benefit individuals who are engaged in increasingly persistent and complex digital relationships; those who spend large portions of each day on a computer, who manage multiple streams of communication with a range of parties. Consider a biometric input alongside other specialized office equipment. The standing desk or lumbar-supporting chair, or an ergonomic mouse. EEG input addresses the ergonomics of the mind, rather than the body.
Reducing the mental tax of increased exposure to digital information and devices through the creation of more information and another device seems counterintuitive. Moreover, creating a numerical description of a person’s mental state to manipulate the behavior of a machine may seem preposterous, at least for everyday interactions. This is the domain of extreme brain-machine interfacing, prostheses, and medical treatment. But history has shown that changing the way we interface with computational ability improves access to those abilities afforded by our machines.

These increases in access have come largely through interfaces which allowed means of communication that increasingly favor the idiosyncrasies of humans over the rules of machines. We have repeatedly looked past existing methods and approaches, abstracted the goals and developed new means. Programming languages, manipulation of visual icons, voice command—this is a trajectory which places greater expectations on machines to translate the actions of humans into forms more native to digital logic. This process has lead us to an unprecedented level of device-saturation.

As we exit the PC era, from what Mark Weiser phrased as ‘one device, one person’, and head toward higher device-saturation, the objectives of human-computer interaction are likely to expand. In a ‘many devices, one person’ model, providing people with improved—more, or easier—access is no longer the sole concern. More effective manipulation of a device and task-completion has been the objective of interfaces for at least the past 60 years. This will not cease to be of concern, but going forward future
interfaces should be capable of addressing the manipulation of a system of devices for the effective completion of many tasks.

This level of change will not be achieved through the implementation of best practices or refinements in current treatments. The very notion of requiring on-going explicit instruction or command to a machine is a fundamental barrier to creating a “calm” or affective experience. This is the realm of implicit, or unconscious communication and the role of biometrics.

This research investigated the use of single-node EEG as a potential source of implicit, affective input. An evaluation of 8 participants in single-task and multi-task processing scenarios suggested that consumer-grade EEG hardware could provide a coarse description of individual reactions to varying attentional and meditative pressures.

The central caveat to this is the need for prolonged use of such hardware coupled with sophisticated processing in order to extract meaning on an individualized basis. Much work has already begun in the automated classification of emotions and mental states, not only through EEG but other biometrics as well, including: skin conductivity, facial analysis, natural language processing, voice analysis and more. The development of biometrics as generalizable input for an interface and interaction model, along with the widespread implementation of such a system, will not be the achievement of any one metric, biological neurological or behavioral.

Reflecting on “A Dao of Web Design”, biometric interfaces are challenged by a need for balance. The certainty and stability with adaptability and fluidity. Understanding the significant characteristics and trends exhibited universally across participants, as well as the ability to identify those which manifest in smaller groups or even at the individual
level, will be essential to creating an interface system with successful or meaningful affect. Continuation of research in this area would need to address both of these notions. A more controlled experimental trial which tests the reliable exhibition and redress of cognitive strain —loss or persistently low values for Attention and Meditation—across participants is critical. Though, this may not necessarily be achieved prior to a system for automated analysis and response. Indeed, due to the high potential for variation across participants, and the amount of processing required to identify significant characteristics in the data, the two may need to be developed in tandem. This follows the notion of interactivity laid out by Dubberly et al. in contrast to reactivity; the development of a system which can change over time, rather than simply react in a fixed, predetermined way.

Beyond the direct and literal application of this research, an underlying objective was to find new approaches and means of thinking. Refining the application of methods is a route toward improvement. But as Whitney suggested, moving forward toward a better version of the same can mean sacrificing a richer perspective and ultimately more robust outcome. Understanding the context—the history of a discipline, as well as the technical underpinnings and inner-workings of it—has provided a more diverse understanding which can be applied not only to the development of some future creation, but to every sort of interface or interaction to be devised and tended in the interim.
References


Berlin, Heidelberg


from: https://www.nngroup.com/articles/mental-models/


