Title
The Complexity of Engagement

Author
Meyer, Timothy

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

The Complexity of Engagement: Measuring Engagement from a Complex Embodied Perspective

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Cognitive and Information Sciences

by

Timothy Scott Meyer

Committee in charge:

Professor Jeffrey Yoshimi

Professor Arnold Kim

Professor Michael Spivey

The dissertation of Timothy Scott Meyer is approved, and it is acceptable in quality and form for publication electronically:

Professor Jeffrey Yoshimi, Ph. D.

Cognitive and Information Sciences, UC Merced

Professor Arnold Kim, Ph. D.

Applied Mathematics, UC Merced

Professor Michael Spivey, Ph. D.

Cognitive and Information Sciences, UC Merced

Dedication

To my wife and children,

without them distracting me I would have gone insane

long before I finished.

Table of Contents

List of Tables and Figures

Acknowledgments

I am deeply indebted to my advisor and friend Jeff Yoshimi. Jeff has been a constant source of guidance and inspiration without whom I am sure I would not have been able to develop this work, though he would never accept that fact.

I also could never have finished this without Arnold Kim who not only taught me the analytical and mathematical skills I needed to complete this research but also provided a steady stream of irreverence that kept me from feeling overwhelmed.

And I am extremely grateful to Michael Spivey who not only helped me develop my experimental skills, knowledge of cognitive science, and laboratory work, but also and perhaps more importantly was always a voice for optimism in my research, even when I couldn't see the forest of good through the trees of hiccups.

I also want to thank the whole Cog Sci department for their intellectual contributions to this work. I am particularly grateful for my amazing lab mates Andrew, Liza, Ben, Brandon, Ayme, Benny, and Tevin who helped me vet and develop my ideas.

And finally, I have to thank my cousin Nate for his constant help and feedback on my ideas. I couldn't have done it without you.

Curriculum Vitae

Education

2022 (Expected) **Ph. D. Cognitive and Information Sciences**, University of California Merced

2016 **Bachelor of Arts, California State University, San Bernardino**
Major: Psychology
Minor: Computer Science

Research Experience

2016-2022 ***Graduate Student Researcher***
Department of Cognitive Science University of California, Merced
I designed and ran cutting edge experiments investigating engagement and developed a new paradigm for analyzing mouse tracking data. I discovered the first domain neutral behavioral measure of engagement, created online applets with restful APIs and built-in data pipelines for analyzing user engagement, and used advanced statistical and analytical techniques to make discoveries for publication in top scientific journals.

2015 ***Honors Thesis, "Alignability and Distinctive Features in Free Categorization"***
Supervisor: Dr. John Clapper
I worked closely with Dr. Clapper to create an experiment examining human categorization. My responsibilities included creating the research instrument for respondents, coding the data, assisting in analyzing the data, and reporting and disseminating the findings.

2014-2016 ***Volunteer Research Assistant***
Department of Psychology, California State University, San Bernardino
I worked with Dr. John Clapper on his experiments on free categorization. Working independently, I helped to code and analyze his data, and helped write articles about the experiments.

2015-2016 ***Volunteer Research Assistant***
Department of Psychology, California State University, San Bernardino
I worked with Dr. Jason Reimer to implement his cognitive training experiments. Working with a team of assistants, I helped recruit participants, manage participants' schedules, and administer cognitive tests and cognitive training.

Teaching Experience

2016-2022 **Instructor/Teaching Assistant**

I taught 11 classes across multiple disciplines, including Cognitive Science, Neural Networks, Philosophy, and Political Science, and I always received above average evaluations for my communication and interpersonal skills. I also contributed to curricular design and lecture instruction of the artificial neural networks course.

2014-2016 ***Writing Tutor***

In a writing lab that specializes in social sciences research writing, I worked one on one with student writers on all aspects of the writing process, from prewriting and drafting to final proofreading. I paid special attention to helping students understand the logic underlying research writing.

2014 ***Teaching Assistant***

Psyc 364: Perception (2 classes). I evaluated students' written and multiple-choice exams.

2009-2012 ***Coach***

I coached for American Youth Soccer Organization, for girls in the under -14 age bracket.

Research Projects

Analysis of Complex Mouse Tracking Data

- Develop an online videogame with a **data collection and analysis pipeline** (used **javascript**, **node,js**, **php**, **css**, and **python**)
- Developed a new paradigm for analyzing cursor data that is non-segmented (combination of **SVD**, **Detrended Fluctuation Analysis** and **Spectral Analysis**)
- New method incorporates over **10000** times more information than standard approaches
- Showed multiscale structure predicts performance during gameplay (**regression**)
- Paper submitted to Behavioral Research Methods

Complexity Matching and Engagement [in preparation for publication]

- Designed observational study which investigated connection between user behavior and user attitudes
- Developed novel interpretable User Analysis techniques for measuring psychological constructs in continuous streams of mouse movements, **solving a long-standing problem** in mouse tracking research
- Found that participant's **psychological engagement** during gameplay was predicted by their multiscale coordination (**Complexity Matching)** using **Multilevel** and **Structural Equation Modeling**
- Developed the first behavioral measure of engagement that is domain neutral: applies to sports, games, websites, classrooms, work, etc.
- Enabled measurement of engagement without user feedback or laboratory devices

saves researchers **thousands of dollars**

Engagement and Mouse Tracking [in preparation for publication]

- **Created experiment** for investigating multiscale coordination (**node.js**, **javascript**)
- Found that humans adjust their movement dynamics at multiple scales to match a stimulus (**Complexity Matching**, **non-parametric Statistics**)
- Showed more engaged participants coordinated more with the stimulus (**regression**)
- Developed method to account for context specificity when measuring **user engagement**
- Advanced scientific understanding of **user engagement** by demonstrating that it is inherently a complex embodied phenomenon

New Theory of Engagement [in preparation for publication]

- Developed a **new theory of engagement** that connects the phenomenological, psychological, and complex embodied cognitive systems components of engagement
- Used our theory to create **novel empirical predictions** which were **supported**
- Precisely identified the roles played by the motivational affordances of an activity and an individual's cares in creating psychological engagement
- Explained the connection between engagement and body dynamics

Alignability and Distinctive Features in Free Categorization [incorporated into larger publication]

- Collected data from participants performing pencil and paper categorization task
- Observed participants and collected feedback from participants

Analysis of Cognitive Training Games

- Ran groups of participants through cognitive training regimens
- Observed participants and collected feedback
- Contributed ideas for future cognitive training tasks

Analysis of Dialysis Clinics

- Scraped clinic care rating and income of population served for dialysis clinics in US
- Showed income is important in converting patients to AV Fistula access (**R markdown**)
- Improves evaluation of dialysis clinics

Analysis of Eye Tracking Dispersion

- Developed measure for creative thinking based on eye movements

Fantasy Football Analysis

- **Regression**, **Structural Equation Modeling**, and **Neural Network** of Fantasy Football Data

---

Publications and Presentations

In prep **Phenomenology of Engagement**.
Tim Meyer and Jeff Yoshimi

In prep **Discovering Engagement**.
Tim Meyer, Arnold Kim, Michael Spivey, Jeff Yoshimi

In prep **Engagement as Spatiotemporal Coordination**.

	Tim Meyer, Arnold Kim, Michael Spivey, Jeff Yoshimi
2021	**Mouse Tracking Performance: A New Approach to Analyzing Continuous Mouse Tracking Data**. Tim Meyer, Arnold Kim, Michael Spivey, Jeff Yoshimi.
2016	**Alignability and Distinctive Features in Free Categorization** (lead author: Dr. John Clapper)
2015	Presented poster, "**Alignability and Distinctive Features in Free Categorization**" at Psychonomics, Chicago
2015	Presented poster, "**Alignability and Distinctive Features in Free Categorization**," at Western Psychological Association Conference, Las Vegas
2015	Presented poster, "**Alignability and Distinctive Features in Free Categorization**" Meeting of the Minds, CSUSB

Skills

Experimental Research

- Experienced in all aspects of the research process including designing experiments, managing participants, handling data, and conducting literature reviews.

Statistics

- Trained in advanced data analysis and statistical methods including linear and nonlinear modeling, Bayesian Statistics, and latent variable modeling
- Analyzed research data using Python, R, and Matlab

Programming

- Developed new Detrended Fluctuation Analysis library in python
- Developed library of analytic techniques for fantasy football website
- Created online experimental platforms with restful APIs and built-in data pipelines
- Created experiments in E-Prime and open sesame for different research labs
- Proficient in Python, R, Matlab, C++, Java, HTML, node.js, SQL, and PHP

Communication

- Trained in research and technical writing
- comfortable explaining technical aspects of research to non-technical collaborators
- Editor for the *Psychology Student Research Journal*

Awards

2022	Dissertation Fellowship
2015	Departmental Honors
2013-2015	Dean's List (every quarter at CSUSB)
2013-2015	Dean's Letter (fall 2013, spring 2014, winter 2015, spring 2015)

Campus Involvement

Co-Chair of Cognitive Science Graduate Student Group

Psychology Student Research Journal
Psychology Club

Other Experience

2010-2013	Dock worker, FedEx, San Bernardino
2009-2010	Cashier and cook, Burger King, San Bernardino
2011	Semi-pro soccer player, CBSG, Riverside
2010-2015	Community organizer for soccer teams and leagues, San Bernardino

Abstract

Research on engagement has exploded in recent years. It is of interest in almost any domain of human interaction from workplaces to schools, to business, to games, and many more besides. However, measuring engagement is difficult. For the most part researchers must ask participants about their engagement (which interrupts the engagement) or bring them into the lab. Here we develop a domain-neutral approach to engagement, drawing on philosophical phenomenology and complex embodied approaches to cognition. We also develop a new paradigm for analyzing mouse traces from this perspective. A phenomenological theory of engagement is presented, and two experiments are described, both of which provide evidence for that theory.

1. Introduction

When I began this research project my initial plan was to look at how different mechanisms of motivation affect a person’s engagement. I was going to choose a quick and easy measure of engagement and then use that to examine how the satisfaction of motivational affordances affects our engagement with different tasks. Instead I spent the majority of my time as a graduate student working on step one of that process, finding a “quick” and “easy” measure of engagement. Engagement, it turns out, is a complicated construct and there were no quick and easy measures of engagement because measuring engagement is quite hard to do.

The problem begins prior to even measuring engagement but with defining it. Though many of us feel that we have a good colloquial understanding of what it means to be engaged, engagement is actually a difficult concept to rigorously define. As such, many researchers operationalize engagement in some task specific way, such as the number of clicks on a link (Zhou, Calder, Malthouse, & Hessary, 2021), or the number of hours spent working beyond the minimum required amount of some project (Schaufeli & Bakker, 2010 p. 14). Though these definitions are fine in context, they generally don’t get us closer to what engagement as a construct is, and how we might be able to measure it in a more domain general way. However, I believe that the reason engagement is easy for us to understand colloquially is that we have some shared phenomenological sense of what it is to be engaged. Therefore, it seems the correct method would be to start with rigorous phenomenological work clarifying how engagement is experienced, and from there seek ways to operationally define and empirically study it.

I begin my phenomenological work, as many do, standing on Heidegger's shoulders, and perhaps even stepping on his face as I use his phenomenology (in ways he might not agree with) to try to capture the phenomenological essence of engagement. In particular we look to Dreyfus’s interpretation of Heidegger, which is known as Dreydeggarean phenomenology[1], and their description of skilled coping. Dreyfus goes to great lengths to explain how what a person cares about grounds their everyday experience. That everyday experience for a human is generally one of using things to get things done, like brushing our teeth with a toothbrush, or getting into and out of a car. The entities we interact with in an everyday way are simply meaningful to us and we just use them to perform activities that are important to us. Dreyfus describes this as a kind of absorption. When we are assembling furniture from Ikea, as long as things are going well, we do not see the hex key as an object with properties, we just us it and are not reflective about it.

When we as people are engaged with activities we aren't thinking about other aspects of our life, we are just participating in the activity and in some sense lose ourselves in the activity. I see this experience as being at the core of engagement, and I think engagement is being absorbed in an activity in this way. This connection between absorption and engagement is quite common. Many theories of engagement or other similar psychological constructs like flow theory, emphasize the feeling of absorption (Bakker, Schaufeli, Leiter, & Taris, 2008; Schaufeli, 2013; Balwant, (2018); Sheldon, Prentice, & Halusic, 2015; Nakamura & Csikszentmihalyi, 2014).

As Heidegger's work is closely linked to embodied cognitive science, one can take phenomenological understandings of engagement and use it to generate empirical hypotheses based on analogues to embodied cognitive science. The premise of embodied cognitive science is that the interactions between the brain, body, and environment give rise to cognitive functions, and embodied cognitive scientists try to study how those, seemingly separate systems, are connected. One way of thinking about their connection is as coordination, where one or many systems adjust their dynamics to maximize the information exchange between systems, and in

[1] There is no consensus on how to spell it, I have chosen Dreydeggar.

essence unify the two systems into a larger system. Horseback riding stands out to me as a very intuitive example of this where both the horse and the rider adjust their dynamics to become a more unified system. Esteemed trainer Carl Nafzger's even went so far as to describe his training method as "becom[ing] one with the horse" (Pedulla, 2020). In this example, the more aligned the rider and horse are the better they will be able to react, predict, and coordinate together to achieve higher speeds. In the case of engagement, I think if we were to hold skill equal, a more engaged rider will be trying harder to coordinate with their horse, trying harder to adjust their dynamics to match that of the horse. We can then think of engagement as information exchange via some form of coordination between a person and the dynamics of the activity they are "engaging" in.

Thinking of engagement as spatiotemporal coordination is powerful because it gives us a way to think about engagement in a domain general way, but still take into account task specificity. In any different task the particular dynamics required may be different, but as long as we can determine what those are for that task, we can determine the extent to which a person is adjusting their dynamics appropriately. We then can potentially study this in any activity that we can determine the appropriate dynamics required.

To study this I chose computer tasks and mouse tracking because they are widely available, and computer tasks that allow us to control the dynamics of the task the person is engaging with. In the horseback riding example, if we could isolate or control the precise dynamics of the horse, as if it were some giant programmable robot, and have people ride it, then monitoring the humans adapting to those dynamics to better ride the horse could be illuminating. Computer tasks allow us to do some of that, in that we can control the environment to a high degree of precision, with minimal overhead. Mouse tracking studies also have the added benefit that we can measure the movements of the cursor with relatively low latency allowing us to have good insight into the dynamics of the person coordinating with the task. If engagement really is related to spatiotemporal coordination, monitoring the adjustments in the dynamics that human's make as they engage or disengage with a system could reveal important insights about engagement.

Our first study simply involved us determining if mouse tracking was an appropriate paradigm for measuring spatiotemporal coordination. Most mouse tracking studies study descriptives of individual movements, like the curvature of the movement or the speed of the movement. Spatiotemporal coordination unfolds at timescales that include multiple movements rendering descriptives like the “amount of curvature” moot as multiple movements combined can traverse the entirety of a page with much more complicated spatial structure than a line or curve. To analyze this, we were forced to develop new tools for mouse tracking that were sensitive to movement patterns that unfold across multiple time scales. We then used these techniques to analyze performance during a simple task, and showed that performance can be described by spatiotemporal structure, and as a corollary, that spatiotemporal structure is measurable, and possibly important in describing mouse movement during activities that take longer than a few hundred milliseconds or one mouse movement.

In a follow up study, we explored the use of these tools on detecting engagement. We operationalized engagement as improvement because engagement is thought to mediate the relationship between practice and performance, and as everyone in the task is getting the same practice (playing the game), engagement should correspond to improvement in performance. We found that engagement was related to how well participants approximated the spatiotemporal structure of the best players, even though they did not have the accuracy of the best players. This is interesting because it means that engaged players in a sense play well, even if they don't have the same results as better players. This also fits with most people’s natural intuition about engagement, that even when not performing well, those who are engaged are behaving in accordance with the activity. They are “on task”. When watching children play soccer, I have an intuition that I can tell which kids are engaged with the game. The players who are chasing the ball or chasing the other players or producing other soccer relevant activity seem to be engaged.

Whereas the players who are picking flowers (something I was notorious for on my first soccer team), or disinterestedly kicking dirt seem to be less engaged.

Finally, in an effort to show that participants are spatiotemporally coordinating with an entity, we conducted an experiment where we had participants track a stimulus on screen with their mouse. The on screen stimulus moved according to a 2d representation of a John Coltrane song which gave the creature multiscale dynamics. We found that participants approximated those multiscale dynamics, and that as participants improved, again indicating engagement, their multiscale coordination increased. This further compounds the evidence that humans adjust their dynamics to match the activity, and that those who are more engaged with the activity make larger adjustments.

The concept that engagement should manifest as spatiotemporal coordination comes directly from the phenomenological theory. It is potentially the first domain general behavioral measure of engagement in that the measures are not task specific like clicks on a link but can apply across task domains. It also has implications, in particular, that: engagement is the product of the matchedness between an activity and our care structure. This is an important concept for those in the multimillion-dollar industry of facilitating engagement. If engagement is a product of matchedness, changing either the care structure or the motivational affordances of the activity alone without reference to the other is driving blind. Work on facilitating engagement needs to focus on aligning the cares of those who are to be engaged and the motivational affordances offered by the activity. It is not enough to simply add points and badges (Chou, 2019).

By the end of this dissertation our work in phenomenology, complex systems, and mouse tracking will converge to reveal important insights about the essence and physical manifestation of engagement. In Chapter 2 I cover relevant theoretical background on engagement, complex systems, and phenomenology. In Chapter 3 I cover our first study which was an exploratory study in using mouse tracking to study phenomena that unfold at timescales longer than individual mouse movements. We found that player performance could be captured with our methods which shows the importance of global structure in user interface data and demonstrates the viability of our methods. In Chapter 4 I develop a detailed phenomenological account of engagement and how that account can be expressed in terms of complex systems theory and embodied cognitive science. In Chapter 5 I cover a second study in mouse tracking where I use that same type of mouse tracking data and the same methods of analysis to investigate improvement in performance over the course of a few minutes. This is of particular interest because improvement of performance is thought to be closely related to engagement, and thus I was inching closer to measuring engagement. We found that improvement and self-reported engagement were related to complexity matching which supports the phenomenological theory we developed and emphasizes the value of our mouse tracking techniques. In Chapter 6 I discuss our final study where we combine the methods learned from the previous mouse tracking research to the theory of engagement to generate empirical hypotheses about engagement, which I then test. We found evidence that engagement is indeed related to spatiotemporal coordination, which supports our theory and has broader implications for understanding humans as complex embodied systems.

2. Background

This chapter covers the theoretical background of the three main research areas of this topic: engagement, phenomenology, and complex systems in cognitive science. The goal is to give readers enough familiarity with these topics that they understand later chapters where they are combined. The primary focus of the dissertation will be on engagement so I will begin by covering the current state of research on engagement. I will then cover phenomenology and complex systems research, but only the limited slice of those two fields that is relevant for engagement. At the end of this chapter, I will briefly discuss how the questions and ideas from these three main theoretical areas are integrated into my empirical and theoretical approach to thinking about engagement.

Engagement

Engagement has become a prized commodity in recent years, with the number of articles published each year growing from about 35,000 in 2000 to about 330,000 in 2021 (see Fig 1). Most organizations benefit from increased engagement on the part of the humans who compose them. More engaged workers are more productive; more engaged students retain material at higher rates and are higher performing; more engaged citizens participate in civic systems more often, etc. (Gaventa, & Barrett, 2012; Reeve, Cheon, & Jang, 2019; Bakker, Albrecht, & Leiter, 2011), because engagement is postulated to be a key mediator between practice and performance (Dotterer & Lowe 2011; Schneider, Yost, Kropp, Kind, & Lam 2018).

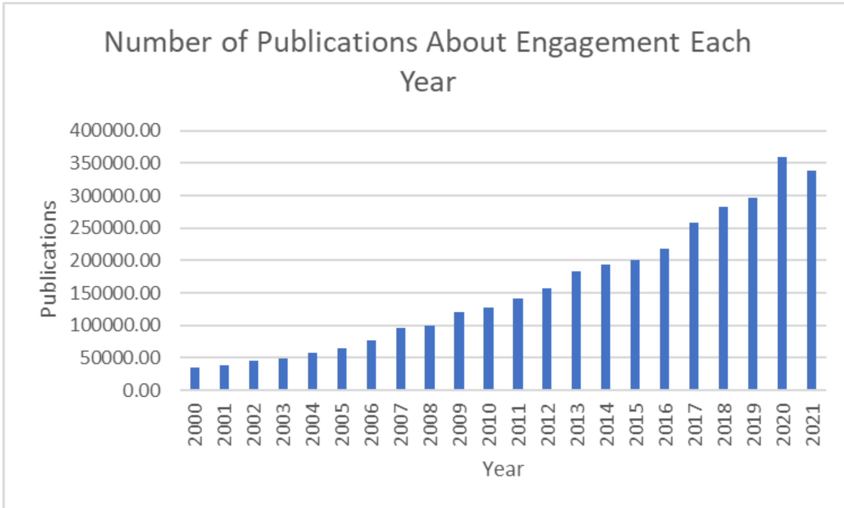

Figure .1 Graph of publications about engagement over the last 20 years. Data for the graph was taken from app.dimensions.ai/analytics/publication/overview/timeline using engagement as the search term.

Given the benefits of high engagement it is no surprise that methods to facilitate engagement have become research areas of significant interest. Education and industry in particular have developed large literatures about increasing engagement of students and employees. The next subsections will review the theoretical literature about engagement, focusing on classroom and workplace engagement as they are some of the largest drivers of interest in engagement research, and I will review how engagement is measured. This review will end by highlighting some relevant gaps in the literature and briefly discuss how this dissertation contributes to filling those gaps using phenomenology and complex systems science.

Classroom engagement

It has been postulated for at least three decades that students learn better when engaged with the learning materials (Collaço, 2017). This is substantiated by studies such as Webber et al. (2013) who found that students who were more engaged spent more time and effort in “curricular and cocurricular” activities and achieved higher GPAs. At least some aspects of engagement, such as time spent on task, have well documented connections to positive outcomes for students (Trowler, 2010; Kuh et al., 2008). Student engagement in educationally purposeful activities is positively related to academic outcomes as represented by first-year student grades and by persistence between first and second year of college (Kuh et al., 2008). In schools engagement is seen as so valuable as to have become a “proxy” for the quality of educational material (Axelson & Flick, 2010; Kuh, 2009a; Lawson & Lawson, 2013; McCormick, Kinzie, & Gonyea, 2013; Vuori, 2014),” a metric for university performance”, a goal for university “competitive advantage”, and a “justification for new teaching techniques” (Baron & Corbin, 2012; Kuh 2009a, p. 313; Vuori, 2014). Some even go so far as to call it the “holy grail” of education research (Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015)

The connection between engagement and student outcomes has led pedagogical theorists to look at ways that they can increase student engagement to enhance student outcomes and has developed a large area of interest and industry around it; for a detailed review see (Trowler, 2010). Though there is much debate around what student engagement is, it is conventionally broken down along three dimensions: behavioral, emotional, and psychological (Fredricks, Blumenfeld, & Paris, 2004). Behavioral engagement is a student’s compliance with behavioral norms such as not being disruptive. Emotional engagement is a student’s emotional reactivity to and investment in the material. Cognitive engagement is the student’s mental participation in their learning. In addition, Trowler (2010) describes each of these dimensions as going from 1 to -1 implying that there are ways to be positively engaged, negatively engaged, and not engaged. For example, with behavioral engagement a student could be positively engaged by politely attending and actively participating while following the interaction guidelines, negatively engaged by deliberately attempting to disrupt the instruction or otherwise derail it, or not engaged by simply not paying attention at all (e.g., being on their phone in class). These three dimensions are also at least somewhat independent as students are believed to be able to have different values for each dimension, i.e., a student could have a positive value for behavioral engagement, but not be emotionally invested in the material and thus have a zero value for emotional engagement.

The methods to increase student engagement involve trying to increase engagement along one or more of those dimensions. Common examples of these methods to increase student engagement range from helping students to rationalize why the material is valuable for cognitive engagement (Jang 2008; Trowler 2010), to interspersing questions for the students in a lecture for a more behavioral account of engagement (Campbell and Mayer 2009), to structuring the information delivery methods to be more enjoyable and more play like for emotional engagement (van Roy, & Zaman 2019).

Workplace engagement

In industry, employee engagement has been a "'hot' topic for about two decades" (Kim, & LePine, 2019); with engagement being viewed as boosting performance and increasing pro-organizational behavior (Harter, Schmidt, Killhan, & Agrawal 2009; Kruse, 2012; Reilly, 2014; as cited in Kim & LePine, 2019). Kahn (1990) first proposed a relationship between psychological engagement and work performance, theorizing that engaged workers would be better performers. According to Alfes, Shantz, Truss & Soane (2013), previous research shows that job engagement promotes organizational citizenship behaviors from employees and employee retention (e.g. Schaufeli and Bakker 2004; Rich et al. 2010). Moreover, job engagement is related to "individual morale, task performance, extra-role performance and organizational performance" (Bailey, Madden, Alfes, & Fletcher, 2017), and predicts "job satisfaction, organizational commitment, organizational citizenship behavior and intentions to quit" (Saks, 2019). Thus, the goals of work engagement research are to better understand the factors that drive work engagement, and to better understand how to positively influence work engagement so that the productive benefits of engaged employees may be attained.

Workplace engagement is often defined as "a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption'" (Schaufeli et al., 2002, p. 74, quoted in Bakker & Albrecht 2018). This "state of mind" is associated with increased productivity, performance, creativity, and pro-organizational behavior (Bakker et al., 2014, as cited in Bakker & Albrecht, 2018). Today, much of the research in work engagement centers around job related resources and how their access enables or hinders employees, with two theories “predominating” the field (Cooper-Thomas, Xu, & Saks, (2018): Job Demands-Resources (JD-R, Bakker & Demerouti, 2017), and Conservation of Resources (COR; Hobfoll, 1989, 2011).

JD-R theory proposes that work engagement is a balance of job demands and job and personal resources, and that employees will be most engaged when in high challenge situations where the appropriate job and personal resources are available. COR proposes that work engagement is determined by an employee's ability to access the collective resources of the company. Bakker and Albrecht (2018) also point out that a current and important trend in the field of work engagement is the recognition of engagement as something that “may fluctuate within persons – across time and situations” (Bakker, 2014; Sonnentag et al., 2010; Bakker & Albrecht, 2018). Xanthopoulou et al. (2009) and Petrou et al. (2012) found that work engagement fluctuates daily as access to resources and job crafting fluctuate.

Measuring Engagement

The most common way to measure engagement is using questionnaires. One of the most common is the User Experience Scale (UES) (Wiebe, Lamb, Hardy, & Sharek, (2014). The UES captures a diverse range of factors relating to engagement including: “the construct of flow, . . . the hedonic aspects of aesthetics, fun/pleasure, and novelty along with pragmatic aspects of usability and a more reactive sense of whether the user would like to re-engage with this experience again in the future" (Wiebe et al., 2014). In addition, the UES is often modified for use in specific situations such as the UESz used in Wiebe et al., 2014), which was modified to be more applicable to computer games.

The next most common way to measure engagement is with physiological and neural data. Many of these studies actually use the flow state from flow theory as a proxy for engagement and thus we included neural and physiological evidence of flow as well. One of the main methods of testing for flow is called difficulty manipulated tests where the difficulty of a task is deliberately manipulated to create distinct conditions of boredom, flow, and overburdened. Participants' brain and bodily states in each condition are then measured and compared.

Ulrich, Keller, Hoenig, Waller, and Grön, (2014) recorded fMRI data of participants performing a difficulty manipulated math task. The researchers found increased activation of the putamen, not the dopamine system. “Putamen is particularly involved in the guidance of goal-directed behavior by coding outcome probability in relation to effort” (Hori et al., 2009; Ulrich et al., 2014). This is particularly interesting because though the dopamine system is often thought of as the reward pathway, it seems to be specifically for extrinsic rewards. The flow theory claims that flow is a type of intrinsic motivation which is supported by this neurological evidence. “From that perspective, flow experiences, although often described as rewarding, cannot be reduced to the mere activity of the human dopaminergic reward system.”

Some other important aspects of flow theory are an increased sense of control and loss of self-awareness. Ulrich et al. (2014) found an “increase of neural activation in the left inferior frontal gyrus” which they claim, “may align with … sense of control.” They also found decreased activation of the “medial prefrontal cortex during the flow condition”. The medial prefrontal cortex is thought to be part of the brain's “default-mode network” (DMN) “which usually shows greater activity during passive states (‘at rest’) than during goal-directed behaviors” (Ulrich et al., 2014). The DMN is commonly thought to be involved in self-referential processing. A reduction in activation of the DMN and thus a reduction in self-referential processing lies remarkably in line with one of the most interesting features of flow, the loss of awareness of oneself. This will come up again when we discuss the phenomenology of absorption.

Katahira et al. (2018) used the same math task as Ulrich et al. (2014) but used EEG instead of FMRI to record brain states to get a more temporally accurate picture of flow. Katahira et al. (2018) found some indication of increased frontal theta which is associated with cognitive control. The authors also suggest that the high frontal theta in the flow condition was indicative of immersion in the task. Finally, the authors also found increasing levels of alpha in the right central area which they claim is indicative of a strain on working memory. The alpha levels were highest during the overload condition and lowest during the boredom condition. Taken together these results imply that a high level of frontal theta and a moderate level of right central alpha are characteristic of the brain when a participant is in the flow state.

Klasen, Weber, Kircher, Mathiak, and Mathiak (2012) performed a more embodied study of flow on participants playing video games in Virtual Reality. These authors found high activation of the motor system "as a neural correlate of internally simulated physical motor activity". The authors claim that "this is a reflection of the participant's feeling of involvement in the game and identification with the first person virtual character." The authors build on this claim, and additionally postulate an influence of flow "on a complex network of sensorimotor, cognitive and emotional brain circuits" implying that the sensory motor system is intimately connected with flow. This belief that flow is an embodied phenomenon is echoed by Rock and Tang (2009) who postulate that engagement, using the threat and reward systems of the brain, "requires coordination and balance between brain and body systems."

Knierim, Rissler, Hariharan, Nadj, & Weinhardt (2019) also used the same task as Ulrich et al. (2014) but instead of recording brain states using FMRI they recorded the peripheral nervous system building on the belief that flow is an embodied phenomenon. Knierim et al. (2019), used ECG to measure heart rate of participants and were able to replicate the results of Ulrich et al. (2014). By looking at the PNS instead of the CNS during a purely mental task (adding numbers) the author's results imply an embodied experience of flow, even in a mental task.

Further emphasizing this embodied perspective, Nacke and Lindley (2008) found additional biometric measures of flow. Nacke and Lindley (2008) collected EMG data on facial muscles and galvanic skin response data while participants played predesigned selections of a videogame (Halflife 2) and found that both were related to the engagement experience of the participants. In addition, Darnell and Krieg (2019) found that students' heart rates during lectures were associated with their reported engagement during lecture, a strongly mental task.

Given the embodied evidence in the physiological data, one might expect strong behavioral measures of engagement as well. Indeed, it might even seem intuitive to many of us that we can tell behaviorally if another person is engaged. However, behavioral data is one of the least common domains for detecting engagement. One area of research that attempts to find behavioral measures of engagement is affective computing research which often targets engagement using cursor, keyboard, and webcam data. In Chanel, Rebetez, Bétrancourt, and Pun (2008) a support vector machine was used to classify boredom, anxiety, and engagement of participants playing tetris. Bixler and D'Mello, (2013) used nine different classifiers "for supervised classification of engagement, neutral, and boredom states during typing”. Arapakis and Leiva (2016) used a random forest model to classify engagement while web browsing. Kołakowska (2013) conducted a meta analysis of affective computing research using (mostly) machine learning to classify engagement and other emotions in cursor and key mapping data. This paradigm is outlined in detail in Cepeda, et al. (2018). It is notable that all of these measures are machine learning methods that don't allow for interpretation and thus cannot be developed into a more general behavioral measure.

It is noticeable then that engagement research lacks a more widely applicable behavioral measure or even a behavioral understanding of engagement, despite the fact that engagement is thought to be an embodied phenomenon. The affective computing measures of engagement are highly context specific and often ambiguous. The common laboratory physiological measures are constrained for use in a lab because they require specialized equipment and monitoring systems.

And, the questionnaires necessarily remove a participant from the state of engagement to fill out the questionnaire. There is an obvious gap in the methodology, where a more generic behavioral measure that could be repeatedly applied in broad contexts would be useful.

Final Thoughts

Part of the difficulty in developing a domain general behavioral measure of engagement is that the definitions of engagement are not domain general. Engagement is a difficult topic to define and thus many studies end up using highly context specific operational definitions (Fincham et al., 2019). The term engagement has become so over burdened with definitions as to cause alarm from some of the researchers who work in the field. In the editorial introduction of a special issue for engagement, Azevedo (2015) wrote “In general, more attention needs to be devoted to the conceptual, theoretical, methodological, measurement, and analytical issues related to engagement; otherwise, the term will lose its value and be replaced with some other construct”. This definitional ambiguity carries over into the measurement context and thus many measures are domain specific, especially the behavioral ones mentioned earlier. However, I think there is core similarity to the concept of engagement as it is used across domains. For example, even though workplace engagement research tends to focus on employee access to resources, while student engagement research tends to focus on making material more relatable or interesting to the students, both domains think of engagement at least in some sense as an absorption into the activity, and they think of engagement as being related to people’s own desires. In later chapters, I will use Heideggerian phenomenology to build an account of engagement that I believe aligns with these domain general aspects of engagement and provides a behavioral model of engagement that can be tested experimentally using complex systems methods.

Phenomenology

In this section we will explain the phenomenology of Heidegger and Dreyfus. Heidegger is notoriously difficult to read. He writes in a strange vernacular sometimes referred to as "Heideggerese" (Braver, 2014). His reason for introducing this language is in part to avoid what he thought were errors associated with the history of philosophy (Dreyfus 1991; Wrathall, 2005). In addition, Heidegger uses familiar words in unusual but specific ways that can be confusing at first. We will do our best here to unpack how he is using these words to characterize certain details of everyday experience.

Dreyfus is probably the most famous Heideggarian scholar and his interpretation of Heidegger has become so prevalent as to spawn its own branch of phenomenology colloquially called "Dreydeggarean Phenomenology" (Londen, in press). Dreyfus also established phenomenological ideas of his own, in particular we use the Dreyfus definition of coping which is an extension of Heidgger's ready-to-hand. Dreyfus also was able to apply phenomenological ideas into more concrete fields such as AI development. We lean heavily on Dreyfus's interpretation of Heidegger in this paper and much of the phenomenology we discuss, including the parts where we are describing Heidegger's ideas, is actually Dreydeggarean, with some of our own spin on it as well.

Phenomenology is the careful study of the lived experience by observation and analysis. It is the study of phenomena, "that which shows itself in itself" and their structure and relationships in experience. Heidegger makes an important distinction between the natural universe, which is physical, and the world we experience. Heidegger does not deny that there are physical things in the world, but says that the world is not just the collection of this physical stuff. Heidegger's particular focus is on the philosophical concept of "Being", which Heidegger defines as "that which determines entities as entities, that on the basis of which entities are already understood" (p. 25-26). This roughly corresponds to our sense of the world, of things as they are for us.[2] Heidegger mines the structure of lived experience for an understanding of human existence and Being, and goes on to show that *care* is central to the Being of humans.

Heidegger begins his phenomenological analysis with *Dasein*, commonly translated to "being there" (*da-sein*), which is something like Heidegger's way of referring to humans. It can roughly be equated to "conscious subject" but that has connotations Heidegger wanted to avoid because he is taking aim at the Cartesian subject-object distinction. Dasein is not a "conscious subject" perceiving a world of objects. Rather, dasein is always "in the world". By "world" Heidgger does not mean world as in the physical earth and its physical objects. Rather, he means "world" in something like the sense of "welcome to my world" or "my world is crazy right now", which is something like a network of meaningful "involvements". The things in the world, the entities, then get their being by their place in the network. A hammer is not just some 22oz of oak and metal. It is a hammer because of its involvement with nails, and with dasein's building of structures. That is what it is for dasein. Likewise dasein is dependent on the network of involvements as the network shapes our thoughts and choices, and thus Heidegger views dasein as necessarily bound up in this world and not some separate entity perceiving objects.

According to Heidegger, *Dasein* is that "being whose Being is an issue for it". In simpler terms this means that *Dasein* is the type of being who must make decisions about its way of

---

[2] "Being" means something quite similar to "phenomena", "lived experience" and even "consciousness", though Heidegger avoids the last few phrases because of their connotation of an internal realm of thoughts and feelings. Heidegger means to emphasize our openness on and engagement with the world with his terminology. Still we will also use the more traditional terms for clarity.

being. Animals and objects may avoid pain, seek food and shelter, and try to reproduce, but they do not make active choices about how they want to be. Dogs do not decide they want to be a dog who is good at caring for their pups and then seek out and learn from other dogs who are good at caring for their pups, or at least we don't commonly think of them as doing such things. On the other hand, a human is born into a world with a “specific range of ways it can be.” This range of ways a human can be does not refer to ways of being like hungry or safe, but ways of being similar to one's sense of identity, like teacher, mother, leader, or soccer player. As Mark Wrathall puts it, “these are potential ways to be, for the sake of which we do everything else” (2005). Depending on the culture we are born into we have an array of options about potential ways to be. If we were born in the late Roman period we could be a gladiator, a slave, or many other things, but not a computer scientist (Wrathall 2005).

At each point in our life we are making decisions, even decisions to do nothing, that determine which way of being we will be and we are aware of this and bear responsibility for it. For example, Dreyfus claims one of his potential ways of being that he is always making decisions towards is being a teacher, he delivers lectures, plans lessons, writes on the chalkboard, etc. all for the sake of being a teacher. At the time of this writing my wife is pregnant, and one of the potential ways of being before me is being a good father. I read parenting books, think about potential life lessons I can teach, and curate a collection of children's books I would like to share with my future children all for the sake of being a good father to my future child. Each of these activities is towards being a good father, and each time I choose to do one of these activities, from among the many that are available to me at any moment, is an active choice of mine about who I want to be. This active role we take in our Being is what Heidegger means as care, the constant selection towards ultimate ways of being, not the “cares of the world”.

To be a *Dasein* is to care in the sense that *Dasein* is always aware that there are options for what it could choose to be and is always acting with respect to those options. *Care* in this sense, as the selection of ways to be, is central to what *Dasein* is. Heidegger goes so far as to say that "*Dasein*, when understood *ontologically*, is care" (p. 84). What it means to be a *Dasein* is to have one's being be an issue for oneself, which means to actively have to make decisions about one's being. Those decisions show up as selection from ultimate ways of beings such as being a teacher or a soldier, the ways of being "for the sake of which" we do everything else. This structure of decisions being made for the sake of ultimate ways of being is what Heidegger means by *care*. We never make decisions in a vacuum but always have our ultimate goals of the kind of person we want to be influencing the decisions we make, even when we aren't explicitly attending to those goals at the moment. Wrathall has a great example about planting a tree. He imagines himself planting a tree and someone asking him: "why are you diggin a hole?" Wrathall responds: "to plant a tree." Again the questioner asks why and Wrathall says "so I can have fresh fruit". Again, the questioner asks why and Wrathall says "so I can provide fruit for my children". Again, the questioner asks why and Wrathall responds "so I can be a good father". The questioner asks "why?" again, and now Wrathall has no further answer. To be a good father is that for the sake of which he is doing everything else in that chain of reasons.

Because care is fundamental to being a dasein, we are never just in a world without care. We always care in the sense that what we do and how we see the world is dependent on the fact that core to our being is that we are always at least loosely aware of and selecting from ultimate ways of being. *Dasein* does things for a potential way of being. As we inhabit our world, we select potential ways of being, and then everything else in the world "shows up in terms of [the] decision[s] I have made about my existence" (Wrathall, 2005). Thus care, that is our selection of ways of being, is fundamental to our world, and what the physical things of the universe are, when they are in the world, is actually dependent on a *Dasein*'s care structure. This is why Dreyfus says that care "unifies all the various structural aspects of *Dasein*'s experience." Everything *Dasein* does is wrapped up in *Dasein*'s care.

Dreyfus described this type of care as having a sort of hierarchical structure. At the highest level of the care structure are the "for-the-sake-of-which", the ultimate ways of being that

Dasein is moving towards. Every action we take is ultimately linked to some for-the-sake-of-which. In the tree planting example, being a good father was Wrathall's for-the-sake-of-which. Or, suppose someone were to ask me why I am writing this paper. I would answer "to fulfill a requirement for my program". They might then ask, "but why fulfill the requirement?" And I would answer "so that I can graduate." To which they could ask again, "but why" and I would answer "so that I can go find a good job". And again, they could ask "but why?" and I would say "so that I can provide for my family." and they could ask why yet again. And I would reply "so that I can be a good father and husband." "At this point if my inquisitive companion asks again 'What for?', I can provide no further answer. I have reached that for the sake of which I do everything else" in the chain of actions (Wrathall, 2005). Being a good father and husband is that for the sake of which I am writing this paper just as being a good father was that for the sake of which Wrathall was planting a tree.

"Towards-which"s are the intermediary goals in realizing some "for-the-sake-of-which". In the previous writing example, each different reason I gave: to fulfill a requirement, to graduate, to get a good job, were all towards the ultimate goal of being a good father and husband. Each of those reasons was a towards-which. When we make things, the products we make are the "towards-which" of our process of making them, as the products are usually involved in some ultimate for-the-sake-of-which. Here is how Heidegger describes it:

> The work to be produced, as the 'towards-which' of such things as the hammer, the plane, and the needle, likewise has the kind of Being that belongs to equipment. The shoe which is to be produced is for wearing (footgear) [*Schuhzeug*]; the clock is manufactured for telling the time. (Heidegger, M., 1927, p. 99)

We make a clock towards the goal of telling time, which we then need to be able to do to realize some other for-the-sake-of-which. For example, one might need a watch to tell the time, towards the goal of being on time to meetings, for the sake of being a successful businessman. The towards-whichs are the intermediary goals, including things like making items we need to realize our ultimate goals of who we want to be. For-the-sake-of-whichs at the top and towards-whichs leading to the for-sake-of-whichs thus comprise the care structure.

The world of dasein is constituted by entities that get their meaning from their involvements with both other entities and with the project of the dasein. For example, when we enter a room we do not see four walls enclosing a 100 sq foot space. We are in a place for reading where we will pick up our book, lie on the sofa and read. This example shows us not only how the meaning of entities are dependent upon each other, but also that that network of involvements is in relation to what dasein wants to do, what dasein *cares* about. The world of involvements is dependent upon the project of dasein. As Heidegger says, Dasein “dwells” in this world.

In Heidegger care grounds out in low level daily activities and equipment. Equipment are the physical things but show up in relation to our care structure. For example, our for-the-sake-of-which could be to be a good parent, and to satisfy that we are building a shelter. Building this shelter at the moment requires that I hammer some nails to attach cross beams to our posts. The hammer, nails and wood are connected to our care structure, and we just use them in the way that is appropriate for that care structure. We don’t even think about them or their properties, we just use them to accomplish the task at hand. We care about building a shelter and thus the piece of wood with metal on the end is meaningful as a hammer that we can use in conjunction with nails to build our shelter to protect us from the elements. If we lived in a time where there was no need for shelter and no need to build using hammers and nails, the hammer would no longer be part of that care structure, would not have the same meaning.

The ready-to hand are the equipment that are directly of use to us as we act towards a current goal, such as a perfectly working hammer when we are trying to hammer some nails into a piece of wood. Equipment that is not suited to our task or goals, but that is meaningful to us in our overall care structure is unready-to hand. This can happen as we are working toward our goal and our tools begin to malfunction such that they become unready-to hand. In the hammering example, if we were in need of a hammer, but the handle on the hammer were to break such that we could no longer just use it as a hammer and work on our shelter, the hammer would become unready-to-hand. Dreyfus calls this process breakdown.

By contrast, "present at hand" or *vorhanden* entities are present to us. We are not involved with them or using them. They are the physical things that are simply there in the world, but are not in our overall way of being. The way the hammer would be if we lived during "a time where there was no need for shelter and no need to build using hammers and nails". When we are attending to an entity that is present at hand we are attending to its properties. These are things that describe or define the entity but do not change with dasein's projects. A red chair is still red whether I see it as a chair to sit in or a stool to step on to change the light bulb. If it is present at hand we simply attend to its properties as a chair thing. Its height, weight, color, etc.

Whether we see the thing as a hammer, an overly heavy hammer, or a piece of brown wood with 22 oz of metal attached to the end is dependent on our care structure, the totality of things we care about. This structure is *a priori* to how we perceive everything else in the world. Heidegger says "The transcendental 'generality' of the phenomenon of care and of all fundamental *existentialia* is … broad enough to present a basis upon which every interpretation of *Dasein* which is ontical and belongs to a world view must move" (Dreyfus 1991). This means care is so transcendent and general that every ontical (i.e. specific example) of *Dasein* doing something belongs under care. Every action, perception or mode of being we are in, is related to our care structure at the time.

For the most part dasein interacts with the world in an involved but ordinary way. Dasein almost always has a project that they are working on and dasein simply makes use of equipment to accomplish that project. As Heidegger, Dreyfus and others point out this is how we usually interact with entities. We spend most of our time just using the things at hand to accomplish whatever we are trying to do, "[w]e walk and read aloud, we get off and on street cars, we dress and undress, and do a thousand useful acts without thinking of them” (Dewey as cited in Dreyfus 1991).

Dasein’s world is a world of involvements. The structure of these involvements is determined by dasein’s projects, that which dasein cares about. The ready-to-hand to are the entities that are part of these involvements and thus being ready-to-hand is necessarily dependent upon care. That we mostly interact with entities that are ready-to-hand emphasizes that we are almost always trying to do something we at some level care about. Even if that something is just watching TV, the TV, the remote, and the couch are ready-to-hand, as we just use them, but don’t really think about any of their actual physical properties.

The primacy of the ready-to-hand is shown when entities become unready-to-hand, which Dreyfus calls “breakdown”. There are three ways this can happen: “conspicuousness”, “obstinacy”, and “obtrusiveness”. Conspicuousness is when the entity that is ready-to-hand becomes nonfunctional in such a way that it is longer something we can just use to accomplish our project, but instead is something we now have to pay attention to; it becomes conspicuous. Heidegger says, “When its unusability is thus discovered, equipment becomes conspicuous” and that “pure presence-at-hand announces itself in such equipment”. When an entity becomes unready-to-hand it becomes conspicuous to us, we attend to it rather than just using it, and when we attend to it we are looking at the way we do something present at hand, seeing its properties. Take the TV remote for example. When we are watching TV and changing the channel we don’t even notice the physical characteristics of the TV remote. However, when the TV remote buttons stop working as we try to change the channel, we suddenly attend the remote that was almost transparent to us before. We might look at the remote to see if any buttons are stuck, or carefully touch the remote to check for stickiness indicating that our child spilled soda on it. The remote has gone from being ready-to-hand, to being something that is more akin to the present-at hand.

The second way in which the ready-to-hand becomes unready-to-hand is obtrusiveness which when an entity that would be ready-to-hand is missing, not only is it “not 'handy' ["handlich"]", but it is not "'to hand' ["zur Hand"] at all". The "more urgently" we find our selves in need of the missing entity, the more unready-to-hand it becomes. In addition, as the missing entity becomes more unready-to-hand, the other entities that are ready-to-hand become more obtrusive, eventually appearing as present at hand.

The third form of breakdown is obstinacy which when an entity is in the way of the work we are trying to do. As Heidegger says "that to which our concern refuses to turn, that for which it has 'no time', is something un-ready-to-hand". Because we are forced to attend to this other entity, what was ready-to-hand again reveals itself as a form of present at hand which will return to being ready-to-hand as soon as we are able to return to our original work.

Breakdown shows us that ready-to-handness is dependent on the work we are doing. As soon as an entity is not something we are using to accomplish our project It changes its mode from ready-to-hand to present at hand. Furthermore, it hints to us that the project is what dictates how we "see" entities. This will become more clear as we discuss anxiety in chapter 6.

Dreyfus extends the idea of the ready-to-hand to his concept of "absorbed coping" or "transparent" coping, where we just use things without thinking about them. For Dreyfus in addition to this absorption being a product of ready-to-hand, it is dependent on our skill level. If we have sufficient skill in our ability to perform the action, then we don't have to think about it and we can simply be absorbed in the activity. Dreyfus borrows the example of a blind man with a cane from Merleau-Ponty. When someone uses a cane to guide themselves, they do not maintain any awareness of the properties of the cane such as its precise weight. Instead they are attending to the world they are experiencing the world at the end of the cane. The cane is entirely transparent to them. This transparency is very much like the ready-to-hand where we simply experience the entity as something for us to use in our work and do not attend to the substance properties of the entity at all.

In later chapters we will use these ideas to build a phenomenological theory of engagement. In particular we focus on how care can be structured, how that structure relates to ready-to-handness, and where that structure might also relate to engagement. In addition, we discuss absorbed coping and why Dreyfus thought absorbed coping might relate to flow, and why we think an additional dimension needs to be added to absorbed coping to create a phenomenological model that can account for flow and engagement. We call this full phenomenon "absorption".

Complex Systems

In this section, I survey theoretical work in complexity and coordination, which occurs in several distinct literatures. The terminology and concepts in these areas are not always used in a consistent way, and the fields have not settled on a unified set of concepts (though many of the mathematical underpinnings are stable). First, I discuss complex systems, what they are and how to measure them, and then develop the concept of coordination which I use here as a general term for two systems being related in time. I will specifically talk about four forms of coordination: correlation, synchrony, complexity matching, and synergizing.

Complex systems are difficult to define. Here I will loosely define them as systems whose components interact, often nonlinearly, and where behavior comes from the interaction of the components. Because behaviors come from interactions among the components rather than from individual components or linear sums of individual components, the causal chain from individual components to behaviors is often hard to disentangle. This means that in a complex system many behaviors cannot be attributed to individual components. Furthermore, in some complex systems, components may be softly assembled, which is when the organization of the interaction between the components is temporary and task dependent – thus further complicating any analysis that tries to decompose the system to analyze its separate parts. Instead of measuring individual components, measures of coordination between components within a system and even measures of cross system coordination provide important insight into how a complex system functions.

There are many ways of understanding the concept of coordination which we previously described as the relationship between entities in time. (1) The simplest measurement of coordination is to simply look for correlation. In this case the time series of behavior from component X or system A, will be correlated with some 0-N lagged version of the time series from some component Y or system B. (2) Synchrony takes correlation a step further and instead of just being correlation in some graded amount is lock-step timing; this would be like having a correlation of 1 at lag 0. (3) Complexity matching is much more abstract than correlation and synchrony. In complexity matching, two time series have a matched fractal structure. In complex systems, it has been suggested that having a matched fractal structure may maximize the efficiency of information-sharing across the two systems (West, Geneston, & Grigolini, 2008). (4) Finally, synergies are when the components of a system or multiple systems operate in an interdependent way. This is measured by estimating the reduction in degrees of freedom compared to what would be expected if the components or systems were acting independently. In other words, we compare the coupled systems with the sum of two systems. This is often measured using dimensionality estimation techniques such as PCA to compare the degrees of freedom for two independent systems vs. two coupled systems. The muscle systems are a good example of this: The number of ways I could move my arm given the different systems that control it are nearly infinite, but I seem to use only a select few as the different muscle systems are interdependent. There is work to be done relating to all of these accounts. We ultimately draw on a generic understanding common to all of them. In addition to the different kinds of coordination, coordination also seems to be a graded process where systems can be more or less coordinated. In fact, even within a single complex system like a person, different subsystems can be more or less coordinated. In the limit of no coordination we just have separate channels. Rigoli, Holman, Spivey, and Kello (2014) provide a perfect deomnstration of this. They examined data from key presses, pupil dilations, and heartbeats in a timing task and found complexity matching between the key presses and the metronome, and between the pupil dilations and the heartbeat, but not across those two subsystems.

Theoretical Accounts of Coordination in Cognitive Science

In cognitive science, a particularly important classification for the mind as a system is whether the system is interaction dominant or component dominant (Van Orden, Holden, & Turvey, 2003). For most of the history of psychology, and still today for many psychological paradigms, the mind has been assumed to be a component dominant system. A component dominant system is a system where the behavior produced by the system is a linear combination of the components. Cognitive components on this view are isolatable and immutable, the Atkinson and Shiffrin (1968) model of memory is a good example where sensory store, short term memory, and long term memory are distinct components that provide separate roles in cognitive processing. In experiments, component dominant models allow experimenters to partition the variance in data and attribute it to the different components.

> The notion of causation implied here is that mind and brain work in a way so that some input for a participant in an experiment is always mediated by the same components in the same (qualitative) manner to link this input to some measured output—simple feedback loops notwithstanding. The causal chain is hard-assembled. (Wallot & Kelty-Stephen, 2018)

On the other hand, in an *interaction dominant system* behavior emerges from the interaction between many components that are “softly assembled” to produce some behavior (Dotov et al., 2010; Wallot & Kelty-Stephen, 2018). Here components do not mean some fixed persistent psychological construct but instead come into existence as part of the activity. For example, though a piece of paper is not normally considered part of my memory, when I jot down notes during a debate the “memory component” involved in maintaining information about the speaker involves a softly assembled system that includes my notebook. Because the memory component does not exist outside of the task, it is difficult to isolate by simply using different experiments or different conditions and thus provide any consistent descriptions, like the magnitude of memory. Instead, interaction dominant systems are generally assumed to be complex systems and are measured according to that paradigm.

In addition, complex systems, being driven by interaction between components, are constituted by multiplicative interactions. This means that the behavior of complex systems also tends to follow power law distributions rather than normal distributions (Kello et al., 2010). Multiplicative interactions produce either log normal or power law distributions depending upon certain mathematical conditions of the interactions (Stephen & Mirman, 2010). Power law distributions are probability distributions where the probability of x is proportional to x raised to some negative power. These distributions are of particular interest because they have very heavy tails compared to the normal distribution, meaning that extreme values are more likely to occur. This has further implications for much psychological and cognitive science research as an underlying normal distribution is thought to be a given in psychological research and many statistical methods used in psychology and cognitive science are dependent on the normal distribution that occurs only in a component dominant system.

As stated earlier, complex systems are systems whose constituent parts can interact in complicated nonlinear ways, as with weather systems. In principle, these systems are governed by a set of differential equations, and instead of solving for a single component we can discover the system of equations that govern the evolution of the system over time, for examples of this see (Schmidt & Richardson, 2008; Haken, Kelso, & Bunz, 1985; Dale & Bhat, 2018). However, sometimes, even if we don’t know all of the details of the differential equations governing the system, due to the nested structure of complex systems, we see patterns that emerge as the system scales up or down, indicating some kind of metastability in how the system is organized (West 2018). These kinds of metastabilities take the form of power laws. Power laws are relationships between variables where one variable is proportional to some other variable raised to some fixed power. An intuitive example of this is bird flocks or schools of fish. It is quite apparent when watching these animals that they are traveling together, and that their behavior is interdependent. While the differential equation governing exactly how they are connected may not be clear, the correlation in orientation between two starlings in a flock declines with distance between the birds as a power law (Cavagna, et al., 2010). This power law implies that correlation only slowly declines with distance and asymptotes. Thus, a bird flock can in theory be nearly infinite in size. In addition, if correlation between birds was governed by some central bird chirping instructions, we would not have a power law. Instead, there would be some threshold at which the birds can no longer hear the leader. The presence of the power law implies that the bird flock is governed by neighborly interactions, and is indeed an interaction dominant system, where behavior of the flock comes from interaction between birds, not instructions from some leader bird.

Now, it is true that systems--both interaction dominant and component dominant--can *coordinate*, at least to some extent. However, as we will see, coordination is more intrinsic to complex systems and thus complex systems produce a wider variety of types of coordination, some of which component dominant systems can’t perform.

In its simplest sense, two systems are considered to be coordinating if they are related in time and the simplest measure of this is for them to be correlated in time. This can be tested by examining the correlation function between the time series of two coordinating systems. For example, when riding a horse, humans coordinate their movements to that of the horse showing correlation in mean relative phase, and mean standard deviation of the relative phase (Lagarde, Peham, Licka, & Kelso, 2005). We think of this as the simplest measure because it tells us very little about how the two systems relate to each other. It sort of minimally suggests information exchange between the two systems and tells us nothing more.

Beyond correlations, systems can also synchronize. This happens when two systems become locked in time, rather than just being correlated in time and as such their time series become identical. This type of coordination occurs between two connected metronomes in a simple case, but can also happen between chaotic systems, and even fireflies (Strogatz, 2012; Stephen and Dixon 2011). Synchrony is much more informative than mere correlation, because it tells us that the two systems have in some sense unified. There are still many ways that the systems could be unified. There could be a central time keeper, like in a computer, or there could be a physical connection between two metronomes that allows disparity in swings to alter the other metronomes swings. However, it still tells us that these two systems are operating as one system and have strong information exchange between them.

Complex systems can also complexity match to coordinate. “The concept of complexity matching (West et al. 2008) states that the exchange of information between two complex networks is maximized when their scaling laws are similar” (Delignières, Almurad, Roume & Marmelat, 2016). In other words, “when two complex systems become coupled; they should attune their complexities in order to optimize information exchange” (Delignières et al., 2016). Complexity matching is analyzed by checking for similar slopes in the power law functions of the behavior of two complex systems. Analyses such as Allan Factor and DFA can detect the precise power law relationship across scales in the behavior. Statistical analysis can then be performed to see if one power law is significantly similar to another.

One caveat to complexity matching is that Deligneres and Marmelat (2014) showed that “a model based on local corrections of asynchronies … was able to adequately reproduce the statistical matching observed in” some complexity matching experiments. Because local corrections can be accomplished by a component dominant system, this means that component dominant systems can, under certain conditions, produce empirical data that is statistically indistinguishable from data produced by complexity matching. Thus, complexity matching, as a measure, does not always indicate that the two systems are complex.

However, component dominant systems cannot perform multifractal complexity matching (Delignières et al., 2016). Multifractal time series are defined by multiple power laws and tend to contain more extreme values but still exhibit some self-similar structure. Multifractal complexity matching involves looking at multiple statistical moments of fluctuation in a time series to capture extreme small and extremely large fluctuations. To rule out spurious results in complexity analysis, researchers have turned to the use of multifractal analysis such as multifractal DFA which is an extension from DFA (there are other methods as well). Using multifractal DFA we are able to determine the q-order scaling exponents which allows for a richer picture of the time series fractal structure. In addition, this multifractal approach allows for more precise measurement of how well two time series' fractal structures are correlated. Using these methods, it has been shown that bimanual coordination and interpersonal coordination are a result of complexity matching (Delignières, Almurad, Roume, & Marmelat, 2016; Ihlen & Vereijken, 2010). In addition, Ihlen and Vereijken (2010) showed that most complexity matching studies performed using mono-fractal analyses were supported when reanalyzed with a multifractal analysis.

Finally, systems can also form synergies and two complex systems can synergize together. A synergy is "the emergence of coordination via reduction in the degrees of freedom in a system of many interacting components" (Abney, Paxton, Dale, & Kello, 2014). This is particularly salient in complex systems where the degrees of freedom of a system are often incomprehensibly large, but the actual number of different behavior patterns produced is much smaller, suggesting that some of the degrees of freedom are interdependent. This is most commonly studied in cognitive science in the motor system, where a massive number of muscle configurations are possible to solve a task, yet humans seem to use a select few of them (Riley, Richardson, Shockley, & Ramenzoni, 2011; Turvey, 2007). Beyond that, synergies between multiple complex systems have been investigated in the "interpersonal interaction domain" where multiple people (complex systems) interact in such a way that as a group they reduce the number of degrees of freedom that would be used if each was operating independently (Riley et al., 2011). One way to imagine this is conversationalists converging on a joint lexicon during a conversation. Rather than simply using the words and phrases one normally does, in a conversation, conversationalists will constrain themselves to fewer words and behaviors (Abney et al., 2014). And beyond language, interpersonal synergies have also been documented among players in team sports. In sports, the actions of a player are often determined by the actions of the other players on their team and even the actions of opponents. For example, in soccer if a teammate dribbles the ball from the center toward the sideline, it can often be prudent for a player who was near the sideline to “open up” by moving either backward or forward to give the dribbling player an open array to pass the ball. Duarte et al. (2013) used a phase clustering algorithm and sample entropy to show that professional soccer players on the same team synergize together, and they also showed with cross sample entropy that the two teams synergize with each other as well.

In terms of interaction dominant systems, these different methods of coordination can be thought of as different principles of organization for the soft assembly of the system. For an interaction dominant cognitive system, if it is simply trying to coordinate with some external object that isn’t trying to coordinate with it, it can simply adjust itself to synchronize or complexity match with that external entity depending on what is appropriate for that object. If the other entity is attempting to reciprocally exchange information, we might expect to see both objects align their complexity or the formation of synergies. One should expect that humans will use any and all of these strategies, using whichever is best suited for engaging with the current entity.

Empirical work on Coordination

Humans coordinate temporally with a variety of other entities ranging from metronomes to horses to other people. One way to think about these temporal relations is as integration into the referential totality, or from a more traditional extended mind perspective those entities are part of the brain body environment system that constitutes the mind.

Some have noted these links already. Dotov, Nie, and Chemero (2010) had participants play a game on a computer using a mouse. The mouse was programmed to start malfunctioning during the game for a short period of time. They collected movement data from an accelerometer placed on a participant's finger. Dotov et al. (2010) then used DFA to analyze both the acceleration data from the accelerometer as a time series. They found that when the mouse stopped working as expected, there was a significant reduction in the hurst parameter (calculated by DFA). They link this to Dreyfus' account of skilled coping, arguing that this parallels the first form of breakdown in ready-to-hand, when an entity becomes unready-to-hand because it is broken or we are no longer able to just use it to accomplish the task at hand. This implies that when an entity is integrated into the cognitive system, the system can be measured at the boundary and will show greater complexity (will be one whole complex system). As the mouse became unready-to-hand, the signature of complexity of the mouse movements changed. One way to think about this is as if that mouse became less subsumed under the dynamics of the person performing the task.

Dotov et al. (2010) show how a person and artifacts can be coordinated into what can be considered a single structure, all subsumed under one "temporal umbrella". The person does not match the dynamics of a hammer, but the hammer becomes integrated with the person and thus the complexity of the person hammer system is somewhat different than the person system. However, that is different from cases where a person coordinates with a separate system that has distinct dynamics, such as learning to ride a horse or tapping along to a metronome as the horse and metronome have their own dynamics. When two complex systems coordinate they will sometimes match complexities in order to maximize the information exchange between them. Stephen and Dixon (2011) found that when humans were instructed to tap along to a chaotic metronome, even though that metronome is necessarily locally unpredictable, the humans were able to approximate that metronome's global temporal structure, as shown by correlated fractal exponents (complexity matching), and anticipate the metronome. Other authors have also shown that humans are able to adjust to coordinate with animals such as horses (Lagarde, Peham, Licka, and Kelso, 2005), and with other people (Abney et al. 2014; Schloesser, Kello, & Marmelat 2020; Almurad, Roume, & Delignières 2017).

Coordinated behaviors between individuals seem to be related to various psychological phenomena such as likeability. For example, Hove and Risen (2009) showed that participants found confederates who coordinated with them in a tapping task to be more likable. Additionally, Ramseyer and Tschacher (2011) found that synchrony of bodily movements between therapist and patient (measured with cross correlation) was a good predictor of therapeutic outcome. They also found that nonverbal synchrony is higher in genuine interactions.

This phenomenon seems to work both ways. Not only does greater coordination increase likability and success, but greater likeability seems to increase participants coordination. Zhao, Salesse, Marin, Gueugnon, and Bardy (2017) found that participants will coordinate better with a more likable confederate. Furthermore, when participants are involved in more collaborative interactions, they seem to become more synchronized. Paxton and Dale (2013) found that participants' bodily movements coordinate better when having a positive discussion than when having an argument. This was found by using cross correlation of movement. Abney, Paxton, Dale, and Kello (2014) also found complexity matching in affiliative conversation but not in arguments using Allan Factor analysis. Arguments instead have more structure at large frequencies indicating greater emphasis on turn taking but not on alignment.

Complexity matching can also be predictive of collaborative performance. For example, (Brennan & Clark, 1996; Brennan & Hanna, 2009)" found that "Complexity matching can correspond with enhanced joint decision making, and by extension with enhanced mutual comprehension as well". Moreover, Marmelat and Deligneres (2012) analyzed whether people in dyads working on a joint task globally synchronize. They found that even though people were not cross correlated in their time series they did produce highly similar fractal exponents according to DFA. “These results suggest that interpersonal coordination, and more globally synchronization of participants with natural environments, is based on non-local time scales”. Complexity matching is thought to boost performance in these cases because complexity matching maximizes the information shared across two complex systems. Schneider, Ramirez-Aristizabal, Gavilan, and Kello (2020), found that bilinguals and monolinguals match complexity and perform lexical convergence in conversations. The separate convergence of complexity matching and lexical matching may imply that there is some more general basis for convergence, such as maximizing the information shared between two complex networks.

The type of complexity convergence that happens between people in dyads is similar to the complexity matching that happens between the limbs in a single person. Schloesser, Kello, and Marmelat (2020) had participants perform a Fitts task as individuals or as dyads. They found that in both the individual and dyad conditions there was complexity matching between the left and right sides indicating that the type of coordination being accomplished in the Fitts task may be similar whether it is the left and right hand of one person or the left and right hand of two different people working together. This result nicely caps off our discussion of complex systems and behavioral coordination because it emphasizes that we complexity match both within ourselves, between our two hands, and to external influences such as another person.

There are even some hints in neuroscience that spatiotemporal synchronization is occurring in the brain. Though neuroscience research on the topic is far from conclusive, the work by Northoff (Northoff, 2020; Northoff, 2016) postulates that our brain is hierarchically divided spatially and temporally in a way that mimics the information it is processing. For example, the frequency bands in sensory regions of the brain tend to be very fast. This allows these frequency bands to carry information about individual stimuli. These high frequency bands then influence the slower frequency bands that could govern higher level cognitive processes such as the frequency bands in the DMN through cross frequency correlation. All of this allows our brains to couple with the environment for maximum information processing.

Two brains also “couple” to each other during information exchange. For example, brains synchronize during social engagement (Dikker et al., 2021). This synchrony is further dependent on bodily coordination and other social factors such as empathy and closeness. Brain synchrony among team members predicts team performance (Reinero, Dikker, & Van Bavel (2021). Brain synchrony between learners and teachers predicts learning (Pan, Novembre, Song, Li, & Hu 2018), and is related to how much the teacher provides mental scaffolding for the learner (Pan, Dikker, Goldstein, Zhu, Yang, & Hu 2020). Shared attention among multiple learners may further drive brain synchrony further implying that spatiotemporal dynamics may be manifested in the brain in some way (Dikker et al., 2017).

In summary, these results show not only that humans coordinate with other entities temporally, but that that coordination often affects other processes that, when viewing the mind from a component dominant perspective, would seem unrelated, such as affective feeling towards said entity or the ability to work together with another complex entity. All of these results when taken together imply to us that when humans are attempting to engage with some entity they try to, by whatever method is most appropriate, coordinate with that entity and that process of coordination seems to, in at least some minimal sense, integrate that entity with the person's consciousness.

Synthesis

In a nutshell we think of engagement as emergent from the connection between the motivational affordances of the task at hand and the Dreydeggarean care structure of the person. The more well matched these are the more “absorbed” the person will be in the activity. This then can be captured in complex systems terminology as the amount of coordination between a person and the dynamics of the activity.

As our approach to engagement is somewhat different than the standard approaches, the specific definitions of “cognitive engagement” or “behavioral engagement”, etc. do not apply perfectly, but we think it is clear that we are roughly talking about the same phenomenon. Motivation is clearly related to what one cares about, and thus engagement as the mediator variable between performance and motivation, is suitably paralleled by the connection between care and task. For example, both our explanation of engagement and other more well-known accounts of engagement would predict that engagement uses significant attentional resources and produces a feeling of absorption (Schaufeli 2013).

Furthermore, we make a few specific predictions that we think could be beneficial to engagement research. First, we offer grounded reasoning that physiological measures such as heart rate or galvanic skin response predict engagement. If one is convinced by the Heideggerian account of phenomenology, and that complex systems and embodied cognition are appropriate empirical extensions of that phenomenological account, then it stands to reason that if we are measuring some part of the brain, body, environment complex system, we should be able to get information about the system as a whole (the mind).

However, our theoretical perspective also extends these arguments suggesting that engagement should be detectable by measuring not just the body as part of the complex system, but by measuring mind, body, and environment boundaries. Just as coping is detectable at the boundary between the body and the environment, we might expect engagement to be detectable here. Part of the phenomenon of coping involves a transparency between the mind and the object in question. One way of thinking about this from a complex systems perspective, is that that object literally becomes part of the cognitive system by becoming part of the brain, body, environment system that is giving rise to cognition. Engagement also involves transparency, but rather than the object becoming transparent to us as we pull it into our cognitive system, we unify ourselves with the spatiotemporal dynamics of the task such that we lose ourselves in the task. Thus we can see how much the task is or isn't being integrated in by how much of its structure is showing up in the complex embodied system which is also directed by the care structure. In this dissertation we will be highlighting our experimental and theoretical work that investigates how coordination is related to engagement, particularly in the domain of mouse tracking

3. Analysis of Complex Mouse Tracking Data

Mouse Tracking

Ongoing, continuous exchange of information between body, mind, and environment suggests that measurements of any of these subsystems will produce information about cognition. This in turn predicts that cursor data, an indirect measurement of the body's movements, will contain information about cognition. Mouse tracking studies can capture this data, and have many additional advantages. The “environment" in a computer task is highly controlled, and environmental variables can be monitored with high precision. The mouse, as a sensor for the body, can be measured with low latency. Computer tasks using mouse tracking data are relatively cheap and easy to produce, and data can be gathered from experiments run remotely over the internet. Most mouse tracking research uses what we will call the `standard paradigm', a “two choice MouseTracker task" (Hehman, Stolier, & Freeman, 2015) in which a participant's cursor begins at a predetermined start location, usually at the bottom-middle of the screen. The participant is then presented with some stimulus and asked to make a choice . They might be shown a picture of an animal and be asked to identify if that animal is a fish or a mammal by moving the mouse cursor to a text box labeled “fish” or “mammal” (See Fig. 1). As the participant moves their cursor towards the target, the (x; y) position of the cursor is recorded. The resulting vector of time stamped positions is a cursor trajectory. These cursor trajectories are then aggregated and analyzed (Hehman et al., 2015; Stillman, Shen, & Ferguson, n.d.; for a broad recent overview of this type of research see Schoemann, O'Hora, Dale, and Scherbaum, 2020) .

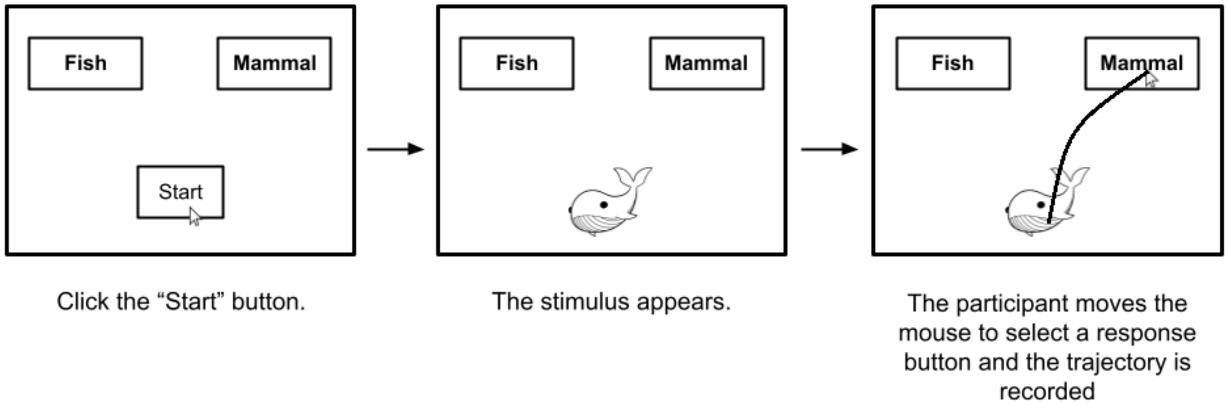

Figure 2. Standard Mouse Tracking Task. Image of a standard mouse tracking task based on the example in (Hehman et al., 2015).

One of the most prominent examples of this type of research asks participants to categorize pictures of animals as mammals or fish (Dale, Kehoe, & Spivey, 2007). The participants are presented with either typical exemplars of mammals such as cats, or atypical exemplars such as whales. Participants' cursor trajectories are found to be more curved during trials with atypical exemplars, suggesting a “graded competitive process" of categorization, rather than a serial model (Dale et al., 2007). Other prominent examples include: evidence for multiple distinct components of inhibitory control in Stroop and flanker tasks (Erb, Moher, Sobel, & Song, 2016), evidence against a dual system of emotion and reasoning in moral reasoning (Koop, 2013), evidence that better self-control facilitates quicker resolution of self-control conflicts shown by earlier changes in curvature of mouse movement (Gillebaart, Schneider, & De Ridder, 2016), evidence for “partial and parallel activation of stereotypes”, implying that “perceptual cues of the face” invoke multiple “simultaneously active stereotypes ... and this mixture settles over time onto ultimate judgments” (Freeman & Ambady, 2009), and “evidence that cursor motion analysis has the capacity to predict emotional experience of the computer users” (Yamauchi & Xiao, 2018).

Despite its value in the study of continuous output during binary decision tasks, this research does have limitations. First, it is limited to serial tasks with a predetermined start and end point for each mouse movement. In addition, because the data are segmented into many discrete trials, dynamic processes which might build over longer timescales are hard to analyze. Thirdly, as is normal in many psychological paradigms, participants are required to follow rigid procedures to obtain clean data. These constraints make it difficult to apply standard mouse tracking techniques to what we will call free mouse tracking data, which is continuous over time, not segmented into individual movements, and not necessarily constrained in terms of starting position, end position, or preferred trajectory. However, it is quite clear to us that engagement is a process that occurs at a longer timescale than individual movements, and mostly occurs in dynamic tasks where the start and end position of distinct movements is unknown. Thus we need to build tools in mouse tracking to analyze information at a much longer time scale and look at information at the global level rather than the individual movement level.

The first step in this process is to try to find indicators about the kinds and amounts of information available at this global level. To do this we performed an exploratory data analysis on a relatively naive, whackamole like, task. Naive in the sense that the task is simple, and doesnt build in our hypotheses. Whackamole is good for this because there are a mix of distinct individual movements, but also many blended movements and many complicated movements that dont follow a simple curve from one location to the next. It is more natural. This allowed us to test for the effectiveness of global measures of interesting variables, and also to test if we could determine good ways to separate individual movements when they are unlabeled.

The Whac-A-Mole game initially has several empty mole hills appear on the screen. Cartoon moles then appear and disappear in the mole hills in a pre-determined sequence. The player's objective is to click the mole before it disappears and reappears in a different mole hill. A mole appears in a molehill for 650 ms before it disappears. If the mole is clicked, it changes to a cartoon picture of an unconscious mole for 350 ms. In both cases a mole then re-appears in another mole-hill. We pre-generated the random sequence of mole appearances so that every participant would experience the same pseudo-random sequence. The game ends after the participant has seen 120 moles (2-3 minutes). Upon completion of the game participants also filled out a short demographics form, and then were thanked and debriefed.

The game was built in javascript and played through the browser at a website and is available here: cogsci.us/TM/clickdamole. During the game we continuously collected participants' cursor data, every 8-12 ms (the maximum polling rate for javascript). We also collected participant’s click locations and recorded their accuracy in the game task.

The experiment was deployed on Amazon Mechanical Turk (MTurk). We required that the MTurk workers have over a 95 percent approval rating, have not participated in one of our studies before, and be in the United States. We collected data from 600 participants in two samples. The first sample was collected on January 14, 2020 and had 300 participants. The second sample was collected on February 19, 2020 and had 300 participants. All participants were paid 35 cents for completing the 2-3 minute study.

Three additional criteria were used to filter the data. First, we focused on mouse tracking data only, and thus removed those who did not report using a mouse (a question asked them what type of device they were using: mouse, trackpad, or other). This removed 83 people from the first sample and 102 people from the second sample. We did not require that they use a mouse explicitly to prevent participants from simply lying and saying they were using one when they were not. Second, we removed people who did not register sufficient cursor movement. Since javascript only polls mouse locations when the cursor is moving, if a participant simply let the mouse sit idle and let the game run or only tried to move the mouse a few times, javascript could not record enough mouse movement for us to analyze their data. We removed 2 participants from sample 1 and no participants from sample 2 for having less than 300 cursor locations reported. This also would filter participants who reported that they were using a mouse but were in fact using a touch screen or some other input device. Third, we included two catch questions in our demographics form: “how many letters are in the english alphabet" and “if you are reading this select the answer 17". This eliminated 12 participants from sample 1 and 19 participants from sample 2. In all we removed 105 participants from sample 1 and 131 participants from sample 2 before analysis.

For each participant, the cursor position was collected throughout the task every 8-12 ms, producing roughly 6000-component vectors of x and y coordinates, which is the mouse trace for that participant. An example of a mouse trace is shown in Fig. 3.

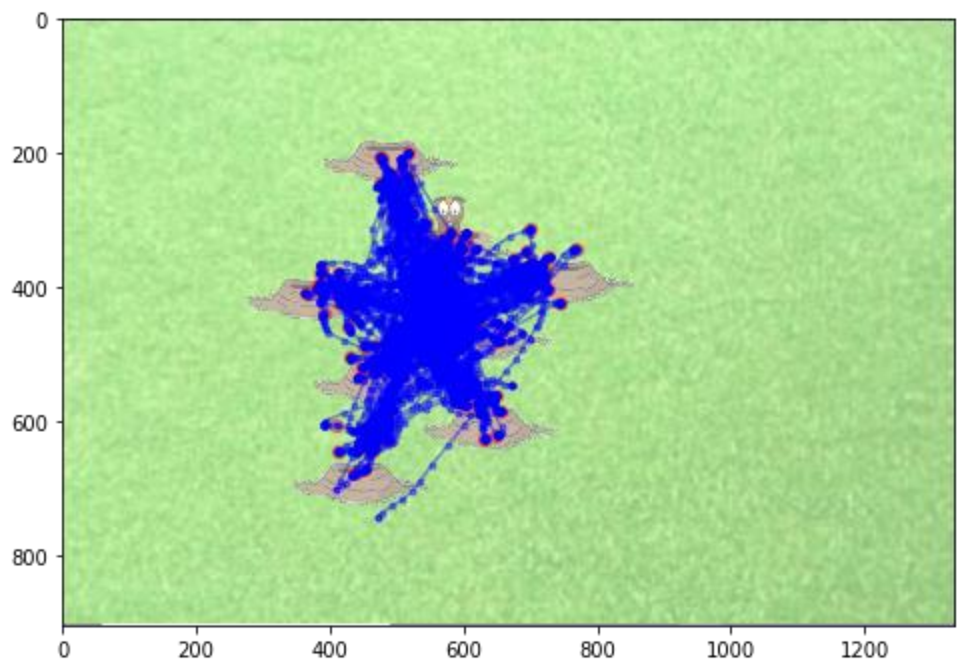

Figure 3. A sample mouse trace for one participant. Red dots correspond to mouse clicks.

Most methods for analyzing time series assume one dimensional data. However, mouse tracking data is inherently two-dimensional since it samples the x- and y-coordinates of the mouse position at discrete times. To accommodate this researchers typically use only one dimension of their mouse trace data e.g., either the x- or y- coordinate over time, which is fine in tasks like binary forced choice where it is clear what the expected mouse path is and what dimensions should have the most relevant information. However, this restriction potentially leads to information loss, especially if one does not have insight into which dimension is likely to carry the most information. As an alternative, we introduce a complex-valued one-dimensional time series z_n where $z = x_n + iy_n, n = 0, 1, 2, \ldots N$ with x_n denoting the *x*-coordinate, y_n denoting *y*-coordinate both at time level n, and $i = \sqrt{-1}$ denoting the imaginary constant. To isolate the *x*-coordinate, we evaluate the real part of z_n and to isolate the y-coordinate, we evaluate the imaginary part of z_n. Thus, by representing the two-dimensional mouse tracking data as a complex time series, we retain the full information available in mouse tracking data in a one-dimensional time series. Hence, this representation opens the opportunity to use time series analysis methods on full mouse-tracking data.

Mouse tracking through a browser has some inherent temporal variability as different computers and different browsers can poll at different speeds. To accommodate this we interpolated all data using the *pandas.resample* method to make sure that all data points were exactly 20 ms apart and then trimmed all time series to the length of the shortest time series in the data set across both samples.

Since the mouse tracking data are resampled to be uniform in sampling rate and length across participants, we are able to compute, analyze, and compare their Fourier spectra. To perform this spectral analysis, we computed the discrete Fourier transform of the complex time series using *numpy.fft.fft*, to produce a spectrum Z_n for each participant satisfying the relation

$$z(t_j) \approx \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} Z_n \, e^{i2\pi f_n t_j}, j = 1, \ldots, N,$$

with $f_n = \frac{n}{N\Delta t}$ denoting the discrete frequencies and Δt denoting the constant sampling rate.

One possible next step to take would be to compute derivatives of the time series so that one can analyze velocity, acceleration, jerk, etc.~which is common in other mouse tracking research. Computing derivatives of Euclidean distance as is done in Kieslich and Henninger

(2017) does not maintain directional information. In contrast, we are able to compute the *k*th derivative of the complex time series through evaluation of

$$z^k(t_j) \approx \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} (i2\pi f_n)^k Z_n \, e^{i2\pi f_n t_j} \; j = 1, \dots , N,$$

and the results will include directional information. However, taking derivatives of mouse tracking data is inherently problematic for several reasons. First of all, computing derivatives of time series amplifies noise. Additionally, mouse tracking data includes discontinuous jumps which technically lead to infinite derivatives. These issues can be dealt with if the number

of samples is sufficiently large and appropriate filters are used and has been done by researchers such as Nazir et al. (2008). Regardless, we show in our results below that we are able to recover valuable insight about information contained in mouse tracking data without needing to compute derivatives.

Analysis and Results

In this section we will cover two analyses that are somewhat less common in cognitive science first detailing how they work and then showing the results from using them on our data. We begin our analysis with the Singular Value Decomposition of the data because it is agnostic and allows us to easily identify important trends in the data. Second, as we are interested in global information from our data, and that approach to cognitive science is often tied to complex systems approaches to cognitive science, we also analyzed our data using Detrended Fluctuation Analysis. Below is a short step by step sketch of our method to provide the reader a mental scaffolding, after which we unpack each method and our results in much more rigorous detail.

Singular value decomposition to analyze performance

Let P denote the number of participants and Z(p) denote the vector of length N containing the values of the discrete Fourier spectrum for participant p. We form the $N * P$ matrix A of complex numbers defined according to $A = [Z_1 \ldots Z_P]$. In other words, the pth column of A corresponds to the discrete Fourier spectrum of the mouse tracking data for participant p. In what follows, we assume that $P < N$ since the number of participants in each study is smaller than the number of discrete Fourier frequencies.

The singular value decomposition (SVD) of the matrix A is $A = U \sigma V^*$ (Demmel, 1997; Trefethen & Bau III, 1997). Here, the superscript * denotes the complex conjugate transpose of the matrix. The columns of the $N * N$ matrix U form an orthonormal basis forC^N, the space of all complex vectors of length N. The columns of the $P * P$ matrix V form an orthonormal basis for C^P, the space of all complex vectors of length P. The $N * P$ matrix σ has non-negative entries along its diagonal called the singular values which we denote by σ_p for $p = 1, \ldots, P$. The non-diagonal entries of σ are zero identically.

In fact, computing the SVD of the matrix A is the same as performing principal component analysis (PCA). Although SVD and PCA are equivalent, we focus on the linear algebra interpretation of the SVD to study the discrete Fourier spectra of mouse tracking data. In particular we use concepts such as projections onto subspaces. By doing so we develop model-free methods that make use of any underlying algebraic structures in these data. All matrices possess a singular value decomposition and the singular values are unique. The singular values are ordered by size,

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \ldots \geq \sigma_P$$

Let $\mathbf{u}_n$ denote the nth column of U and $\boldsymbol{v}_p$ denote the pth column of V. We can rewrite the SVD of A as

$$A = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \ldots \sigma_3 u_3 v_3^* + \sigma_P u_P v_P^*$$

By writing A as this sum, we see that the singular values give a relative rank of the importance of the corresponding columns of U and V in the data -- the first term gives the largest contribution, the second term gives the next largest, and so on. Additionally, we can consider approximations by truncating the sum above after some specified amount of terms. This approximation corresponds to the projection onto the subspace spanned by the vectors included in it.

Suppose we only use the first k singular values. Let Ũ denote the $N * k$ matrix formed by taking the first k columns of U and removing the rest. The columns of Ũ form an orthogonal basis for a subspace of A, which we denote by Ũ. Now consider an individual participant's discrete Fourier spectrum, $\boldsymbol{Z}_p$. The projection of $\boldsymbol{Z}_p$ onto Ũ is given by $\tilde{U}\tilde{U}^* \boldsymbol{Z}_p$. The length of this resulting vector is $|| \tilde{U}\tilde{U}^* \boldsymbol{Z}_p ||$ with $|| \cdot ||$ denoting the Euclidean norm. When we compute

$$\eta = \frac{||\tilde{U}\tilde{U}* \boldsymbol{Z}p||}{||\boldsymbol{Z}p\ ||}$$

the value $0 \leq \eta \leq 1$ gives the fractional amount of $\boldsymbol{Z}_p$ lying in the subspace Ũ. When $\eta = 1$, $\boldsymbol{Z}_p$ lies entirely in Ũ. When $\eta = 0$, none of $\boldsymbol{Z}_p$ lies in Ũ. We explain below how we use η to study performance.

To analyze performance we first operationalized performance as accuracy in the game -- the higher the percentage of moles clicked, the higher the accuracy, and the better the performance. To investigate performance operationalized as accuracy, we first partitioned our data into “accurate" and “inaccurate" groups, where “accurate" participants scored above 50.5 percent, and “inaccurate" participants scored below 12 percent. These numbers were chosen to keep the group sizes of more accurate and less accurate players about the same across two samples.

We collect the discrete Fourier spectra of accurate players from Sample 1 and form the matrix A with them. Upon computing the SVD of A, we determine how many singular values are important in explaining the data. The singular values σ_p for $p = 1, \ldots, P$ are proportional to the square root of the variance accounted by the corresponding column of *U*. Thus, the cumulative sum of squares of the singular values is the cumulative sum of variance explained. This cumulative sum is shown in Fig. 4 and we determine from these results that $k = 9$ explains 50 percent of the variance. We call the resulting subspace Ũ by considering the first 9 columns of *U* the *accuracy subspace*. By computing η we determine the fractional amount a given discrete Fourier spectrum lies in the accuracy subspace. Consequently, η is a measure of fitness to high performing players.

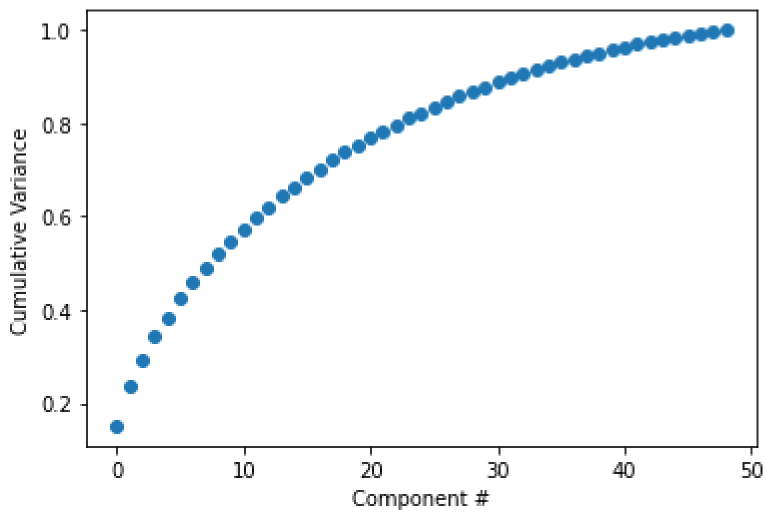

Figure 4 Cumulative Variance Plot. The cumulative variance accounted for by each component in the accuracy space. Notice that the first 8 components account for about 50% of the variance.

We consider the results of Sample 2 to test how well the accuracy subspace from Sample 1 generalized to new, out-of-sample participants. We identified accurate and inaccurate players in Sample 2 using the same criteria that we used to determine accurate and inaccurate players for Sample 1. To test out-of-sample performance, we computed η for accurate and inaccurate players in Sample 2. The results of this computation are shown in Fig. 5. These results show that the two groups: accurate and inaccurate, are almost completely separable. They are shown to be significantly different according to a Welch's t-test ($p<0.0001$). These results demonstrate the existence of structural features in the discrete Fourier spectra of accurate players that are not shared by less accurate players. The accuracy subspace contains algebraic structures inherent in accurate players that are not shared by less accurate players. Therefore, testing the extent to which a player's discrete Fourier spectrum aligns with the accuracy subspace provides a diagnostic method for performance. Moreover, these results suggest that these structural differences persist across different samples.

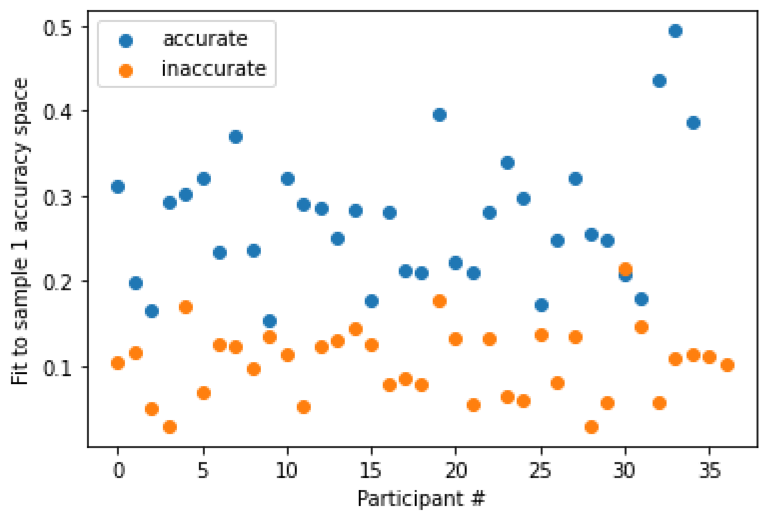

Figure 5. Projection of accurate and inaccurate participants in sample 2 to the accurate space from sample 1. Blue dots correspond to accurate players; orange dots to inaccurate players. The accurate players fit the space of accurate players from the earlier sample better with higher η (which corresponds to how well a participant fits to a space). The two groups are also significantly different ($p<0.0001$). This shows that, within a sample, accurate and inaccurate players are separable using SVD, without having to know their accuracy.

Next, we computed η given in equation 3.1 for all players in Sample 2. In doing so, we found that there was a significant relationship ($p<0.01$) between accuracy and the value of η (see Fig. \ref{sample2_regression}) with $R^2 = 0.61$. Accuracy and fit both range from 0 to 1, and the unstandardized β coefficient for fit regressed on accuracy was β = 0.3 with standard error of .02, which means that as accuracy increases fit increases. These results show that the relationship does not just separate two groups of accurate and inaccurate players in the sample but explains degrees of accuracy throughout the sample.

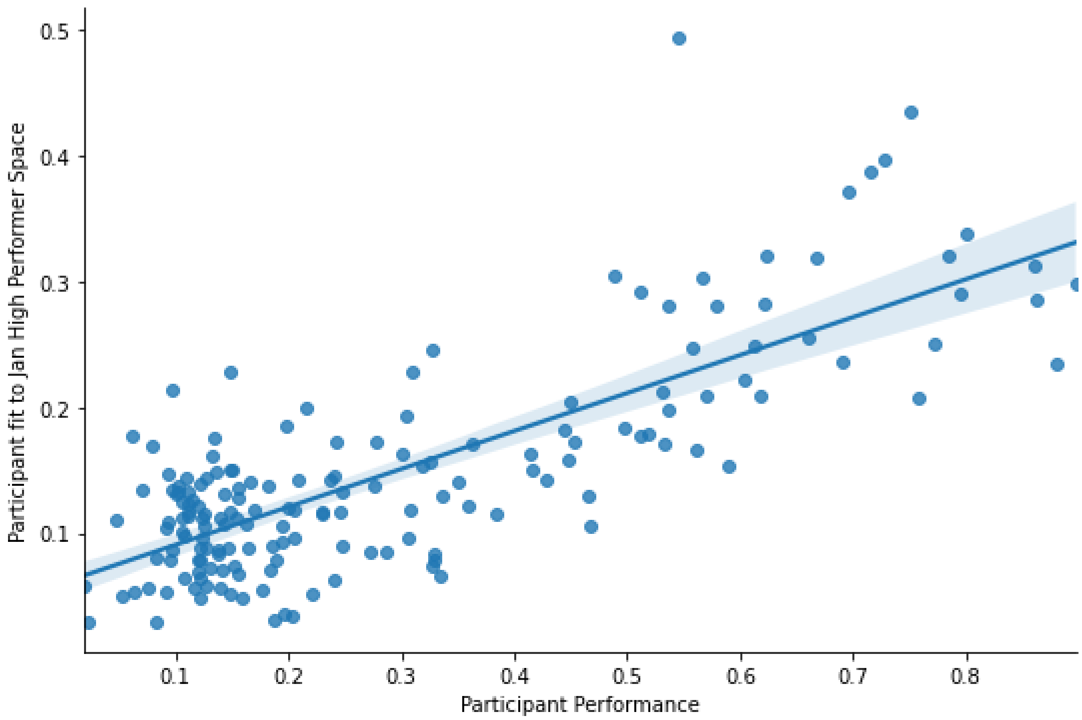

Figure 6 The relationship between accuracy and fit to the accurate space of sample 1 for participants from sample 2. Accuracy ranges from 0 (no mole hits) to 1 (every mole was hit). We can see that accuracy is related to fit to the accuracy space. The unstandardized B is 0.30 with SE B 0.02. R2 = 0.61 and p<0.01

Representing two-dimesional mouse tracking data as a one-dimensional complex time series opens access and opportunity for novel methods of analysis. For example, we have been able to perform SVD/PCA directly on the discrete Fourier spectra of the full mouse tracking data. In doing so, we have been able to identify structural differences in the discrete Fourier spectra between accurate and less accurate players. We have found that these structural differences are statistically significant and persist in out-of-sample results.

DFA Analysis of Performance

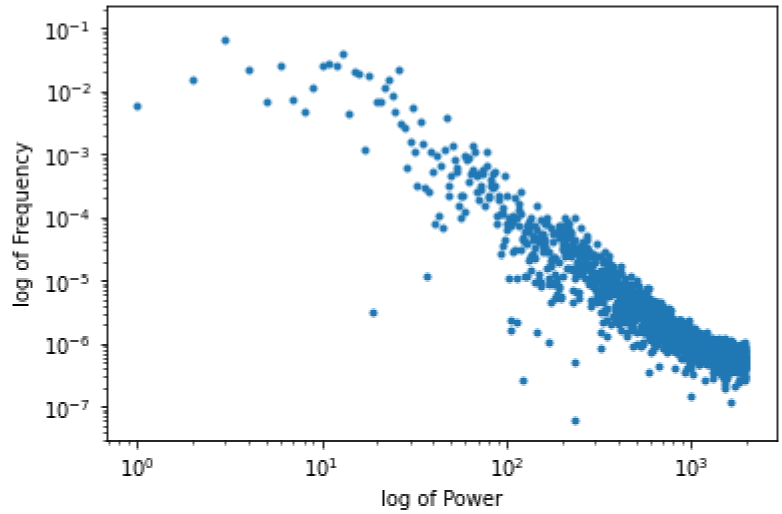

Figure 7. The most important component used for creating the accurate player space, plotted in log-log scale.

To investigate the accuracy subspace Ũ further, we consider the power spectrum of the columns of Ũ corresponding to the absolute value squared of each entry of a column of Ũ. We then examine the shape of the power spectrum on a log-log scale. We have performed this analysis on the first 8 columns of Ũ. Figure 7 shows the power spectrum for the first column of Ũ plotted in log-log scale. Over the first 8 columns of Ũ, we observe a consistent linear structure to these power spectra. A linear trend for frequencies plotted in log-log can be indicative of a power law, which would imply long-range correlation over time in a complex system. To investigate possible long-range correlations, we use detrended fluctuation analysis (DFA) (Peng, Havlin, Stanley, & Goldberger, 1995).

DFA is a measure of the relationship between variance within windows of a time series and the size of those windows which, in turn, provides a measure of the amount of long range correlation in a time series (Stergiou & Decker, 2011). DFA has been applied in several areas of cognitive science (and extensively in other fields) as a tool to measure complexity in a time series. Notably, Dotov et al. (2010) use DFA to identify a change in the complexity of motor movements corresponding to a change from smoothly using an interface to cases of "breakdown" where the user interface is perturbed so that it no longer behaves as the user expects it to.

To compute DFA we start with our complex time series:

$$z_1, z_2, \dots, z_n$$

We "center" the data by subtracting the mean

$$\xi_n = z_n - \bar{z}, n = 1, \dots, N$$

and then compute the cumulative sum,

$$C_n = \sum_{j=1}^{n} \xi_j = 1, \dots, N$$

We then partition the time series $C_{(n)}$ into windows of size 4 < s < N. Using the smallest possible value, s = 2, is often not advised (Bryce & Sprague, 2012). In our study, we have set s = 4 as the minimum window size. For a fixed window of width s starting at n, we compute a least-squares regression model satisfying,

$$P_d(t_{n+j-1}) \approx C_{n+j-1}, j = 1, \dots, s$$

and then calculate the residuals,

$$r_i = C_{n+j-1} - P_d(t_{n+j-1}), j = 1, \dots, s$$

Here $P_d(z)$ denotes the fitted polynomial of a degree d (Shao, Gu, Jiang, Zhou, & Sornette, 201). Linear fits are usually used so that in most applications (including ours), *d*=1.

We compute the root mean square of the residual for each window size *s*, to obtain the fluctuation value (Shao et al., 2012),

$$F_s = (\frac{1}{N} \sum_{t=1}^{N} |r_n|^2) \frac{1}{2}$$

Note that we have used the absolute values of the residuals, $|r_n|$ rather than the residuals themselves, since we work with complex-valued time series.

We then fit a line to the relationship between the log-scaled F_s and the log-scaled s (Shao et al., 2012). The slope of this line is α, which is taken to approximate the Hurst parameter H. The Hurst parameter is a measure of fractal dimension in a time series. If $H < 0.5$ the process is considered to be anti-correlated in time such that high values tend to be followed by low values and vice versa. If $H = 0.5$ the process is not correlated in time, and if $1 > H > 0.5$ then the process is said to be positively correlated in time (Ihlen 2012, Nolds Module – 0.5.2 Documentation, n.d.). To interpret α we treat it as H such that an $1 > \alpha > 0.5$ implies positive correlation in time, for example. However, if $\alpha > 1$ the process is non stationary and can be modeled as fractional Brownian motion where the hurst parameter of the systems is approximated by $H = \alpha - 1$ instead of $H = \alpha$ (Hardstone et al., 2012, Nolds Module – 0.5.2 Documentation, n.d.).

DFA is a frequently used to analyze complex systems to determine the amount of long-range correlation in the data which some contend is indicative of the degree of fractal structure in the system. The Hurst parameters provided by DFA were significantly predictive of accuracy $p<.01$. However, it was a much less powerful model with $R^2 = .08$. Virtually all participant's Hurst Parameters indicated some degree of nonstationairity in their fractal structure as all participants were $H > 1$. The negative relationship $B = -.75$ implies that lower performing participants actually have more long range positive correlation in time than high performing participants. This might seem strange at first, given that previous literature suggests that long range correlations are positively related to performance. However, given that the moles' next position is pseudorandom, it could also mean that participants whose Hurst parameters are below 1.5 are actually better approximating the target. In a similar fashion, expert Tetris players have been shown to rely heavily on the rotate button, allowing them to offload some of their cognition (Kirsh & Maglio, 1994). High-performing participants in this task appear to be developing a pattern of interfacing with their environment that is substantially different from how the low-performing participants are interfacing with their environment. By exhibiting a lower fractal dimensionality in their time series, these high-performing participants are using up less of the "real estate" in the two-dimensional playing field and generating more efficient mouse traces that can be characterized in fewer fractal dimensions.

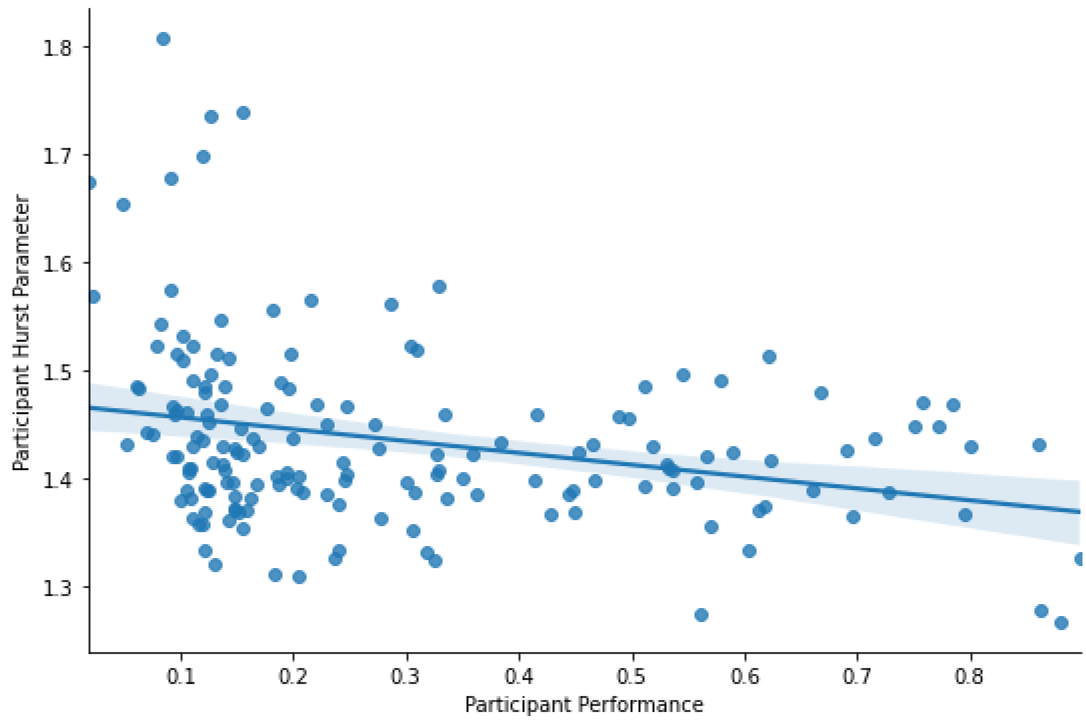

Figure 8. The Hurst Parameter value for each participant's time series regressed on participant performance. The negative slope B = -.75 implies that the fractal dimension of the time series is decreasing as performance goes up.

Discussion

We designed a simple Whac-A-Mole like video game which participants played for a few minutes, during which mouse position data were captured. We did not know what might predict performance in this data but were able to systematically explore it and find quantitative structures that were highly predictive of performance. Our results show that accurate players play differently than inaccurate players and that this difference is detectable in the mouse tracking data. Overall our efforts provide a good case study of how to explore and analyze unconstrained mouse tracking data. Even if a measure of accuracy were not available for a computer mouse task, these results show that it could nicely be approximated by the mouse movement alone as long as a PCA could be generated from known experts. In addition, the DFA results suggest that information exists at the mouse trace level that does not exist at the level of individual mouse movements.

These methods and results provide two contributions to mouse tracking research. First, to our knowledge the only existing mouse tracking analyses that consider whole series of movement are machine learning techniques (Kołakowska 2013; Liu, Fernando, & Rajapakse, 2018), which do not, however, provide interpretable results or generalizable insight into the mouse movements of the participants. In addition, other researchers, such as Hehman, Stolier, and Freeman (2015), have indicated interest in seeing techniques such as PCA, and derivatives of motion analyzed, but as far as we know this has not yet been done. We have provided an initial guide to exploring these types of data in a way that is systematic, and can produce interpretable results.

Second, mouse tracking data naturally lies on a two dimensional plane, but most time series analyses require a one dimensional time series. The problem is usually solved by taking a time series of one dimension, such as the *x* dimension (Schulte-Mecklenbeck, Kühberger, & Johnson, 2019). This solution works for binary forced choice tasks, where participants are making single movements and most of the information exists along one axis. However, even in these situations the one-dimensional analysis ends up neglecting potentially meaningful information in the other dimension, or in the combination of the dimensions. To address this we instead fit the data to the complex plane, which allows us to take 2 dimensional coordinates and express them as single numbers. We are then able to convert our time series directly into the frequency domain. In addition, by doing this we were able to compute the approximate derivatives of our time series easily in the frequency domain. We were also able to maintain integrity in our data while we compute its Fourier transform or calculate its *n*-th derivative. To our knowledge converting mouse coordinates to the complex plane has not been done before.

We also want to clarify that we do not see this as an overhaul of mouse tracking analysis. We are specifically targeting the problem of unconstrained mouse tracking data as at the time we began this work, there were no methods for analyzing it. Conventional mouse tracking methods are valuable and are powerful in simplicity, elegance, and interpretability in work where they are applicable. However, they simply aren't appropriate for unconstrained data, e.g., what might it mean to measure the area under the curve when you have no sure way of knowing where the curve starts and stops and therefore the area measure changes wildly depending on where the scientist decided to place those end points?

Outside of advancing methods in mouse tracking, we believe that our results help to characterize how high performers use their mouse. Our results indicate that behaviors associated with accurate game play produce long-range anti correlations in the mouse movements. This result contrasts with some existing literature, which has found examples where long-range correlation are indicative of health or performance in human systems (Voytek et al., 2015; Hausdorff, 2007; Diniz, 2011). However, we could interpret accurate play as attempting to synchronize spatially with a pseudo-random target. In Stephen and Dixon (2008) participants who were attempting to synchronize with a chaotic metronome were able to coordinate with the metronome globally. Given that our mole is pseudo-random and doesn't have a high-level structure for a participant to map to, it is possible that self-correction could produce anti-correlation in time. Additional hypothesis testing is needed to confirm this. We believe this analysis could serve as a foundation for future research. For example, it may be that more subtle phenomena, such as affective states or engagement, are detectable in this type of data using these types of methods.

In future research we plan to investigate engagement and other affective states using the methods outlined in this paper. Affective states, and a great deal of other information one might be interested in about a participant, will often unfold at temporal scales much longer than individual mouse movements, thus necessitating analyses like the one we have developed here. Indeed, the present analysis was developed with an eye towards engagement. In addition, in practice there are many areas where it is important to be able to infer information about computer users. In almost any natural setting user data can’t be precisely segmented into individual movements, and therefore, any analysis of individual movements, even if it could detect information about affective states, would be useless in natural settings where participants are doing things other than strictly controlled experimental tasks. We also plan to work on further developing this method, and investigating potential additions to this method, such as best practices for analyzing derivatives in mouse tracking data.

Conclusion

Using a simple online game, we studied what information is available in unconstrained mouse traces. We used SVD to analyze the data, which allowed us to systematically explore it, while maintaining interpretability and building toward concrete hypotheses for future experiments. We found that the time series of mouse movements can reveal that accurate players play systematically differently than non-accurate players. In addition, the components of our SVD matrix factorization revealed that components which best-described accurate players had a power law structure. We then applied DFA and found that the behavior of accurate players is indeed characterized by a hurst parameter that differs from that of inaccurate players. These findings also confirm the existence of high-level information in mouse trace data. Thus there could be value in associating this type of data with more subjective and subtle states, such as levels of engagement, motivation, or affect.

4. Phenomenology of Engagement

In this chapter we will build on Dreydeggerean phenomenology to develop an account of engagement. We will begin with a deeper discussion of the important Dreydeggarean concepts surrounding absorbed coping which we will draw on for our analysis: for-the-sake of, ready-to-hand, breakdown and care. We will also use the work of Hatab to help emphasize the breadth of the phenomenon of absorption. Hatab details how Heidegger's emphasis on tools tends to conceal the true nature of the ready-to-hand which is simply prereflective interaction. This broader interpretation of the ready-to-hand sets up our discussion of how anxiety and breakdown are connected, something Heidegger only hinted at (but that is discussed in Ward 2020).

Anxiety has a formal structure that parallels the phenomenon of breakdown, except that it is not associated with any particular entity. We argue that what has broken down in this case is the care structure. This then shows us that the main phenomena of breakdown, which is to leave the prereflective mode of interaction, can be accomplished by changing entities or changing care, suggesting a two-dimensional model of absorption in which stability of care and the proficient use of entities are both important.

The first dimension of absorption is the ability to perform a task (coping, proficient use of entities). This component of absorption is what Heidegger mostly focuses on. The second dimension is the ability of a task to satisfy the cares of Dasein. We call this dimension engagement. We refer to the case where a person is both proficient in using entities and engaged (high on both dimensions) "absorption".

In addition, we see both dimensions as continuous rather than categorical, *contra* Dreyfus' concept of breakdown which suggest that breakdown occurs in a discrete way. We argue that there is no categorical "breakdown", rather there are different degrees of coordination with the entities you use (dimension 1) and degrees of matchedness / alignment between the motivational affordances of the activity and your cares (dimension 2). As you decline along either dimension something like breakdown occurs: in the one case you get worse and worse and using a tool, in the other case you become more and more disengaged.[3]

Dreydeggarean phenomenology is broadly associated with embodied/extended approaches to cognitive science, and we use these connections to derive empirical analogues for these concepts. In doing so we build on work done by Dotov et al. (2010) who developed initial empirical measures of Dreydeggarean coping. They have an online game and measure Dreyfus style breakdown via DFA, comparing the case where the controller works vs. where it breaks. DFA is reduced as the game controller breaks down. We build on this work, with a focus on the engagement dimension of absorption. In subsequent chapters we use this theory to inform our analyses and experimental design. By thinking of engagement as coordination we argue that we should be looking at complexity matching instead of just complexity, a claim we test in the next chapter. Furthermore our interpretation of Heidegger highlights a potential conflict in the work of Dotov et al. (2010) who added white noise to cursor positions to simulate breakdown. An alternative explanation to participants simply entering breakdown is that participants are engaged and still performing the task, but that the appropriate dynamics for the task are now different due to the added white noise. We will test this explanation, and by extension our interpretation of Heidegger in chapter 6, by testing the effect of adding different kinds of noise to cursor positions.

In the final subsection of this chapter, we explain how these ideas link to contemporary theories of psychological engagement, with a particular emphasis on flow theory. Dreyfus has in the past attempted to make connections between absorbed coping and flow theory but we believe

---

[3] We believe this also better matches Heidegger, who identifies three examples of ways dasein can be pushed away from the pre reflective mode of interaction, but none of which really involve "falling out of the activity".

he was unsuccessful, for the reasons outlined in Breivik (2007). However, we think many of the concerns outlined in Breivik (2007) are addressed when we include engagement as a second dimension to absorption. In addition, we think this is important because in our view Dreyfus was correct to notice the phenomenological commonalities that Flow and other theories of engagement share with Dreydeggarean absorption.

Basic Concepts

In this section we survey Dreyfus' reading of Heidegger, the account of skilled coping and how it is disturbed in breakdown, and then an account of care, whereby all our lives are organized around a hierarchy of concerns. Anxiety emerges as a kind of breakdown in our overall care structure which parallels breakdown in skill. We are not orthodox Heidegger or Dreyfus scholars, and freely draw on cognitivist language which both authors rejected, e.g. treating care as a kind of hierarchy of goals, where goals need not be explicitly formulated. We will begin by summarizing the basics of "Dreydeggarean" phenomenology and want to be clear that this is our interpretation and may depart in several ways from Heidegger scholarship.

Heidegger uses dasein to refer to a particular kind of consciousness that he thinks humans have. This kind of consciousness is defined by the fact that it "cares" about what kind of being it has (Heidegger 27). By kind of being here we mean something like the kind of thing you might think of if someone were to ask you "who do you want to be when you grow up", and by care here we mean that dasein is aware that there are multiple choices for its kind of being and is aware that the choices it makes in life affect the kind of being it has, and is therefore always making decisions in the context of who it wants to be (Wrathall 2005, Dreyfus 1990). This kind of being is called the for-the-sake-of-which, because it is that for the sake of which we do everything else (Wrathall 24). As we strive toward a for-the-sake-of-which there are sub goals that need to be fulfilled, like planting a tree ultimately for the sake of being a good parent, those are called “towards which” or “in-order-to” (Wrathall p 25, Dreyfus p 105).

We take a step away from Heidegger to visualize these involvements in a hierarchical network or “care structure”. Where for-the-sake-of-which is the base and the branches down are the subcares of the towards which and the in order to. Though this departs from Heidegger (who would deplore talk of explicit goals let alone a hierarchical organized tree-structure of goals), this allows us to imagine a sort of gradation in care. I might care a little about an activity that is not as connected to my for-the-sake-of-which and more for an activity that is more directly connected. This also gives a grounding for how concepts like immersion or absorbed coping could have a gradation. We will be less immersed in projects that are less strongly connected to our for-the-sake-of-which.

In addition, we view this structure as dynamic. Though the for-the-sake-of-which at the top is mostly stable throughout one’s life, the intermediary goals change continuously as dasein satisfies objectives, or gathers new information that change the objectives. If one’s for-the-sake-of-which were to be a teacher, it would seem trivially true that as new information about effective teaching methods comes out one should be updating how they want to be a teacher. Likewise once one has mastered some particular skill it seems clear that one does not continue wanting to be better at that skill out of context of where that skill is used rather one might change to practicing a new skill that combines well with the former.

This care structure implicitly guides everyday activities such as using hammers, digging holes or even writing. Dasein’s world is a world of involvements where the entities in the world get their being from the care structure of dasein (Dreyfus 97). When dasein is building a treehouse to be a good parent and they need to hammer some nails, the wooden stick with the 22oz of metal on the end is meaningful as a tool for hammering nails and is seen that way. Dasein doesn’t even attend to the hammer's physical properties, it just uses the hammer to hammer the nails. The hammer when it is being used in this way is ready-to-hand, it is available to us to be used for our project and requires little or no extra thinking on our part.

If the hammer were to break or become unavailable to us for whatever reason, it would become “unready-to-hand” (Dreyfus 70). This transition from an entity being available to us, to becoming unready-to-hand, is what Dreyfus calls breakdown as it represents a breakdown in our everyday coping with objects. When this happens the equipment stops being something we can just use, and instead the physical properties of the unready-to-hand entity “obtrude” as we are forced to attend to the entity (Dreyfus 71-73). Instead of just using the object as a hammer we attend to its too short handle, or too heavy weight. The hammer becomes similar to the present at hand.

Dreyfus added coping to the distinction between ready-to-hand and unready-to-hand, noticing that it's not just that a hammer may break, but also that we may be bad at using a hammer. Absorbed coping as Dreyfus calls it is using something to perform the task at hand, but being skilled enough that we are not distracted by the work, we are able to simply be absorbed in the project (Dreyfus lectures).

Present at hand entities are those which are not at all connected to our care structure. For present at hand entities dasein only attends to the physical properties (Dreyfus). For example, if the hammer were not part of our care structure at all, say dasein lived in a post nuclear apocalypse where humans are not building things with hammers, and dasein were to discover a hammer, to this dasein the hammer would be merely a physical object present in the environment, just a piece of wood with a heavy chunk of metal at one end (Wrathall p 25).

To summarize, we have entities that are present at hand, who are not connected to our care structure, and thus we interact with those entities simply through their physical nature. We also have entities that are ready-to-hand that are directly connected to what we care about. These entities we simply use as their being is defined by dasein’s care structure, and their physical properties are “transparent” to us. When ready-to-hand entities break they become something in between the ready-to-hand and the present at hand, where we are forced to attend their properties but still recognize their being in relation to our care structure. However, the more we are forced to focus on the entity and our ability to use it breaks down, the more the entity becomes like something present at hand.

Hatab's new vocabulary and focus

Hatab (2018) has developed ideas similar to the ones we will, describing a new, more unified, and open vocabulary for Heidegger. First Hatab clarifies Hedigger's concept of ready-to-hand. Hatab notes that the concept does not only apply to tools. When "explicating [ready-to-hand] Heidegger talks of 'concernful absorption' … which I take to be central to what [ready-to-hand] is meant to open up". Heidegger's concept of absorption which is shown in the ready to the hand goes beyond just entities. He says: "Heidegger usually talks of *Zuhandenheit* in terms of "entities" (ga 2: 117/sz 88), and here tools certainly apply. But his own concept of absorption is something distinct from entities per se." According to Hatab, this absorption can occur with non entities like writing. I think Heidegger would agree with this, though as Hatab says it is under emphasized. For example, Heidegger says that the towards which and in order to also bear the character of the ready-to-hand, even though it is clear that some towards which could be non-entities.

Hatab introduces his own concept of immersion that he thinks is a better term for what Heidegger really means by the ready-to-hand. Hatab defines immersion as "simply non-reflexive performance or intimation without directed attention or analysis". Hatab explicitly says: "In Basic Problems of Phenomenology, *Zuhandenheit* extends to the whole milieu of concernful dealings, including things like house and yard, forest, sun, light, and heat (ga 24: 153/108, 431–45/303–313). It seems to me that engaged immersion better extends to such phenomena as modes of pre-reflective dwelling than does the sense of practical usage that I believe is overemphasized in

Heidegger's discussion of *Zeug* and *Zuhandenheit*." Specifically Hatab thinks that "Heidegger's emphasis on tool use tends to conceal" this "richer sense of *Zuhandenheit*" where it applies to anything we do "pre-reflectively", and this is "opened up by [using the term] immersion" (p 4).

In contrast to immersion, Hatab uses the term "exposition" to describe the kind of thinking we do when we are reflecting on things. This parallels what Heidegger calls disclosing. In exposition we are not simply using things, but thinking about them, taking note of their properties, etc. Again Hatab is focusing on a broader phenomenon than is sometimes interpreted from Heidegger. Here exposition is not just focusing on tools or entities, but rather on anything that we can reflect on and think about separate from simply doing it. In addition, similar to Heidegger, Hatab points out that exposition is not the default mode, and that we enter into exposition by undergoing breakdown or as he calls it contraposition.

Generally Hatab's terminology emphasizes broader concepts that are less linked to tools and entities specifically and are more focused on reflective and pre-reflective modes of interaction. These broader interpretations fit well with our own work here in explicating the phenomenology of engagement. The broader concept of immersion aligns with our view of transparency, and how engagement and skill are both important underlying variables in the phenomenon of transparency, which we will detail later. Further, because of his focus on broader interpretations of Heidegger, Hatab also finds that "contraposition extends all the way to anxiety", which supports our interpretation of a parallel between breakdown and anxiety.

How anxiety parallels break down

Heidegger distinguishes anxiety from fear, where fear is more directed and local and anxiety is more global and holistic (Dreyfus, Heidegger 231). The pervasive nature of anxiety, which dissolves all involvements into a kind of existential nothingness, sheds light on the pervasive nature of care, which encompasses all we do and the network of involvements associated with dasein's "world" (Heidegger (231) [186], Dreyfus 178). For Heidegger, fear is when we are perceiving some direct threat. Heidegger says "our Interpretation of fear as a state-of-mind has shown that in each case that in the face of which we fear is a detrimental entity within-the-world." We might fear a tiger as a detrimental entity (Heidgger 231).

Anxiety on the other hand, though it may feel like fear does not happen when there is some present danger. As Heidegger says, "that in the face of which one has anxiety is not an entity within-the-world. Thus it is essentially incapable of having an involvement" (Heidegger 231). The fact that anxiety occurs when there is nothing to fear, tells us that anxiety does not have to do with some particular entities, and that anxiety as a phenomenon does not have "an involvement" as Heidegger puts it. What he means here is that unlike other phenomena like equipment and the ready-to-hand, anxiety is not defined by its connection to other aspects of dasein. ready-to-hand is connected to our projects. Fear is connected to entities that are detrimental to us. But anxiety is unique in that it is not connected to anything else because it doesn't have involvements. Heidegger says

> That in the face of which one is anxious is completely indefinite. Not only does this indefiniteness leave factically undecided which entity within-the-world is threatening us, but it also tells us that entities within-the-world are not 'relevant' at all. Nothing which is ready-to-hand or present-at-hand within the world functions as that in the face of which anxiety is anxious. (p 231)

Anxiety's lack of involvements may actually be what gives rise to it. When we are anxious we are experiencing something that is not involved, and as such it reveals to us a lack of involvements. As Heidegger puts it, "the 'It is nothing and nowhere' becomes manifest" (231). Heidegger goes on to say, "the obstinacy of the 'nothing and nowhere within-the-world' means as a phenomenon that the world as such is that in the face of which one has anxiety." The "obstinacy" with which the nothingness persists shows us that it is the world of nothingness, or the world of no involvements that causes anxiety.

Recall that obstinacy is one of the forms of break down. It is when some other entity is in the way causing you to lose the ready-to-hand, which is dependent on involvements (Dreyfus 72). Thus when we experience anxiety we lose access to the ready-to-hand as the "nothingness" is obstinately standing in the way of involvements that give rise to the ready-to-hand. Since the involvements that give rise to the ready-to-hand are determined by the projects of dasein, then when the nothingness precludes us from experiencing entities as ready-to-hand, this must be related to dasein's projects, or lack there-of. Heidegger goes on to say that his analysis of anxiety shows us that "the referential totality of significance (which as such is constitutive for worldhood) has been 'tied up' with a 'for-the-sake-of-which'." In other words Heidegger says that his analysis of anxiety shows that the dasein's world is connected to its care. To us this also implies that changes to care can affect dasein's world, which our previous discussion of the ready-to-hand did not show us. If dasein can become anxious, and anxiety is reflective of a change in care that undermines the involvements of the world, then it would seem to imply that just as our experience changes from when objects become unusable and our coping "breaks down" our experience of the world changes in a parallel way when we change or our care structure changes.

The ready-to-hand are entities we as dasein are using to accomplish our current project. The current project is some node at the bottom of our care structure. Our interactions with objects are then defined by that tree diagram. This is what Heidegger means when he says we are care. Everything else that we experience phenomenologically is determined by its connection to, or lack of connection with our care structure. Whether we experience some rod with a lump of metal on the end as a hammer, a club, or something in the way is dependent on what we are trying to do at the moment, what we care about. And all of this is connected by some number of degrees to our for-the-sake-of-which, that ultimate way of being we are striving towards.

Breakdown occurs when the entity no longer matches what we care about. If the hammer breaks such that it can no longer be used for us to accomplish our goal, the hammer loses its ready-to-hand because it can no longer be connected to our care structure, it is now like a present at hand entity. When an entity we need for the work is missing, the present at hand obtrudes, masking the involvements that make the ready-to-hand ready-to-hand. And when some other entity is blocking us from accomplishing our work, again the ready-to-hand loses its significance because we can longer work on the project that gives it its significance. The ready-to-hand becomes like the present at hand until we can clear the blockage and resume our work.

The parallel to anxiety is now obvious. While break down occurs when in the face of some entity, anxiety occurs with no involvements, in face of nothing. If breakdown occurs as a result of something in the world happening, anxiety seems to come from the other direction. It is dasein who has somehow changed or is seeing things differently and thus produces something that is phenomenologically similar to break down. In our way of thinking, if the absorption or immersion as Hatab would call it is created by the connection between something and the care hierarchical network, then that absorption can deteriorate from changes to the object (breakdown) or changes to the care structure (disengagement).

Anxiety happens when we find that we are living inauthentically and thus the projects we have been doing are not our own and not attached to our care structure anymore. Thus we experience a loss where we feel as though most of the world we are currently in, which was defined by some care structure we no longer are, is not our own, we are not connected to it, and everything becomes like the present-at-hand.

Anxiety shows us that everything is grounded in our care. We already saw that the ready-to-hand and breakdown show us that the “referential totality” is related to our care structure. Anxiety shows us that it is in fact fully depended upon it and when that care structure is radically altered in some way, like realizing that one has been absorbed in the world of the they and their cares are not their own, the ready-to-hand evaporates along with the previous care structure.

Thus we are able to see everything is connected to care. Heidegger goes so far as to say "dasein is care". Dasein's experience of the world is fundamentally tied to dasein's care structure. The care structure underlies the referential totality that is dasein's "world". Changes to this care structure, or to the entities in the world and their connection to the care structure alter the referential totality, the network of connections, and produce phenomenological changes to dasein's experience.

Engagement as coordination between cares and tasks.

In the literature on engagement there are many ways engagement is operationalized, with some researchers treating it as synonymous with flow, while others use it to mean the number of clicks per hour, or the number of extracurricular minutes invested in a project (Khan & Ahmad, 2021; Shernoff, Csikszentmihalyi, Schnieder & Shernoff 2014; Axelson & Flick, 2010). But, in a sentence engagement is generally thought of as how much a person's full cognitive resources are dedicated to a task, and this is generally thought to be related to how rewarding the task is, how challenging it is, etc. (Schaufeli, 2013; Axelson & Flick 2010). While engagement is hard to define, or behaviorally operationalize, boredom is often used as a foil for engagement, such that engagement is measured as the opposite of boredom, or experiments will be designed to place participants in an engaging condition vs a boring condition (Ulrich et al., 2014).

While engagement is not defined in Dreydeggarean, boredom is addressed and thus we similarly will use boredom as our concrete starting place when thinking about engagement. Heidegger uses an example of a woman sitting at a train station waiting for a train (Ciocan 2010). The woman has a book but she doesn't feel like reading. In every other way the woman is simply stuck waiting for the train. In this situation the woman has full access to her care structure, she cares about where she is going and why, but has no access to any activities that are relevant to her care structure. Interestingly, the phenomenology Heidegger lays out here strongly resembles that of breakdown and anxiety. In the case of the bored woman, she notices the physical properties of the things in her environment the way she would something present at hand (Ciocan 2010). She also notices the flow of time almost as a present at hand entity, feeling the seconds tick by (Ciocan 2010).

Again we see that the world dasein experiences is related to the care structure. Boredom emphasizes how dynamic that relationship is. In particular, we can see that only when our care structure and environment/ability are well matched do we see the kind of everyday coping that Dreyfus and Heidegger talk about. We can also see that changes to the environment such as breaking a tool or breaking our cares such as a moral epiphany produce similar phenomenology.

As we stated earlier, we believe the care structure to be a dynamic thing, where the for-the-sake-of-which can change, but is usually stable, and the lower order cares on the structure being more variable. This allows us to think about breakdown in a more dynamic way. The type of breakdown focused on by Dreyfus and Heidegger focuses on changes to the entities in the world. Sometimes those entities change in such a way that they interrupt our ability to perform a task, in which case we see a breakdown in dasein's ability to cope. In this side of breakdown, the changes happen externally to dasein, the hammer breaks, or goes missing, etc. However, we feel that it is equally true that the care structure can change such that a task becomes meaningless or unnecessary. Anxiety is the extreme example of where the for-the-sake-of-which itself has changed such that projects dasein was involved with have become unnecessary, and thus dasein's whole world becomes like the hammer that is unready-to-hand. Boredom represents a similar change but a smaller scale where dasein's current cares are unattainable at the moment.

We want to emphasize that similar to the breakdown described by Dreyfus and Heidegger, phenomenologically what happens in boredom and anxiety is an experience of a loss of meaningfulness, only instead of it being towards an individual entity, it is directed at all of the entities as it is the care structure, which the network of involvements is based on, that changes. This reveals a new point of focus: the ability to satisfy cares which is an additional dimension, creating a two dimensional system which we simply call absorption. Dimension one is the ability to perform the task, this is the dimension focused on by Dreyfus and Heidegger, and Dimension two is the ability of the task to satisfy the cares of dasein.

We think of this second dimension as engagement. We described engagement earlier as using all of one's cognitive resources on a task. This is deliberately similar to the way that Dreyfus described absorbed coping, which Dreyfus thought was somewhat synonymous with flow, and as we have already stated flow is generally thought to be similar to if not somewhat synonymous with engagement (Dreyfus 2007). Thus we think it is fair to try to examine engagement as a state, similar to absorbed coping, but to recognize that there are two ways that absorption can breakdown, while Dreyfus only emphasized one's ability to perform the task (similar to flow). We will use absorbed coping as Dreyfus intended to describe one's ability to perform the task (skill, usability of equipment, etc.), and we will use engagement to describe how well a task is matched to the care structure of dasein. Since this is now a problem with 2 continuous dimensions we can think of each quadrant as a somewhat discrete case for simplicity and we get four cases: the unengaged and not coping, the coping but not engaged, the engaged but not coping, and the coping and engaged.

In the first case, where the task is not connected to dasein's care structure and is not a task dasein is interested in, we expect low engagement and low coping. Thinking of this always reminds me of times when I was bored in PE class, and I started attending to the greenness of the grass, the shape of the clouds or the shininess of my Mr. Teacher's bald spot. As compared to case 2 where one is unengaged but is coping well. This is more like brushing my teeth, where the tooth brush is transparent to me, but the task is not. In fact I have to continuously remind myself to keep doing it so that I don't get bored and just stand there with a toothbrush hanging out my mouth while I think about something else. In case three, high engagement but low coping, could occur when we are learning to do something new that we are interested in, and in this case we would expect nearly the opposite. When we are learning something new that we care a lot about, we don't have to attend to performing the tasks, or try to keep our mind on task, in a sense we get absorbed in the activity, but the particular tools or skills are not transparent to us. Instead I am completely absorbed in thinking about how to get my leg to move exactly how I want it to, to kick the ball. Finally, when we are both engaged and coping well we get full absorption where we are completely immersed in the activity, similar to the flow state as we mentioned earlier.

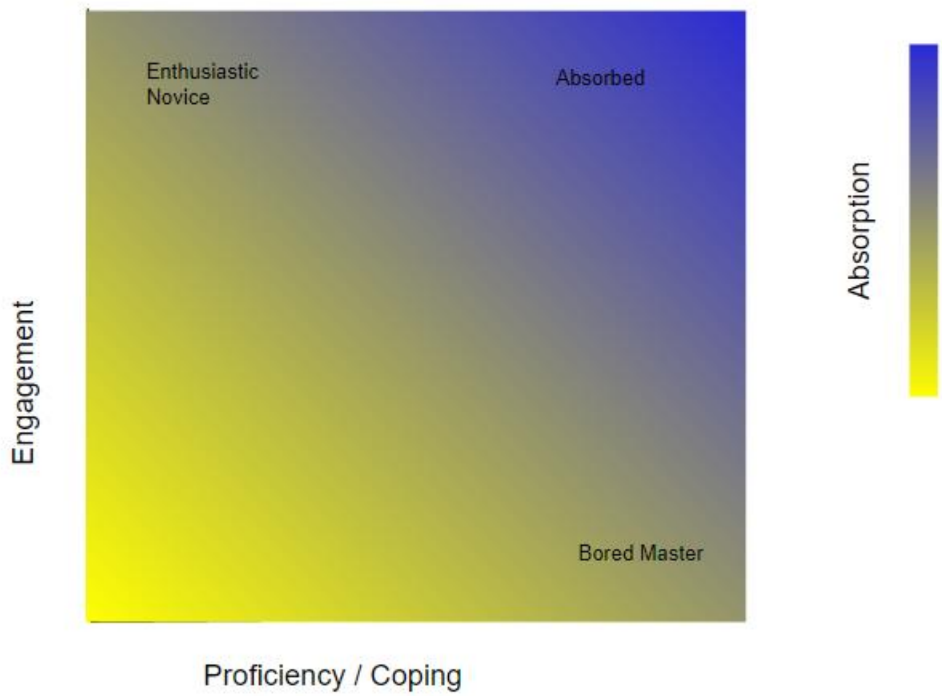

Figure 9. Two-dimensional Absorption. The amount of absorption is depicted with color where yellow is low absorption and blue is high absorption.

Empirical Analogue

When entities are ready-to-hand they are available to us to just use them, and we don't consider their properties. Dotov et al., (2010) give a complex systems account of this and show that complexity at the body tool boundary decreases when entities become unready-to-hand. In this case the body is coordinated with artifacts, and for Dotov et al., (2010) that implies that the artifact is part of the cognitive system at that time.

Given that Dotov et al., (2010) showed that when ready-to-hand an entity is integrated into the human experience by connection to our complex system, and that in breakdown we can see that connection decrease, we might expect that a similar phenomenon like anxiety would have some kind of similar measure. However, anxiety operates at a larger scale, and thus we would expect larger shifts or shifts that are not connected to a single entity, but something like a system of entities such as an entire project/task. Here in particular we think it is useful to think about the concept of coordination, for while we do not “coordinate” with individual static objects, in a mutual sense, it might be reasonable to say that we coordinate with the dynamics of the activity we are trying to perform, and thus we are coordinating with all of the entities involved as they work together as a system. We are coordinating with the combination of entities in the game or with the other players on the team or in the soccer game.

These ideas can be extended to the case of engagement, where cares are aligned with tasks. But cares and tasks are abstract. To make this concrete, we measure something in the body of the person with cares, and something in the task they are engaged in, and determine the amount of coordination between.

At a high level this is extremely intuitive. Many of us would assume that we can tell by watching someone if they are engaged in what they are doing or if they are bored. An example is watching young kids play soccer. In soccer we would expect a person's bodily movement to correlate or even synergize with the other players (Vilar, Araújo, Davids, & Bar-Yam, 2013; Araújo, Silva, & Davids, 2015; Araújo & Davids, (2016); Araújo, Ramos, & Lopes, 2016). As Duarte et al., (2013) demonstrated, soccer players on the same team form a synergy. In younger children interest in the game varies, and so we would expect to see significant variability in how much certain players are part of the synergy. Some children might be picking flowers, while others chase the ball, while others attempt to play their position and be open for passes. Other examples might include people playing video games, having a conversation, even performing or watching a performance.

In phenomenological terms if the cares and motivational affordances of the task align we expect greater absorption into the task. In folk psychological terms this corresponds well to engagement, in which case, engagement is then a function of how connected a task is to a Dasein's care structure. Is the task closely related to that for the sake of which dasein lives (engagement) or are there many degrees of separation or weak connections between the task and that for the sake of which it lives (disengaged)? In complex systems terms then engagement predicts more spatiotemporal coordination between people (e.g., a single soccer player) and the tasks they are engaged in (the soccer game as a whole). A person who does not care about a video game they play will not correlate well with the game entities, as they lazily push the controls around. A person engaged with the game will make finger movements that correlate well with the dynamics of the entities in the game, and a person who is engaged and good at the game may use a different/higher form of coordination, like forming a synergy with the entities in the game. For example, expert tetris players rotate pieces on screen, offloading the mental rotation to the game. Though this wasn't discussed as a synergy in the original article it seems to be a good example of player integration with game dynamics (Kirsh & Maglio, 1994).

Applications of the theory

It is also worth considering that our description of engagement, and our theoretical structure predict the flow phenomenon. However, rather than predicting the phenomenology of flow to be purely dependent on a balance between skill and challenge, we predict that it will be dependent upon the combination of how much the motivational affordances of a task relate to our care structure and how adept we are at coping with the environmental variables involved. The contribution of genuinely caring about the activity is well documented in flow but is accounted for by the "autotelic personality" (Baumann, 2021). Our theory would seem to predict that rather than that people have autotelic personalities, some tasks are simply more connected to the person's care structure and thereby more interesting to that person, and autotelic personality will be more task dependent than person dependent. There is mixed support for this with Mosing et al., (2012) finding that some parts of flow proneness were consistent across domains (work, maintenance, and leisure), while others were domain specific.

One additional benefit of our account is that it addresses some of the issues philosophers have with Dreyfus's attempts to extend coping to flow. Breivik (2007) specifically targets Dreyfus' model of absorbed coping. First, they point out that Dreyfus model of absorbed coping seems to leave out how people learn skills initially in which we are very actively attending to the thing we are using whether it be an object or our limb. Breivik (2007) specifically uses children as an example when they say: "when children are in a learning situation they often are consciously and deliberately trying to perform the intended movements. They consciously try to get better, to learn to perform well, and pick up new skills by mindful imitation." In these situations entities being used to acquire the new skills are not something we can be transparently coping with as they require our direct attention. However, they also do not appear to be simple present at hand entities. Breivik (2007) is pointing out that there seems to be a gap in which some entities are not something we can transparently cope with because are forced to focus on them, these entities are also not simply ready at hand for the child to use as the child is still learning to use them, and they are not present at hand as they clearly have meaning. Breivik (2007) goes on to compound the point saying that "Even the Heideggerian carpenters are not totally absorbed in their activities where the equipment disappears into the equipmental context and only the work to be done is in focus. The carpenter needs to design the house, find a suitable place to build it, think about materials, and collect the necessary tools and equipment. All this is very thematic, very conscious, very deliberate, and even self-referential." Here Breivik (2007) is really arguing for a more dynamic, flexible model of the referential totality. A model we fully support and think is dynamically shaped by a dynamic care structure.

Breivik (2007) also points out that even experts can actively focus on entities that would be ready-to-hand if they want to. Breivik (2007) believes this stands as evidence against Dreyfus ideas of absorbed coping, because in this case, in spite of mastery, masters can still engage in non-absorbed use of things. Breivik (2007) specifically uses the example of expert runner Bernd Heinrich who in an interview states: “I then focused attention on my calves, thighs, arms, trying to relax them even during training runs, so that the most essential running muscles would be exercised.” Here the expert runner is clearly actively focusing on their body and are not simply absorbed in running.

Finally Breivik (2007) argues that flow is not a state of absorbed coping. “But normally hammering with a hammer or performing normal daily activities are not experienced as states of flow. Here I think Dreyfus is on a wrong path. States of flow are reached only under special circumstances, like when our minds are focused, when we perform well and experience a sort of emotional elation” (pg 8). This shows that even though flow and absorbed coping are similar, they are quite clearly distinct. This distinction again shows that the absorption that both flow and absorbed coping have in common is not purely dependent on skill level because flow can occur even in lower skill level people as long as their mount of challenge is well matched, and because flow generally is tinged with an emotional reaction that is not fully accounted for just by having someone do something they are good at.

However, it is hard to simply dismiss Dreyfus's intuition that flow is related to absorbed coping. The phenomena are too close for coincidence one might say. We believe that by simply adding the relationship between care and task, we are able to provide the rest of the puzzle that when combined with Dreyfus' idea of absorbed coping becomes something that is phenomenologically like flow. Though we of course would not argue that it is the flow model, rather it is maybe a similar but competing model of how the phenomenological state of flow arises.

Furthermore, by allowing the care structure to be dynamic we also are able to address Breivik's issues with coping. It is not that we can't ever be absorbed and not focus on our limbs, but our web of involvements which are the ready-to-hand or the things we are absorbedly coping with can change dynamically depending on what our concern is at present. To address the running example, if our goal were to run to a place so that we could be on time for a meeting we wouldn't think about our body (unless its as out of shape as mine and goes into break down). However, if we were a racer and in our care was to win the race a subgoal might be to run as efficiently as possible, in which case it would make sense for us to focus on our legs and actively control those muscles. However, in that case the person absorbed in running is not wind wandering, or simply unable to see past the obtrusive physical properties, they are engaged, but not purely absorbed coping, which we think Breivik (2007) might agree with.

Throughout the rest of this document we will use the theory we have established here to inform our analyses and experimental design. By thinking of engagement as coordination we realized that we should be looking at complexity matching instead of just complexity, which we will test in the next chapter. Furthermore, our interpretation of Heidegger highlights a potential conflict in the work of Dotov et al. (2010) who added white noise to cursor positions to simulate breakdown. An alternative explanation to participants simply entering breakdown is that participants are engaged and still performing the task, but that the appropriate dynamics for the task are now different due to the added white noise. We will test this explanation, and by extension our interpretation of Heidegger in chapter 6, by testing the effect of adding different kinds of noise to cursor positions.

5. Complexity Matching and Engagement

As we said in chapter 3, most mouse tracking studies today use very controlled conditions to get testable data. Probably the most common mouse tracking paradigm is forced choice tasks where the user must move their mouse to select between one of two or more options, such as the task described in (Maldonado, Dunbar, & Chemla, 2019; Rheem, Verma, & Becker, 2018; Kieslich, & Henninger, 2017; Kieslich, Henninger, Wulff, Haslbeck, & Schulte-Mecklenbeck, 2019; Stillman, Shen, & Ferguson, 2018; Hehman, Stolier, & Freeman, 2015). In this kind of task the cursor movement start and end positions are determined, and each movement is thus a single discrete movement that is easily distinguished from other movements. This affords the use of analysis of conventional mouse trace descriptives such as curvature in mouse movement, xflips, etc. However, when the movements are less strictly controlled, as is the case in almost all normal use of computers, it is often difficult to impossible to tell how to segment the total mouse movement into individual mouse movements and then most of these analyses are untenable (Calcagnì, Lombardi, D'Alessandro, & Freuli, 2019).

In this chapter we will reapply and extend the reconstruction (SVD) and multiscale (DFA) methods we developed in chapter 3 to analyze improvement in performance and engagement. Improvement in performance is still concretely measurable but is a more subtle variable that is thought to be related to engagement. To do this we will first characterize the performance time series of each player. Then we will use our previous methods from complexity science and linear algebra to mathematically describe how improvers play. Finally, we will use what we learn from exploring the way improvers play to formalize a structural equation model of engagement.

Our previous work in mouse tracking made progress on analyzing unconstrained mouse traces. First rather than only using the x or y dimension, we preserve information from both dimensions by mapping the mouse coordinates to the complex plane, thus converting a time series in a 2-dimensional space to a one dimensional time-series which the complexity algorithms we deal with can accept. Second, in more natural mouse trace data, the mouse traces are much different from the movements studied in traditional mouse tracking. There are no simple beginning and end points to delineate individual movements relative to which curvature could be defined. We could try to partition the motion into movements, but there is no natural principle for doing that. By utilizing techniques that look at the whole mouse traces instead of focusing on individual movements we are avoiding the arbitrary decisions that would be necessary to press this kind of data into a more traditional paradigm. To do this we applied a technique from complex systems research, Detrended Fluctuation Analysis (DFA) (Peng et al., 1995). We also updated DFA to be usable with a complex valued time series. We found that DFA significantly predicted performance in our task.

Finding that DFA is predictive of performance demonstrates the value of looking at mouse trace data from a complex systems perspective. Complex systems science has some explanatory power when it comes to mouse trace data. DFA gives us a sense of the multiscale structure of the movements instead of focusing on exactly how individual movements are made. This shows that even when individual mouse movements can’t be isolated, there is still information available for analysis at the global level for mouse traces.

Though our previous results imply the potential of looking at this type of mouse trace data, they don’t demonstrate its explanatory power as performance is not as subtle of a target variable as something like affective state. Performance was ideal as an outcome variable for exploratory work due to its objective nature and its precise measurement. However, a more subjective variable, that can't be tested with conventional methods, would make for a good use case of our methods. Here we will begin work on investigating engagement, using and building on the methods we have so far developed.

As discussed in chapter 2, engagement is a complicated and subtle variable but is gaining interest in a variety of settings (Bakker, Albrecht, & Leiter, 2011; Porter 2006). Schools, businesses, pretty much anywhere that it is important that people interact with something external, the engagement of the people involved is of interest. As technology has progressed the potential for understanding this variable has excited a massive growth in research on the topic, with research articles discussing engagement growing by an order of magnitude since 2000 (see Fig 1).

The most common way to measure engagement is using questionnaires such as the User Experience (UEX) scale Wiebe et al. (2014). The next most common way to measure engagement is with physiological data, and several physiological markers for engagement have been found (Katahira et al., 2018; Nacke and Lindley 2008; Darnell and Krieg 2019). But behavioral data is one of the least common domains for detecting engagement but research on discovering behavioral measures of engagement is common in affective computing research. Affective computing research often targets engagement using cursor, keyboard, and webcam data. Kołakowska (2013) is a meta-analysis of affective computing research using (mostly) machine learning to classify engagement and other emotions in cursor and keytapping data.

It is noticeable in the literature that engagement research lacks a widely applicable behavioral measure. The affective computing measures of engagement are highly context specific and often meaningless outside of the particular task they are developed for. The common laboratory physiological measures are constrained to use in a lab because they require specialized equipment and monitoring systems. And the questionnaires necessarily remove a participant from the state of engagement to fill out the questionnaire. There is an obvious gap in the methodology, where a more generic behavioral measure that could be repeatedly applied in broad contexts would be useful.

One mainstay of engagement is that a large part of its attractiveness is that it is supposed to mediate the relationship between performance and the factors that improve performance such as practice or access to resources. For example, Dotterer and Lowe (2011) found that "psychological and behavioral engagement mediated the link between classroom context and academic achievement". In work engagement Schneider et al., (2018) showed that engagement mediates the connection between organization aspects such as "company organizational practices" and job performance. We think a simple intuitive interpretation of these findings that could be applied to our work is that engagement is related to improvement in performance, that people who are engaged in what they are doing will improve at it more than people who are not engaged in doing it. For example, in the student engagement model, it is thought that students who are more engaged process the material more, and thus learn/improve more (Lee, 2014). Thus it seems like a step in the right direction for finding a mouse trace measure of engagement would be to apply our previously developed methods to improvement in performance instead of just performance.

In addition, in the previous chapter we outlined a new model of engagement. In this model we defined engagement to be the matchedness between a person's cares and the motivational affordances of the task. We determined that when people are engaged with a task they should adjust their spatiotemporal dynamics to be appropriate for the task, and it is this coordination of spatiotemporal dynamics in the activity that produces the phenomenon of absorption into the activity.

This implies that spatiotemporal coordination should be related to the absorption aspect of engagement. In our task we expect that participants who are engaged will produce dynamics that are appropriate for the task, even if they are not necessarily performing well. This study is exploratory and thus we don't formulate any explicit hypotheses here about how exactly that will manifest. However, based on our previous work in mouse tracking we expect that the multiscale structure and the global structure of the time series will be important indicators of the task appropriateness of a participant's overall dynamics. We will then compare those measures we previously developed to participant improvement and engagement (measured with a questionnaire).

Methods

Design

In this study we explore cursor data during play to find dynamics that relate to improvement in performance. To do this we first designed a whack a mole like game that participants play for 3 minutes. During the game we collected their cursor data, which we then compared and conducted several exploratory analyses on. As we have argued, in previous papers with continuous mouse tracking data like this, standard measures such as reaction time, and curvature are untenable, and thus we begin our exploration with PCA, and signal processing to get a sense of general trends that players follow. Based on our findings we proceed to use DFA, a measure of fractality, to analyze scaling structure in the data as we believe that might be related to engagement, operationalized in this case as improvement in performance. In the following subsection we will describe the task, participants, and analyses in much greater detail.

Task

This game is based on the game used in Ch 3. There are 8 preset mole hills which appear on the screen and stay fixed throughout the game. Then a cartoon mole will appear in one of the mole hills and will stay there until they are clicked on or for a maximum of 1200ms. This timing was determined in piloting because it seemed to best split participants where 50% accuracy was the median value for participants. If the mole was clicked a cartoon mole that appeared incapacitated was shown over the original cartoon mole for 350ms. After the mole was clicked or the time ran out and it disappeared, a new mole would appear in a randomly selected mole hill.

The sequence of mole locations was randomly generated ahead of time so that all participants experienced the same sequence. At the end of the task participants were asked a series of questions, thanked, and debriefed. The game and questions are available online at the Yoshimi Lab website and the source code is available upon request.

Participants

We recruited our participants on mturk. We filtered mturkers based on their location, age, rating, and if they had ever participated in one of our studies before. All participants fully consented to the research before beginning the experiment and were treated in accordance with IRB. We collected 150 participants. We then eliminated participants if they failed to answer our catch question correctly, or if they selected an input device other than the mouse. This left us with 82 participants.

Data

From each participant we collected their mouse coordinates (time stamped), clicks (time stamped), their answers to the questions, and the participants' score as a time series of ratios. The ratio was simply the ratio of clicked moles out of the total number of moles shown.

We performed several steps of data pre-processing. First, browsers only poll the mouse when it's moving and thus we had to forward fill (taking the last member of the series and using it to fill in subsequent gaps) for times when the mouse was sitting still and to fill in large gaps in time between mouse locations. We also interpolated and resampled all participants to have exactly 9053 data points. Once the mouse movements were forward filled and linearly interpolated, we transformed the data into the frequency domain. To do this we first made complex valued coordinates out of the real valued coordinate pairs such that $c_t = x_t + i * y_t$.

We also had to create a measure of improvement during the task. We used least squares to fit a line to the time series of scores for each participant. We then used the slope of that line as a measure for improvement.

Analysis and Results

Our goal in this research was to better understand how engaged players might play differently than unengaged players. To begin we start out analyzing improvement instead of engagement because the two are thought to be related and we thought that during exploratory work, a single concrete variable would be easier to analyze. The first step we took was to use our SVD analysis from mouse tracking study 1, to see if improved players have any particular spatiotemporal structure to their play. We created a space of improved players from the first from the first half of the sample and found that in the second half of the sample, fit to the space of improved play was predictive of improvement. We took this as evidence that there was a spatiotemporal structure to improvement. We proceeded to check if complexity was related to improvement, as it was for accuracy, and we did not find evidence to support this. Then we tried complexity matching instead of complexity as that better fits with our overall theory of engagement as described in Chapter 4. We did find evidence for this. Finally, we created a structural equation model of engagement predicting complexity matching as fits our overall theory of engagement being spatiotemporal coordination rather instead of just complexity, and we found that self-reported engagement did predict complexity matching. In the following section we go through each of these results in more detail.

SVD of Improvement

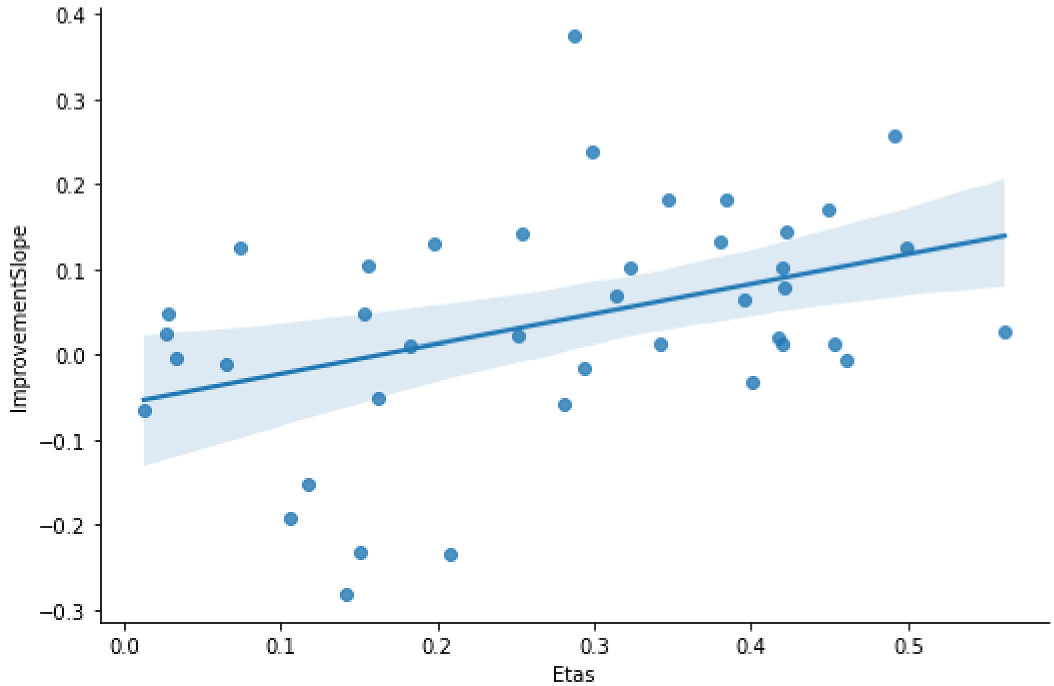

Figure 10. The degree to which improvement is predicted by fit to the space of improved players. This can be interpreted as how much play style actually predicts improvement, or how much a participants improvement is actually predicted by how much they play like an "improver".

The first analysis we conducted was a test to determine if players who improve have a particular structure to the way they play. We tested this by splitting the data in half, into a training and a test sample. Then we used the most improved players, the players with an improvement slope above .1, meaning that their accuracy went up by more than 10% over the course of the game, from the training half of the sample as our target population. We took the singular value decomposition of the frequency domain representations of their mouse movements. We then regressed all of the participants from the test sample onto the space created from most important components in the the singular value decomposition. In figure 10, you can see that there is a linear relationship (*B*=.35) between a player's fit to the space of improved players, and that player's own improvement. The *SE B* = .13 and the R^2 = .16 *p* = .009 , which we take to be sufficient evidence for a relationship between improvement and fit to the improvement space. In other words, how much someone actually improves during the game is related to how much they play like an improver.

However, we also want to clearly state that in subsequent random splits this finding was not always stable. We attribute this to having too little data. As there are only sixteen improvers in the entire sample, some of the splits do not split that evenly, and some of the random splits contain as few as three improvers. We think that in a larger sample this result would stabilize, but want to be clear that this should be taken with a grain of salt.

Hurst Parameter of Improvement

In chapter 3 we found that the Hurst parameter was predictive of performance. Therefore, our next step in looking at this data was to calculate participants' Hurst parameters with DFA and see if there is any evidence of Hurst parameters being predictive of improvement.

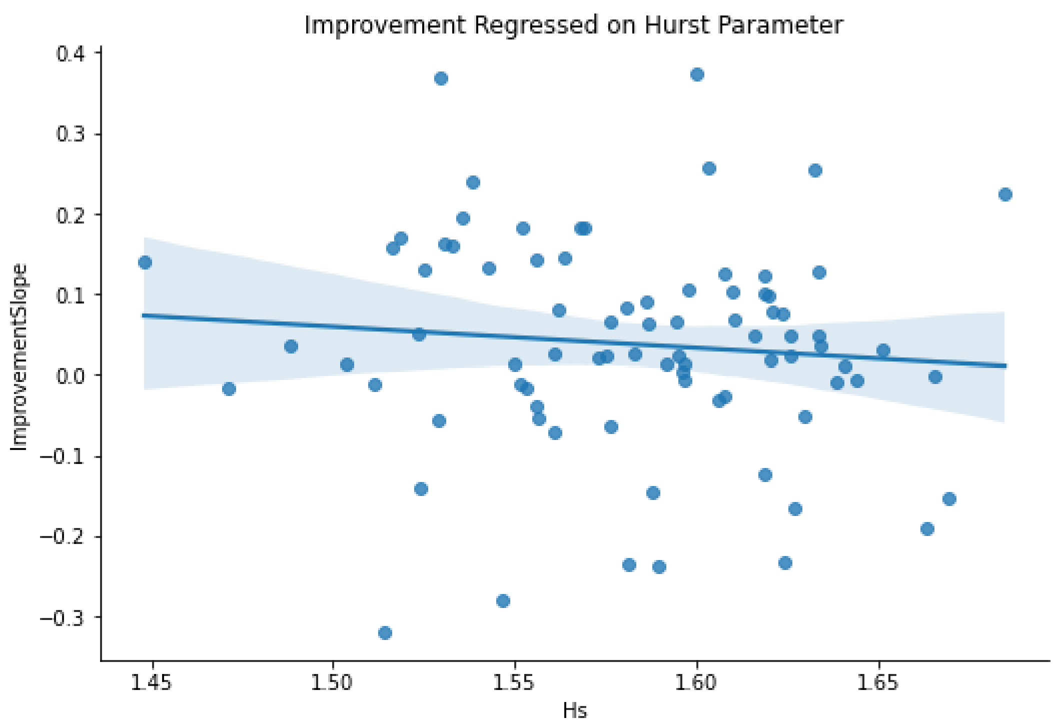

Figure 11. Improvement Regressed on Hurst parameter. Shows that the improvement of a player is not linearly related to the hurst parameter.

We regressed improvement on Hurst parameter (Hs) plotted in Fig 11. Beta = -.25 with a SE of .31, and an R^2 of .008 p = .39. The miniscule R^2 as well as the fact that the standard error of the slope is larger than the absolute value of the slope itself lead us to conclude that there is no evidence of a relationship between the Hurst parameter and improvement.

Complexity Matching as Improvement

Our theory of engagement is that people spatiotemporally integrate with the task they are engaged with. Therefore, instead of there being a particular trend in the relationship between complexity and improvement, improvement may be related to how well matched a person is to complexity of the task they are playing, or in other words, how well they are spatiotemporally coordinating with the activity.

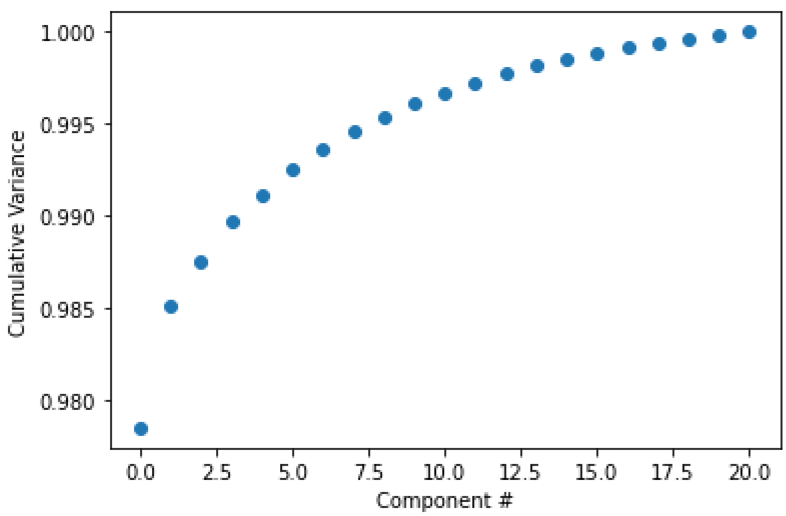

Figure 12. Cumulative Variance Explained in Mouse Coordinates. Shows that about 98% of the variance in mouse coordinates for good players is accounted for by the first component.

Complexity matching is usually computed between entities that produce their own complex behavior, such as two humans in a dyad (Abney et al., 2014), or even the separate hands of a single human (Schloesser et al., 2020). The logical extension in our case would be to determine if participants' complexity converges onto the complexity of dynamics involved in performing the task. As such, what we need is the complexity of the dynamics of the task. However, this is difficult to obtain because it is not clear that even if we were to simulate perfect play, that that would be the optimal way for humans to play, and therefore reflect the appropriate dynamics of a human playing the task. To solve this problem, we took the SVD of our best players' (all players with over 75% accuracy) raw mouse data. The vast majority of the variance (97.85%) is accounted for by the first component (as shown in Fig 12), and therefore we used only the first component to create a synthetic "golden player". To be clear, the golden player does not correspond to any actual player of the game, but rather to the amalgamation of the best players' movements which we capture as the first component of the SVD. These are then an approximation of the way that the best players play, which we use as the "ideal way to play". We then calculated the complexity matching between each participant and our synthetic "golden player".

Complexity matching, as defined by West et al. (2008), is the similarity of two systems' power law indices. According to West et al. (2008) information exchange between complex systems is maximized when these power law indices are equal. Here power law indices describe the relationship between scales in the system, such that those scales are related by an inverse power law. Abney et al. (2014), proposed calculating complexity matching (D) as:

$$D_{a,b} = \sum_{s} log|A(S_a) \; - A(S_b)| + c$$

Here A(S) is the Allan Factor function applied to the system at scale S. Allan Factor variance measures the amount of clustering at scales (Ramirez-Aristizabal, Médé, & Kello, 2018), and is similar to DFA, which we discussed earlier. Thus in the measure of complexity matching used by Abney et al. (2014), complexity matching is the log of the absolute value of the difference in the Allan Factor function of two different time series a and b at scale S and summed for all such scales. This value is then multiplied by negative -1 so that greater complexity matching results in larger numbers, and a constant c is added to make all numbers larger than 0.

To compute complexity matching we slightly modified the technique used in Abney et al. (2014). Instead of using Allan Factor to calculate the variance, we altered their algorithm by using the calculations of the variance at each scale that we obtained from a DFA.

$$D_{a,b} = \sum_{s} log|F(S_a) \ - F(S_b)| + c$$

The final step in calculating the Hurst parameter with DFA is to fit a line to the log scaled Fs and log scaled S. Here F(S) represents:

$$F_s = (\frac{1}{N}\sum_{t=1}^{N}|r_n|^2)\,\frac{1}{2}$$

which is the calculation of the fluctuation at each scale, which is the penultimate step performed in DFA. In this case the F(S) is the square root of the average squared residual around a trendline fitted to each window, and the A(S) is the “expected value of the squared differences, normalized by mean counts of events per window” at each scale S. We think of these as relatively equivalent in this context. We also added a constant of 100, which sometimes was not enough to make all values greater than 0. We chose to use the same constant in all cases so that differences in complexity matching in different analyses would be comparable.

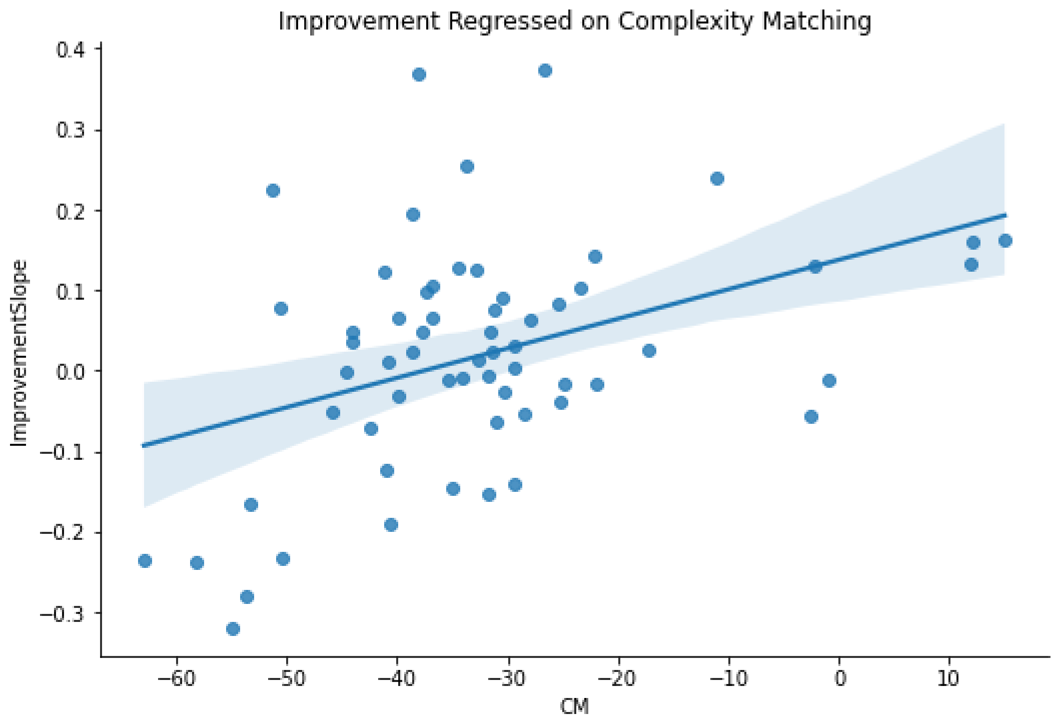

Figure 13. Participant improvement regressed on their complexity matching with the golden player. Shows that improvement is significantly related to complexity matching.

First, we computed the complexity matching for all individuals with our golden player. We then regressed the improvement of all players who had less than 75% accuracy onto their complexity matching (CM), plotted in fig 13. The *B* was .0037, with *SE B* .001, and R^2 .173 $p \le .001$. We take this to be evidence that there is indeed a positive linear relationship between complexity matching and improvement. For this and the subsequent regressions we only included data from participants who finished with below 75% accuracy for 2 reasons. The first is that the players with over 75% accuracy were included in the SVD to create the golden player. Thus fitting them certainly introduces bias. Second, many of the players who finished with over 75% accuracy did not improve during the game but simply played well throughout. We do not think of their slopes of no improvement to have the same meaning as those who played poorly and did not improve. Therefore, we excluded them from these regressions.

We also decided to compute complexity matching over shorter intervals the data into thirds. We did this because we were initially hoping that complexity matching very early on might predict improvement before it has occurred. So, we calculated the complexity matching in each third of the data: CM1 for the first third, CM2 for the second third, and CM3 for the final third.

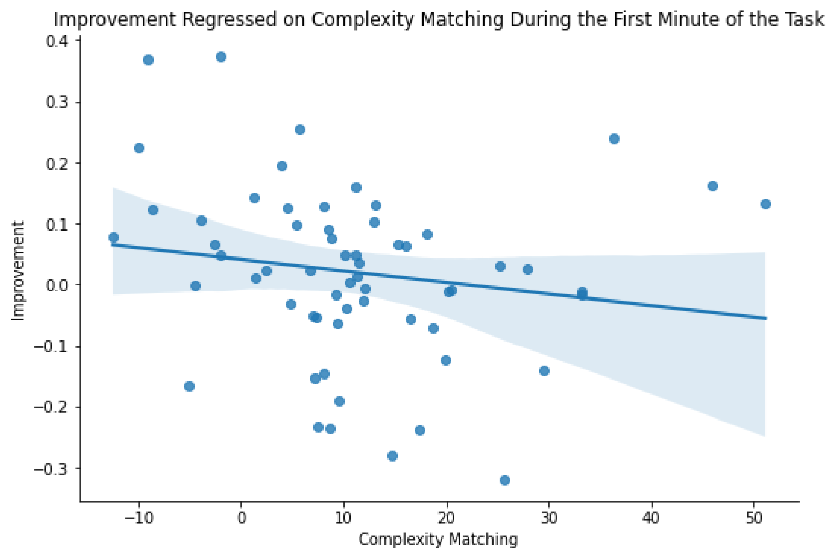

Figure 14. Improvement Regressed on Complexity Matching in only the first third of the task. Shows that complexity matching in the first third of the task does not predict improvement.

In the first third the regression of Improvement on complexity matching is negative, with a *B* of -.0019 and *SE B* of .01, displayed in fig 14. However, in the middle and final thirds (Fig. 15 & Fig. 16), CM2 and CM3, respectively, complexity matching is positively related to improvement. For complexity matching in the middle (CM) regressed on improvement the *B* =.0048, *SE B*= .001, R^2 = .251 and p <= .000. For complexity matching at the end regressed on improvement *B* =.0036, *SE B* = .001., R^2 = .269 and p <= .000.

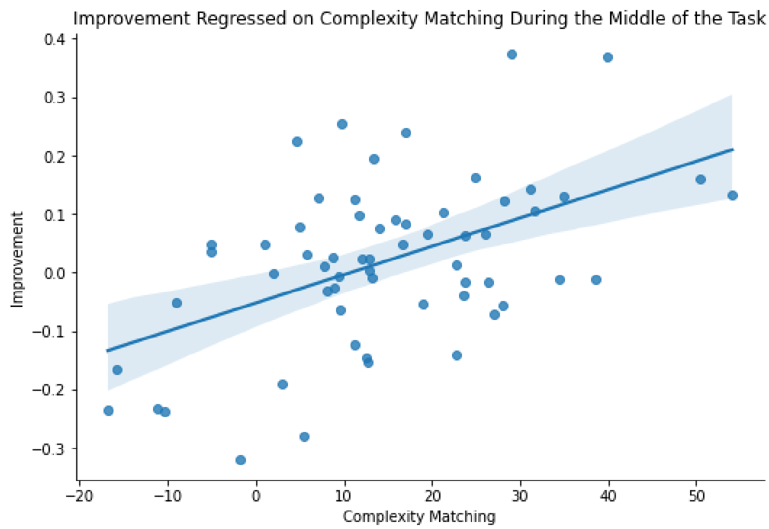

Figure 15. *Improvement regressed on complexity matching in the middle third of the task. Shows that complexity matching does predict improvement in the middle third of the task.*

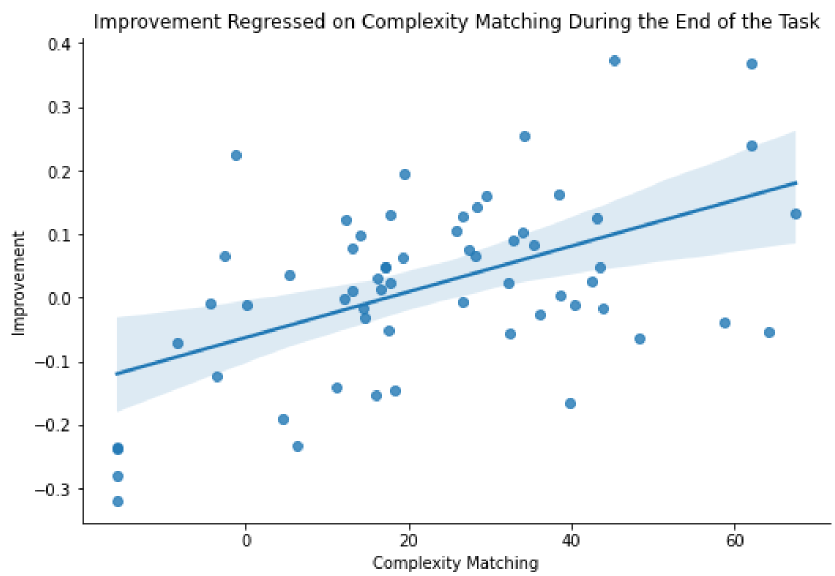

Figure 16. *Improvement regressed on complexity matching in the final third of the task. Shows that complexity matching does predict improvement in the final third of the task.*

We found the lack of a relationship at the beginning to be an interesting finding and to investigate it further we examined the time series data for our best improvers, and we found that in the beginning they often were simply not playing yet (hence why they improved so much) or seemingly only moving their mouse very infrequently. Thus, their complexity matching towards the best players, who were playing and playing well is very low. Likewise we looked at the end of the times series data and found that our least improved players, those who actually declined in score over the course of the game, often didn't play for portions of the final third, which brought their complexity matching down and slightly biased the result there. Given this information we find the complexity matching in the first and final thirds of the data to be somewhat artificial.

However, during the middle third, all of the players played, and there was no noticeably systematic relationship between any who had periods of stagnancy and their improvement which could falsely inflate how related improvement and complexity matching seem to be. We therefore take the positive linear relationship shown in Fig 15, to be evidence that there is a relationship between improvement and complexity matching and to be indicative of how that relationship works.[4]

Engagement as complexity matching

Recall that we used improvement as an operationalization for engagement. However, we did also collect questionnaire data for engagement. The questions used were a subset taken from Weibe et al. (2014) corresponding to their factor Focused Attention. This data is harder to explore as it requires fitting a structural equation model to model the latent variable engagement. Thus we performed only one test of this, a structural equation model of engagement predicting complexity matching during the middle of this experiment, displayed in Fig 17. We used the middle of the experiment as that was where we found the improvement data to be the most reliable in the analysis of improvement. In addition, we removed the participants who stopped playing before the end of the game as we believe that their answers to the questionnaire could also be inaccurate.

[4] We also want to note that we tried the DFA for Hurst parameters on the same more segmented versions of the data and did not find any evidence of the Hurst parameter being predictive.

The structural equation model revealed a significant relationship between engagement and complexity matching with a β of .261 and a SE of .093. We also plot a scatter plot of the predicted values of engagement given the loadings in the SEM and the values for complexity matching in Fig 17. However, this is merely a visual aid, and should not be interpreted too strongly as those are predicted values of engagement based the questionnaire data and therefore should have their own error bars, which we do not show in the scatter plot.

Finally, model fit is not entirely appropriate for discussion with the model as we are not making claims about model fit compared to other theoretical models. However, in the interest of transparency we will report it. Our model fit was somewhat mixed. The CFI for the model (compared with baseline) was .933 which is good, but the RMSEA was .114 which can be indicative of poor fit. We think this is mostly due to a lack of participants. We only had 45 participants in our SEM, and simulation work by Wolf, Harrington, Clark, & Miller, (2013) recommend at least 60 participants for a CFA with 8 indicators and one latent factor and RMSEA is known to be particularly sensitive to low power (Taasoobshirazi & Wang, 2016). The relationship between the latent variable engagement and complexity matching does seem to be relatively strong though. Therefore, while we do think this is evidence that engagement is related to complexity matching, we also recognize another test with a larger sample is necessary here to draw any strong conclusions.

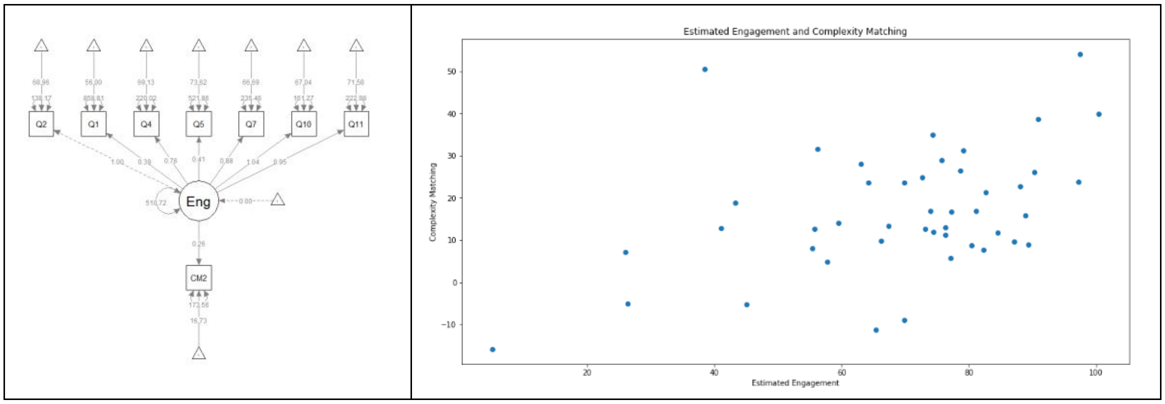

Figure 17. Structural Equation Model of Complexity Matching Regressed on Engagement (left) and Scatter plot of Estimated Engagement and Complexity Matching (right).

Discussion

To summarize we found evidence that complexity matching may be related to improvement and engagement. We first looked at improvement as it is generally thought to be related to engagement, and improvement is an easier variable to use for an exploratory analysis. We found that improvement was not related to complexity but was related to complexity matching. We also found some weak but interesting evidence that complexity matching at a more fine grain scale can predict improvement before it happens, suggesting that even before people are able to perform well they move with appropriate dynamics. After using improvement to refine our exploratory path, we used a structural equation model to test if complexity matching was predicted by self-reported engagement using the questionnaire from Weibe et al. (2014). We found evidence to support this as well. However, we want to be clear here that this was an exploratory study and not a hypothesis test, and as such further study is required. Still, in general we found that complexity matching was related to engagement, both when it was operationalized as improvement, and when it was measured with self-report data.

We think of engagement as spatiotemporal coordination with the dynamics of the activity which a person is engaging with. Under this paradigm, one measure of that would be complexity matching where humans adapt their behavior to have a multiscale structure that is appropriate for the activity. This is somewhat different from other measures of spatiotemporal synchronization in that the multiscale structure (DFA) can be "good" even when a player is scoring poorly (low accuracy). In this exploratory study we collected data from people performing a relatively normal activity, as compared to binary choice tasks, and examined the scaling structure of their data. We found that though complexity itself was not predictive, i.e., neither increases nor decreases in scaling were predictive of good engagement (operationalized as improvement), complexity matching, how close people were to the ideal scaling structure, was a good indicator of engagement. In this case we think of complexity matching as a sort of convergence where, as participants play, those who are engaged converge on appropriate dynamics for playing this game.

According to West et al. (2008), the information exchange between two complex systems is maximized when they have matched power law indices. Thus here we have an underlying assumption that increased information exchange, in this case information gained from the game by the person, is related to performance. To maximize their performance, participants attempt to adapt their behaviors so that they are of appropriate structure for the game. In this sense, those who are not as accurate as other players, but are trying to maximize their performance, have spatiotemporal structure that is more appropriate for the game. Colloquially, one might call this something like behaviors that are "on task", and in general we as humans tend to intuit that people who are noticeably on task, regardless of their performance, are trying their best and are engaged.

As we discussed earlier, there are two primary ways of measuring engagement. The first is with questionnaires such as the User Engagement Scale and the Flow State Scale (O'Brien, Cairns, & Hall, 2018; Jackson & Marsh, 1996). These questionnaires measure how engaged a person felt they were with the task in question. While the insight from these questionnaires has been valuable in understanding engagement, the questionnaires have a major drawback in that one must interrupt the engaging experience in order to ask questions about it.

Physiological measures have been developed to address some of the problems with immediacy and reliability in feedback from questionnaire data. Measures such as heart rate, pupil dilations, and galvanic skin response have all been found to be useful measures of engagement. However, these measures have their own weakness in that they generally require specific equipment that limits the kinds of activities they are usable in. It is rather hard to look at pupil dilation and eye tracking during sports for example.

The drawbacks of these types of measurement emphasize the value that a behavioral measure of engagement, such as spatiotemporal coordination, would provide. Spatiotemporal dynamics can be measured in a variety of settings with relative ease. For example, Marghetis and Núñez (2013), analyzed bodily movement and writing patterns in videos of mathematicians to show that their spatiotemporal dynamics reflected the dynamics of what they were metaphorically reasoning about. Duarte et al., (2013) was able to analyze recorded video of soccer players to discover that soccer players form synergies. And in this paper and in the previous paper, looking at the cursor movement data alone we were able to find important spatiotemporal structure. Clearly this might not be feasible for all activities, reading for example might present an area where spatiotemporal body dynamics are hard to measure, as the person is mostly sitting still.

Furthermore, spatiotemporal coordination offers an explanation as to why some of the physiological measures work in the capacity that they do. Rigoli et al. (2014) found that different subsystems in people can coordinate separately. They found for example that pupil dilations and heartbeat intervals were coupled, but separate from key-press durations and timing deviations in tapping. One interpretation of those results is that different subsystems are coordinating with the environment as befits the person's interaction with the environment. A concept discussed thoroughly in Northoff (2016; 2020). In this paradigm we might expect that the physiological measures that correspond to engagement are often a form of spatiotemporal coordination between the body, brain and environment as an embodied extended cognitive system. For example, we might predict that a time series of pupil dilation might relate to the structure of the activity the person is engaging with when pupil dilations are important to spatiotemporal coordination with the activity, such as was discovered by Wohltjen and Wheatley (2021) who found that pupil dilations and eye contact played important roles in shared attention and were temporally related to shifting dynamics in conversations.

Spatiotemporal coordination, as a measure for engagement, could dramatically expand the domain of measurements of engagement. It could allow researchers to measure engagement in a wider variety of environments than physiological data, with much lower latency than questionnaire data, and is relatively cheap and easy to deploy. However, it is not without any drawbacks. The first and most obvious drawback being that it is somewhat ad hoc. It is up to the researcher to determine which behavioral dynamics are of interest in a particular activity, and which type of coordination they should be investigating. No doubt with time and increased adoption this will decline as behavioral measures of engagement become more widespread and standardized, but as of this writing this will pose a significant early challenge to behavioral measures of engagement.

In addition to advancements in measuring engagement, our results have a significant impact on theoretical accounts of engagement, and on cognitive science. First our results emphasize the need to consider engagement as an embodied phenomenon. Even theories of engagement that include behavioral accounts of engagement, such as Fredricks et al. (2004), emphasize the separation between the cognitive, emotional, and behavioral aspects of engagement. While we agree that there is useful distinctiveness between these aspects of engagement, we think our results imply important interactivity between these aspects of engagement.

Second, we feel that our results support our Dreydeggarean account of engagement. In our account of engagement we stipulated that engagement should be reflective of how well the motivational affordances of an activity satisfy the “care structure” of a person. This engagement should then show itself as a transparency between the person and the activity in which the person “loses themselves” in the activity. In empirical terms, we would predict that people are more likely to be engaged when their cares are being satisfied, and that that engagement will manifest as a spatiotemporal coordination between the person and the activity. In the present study we did not analyze if satisfaction of care plays a role in engagement (though this is generally consistent with other research). However, we did find evidence to support the idea that spatiotemporal coordination is related to engagement.

The support for our account of engagement has further implications. First our account is somewhat unique in that because it is derived from Dreydeggarean phenomenology we offer an account of why motivational, situational, and phenomenological aspects of engagement are related. For example, our account of engagement predicts a very similar phenomenological experience to the Flow state. However, we also account for why that phenomenon should be related to engagement, namely that it is our network of involvements that produces that phenomenon (for a detailed explanation see Ch 4). In addition, our theory of engagement is domain general in that it fits with the JDR model of engagement, Flow theory, and the multidimensional model of classroom engagement. The relationship between the care structure and the absorption, as explained by Dreyfus, shows why the relationships between skill and challenge; between jobs demands and resources; and various aspects of intrinsic motivation, are all related to the kind of absorption that is characteristic of engagement (Nakamura & Csikszentmihalyi, 2009; Hakanen, Schaufeli, & Ahola, 2008, Kowal & Fortier, 1999). We believe at a minimum these results show that thinking of engagement from a Dreydeggarean, and embodied cognitive systems perspective is important, and we believe it can potentially provide more unifying theories of engagement, and open measurement of engagement up to behavioral measures.

As this was an exploratory study, we recognize that a hypothesis driven experimental study should be conducted to confirm these results. In addition, the evidence discovered in this study elucidates the need for more studies to flesh out a complexity matching account of engagement. In the current paper engagement is thought to be related to complexity matching, but the type of matching shown is more of a convergence toward an optimum than a matching because the game itself does not have a definitive structure for researchers to use in calculating matching. A study of a different game that has complex dynamics that a human could explicitly match to would produce a fruitful experimental paradigm for further examining behavioral measures of engagement.

Finally, we believe this only scratches the surface of the potential for measuring engagement as a spatiotemporal coordination. Follow up studies in research on engagement should be performed in a variety of domains to determine the limits of the efficacy of measuring engagement as spatiotemporal coordination of a complex embodied system. Both different kinds of spatiotemporal coordination (correlation, synchrony, complexity matching, synergy, etc.) and different domains of measurement (school, sports, games, work, etc.) need to be investigated to determine the true effectiveness of this paradigm.

6. Engagement and Mouse Tracking

In Ch 4. we outlined a phenomenology of engagement, rooted in Heideggerian phenomenology. In Heideggerian phenomenology the world we experience is one of relations, or "involvements". We do not experience the hammer as just an object but rather a tool for hitting nails and building things (Dreyfus 1991). This network of involvements is dependent on what we call the care structure. The care structure is a sort of a hierarchy of goals, with many subgoals all of which lead towards our for-the-sake-of-which which is the ultimate way of being for that person, the goal for which we do everything else, it is that "for the sake of which" I do everything else in my life (Wrathall 2005). For the most part when we are just doing things we only experience the entities in the world based on their "involvement" (Dreyfus 1991). For example, I have no idea how much my toothbrush weighs, I just use it to clean my teeth. However, this experience can "breakdown" when we are no longer able to use the entity as we want, at which point we will experience the entities as simply as objects with physical properties (Dreyfus 1991). We argue that care plays an important role in certain types of breakdown, namely that not only can an object become unusable and present-at-hand because it breaks, but it can also happen because your desires shift and the involvement upon which that entity was revealed to you disappears.

Thus we argue that the absorption that Dreyfus describes is actually two dimensional. One dimension is skilled coping, which is governed by one's ability to use an entity to serve their purpose (Dreyfus, 1991). The other dimension is the matchedness between your cares and how well the task at hand affords you the satisfaction of your cares. This second dimension of absorption is what we call engagement. Thus we expect engagement to feel like being absorbed in a task. Just as skilled coping implies that one is absorbed in the work and just using the objects at hand, because they are sufficiently skilled at using those objects and those objects are sufficiently good at filling one's needs, engagement implies absorption in the task dynamics because one "cares" enough about the task. In our account we would predict that people are integrated with the tasks they are engaged with, similar to the way that a person and the entities they are absorbed coping with are integrated. A complex systems perspective on cognitive science provides a very satisfying empirical analogue to the ideas of integration we have discussed so far.

In recent years there has been significant growth in thinking of cognition as an embodied/extended complex dynamical system (Shapiro & Spaulding, 2021). This is described much more thoroughly in the beginning of the dissertation but briefly, this means thinking of cognition as an emergent property that comes from the nonlinear interactions of the brain, body, and environment (Dotov et al. 2010; Wallot & Kelty-Stephen, 2018). In complex systems the behavior of the system is driven by the interaction of the variables and because the variables are interdependent and nonlinearly related it is impossible to decompose the system into smaller causal components (Wallot & Kelty-Stephen, 2018). Instead complex dynamical systems are often analyzed more holistically by looking at the behavior of the system over time and examining things such as the scaling properties of the behavior (West 2018).

Heideggerian phenomenology emphasizes the holistic nature of dasein, and dasein's past, present, future, and environment and has close ties to embodied cognitive science, and indeed was inspirational to the original development of the paradigm (Gallagher 2014). This is somewhat intuitive. In Heideggerian phenomenology, consciousness, though Heidegger would never use that word, comes from the interactions between the person and their environment. It shows up in the network of involvements we described earlier which is dependent on the person's care, and on the actual objects in the environment. For Heidegger dasein is always "in the world" and there is no separate subject perceiving objects, the way Descartes, and many non-embodied cognitive scientists think. However, while embodied cognitive science has enjoyed many breakthroughs and now has a stockpile of evidence to support its ideas, direct empirical connections to Heideggerian phenomenology have been very rare.

Dotov et al. (2010) provide one of the few empirical tests of Heideggerian phenomenology. They had participants play a game on a computer using a mouse and an accelerometer on one hand. Part way through the game, white noise was added to the cursor movements to simulate the breaking of a tool. They showed that the participants' movements showed a decline in complexity when the cursor was “broken”. They claim that this is evidence that when the mouse is ready-to-hand the participant is in a larger complex dynamical system and thus the behavior is more complex, but once the cursor starts misbehaving, the mouse is no longer part of the complex system, and thus the complexity of the interaction between the person and the mouse is reduced.

We find this study very compelling. However, we have a different interpretation of Heidegger and from that we draw slightly different conclusions from this study. First, recall that we believe that there are two channels to absorption and ready-to-handness is just one of those. We think that an engaged player is spatiotemporally coordinating/integrating with the task. In this case, if the person stays engaged with the task, they should maintain some spatiotemporal coordination. However, the task changes when the white noise is added. In fact, we suspect, given that white noise is of low complexity, adding white noise probably reduces the complexity of the task, and thus an engaged spatiotemporally coordinated player adjusts their behaviorally appropriately and we see the downward shift in complexity discovered by Dotov et al. (2010). We interpret the findings to be an example of complexity matching during engagement, rather than a simple decline in complexity during breakdown. Our ideas are somewhat supported by Bennett, Roudaut, & Metatla, (2022) who performed a follow up study on Dotov et al. (2010) and found that complexity was linked to task engagement. However, they did not test how different forms of disruption affect break down, or the role of complexity matching.

The mouse tracking tools we developed are well suited to this paradigm. We developed ways to look at continuous 2 dimensional mouse traces and apply complex systems techniques to them. Those complex systems techniques, mainly DFA, are the same techniques used by Dotov et al. (2010) and Bennett et al. (2022) on accelerometer data. We think it seems viable and reasonable to apply them to cursor position data in a similar task.

As we are interested in complexity matching rather than just complexity, we will have to adjust the task the participants perform somewhat. Stephen and Dixon (2011) showed that humans complexity match to a chaotic metronome when attempting to tap along to that metronome. A mouse tracking analogue for the chaotic metronome would be a target tracking task with an onscreen stimulus whose movements are complex. We can then measure the complexity matching between participants and the complex stimulus. In addition, we can make this task have a similar structure to Dotov et al. (2010) where we introduce noise to the cursor. However, we will add two kinds of noise to the cursor and see if complexity matching is different between the two groups. If the shifts in complexity matching are the same and negative, then people are simply reducing complexity in breakdown in accordance with the interpretation provided by Dotov et al. (2010). However, if the shifts are different, then people may be complexity matching to the task as it is changing and the shifts are indicative of our view of engagement as integration of person with task via spatiotemporal coordination.

Methods

Design

We created a target tracking game where a player attempts to keep their cursor on top of a moving target. We describe the target as a small creature and the participant as a scientist whose job is to keep a magnifying glass on top of the creature. The creatures' movements are determined by a song, using a mathematical transformation of an audio signal into 2-dimensional movement. The participants are told to try to keep their magnifying glass on the creature as much as possible as the creature moves around the screen. The participants' scores are displayed in the top right hand corner of the screen and go up by 1 every 20ms if the scanner is on the creature, and go down by .1 otherwise. Overall, the task takes three minutes and is broken into three one-minute segments:

1) Normal game play

2) Noise is added to the cursor movements, following the same procedure as Dotov et al. (2010).

3) The cursor reverts back to normal for the final minute of the task.

The two conditions of the experiment correspond to two kinds of disruption in the second segment. For the control condition we added white noise, which is the kind of noise added in Dotov et al. (2010). In the alternate condition we added noise based on the percussive part of the song which was used to generate the creature's movements. The cursor adjustments--which seem arbitrary to casual inspection--are actually complementary to the creature's movements, and thus keeping the cursor on the creature is a more unified task than just fighting against added randomness.

Participants are then directed to a questionnaire about their engagement, which we designed by sampling questions from Wiebe (2014). We did not use all of the questions, because many of the 28 questions target definitions of engagement that we are not interested in, such as definitions based on aesthetic appeal. Instead we subsampled one of the 5 factors in their questionnaire: focused attention, which gives us 8 questions. Finally, the participants fill out a short demographics questionnaire and are thanked and debriefed. All together this should take between 5-10 minutes. The questions, source code, and a playable version of the task are available on the Yoshimi Lab website.

Creature Movement

We wanted the on-screen stimulus to move in such a way that it would have a complex time series. However, we also thought it would be best if that complex time series was rooted in human activity. We chose jazz music, because jazz music (unlike other types of music) has been shown to produce complex structured time series that are similar to conversation (Kello, Bella, Médé, & Balasubramaniam, 2017), which is a good exemplar of the complexity profile of human behavior. We used the first 3 minutes of *Ascension* by John Coltrane as the basis for the bug movements.

To transform *Ascension* into a 2d time series for movement on a computer screen we first took the wavelet transform of the song. This gives us precise amplitudes of the frequencies in the song at time indices. Next we downsampled the song to 100 hz as that is the maximum refresh rate of the code for the task. We then created what we call the *pitch driver algorithm*.

In the pitch driver algorithm, we use change in pitch, and change in amplitude of pitch to control the movements of the bug. At each time point *t* we take the amplitude of the most dominant frequency. We calculate the change in pitch, and change in amplitude between each time t, and time t +1. We convert the change in pitch to radians and the change in amplitude was scaled down to become magnitude giving us polar coordinate steps for movement. We then cumulatively sum the thetas so that the time series of movement are contiguous such that the end of each movement is the starting place for the next polar step. Finally, this time series is converted into cartesian coordinates for displaying the on-screen stimulus position and the thetas are preserved for on-screen stimulus orientation. When done this way, the stimulus is visually always rotating and then taking steps forward. An example trace of stimulus movement is shown in Fig. 18.

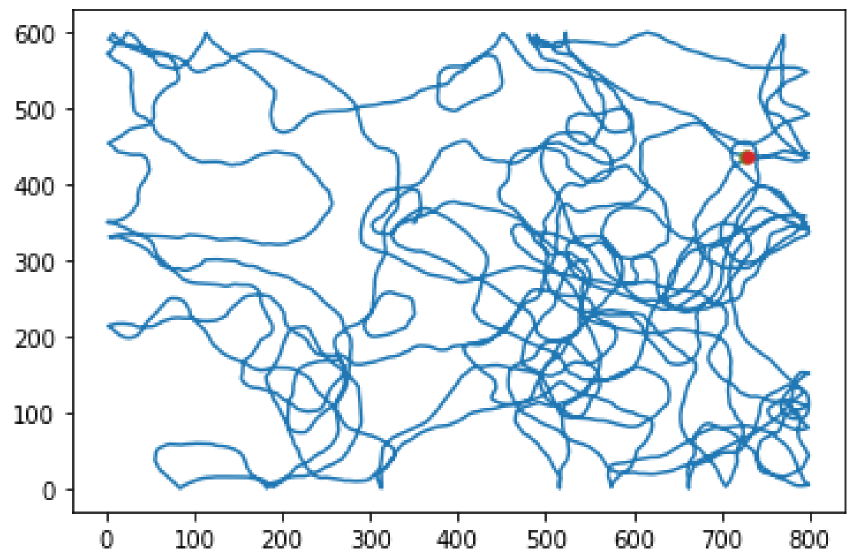

Figure 18. Stimulus Path. The path that the stimulus follows starting at the red dot. The numbers on the axis correspond to the pixel coordinates of the participant's screen.

Disruptions

We created two different kinds of disruptions for this experiment. The first disruption we added was a white noise disruption. In this we generated white noise to be added to the cursor position. In the second disruption we used noise generated from the same song the creature was moving according to. The librosa library provides a function to split music into harmonic and percussive components. During the disruption, for both conditions, the creature is moving according to only the harmonic component. The percussion noise is noise generated the same way the creature movement is generated, but only from the percussive portion of the song. The amount of disruption, which we call the displacement is the average distance the visible cursor is from the actual cursor, or the average magnitude of the Euclidean distance of the offset added to the visible cursor. The two time series are controlled to have the same (25 px) average offset. Therefore, the primary difference is that the percussion noise is complementary to the creature movement, while the white noise is truly random.

Participants

Participants were collected through Amazon Mechanical Turk. They must live in the U.S., be over 18, have a 95% approval rating, and have over 500 approved HITs.

We recruited 149 participants. We then eliminated participants for not producing enough data because they were not moving their mouse enough, not answering our catch questions correctly, or not using a mouse. We removed 12 participants for not producing enough data. We removed 39 participants for not using a mouse or not answering the catch question correctly.

During data analysis we also discovered that we had to eliminate participants who had long periods of inactivity. Long period of inactivity will cause the complexity analyses we use to produce inappropriate (Nan) values which cannot be analyzed. We removed anyone who did not move their mouse for more than 50 seconds, and in this process, we removed an additional 4 people. Finally in our previous study we learned that it was important to eliminate people who are no longer playing at the end of the game as their questionnaire data is potentially also suspect. So, we eliminate any players whose score during the final third of the game was a net negative. This means that the participant had their cursor on the stimulus less than 10% of the time. We found during play testing that even placing your cursor in the general area and of the creature and making only minimal movement can generally get scores of greater than zero just from the stimulus moving across the cursor and thus we feel comfortable removing these participants but recognize that we did not specify this ahead of time. We removed 12 participants for this reason.

Data

From each participant we collected their mouse coordinates (time stamped), the creature coordinates and timestamps as displayed on their computers, their answers to the questions, and their score both as a final value and as a timeseries.

We performed several steps of data pre-processing. First, browsers only poll the mouse when it's moving, and thus we had to forward fill (taking the last member of the series and using it to fill in subsequent gaps) for times when the mouse was sitting still to fill in large gaps in time between mouse locations. We also interpolated and resampled all participants to have exactly 9000 data points. Once the mouse movements were forward filled and linearly interpolated, we transformed the x and y coordinate data into complex valued data such that:

$$c_t = x_t + i * y_t$$

Analysis and Results

Hypothesis 1: Change in complexity when coping is disrupted

Our first hypothesis was that we would see a change in complexity when coping was disrupted. This follows from the previous work by Dotov et al. (2010) who found that complexity decreased when they added white noise to the cursor movements. Here we tested this with a repeated measures anova. We found support for this hypothesis, that there were differences between the three different game phases: beginning, middle (disruption), and end (N=81, f=24.9, $p < .01$), and using a Tukey's post hoc found that the significant difference was between the beginning and middle phase, which supports our hypothesis.

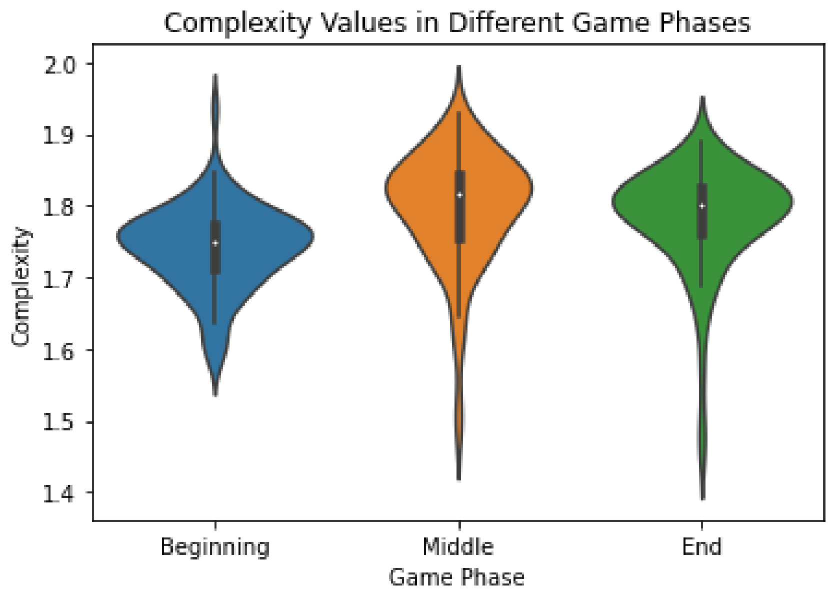

Figure 19. Violin Plot of participants hurst parameters (complexity) during the different portions of the game, beginning, middle (disruption), and end.

Hypothesis 2: Different effects for different disruptions

Our second hypothesis was that the complexity of the players would be different in the different disruption conditions. We tested this with a Welch's t-test. The 43 participants who were in the percussive noise disruption condition ($M = 1.78$, $SD = .07$) compared to the 38 participants in the white noise condition ($M = 1.82$, $SD = .08$) demonstrated significantly different complexity, $t(73.4) = -2.59$, $p < .05$. In addition, the complexity was lower in the percussive noise condition, implying that the disruption was less impactful.

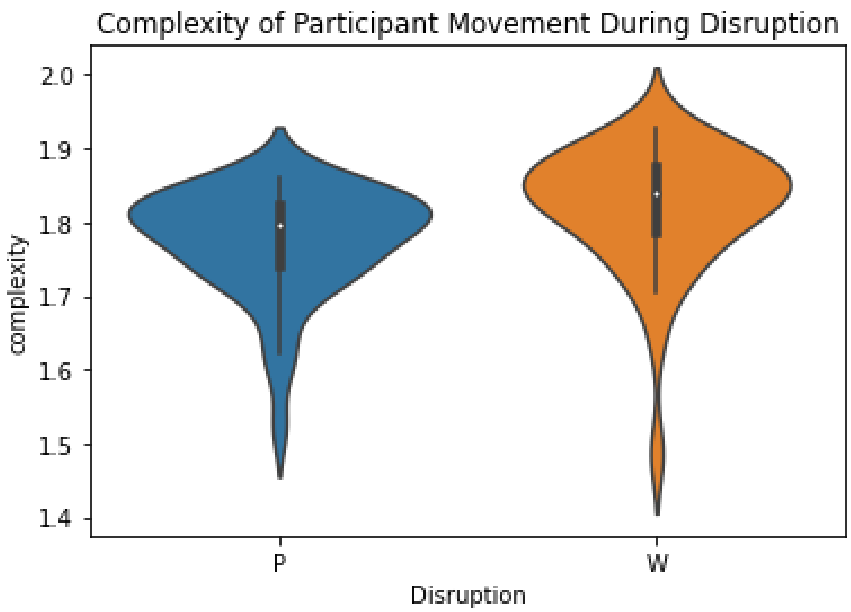

Figure 20. Violin plot of the two different kinds of disruption with the white noise corresponding to the orange violin and the percussive noise corresponding to the blue violin. The two groups were significantly different p=.03

Hypothesis 3: Engagement as Complexity Matching

Our third hypothesis was that complexity matching would be related to engagement. We tested this with a structural equation model. We did not find direct support for this hypothesis. CFI and RMSEA both indicated very poor fit for our model. In addition, we performed a subsequent confirmatory factor analysis (CFA) and found that that also had very poor model fit, indicating that we are not adequately measuring our latent variable. Furthermore, the average values for reported engagement in this study are generally higher than those in the previous study, with a mean of 461 in the previous work, and a mean of 572 in this study (with a maximum possible of about 700). We believe this might be affecting the variance in our latent variable. We added several gamification techniques to this task including narrative manipulation and a leader board. In hindsight we may have accidentally created a high engagement condition, with no corresponding low engagement conditions and as such much of the remaining variance that is being measured as our latent variable could be noise in interpretation of the questions with most people having high engagement. In future studies a better test of the effect of complexity matching here might be to have two experimental conditions corresponding to high and low engagement and determine if complexity matching is different in those two conditions.

In the interest of better understanding the results we did perform one additional analysis that wasn't preregistered using improvement as a proxy for engagement based on our analyses in previous studies. In our first regression we found that extreme values for complexity matching and improvement were producing a violation of heteroscedasticity of variance. Fig 23 shows the RVF plot for the linear regression and we can clearly see that there are a few very large residuals. Fig 24 is the QQ plot for the model. Again here we see several extreme values and in the case of the qq plot a possible violation of normality.

Figure 21. Residual plot for complexity matching and improvement. Shows the residuals from the model. At least 2 outliers are clearly visible.

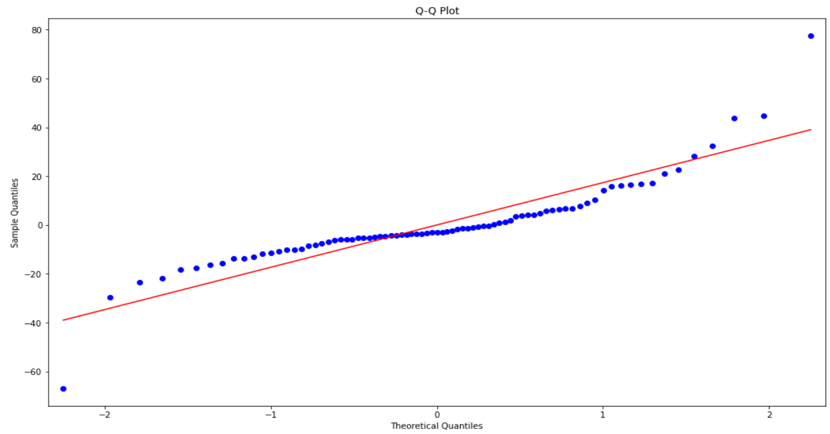

Figure 22. Quantile Quantile plot. Shows that we have at least 2 outliers and heavy tails.

We removed outliers whose values were more than 2 standard deviations away from the mean in either variable and again fit a regression model. In this case we found that change in complexity matching was related to change in improvement: β=.27 se=.10 R^2=.10 and $p < .01$. However, this model still contains one outlier with a standardized residual > 3. If we remove this outlier $\beta = .36$ with se .10 $R^2 = .15$ and $p < .01$. This implies that there may be a spatiotemporal coordination aspect to improvement in this game as well, even if our engagement questionnaire data was very noisy. However, we also recognize that the necessity of removing outliers and evidence that the underlying distribution in the original QQ plot is non normal are evidence that there may be more going on here. One possibility is that the measure of complexity matching is overly sensitive to the extreme values; Or, it could be that the underlying distribution is actually heavy tailed. As this analysis (removal of outliers and regressing on improvement rather than directly on engagement) was not pre-registered, and the data show some signs following a heavy tailed distribution, a follow up study should be conducted to ensure the validity of the finding. Still, we think it would be a mistake to ignore the trend that is clearly visible once we remove the outliers, especially as the practice of removing outliers who are more than two standard deviations away from the mean is relatively common practice.

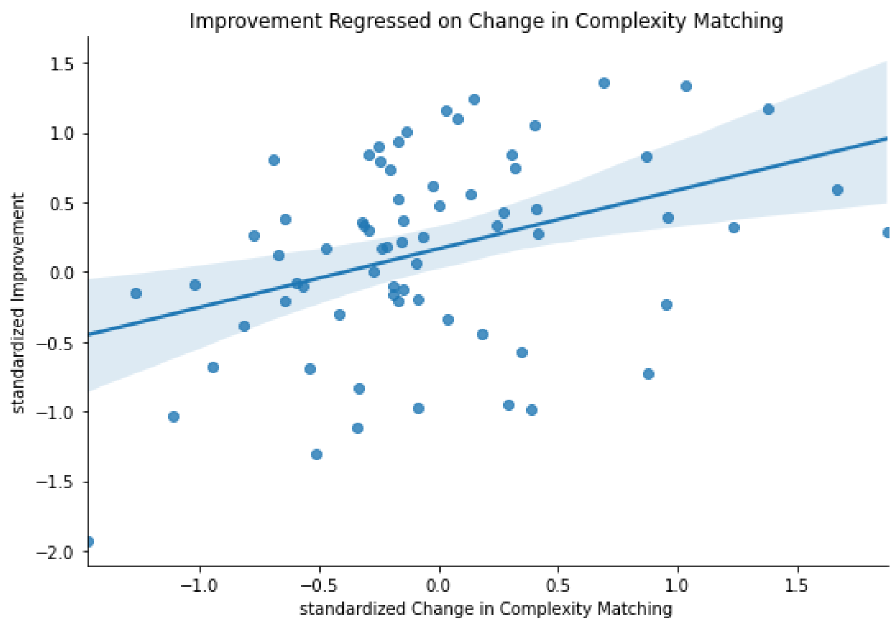

Figure 23. Improvement regressed on complexity matching. Shows a significant relationship between complexity matching to the stimulus and improvement p < .01.

Discussion

To summarize our results, we found support for hypotheses one and two, and mixed support for hypothesis 3. We found a shift in complexity when coping is disrupted, which supports the findings of Dotov et al (2010). We also found that the shift is related to the type of disruption which implies that the shift may not be a simple reduction in complexity due to a degraded connection between the person and the tool, as described by Dotov et al. (2010). Finally, we did not find support for the claim that engagement is related to complexity matching. We believe that had to do with excessive error in the measurement part of the model of engagement. Again, the reason may be that we have created a high engagement game and the differences in reported engagement become more about differences in the interpretation of the questions than actual differences in the level of engagement, since everyone is relatively highly engaged. When we substituted improvement as a proxy for engagement, we did find a positive relationship with improvement and a person's change in complexity matching, i.e., those whose complexity matching increased, also improved more at the game.

The evidence in support of hypothesis one showed a shift in the Hurst parameter when coping was disrupted. Earlier work by Dotov et al. (2010) found that the Hurst parameter declined during “breakdown”. Here we actually find an increase in the Hurst parameter during breakdown. This could simply be due to the different modalities of measurement. Dotov et al. (2010) used an accelerometer attached at the wrist and analyzed acceleration data while we used the actual cursor data. However, we do think the general trend of disruption shifting the Hurst Parameter of cursor data and acceleration data provides evidence that we are measuring a “connected” interaction dominant system. In this case we are now measuring the same extended cognitive system but from a different point. Dotov et al. (2010) are measuring the “hand tool boundary” while we are measuring the dynamics of the part of the environment that the human is “moving” in the human environment system. The fact that these results align, we think, is evidence of the connectedness of these systems. Rigoli et al. (2014) also showed that different but connected parts of a system have convergent behavior patterns that do not necessarily cross converge with other less connected parts of the system. More broadly the concurrence of our results with those of Dotov et al. (2010) support the idea that humans form a complex system with the environment and that system can be measured from different points in the system.

The evidence in support of hypothesis 2 showed that there was a difference in the shift in complexity in the two different forms of breakdown. While we think our evidence is generally in line with more broad interpretations of Dotov et al. (2010), that our environment is part of the interaction dominant system that constitutes the mind, here we show that the mechanisms may be more subtle than increases or reductions in complexity. Dotov et al. (2010) postulate that the decline in complexity is due to the reduction in the coupling between the person and the tool. However, as they test this hypothesis by injecting white noise, it seems plausible that participants could simply be shifting their dynamics to be more appropriate for the task, and it is actually the task that has changed and been reduced in complexity by adding white noise which is of low complexity (Ihlen 2012). Using two different kinds of disruption we found that the response (shift in complexity) is indeed dependent upon the type of disruption. This might imply that rather than a disintegration of the system, we are seeing a shift in the dynamics of the system that mirrors the shift in the dynamics of the task.

This process of shifting dynamics to mirror the dynamics of the task implies that, while humans may indeed form a complex system with ready-to-hand entities, they also are organizing their behaviors to fit the activity they are performing. This implies that when thinking about the mind as an IDD from a Dreydeggarean perspective, there is at least one dimension other than coping. Coping only describes the connection between the person and the entities, which is what Dotov et al. (2010) were attempting to measure. Our results imply that there are also connections between the person and the task. We find this to be evidence for our alternative version of absorption which is two dimensional with integration between the body with individual entities representing coping, and integration of the dynamics between the brain, bodily, and task representing engagement.

The evidence in support of hypothesis 3 was somewhat mixed but we feel that overall it showed support for engagement, at least when operationalized as improvement, being related to complexity matching. Complexity matching is a form of integration between a person and a task in terms of their dynamics. In previous research we saw that complexity matching was related to improvement (Ch 5). In general we think of engagement as being theoretically linked to improvement because engagement is supposed to mediate the relationship between practice and performance (Ch 5). In other words, practice alone does not improve performance, it requires engagement for it to be effective. We find the relationship between improvement and complexity matching to be compelling evidence in support of our model of engagement.

Our model of engagement is that phenomenological engagement is integration between a person and a task, and this happens when the person’s care structure and the motivational affordances of the task are well matched. Integration between people, entities and tasks is then measured as spatiotemporal coordination. Integration between people and entities, such as ready-to-hand entities which Dotov et al. (2010) argues people are integrated with, empirically produce changes to the scaling structure of the human-entity system. However, when we are integrating with a task, which is what we call engagement, humans cannot simply subsume the task. Instead, the engaged parties adapt their dynamics to maximize their integration with the system. We think of this change as spatiotemporal coordination, and one form of that coordination is complexity matching. By showing that the amount of improvement participants made was indeed related to those participants own shift in complexity matching we have shown that participants may indeed be adjusting their dynamics to match the activity they are engaged in.

Recall that research on engagement generally relies on self-report and on physiological measures (Henrie, Halverson, and Graham, 2015; Greene, 2015; Mandernach, 2015; Hookham, & Nesbitt, 2019). Here we have provided a potential behavioral measure of engagement: spatiotemporal coordination. This could be very impactful for research on engagement as it can be applied in a wider variety of settings than the equipment dependent physiological measures, and it can be applied instantaneously without the need to interrupt a participant from their engaging activity.

Furthermore, these results provide some insight for theories of engagement. First, theories of engagement that dissociate behavioral and cognitive engagement, such as theories of classroom engagement described in Fredricks et al. (2004), should be careful how they stipulate the independence of those constructs. While we certainly would agree that there is distinctiveness to those constructs and they can be effectively modeled separately, we also point out that our results (and embodied cognitive science in general) indicate that those constructs are probably more interconnected than they at first seem.

In addition, we think our results provide interesting insight into the flow phenomenon. Starting from a completely different direction, our phenomenological model of engagement predicts a phenomenology that is similar to the flow phenomenon. However, our model of engagement has the added benefit of explaining why the phenomenon of engagement should feel the way it does. In addition, our model offers an empirical behavioral prediction for the phenomenon of absorption, which is that it should be accomplished through spatiotemporal coordination. In future research it could be valuable to determine how often the flow phenomenon is actually related to spatiotemporal coordination.

Our results also have implications for mouse tracking. As we stated previously, mouse tracking research in cognitive science has been largely dominated by binary forced choice tasks such as those used in (Kieslich, Henninger, Wulff, Haslbeck, & Schulte-Mecklenbeck, 2019), (Stillman, Shen, & Ferguson, 2018), and (Hehman, Stolier, & Freeman, 2015). However, our results among others such as Calcagnì et al. (2019) are beginning to show the richness that less rigid techniques for analyzing mouse tracking data can provide to cognitive science. In order to fully utilize mouse tracking data, we will have to go beyond simple descriptives such as reaction time, curvature, or velocity.

Finally, we think these results have implications for cognitive science in general. First, these results are in line with a growing body of evidence that it is important to analyze human data from a complex systems, embodied cognition, perspective (Dotov et al. 2010; Spivey, 2008; Wallot, & Kelty-Stephen, 2018; Kello, et al. 2010; Van Orden, et al., 2003). Moreover, these results specifically imply that spatiotemporal coordination plays an important role in “cognitive” activities. Others such as (Kang & Wheatley, 2017), Abney et al., (2014) and Proksch, Reeves, Spivey, and Balasubramaniam, (2022) among many others (for a thorough review see (Spivey, 2020) have found similar evidence, which when taken together we think implies that humans, as embodied cognitive systems, interface with the world spatiotemporally. Northoff (2016, 2020) in particular has been developing a paradigm to uncover how spatiotemporal coordination between brain dynamics and environmental dynamics plays a role in cognition.

Though we clearly find these results impactful there is much yet to be done. First and foremost, it will be important to test the predictions of engagement as spatiotemporal coordination in a wider variety of domains. One potentially interesting example would be to build on the research of Duarte et al. (2013) who found that soccer players form synergies during game play. As the flow phenomenon has a history in sports science (Jackson & Csikszentmihalyi, 1999), it might be interesting to see if the players who more closely synergize with their teammates also report the most engagement or highest flow. In addition, in the mouse tracking / user interface domain it would be interesting to test these predictions in naturally acquired data of participants playing commercial video games or performing actual work-related tasks.

7. Conclusion

In this final section I will wrap things up by giving a quick summary of each of the sections so that we can see the results in the context of each other. Then I will try to highlight what I think those results mean as a group, which might have been less noticeable when we discussed them in isolation. And finally I will try to show what I think the impact of these results should be, and discuss some potential avenues of future research.

In general this dissertation is about engagement. As I discussed at several points, in the scientific literature engagement is somewhat loosely defined and definitions can vary across fields quite a lot. However, most definitions include some description of absorption or cognitive focus (Bakker, Schaufeli, Leiter, & Taris, 2008; Schaufeli, 2013; Balwant (2018); Sheldon, Prentice, & Halusic, 2015; Nakamura & Csikszentmihalyi, 2014). I take these concepts as being core to what engagement is, irrespective of the field or the task in which it is being measured.

Measurements of engagement vary, partly as an artifact of the variability of definitions of engagement, and partly as convenience because it is often easier to provide some operationalized, task specific, measure of engagement, like clicks on a link (Zhou, Calder, Malthouse, & Hessary, 2021). Part of the allure of these task specific measures of engagement is that they are easy to make, implement, and interpret. Outside of these task specific behaviors, the majority of measures of engagement are either self-report questionnaires, or physiological measures, both of which have significant drawbacks in implementation (Lalmas, O'Brien, & Yom-Tov, 2014).

Over the course of this work, I developed a theory of engagement that I think helps to unify the common components of different definitions of engagement and provides a foundation for measuring engagement from a behavioral perspective in a way that is generalizable, but also incorporates the task specific nature of engagement. We think of engagement as the matchedness between a person’s cares and the ability of the activity to satisfy those cares. We think people will be more engaged when the activity they are performing is one that satisfies their own interest, which aligns with most other theories of engagement. However, by thinking about this from a Dreydeggarean perspective and grounding our theory in the care structure we are able to get powerful new insights about how engagement works.

Dreydeggarean phenomenology provides a philosophical foundation for a theory of why engagement would feel the way it does, what might increase engagement, and what kinds of behavior are associated with engagement. For example, Dreydeggarean phenomenology shows us how our care structure is connected to our experience of the world. We are always situated in a meaningful world that is characterized by our network of involvements which is in turn dependent on our care structure. Things that are directly connected to our network are "transparent to us" in that we simply see them as meaningful entities and use them appropriately (Dreyfus 1990). We can use this insight to explain why activities that are connected to our care structure produce the kind of absorption that is characteristic of engagement in other theories such as Flow (Nakamura & Csikszentmihalyi, 2014), Classroom engagement (Balwant, 2018), or workplace engagement (Schaufeli, 2013). Because our experience of the world is dependent on our care structure, activities that satisfy our care structure we become absorbed in. The more important an activity is to us, the more directly connected it is to our for-the-sake-of-which, the more absorbed we are in it.

In addition, Dreydeggarean phenomenology is theoretically aligned with an embodied, extended approach to cognitive science (Gallagher 2014). The embodied extended approach to cognitive science in essence claims that the body and environment are actually part of the cognitive processes that many people think of as happening purely in our head (Gallagher 2014). Just as dasein is a network of involvements including the entities in the environment, which is inherently connected to dasein's care structure, embodied cognitive science stipulates that the body and environment are part of the mind. This idea of examining cognition in the entire brain body environment system invites analyses from complex systems science that examine the system as a whole, since the brain, body, environment system is non decomposable (Wallot & Kelty-Stephen, 2018).

Dotov et al. (2010) showed that when people are coping with an entity the complexity of their movement increases. However, recall that we argue that absorption is two dimensional, with engagement being the second dimension. In the case of coping it makes sense to think of simply adding an entity to the network of involvements which then creates a slightly more complicated network. However, in the case of engagement we are looking at the integration of the person with the dynamics of an activity rather than the integration of a single entity into the network of involvements. Earlier we showed many instances where the dynamics of people's behaviors reflects the dynamics of the activity they are engaged in such as jockeys riding horses, patients undergoing therapy, or soccer players during gameplay. In all of these cases people who were engaged in the activity coordinate with the other entities involved in the activity. We think this implies that while coping involves including an entity in the system, engaging involves shifting the dynamics of the system so that it can better coordinate with the other entities involved in the activity. Therefore, we determined that the appropriate thing to measure for engagement was the spatiotemporal coordination between the person and the other entities, or the appropriate dynamics of the activity.

We began our empirical work working on mouse tracking research. In cognitive science most mouse tracking analyses are performed on individual movement descriptives such as curvature, reaction time, or speed. However, we argue that the kind of information that would be interesting for constructs like engagement, which we think of as spatiotemporal coordination, will unfold across multiple movements. As such we developed new tools for analyzing cursor movements across multiple scales. We also showed that these tools were successful at capturing information related to the participants performance in a simple mouse tracking task that conventional mouse tracking measures would not have been applicable to.

For our next study we took those same tools and applied them to engagement (operationalized as improvement or modeled as a latent variable). We found that the degree to which a player approximated the spatiotemporal structure of good play, regardless of their actual score, was indicative of how much they reported engagement. This is particularly interesting because it is the first potentially domain neutral measure of engagement. In addition, this fits with our theory of engagement, that when people are engaged in some activity we should expect to see people's behavior be more related to that activity. In this case, participants shifted their behavior to better reflect the appropriate dynamics of the activity. Though this is a rather intuitive explanation of engagement, those who are more engaged behave in ways that reflect the activity they are engaged in, we think it is powerful that we have potentially discovered a way to quantify it.

We followed that experiment up with another experiment in which we also found that participants spatiotemporally coordinate with the dynamics of that which they are engaged with. In this task we created a stimulus that had multiscale dynamics and showed that humans spatiotemporally coordinate with the multiscale dynamics of the stimulus in the task. In addition, we showed that participants who improved more at the task, which we use as an operationalization of engagement, also increased their coordination more, suggesting again that spatiotemporal coordination with an activity is related to engagement.

In the immediate future, there is much work to do building on the findings discussed here. First work should be done confirming the usefulness of spatiotemporal coordination as a measure of engagement which would involve more experiments and a wider variety of mouse tracking tasks. The two experiments conducted in this work imply that spatiotemporal structure in behavior can be used to make inferences about engagement. However, these represent only two tasks. It would seem imperative that more experiments using a variety of mouse tracking tasks should be conducted. For example, Nalepka, Riehm, Mansour, Chemero, & Richardson (2015) studied participants performing a computerized sheep herding task with another participant.

Successful participants eventually form a synergy. It seems that spatiotemporal coordination of behavior, especially before the participants form a synergy could be related to engagement in the task. This would also help to explain why some participants don't seem to ever synergize and succeed at the task.

Beyond just mouse tracking, the ideas about spatiotemporal coordination need to be examined in other contexts such as sports or actual work environments. The true power of our theory of engagement is that it potentially explains engagement in any activity that has observable behavior. However, clearly this needs to be tested. Data from other domains, such as data from sports, or workers, or even just people performing any natural activity needs to be analyzed for spatiotemporal structure that implies coordination, and that needs to be compared with the engagement reported by those participants. One obvious study to perform would be to examine soccer data like that used in Duarte et a. (2013) and determine the extent to which individual players synergize with their teammates and how related that synergy is with engagement.

Finally, future experiments need to be done examining the role that different types of coordination (correlation, synchronization, complexity matching, synergizing, and potentially others) play in engagement, and in what context they are most applicable. It seems plausible that the ways humans spatiotemporally coordinate are task specific. For example, in our second experiment, the participants were asked to track an object on screen, and thus they produced behavior that has the same spatiotemporal structure as the stimulus. However, in work where participants were asked to collaborate with others as a group, or maybe to even behave antagonistically to a stimulus could require other forms of coordination such as synergy or antiphase synchronization. Furthermore, future experiments where participants collaborate with other participants or other agents might show different modes of coordination between individual participants and the dynamics of the rest of the group where some participants are synergizing together while others might only be slightly correlated with some of their neighboring agents. It seems plausible that there might be a continuum of integration that underlies coordination and different forms of coordination could be related to meaningful differences in engagement.

As it stands, these results emphasize the power of looking at mouse movements in aggregate rather than as individual movements. One of the important implications of the interaction dominant approach to cognition is that behaviors do not happen in isolation, rather behaviors such as mouse movements are often influenced by the behaviors that happened before them. To not account for this structure both limits the kinds of phenomena that can be studied, and also potentially introduces error into current studies (Wallot & Kelty-Stephen 2018).

Furthermore, these results demonstrate the value mouse tracking can add to cognitive science, beyond binary forced choice tasks. Mouse tracking, and other measurements of a user interface allow us to conduct research where we are able to control and manipulate part of the environment to see how the rest of the brain, body, environment system changes in response. Though this is certainly the idea in the seminal Dale et al. (2007) paper, adding tools like ours allows us to apply this paradigm to a much wider variety of phenomena, including constructs like engagement.

In addition, these results also provide a rare instance of a connection between phenomenology and empirical work and emphasize the power of taking a unified approach. Our phenomenological foundations for our model make it much more robust in what it can say about cognition. They provide us with a basis to explain not just what mechanistically happens when people are engaged, but why engagement should feel the way it does. Something that we find to be noticeably missing from Flow theory and other competing theories. In addition, our empirical results also provide feedback and provide new insights into phenomenology, that absorption and breakdown are not purely dependent on the availability of entities. In fact, one could also use these results to claim that breakdown and absorption are not distinct categories, but rather something that happens by degree and is dependent upon skill, availableness, and matchedness to care. This would also be, to my knowledge, a rare instance of empirical studies being used to modify existing phenomenological theory.

Finally, these results as a whole are important for several aspects of engagement. First, as our results support our phenomenological model of engagement, which we think implies the importance of Dreydeggarean phenomenology in understanding engagement. In our model of Dreydeggarean phenomenology absorption is characterized by how much one is able to perform some activity, and, according to us, how well that activity satisfies our care structure. The idea that engagement is dependent on how well an activity connects to our cares aligns well with many motivational models of engagement such as Flow or SDT theory (Baumann, 2021; Reeve, 2012). However, our model uniquely explains why humans feel the way they do when they are engaged and makes specific empirical predictions about what engaged behavior in general should be like. Our conclusions reveal two further insights, one: it is possible to measure engagement from a behavioral perspective, and two: that it is unrealistic to unilaterally increase engagement.

As far as measuring engagement, we show that engaged behavior should be task specific. The many instances of studies using task specific measures of engagement like clicks on a page should show that others have similar intuitions about engagement (Zhou et al., 2021). However, our model and results begin to show that engagement, even though it is context specific, also can be modeled in general as degree of spatiotemporal coordination or approximation of the appropriate dynamics of performing the task, which is a general concept but accounts for context specificity.

Measurements of spatiotemporal coordination can be done in a variety of ways. Although here we focused on scaling dynamics, several other methods such measures of synchrony, or synergy could certainly be appropriate. The impetus will be in the research to determine the modality through which the person is coordinating with the activity, then from there determine how best to measure that coordination. The added benefit here is vast as measures of spatiotemporal coordination are potentially cheaper, more easily applied, and more synchronous than many of the current alternative measures.

In addition, it is worth noting that this work also has implications for those who wish to facilitate engagement. In our model engagement is a product of the matchedness between the cares of the engaged person and the motivational affordances of the activity. This means that each combination is unique, and it would be unrealistic to unilaterally increase engagement. Instead, efforts to increase engagement should focus on discovering what the care structures of the target population are and try to modify the activity to better approximate those care structures. There is no magic bullet, badge, or leader board, something that has been increasingly reported in research on gamification (Chou 2019).

As a broad theoretical work, the work here should contribute to understanding engagement, measuring engagement, or facilitating it. But, in terms of practical application I believe we have only scratched the surface of what these ideas can be used for. And I am excited for the future contributions to this paradigm.

References

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2014). Complexity matching in dyadic conversation. Journal of Experimental Psychology: General, 143(6), 2304.

Alfes, K., Shantz, A. D., Truss, C., & Soane, E. C. (2013). The link between perceived human resource management practices, engagement and employee behaviour: a moderated mediation model. The international journal of human resource management, 24(2), 330-351.

Almurad, Z. M., Roume, C., & Delignières, D. (2017). Complexity matching in side-by-side walking. Human Movement Science, 54, 125-136.

Alviar, C., Dale, R., & Galati, A. (2019). Complex communication dynamics: Exploring the structure of an academic talk. *Cognitive science*, *43*(3), e12718.

Arapakis, I., & Leiva, L. A. (2016, July). Predicting user engagement with direct displays using mouse cursor information. In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval (pp. 599-608).

Araújo, D., Silva, P., and Davids, K. (2015). "Capturing group tactical behaviors in expert team players," in Routledge Handbook of Sport Expertise, eds J. Baker and D. Farrow (London: Routledge), 209–220.

Araújo, D., & Davids, K. (2016). Team synergies in sport: theory and measures. Frontiers in psychology, 7, 1449.

Araújo, D., Ramos, J., and Lopes, R. (2016). "Shared affordances guide interpersonal synergies in sport teams," in Interpersonal Coordination and Performance in Social Systems, eds P. Passos, K. Davids, and J. Y. Chow (London: Routledge), 165–178.

Axelson, R. D., & Flick, A. (2010). Defining student engagement. Change: The magazine of higher learning, 43(1), 38-43.

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. Educational psychologist, 50(1), 84-94.

Bailey, C., Madden, A., Alfes, K., & Fletcher, L. (2017). The meaning, antecedents and outcomes of employee engagement: A narrative synthesis. International Journal of Management Reviews, 19(1), 31-53.

Bakker, A. B., Schaufeli, W. B., Leiter, M. P., & Taris, T. W. (2008). Work engagement: An emerging concept in occupational health psychology. Work & stress, 22(3), 187-200.

Bakker, A. B., Albrecht, S. L., & Leiter, M. P. (2011). Key questions regarding work engagement. European journal of work and organizational psychology, 20(1), 4-28.

Bakker, A. B., & Costa, P. L. (2014). Chronic job burnout and daily functioning: A theoretical analysis. Burnout Research, 1(3), 112-119.

Bakker, A. B., & Demerouti, E. (2017). Job demands–resources theory: taking stock and looking forward. Journal of occupational health psychology, 22(3), 273.

Bakker, A. B., & Albrecht, S. (2018). Work engagement: current trends. Career Development International.

Baron, P., & Corbin, L. (2012). Student engagement: Rhetoric and reality. Higher Education Research & Development, 31(6), 759-772.

Baumann, N. (2021). Autotelic personality. Advances in flow research, 231-261.

Balwant, P. T. (2018). The meaning of student engagement and disengagement in the classroom context: Lessons from organisational behaviour. Journal of Further and Higher Education, 42(3), 389-401.

Bennett, D., Roudaut, A., & Metatla, O. (2022, April). Multifractal Mice: Operationalising Dimensions of Readiness-to-hand via a Feature of Hand Movement. In CHI Conference on Human Factors in Computing Systems (pp. 1-18).

Bixler, R., & D'Mello, S. (2013, March). Detecting boredom and engagement during writing with keystroke analysis, task appraisals, and stable traits. In Proceedings of the 2013 international conference on Intelligent user interfaces (pp. 225-234).

Bogost, I. (2015). Why gamification is bullshit. The gameful world: Approaches, issues, applications, 65.

Braver, L. (2014). Heidegger: thinking of being. John Wiley & Sons.

Breivik, G. (2007). Skillful coping in everyday life and in sport: A critical examination of the views of Heidegger and Dreyfus. Journal of the Philosophy of Sport, 34(2), 116-134.

Bryce, R. M., & Sprague, K. B. (2012). Revisiting detrended fluctuation analysis. Scientific reports, 2(1), 1-6.

Calcagnì, A., Lombardi, L., D'Alessandro, M., & Freuli, F. (2019). A state space approach to dynamic modeling of mouse-tracking data. Frontiers in psychology, 2716.

Cavagna, A., Cimarelli, A., Giardina, I., Parisi, G., Santagati, R., Stefanini, F., & Viale, M. (2010). Scale-free correlations in starling flocks. Proceedings of the National Academy of Sciences, 107(26), 11865-11870.

Cepeda, C., Rodrigues, J., Dias, M. C., Oliveira, D., Rindlisbacher, D., Cheetham, M., & Gamboa, H. (2018, August). Mouse tracking measures and movement patterns with application for online surveys. In International Cross-Domain Conference for Machine Learning and Knowledge Extraction (pp. 28-42). Springer, Cham.

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2008, October). Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. In Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era (pp. 13-17).

Chou, Y. K. (2019). Actionable gamification: Beyond points, badges, and leaderboards. Packt Publishing Ltd.

Ciocan, C. (2010). Heidegger and the Problem of Boredom. Journal of the British Society for Phenomenology, 41(1), 64-77.

Collaço, C. M. (2017). Increasing student engagement in higher education. Journal of Higher Education Theory and Practice, 17(4).

Dale, R., & Bhat, H. S. (2018). Equations of mind: Data science for inferring nonlinear dynamics of socio-cognitive systems. Cognitive Systems Research, 52, 275-290.

Darnell, D. K., & Krieg, P. A. (2019). Student engagement, assessed using heart rate, shows no reset following active learning sessions in lectures. PloS one, 14(12), e0225709.

Delignières, D., Almurad, Z. M., Roume, C., & Marmelat, V. (2016). Multifractal signatures of complexity matching. Experimental brain research, 234(10), 2773-2785.

Demmel, J. W. (1997). Applied numerical linear algebra. Society for Industrial and Applied Mathematics.

Diniz, A., Wijnants, M. L., Torre, K., Barreiros, J., Crato, N., Bosman, A. M., ... & Delignières, D. (2011). Contemporary theories of 1/f noise in motor control. Human movement science, 30(5), 889-905.

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., ... & Poeppel, D. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. Current biology, 27(9), 1375-1380.

Dikker, S., Michalareas, G., Oostrik, M., Serafimaki, A., Kahraman, H. M., Struiksma, M. E., & Poeppel, D. (2021). Crowdsourcing neuroscience: inter-brain coupling during face-to-face interactions outside the laboratory. NeuroImage, 227, 117436.

Dotov, D. G., Nie, L., & Chemero, A. (2010). A demonstration of the transition from ready-to-hand to unready-to-hand. PLoS One, 5(3), e9433.

Dotterer, A. M., & Lowe, K. (2011). Classroom context, school engagement, and academic achievement in early adolescence. Journal of youth and adolescence, 40(12), 1649-1660.

Dreyfus, H. L. (1991). Being-in-the-world: A commentary on Heidegger's Being and Time, Division I. Mit Press.

Dreyfus, H. L. (2007). Response to McDowell. Inquiry, 50(4), 371-377.

Duarte, R., Araújo, D., Correia, V., Davids, K., Marques, P., & Richardson, M. J. (2013). Competing together: Assessing the dynamics of team–team and player–team synchrony in professional association football. Human movement science, 32(4), 555-566.

Erb, C. D., Moher, J., Sobel, D. M., & Song, J. H. (2016). Reach tracking reveals dissociable processes underlying cognitive control. Cognition, 152, 114-126.

Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J. P., & Gašević, D. (2019, March). Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. In Proceedings of the 9th international conference on learning analytics & knowledge (pp. 501-510).

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of educational research, 74(1), 59-109.

Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. Psychological science, 20(10), 1183-1188.

Gallagher, S. (2014). Phenomenology and embodied cognition. The Routledge handbook of embodied cognition, 9-18.

Gallagher, S., & Zahavi, D. (2020). The phenomenological mind. Routledge.

Gaventa, J., & Barrett, G. (2012). Mapping the outcomes of citizen engagement. World development, 40(12), 2399-2410.

Granger, C. W., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. Journal of time series analysis, 1(1), 15-29.

Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. Educational Psychologist, 50(1), 14-30.

Hakanen, J. J., Schaufeli, W. B., & Ahola, K. (2008). The Job Demands-Resources model: A three-year cross-lagged study of burnout, depression, commitment, and work engagement. Work & stress, 22(3), 224-241.

Haken, H., Kelso, J. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. Biological cybernetics, 51(5), 347-356.

Hardstone, R., Poil, S. S., Schiavone, G., Jansen, R., Nikulin, V. V., Mansvelder, H. D., & Linkenkaer-Hansen, K. (2012). Detrended fluctuation analysis: a scale-free view on neuronal oscillations. Frontiers in physiology, 3, 450.

Hatab, L. J. (2018). Redescribing the Zuhanden-Vorhanden relation. Gatherings: The Heidegger Circle Annual, 8, 21-35.

Hausdorff, J. M., Purdon, P. L., Peng, C. K., Ladin, Z. V. I., Wei, J. Y., & Goldberger, A. L. (1996). Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations. Journal of applied physiology, 80(5), 1448-1457.

Hausdorff, J. M. (2007). Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking. Human movement science, 26(4), 555-589.

Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. Group Processes & Intergroup Relations, 18(3), 384-401.

Heidegger, M., Macquarrie, J., & Robinson, E. (1962). Being and time.

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. Computers & Education, 90, 36-53.

Hookham, G., & Nesbitt, K. (2019, January). A systematic review of the definition and measurement of engagement in serious games. In Proceedings of the australasian computer science week multiconference (pp. 1-10).

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. Social cognition, 27(6), 949-960.

Ihlen, E. A., & Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: Beyond $1/f\alpha$ fluctuation. Journal of Experimental Psychology: General, 139(3), 436.

Ihlen, E. A. F. E. (2012). Introduction to multifractal detrended fluctuation analysis in Matlab. Frontiers in physiology, 3, 141.

Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. Journal of sport and exercise psychology, 18(1), 17-35.

Jackson, S. A., & Csikszentmihalyi, M. (1999). Flow in sports. Human Kinetics.

Jackson, S. A., & Csikszentmihalyi, M. (1999). Flow in sports. Human Kinetics.

Kahn, W.A. (1990), “Psychological conditions of personal engagement and disengagement at work”, Academy of Management Journal, Vol. 33, pp. 692-724.

Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. Journal of Experimental Psychology: General, 146(4), 569.

Katahira, K., Yamazaki, Y., Yamaoka, C., Ozaki, H., Nakagawa, S., & Nagata, N. (2018). EEG correlates of the flow state: A combination of increased frontal theta and moderate frontocentral alpha rhythm in the mental arithmetic task. Frontiers in psychology, 9, 300.

Kello, C. T., Brown, G. D., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. Trends in cognitive sciences, 14(5), 223-232.

Kello, C. T., Bella, S. D., Médé, B., & Balasubramaniam, R. (2017). Hierarchical temporal structure in music, speech and animal vocalizations: jazz is like a conversation, humpbacks sing like hermit thrushes. Journal of The Royal Society Interface, 14(135), 20170231.

Khan, S. I., & Ahmad, B. (2021). Tweet so good that they can't ignore you! Suggesting posting strategies to micro-celebrities for online engagement. Online Information Review.

Kieslich, P. J., Henninger, F., Wulff, D. U., Haslbeck, J. M., & Schulte-Mecklenbeck, M. (2019). Mouse-tracking: A practical guide to implementation and analysis 1. In A handbook of process tracing methods (pp. 111-130). Routledge.

Kieslich, P. J., & Henninger, F. (2017). Mousetrap: An integrated, open-source mouse-tracking package. Behavior research methods, 49(5), 1652-1667.

Kim, J. K., & LePine, J. A. (2019). Employee Engagement: The Past, the Present, and the Future. The SAGE Handbook of Human Resource Management, 371.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. Cognitive science, 18(4), 513-549.

Klasen, M., Weber, R., Kircher, T. T., Mathiak, K. A., & Mathiak, K. (2012). Neural contributions to flow experience during video game playing. Social cognitive and affective neuroscience, 7(4), 485-495.

Knierim, M. T., Rissler, R., Hariharan, A., Nadj, M., & Weinhardt, C. (2019). Exploring flow psychophysiology in knowledge work. In Information Systems and Neuroscience (pp. 239-249). Springer, Cham.

Kołakowska, A. (2013, June). A review of emotion recognition methods based on keystroke dynamics and mouse movements. In 2013 6th International Conference on Human System Interactions (HSI) (pp. 548-555). IEEE.

Kowal, J., & Fortier, M. S. (1999). Motivational determinants of flow: Contributions from self-determination theory. The journal of social psychology, 139(3), 355-368.

Kuh, G. D. (2008). Excerpt from high-impact educational practices: What they are, who has access to them, and why they matter. Association of American Colleges and Universities, 14(3), 28-29.

Kuh, G. D. (2009). The national survey of student engagement: Conceptual and empirical foundations. New directions for institutional research, 2009(141), 5-20.

Lalmas, M., O'Brien, H., & Yom-Tov, E. (2014). Measuring user engagement. Synthesis lectures on information concepts, retrieval, and services, 6(4), 1-132.

Lagarde, J., Peham, C., Licka, T., & Kelso, J. S. (2005). Coordination dynamics of the horse-rider system. Journal of motor behavior, 37(6), 418-424.

Lalmas, M., O'Brien, H., & Yom-Tov, E. (2014). Measuring user engagement. Synthesis lectures on information concepts, retrieval, and services, 6(4), 1-132.

Lawson, M. A., & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. Review of Educational Research, 83(3), 432-479.

Lee, J. S. (2014). The relationship between student engagement and academic performance: Is it a myth or reality?. The Journal of Educational Research, 107(3), 177-185.

Liu, H., Fernando, O. N. N., & Rajapakse, J. C. (2018, November). Predicting affective states of programming using keyboard data and mouse behaviors. In 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV) (pp. 1408-1413). IEEE.

Londen, P. "Beyond 'Dreidegger': The Future of Anglo-American Existential Phenomenology. Horizons of Phenomenlogy." Springer. Londen, Walsh, and Yoshimi ed. forthcoming

Maldonado, M., Dunbar, E., & Chemla, E. (2019). Mouse tracking as a window into decision making. Behavior Research Methods, 51(3), 1085-1101.

Mandernach, B. J. (2015). Assessment of student engagement in higher education: A synthesis of literature and assessment tools. International Journal of Learning, Teaching and Educational Research, 12(2).

Marghetis, T., & Núñez, R. (2013). The motion behind the symbols: A vital role for dynamism in the conceptualization of limits and continuity in expert mathematics. Topics in cognitive science, 5(2), 299-316.

Mosing, M. A., Magnusson, P. K., Pedersen, N. L., Nakamura, J., Madison, G., & Ullén, F. (2012). Heritability of proneness for psychological flow experiences. Personality and Individual Differences, 53(5), 699-704.

Nacke, L., & Lindley, C. A. (2008, November). Flow and immersion in first-person shooters: measuring the player's gameplay experience. In Proceedings of the 2008 conference on future play: Research, play, share (pp. 81-88).

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. Handbook of positive psychology, 195, 206.

Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In Flow and the foundations of positive psychology (pp. 239-263). Springer, Dordrecht.

Nazir, T. A., Boulenger, V., Roy, A., Silber, B., Jeannerod, M., & Paulignan, Y. (2008). Language-induced motor perturbations during the execution of a reaching movement. Quarterly Journal of Experimental Psychology, 61(6), 933-943.

Nalepka, P., Riehm, C., Mansour, C. B., Chemero, A., & Richardson, M. J. (2015). Investigating Strategy Discovery and Coordination in a Novel Virtual Sheep Herding Game among Dyads. In CogSci.

Nolds module | 0.5.2 documentation. (n.d.). Retrieved 2021-03-02, from https://nolds .readthedocs.io/en/latest/nolds.html#detrended-fluctuation-analysis

Northoff, G. (2016). How do resting state changes in depression translate into psychopathological symptoms? From ‘Spatiotemporal correspondence’to ‘Spatiotemporal Psychopathology’. Current opinion in psychiatry, 29(1), 18-24.

Northoff, G. (2020). Anxiety disorders and the brain’s resting state networks: from altered spatiotemporal synchronization to psychopathological symptoms. Anxiety Disorders, 71-90.

O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. International Journal of Human-Computer Studies, 112, 28-39.

Pan, Y., Novembre, G., Song, B., Li, X., & Hu, Y. (2018). Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. Neuroimage, 183, 280-290.

Pan, Y., Dikker, S., Goldstein, P., Zhu, Y., Yang, C., & Hu, Y. (2020). Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. NeuroImage, 211, 116657.

Parsons, J., & Taylor, L. (2011). Improving student engagement. Current issues in education, 14(1).

Pedulla, T. (2020, November 16). Carl Nafzger: Listening to the horse. America's Best Racing. Retrieved May 23, 2022, from https://www.americasbestracing.net/the-sport/2020-carl-nafzger-listening-the-horse

Peng, C. K., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos: an interdisciplinary journal of nonlinear science, 5(1), 82-87.

Petrou, P., Demerouti, E., Peeters, M. C., Schaufeli, W. B., & Hetland, J. (2012). Crafting a job on a daily basis: Contextual correlates and the link to work engagement. Journal of Organizational Behavior, 33(8), 1120-1141.

Porter, S. R. (2006). Institutional structures and student engagement. Research in Higher Education, 47(5), 521-558.

Proksch, S., Reeves, M., Spivey, M., & Balasubramaniam, R. (2022). Coordination dynamics of multi-agent interaction in a musical ensemble. Scientific reports, 12(1), 1-14.

Ramirez-Aristizabal, A. G., Médé, B., & Kello, C. T. (2018). Complexity matching in speech: Effects of speaking rate and naturalness. Chaos, Solitons & Fractals, 111, 175-179.

Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: coordinated body movement reflects relationship quality and outcome. Journal of consulting and clinical psychology, 79(3), 284.

Reeve, J. (2012). A self-determination theory perspective on student engagement. In Handbook of research on student engagement (pp. 149-172). Springer, Boston, MA.

Reeve, J., Cheon, S. H., & Jang, H. R. (2019). A teacher-focused intervention to enhance students' classroom engagement. In Handbook of student engagement interventions (pp. 87-102). Academic Press.

Reinero, D. A., Dikker, S., & Van Bavel, J. J. (2021). Inter-brain synchrony in teams predicts collective performance. Social cognitive and affective neuroscience, 16(1-2), 43-57.

Rheem, H., Verma, V., & Becker, D. V. (2018, September). Use of mouse-tracking method to measure cognitive load. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 62, No. 1, pp. 1982-1986). Sage CA: Los Angeles, CA: SAGE Publications.

Rigoli, L. M., Holman, D., Spivey, M. J., & Kello, C. T. (2014). Spectral convergence in tapping and physiological fluctuations: coupling and independence of 1/f noise in the central and autonomic nervous systems. Frontiers in Human Neuroscience, 8, 713.

Riley, M. A., Richardson, M., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. Frontiers in psychology, 2, 38.

Saks, A. M. (2006). Antecedents and consequences of employee engagement. Journal of managerial psychology.

Saks, A.M. (2019), "Antecedents and consequences of employee engagement revisited", Journal of Organizational Effectiveness: People and Performance, Vol. 6 No. 1, pp. 19-38.

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. Journal of Happiness studies, 3(1), 71-92.

Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 25(3), 293-315.

Schaufeli, W. B., & Bakker, A. B. (2010). Defining and measuring work engagement: Bringing clarity to the concept. Work engagement: A handbook of essential theory and research, 12, 10-24.

Schaufeli, W. B. (2013). What is engagement. Employee engagement in theory and practice, 15(321), 9780203076965-10.

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In Coordination: Neural, behavioral and social dynamics (pp. 281-308). Springer, Berlin, Heidelberg.

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2014). Student engagement in high school classrooms from the perspective of flow theory. In Applications of flow in human development and education (pp. 475-494). Springer, Dordrecht.

Schloesser, D. S., Kello, C. T., & Marmelat, V. (2019). Complexity matching and coordination in individual and dyadic performance. Human movement science, 66, 258-272.

Schneider, S., Ramirez-Aristizabal, A. G., Gavilan, C., & Kello, C. T. (2020). Complexity matching and lexical matching in monolingual and bilingual conversations. Bilingualism: Language and Cognition, 23(4), 845-857.

Schneider, B., Yost, A. B., Kropp, A., Kind, C., & Lam, H. (2018). Workforce engagement: What it is, what drives it, and why it matters for organizational performance. Journal of Organizational Behavior, 39(4), 462-480.

Shapiro, Lawrence and Shannon Spaulding, "Embodied Cognition", The Stanford Encyclopedia of Philosophy (Winter 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2021/entries/embodied-cognition/>.

Shao, Y. H., Gu, G. F., Jiang, Z. Q., Zhou, W. X., & Sornette, D. (2012). Comparing the performance of FA, DFA and DMA using different synthetic long-range correlated time series. Scientific reports, 2(1), 1-5.

Sheldon, K. M., Prentice, M., & Halusic, M. (2015). The experiential incompatibility of mindfulness and flow absorption. Social Psychological and Personality Science, 6(3), 276-283.

Schulte-Mecklenbeck, M., Kühberger, A., & Johnson, J. G. (Eds.). (2019). A handbook of process tracing methods. Routledge.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. Educational psychologist, 50(1), 1-13.

Spivey, M. (2008). The continuity of mind. Oxford University Press.

Spivey, M. J. (2020). Who you are: The science of connectedness. Mit Press.

Stephen, D. G., & Dixon, J. A. (2011). Strong anticipation: Multifractal cascade dynamics modulate scaling in synchronization behaviors. Chaos, Solitons & Fractals, 44(1-3), 160-168.

Stergiou, N., & Decker, L. M. (2011). Human movement variability, nonlinear dynamics, and pathology: is there a connection?. Human movement science, 30(5), 869-888.

Stillman, P. E., Shen, X., & Ferguson, M. J. (2018). How mouse-tracking can advance social cognitive theory. Trends in cognitive sciences, 22(6), 531-543.

Strogatz, S. H. (2012). Sync: How order emerges from chaos in the universe, nature, and daily life. Hachette UK.

Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. Group Processes & Intergroup Relations, 18(3), 384-401.

Taasoobshirazi, G., & Wang, S. (2016). The performance of the SRMR, RMSEA, CFI, and TLI: An examination of sample size, path size, and degrees of freedom. Journal of Applied Quantitative Methods, 11(3), 31-39.

Trefethen, L. N., & Bau III, D. (1997). Numerical linear algebra (Vol. 50). Siam.

Trowler, V. (2010). Student engagement literature review. The higher education academy, 11(1), 1-15.

Turvey, M. T. (2007). Action and perception at the level of synergies. Human movement science, 26(4), 657-697.

Ulrich, M., Keller, J., Hoenig, K., Waller, C., & Grön, G. (2014). Neural correlates of experimentally induced flow experiences. Neuroimage, 86, 194-202.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. Journal of Experimental Psychology: General, 132(3), 331.

van Roy, R., & Zaman, B. (2019). Unravelling the ambivalent motivational power of gamification: A basic psychological needs perspective. International Journal of Human-Computer Studies, 127, 38-50.

Varela, F. J., Thompson, E., & Rosch, E. (2016). The embodied mind: Cognitive science and human experience. MIT press.

Varga, Somogy and Guignon, & Charles (2020). Authenticity. The Stanford Encyclopedia of Philosophy. Spring 2020

Vilar, L., Araújo, D., Davids, K., & Bar-Yam, Y. (2013). Science of winning soccer: Emergent pattern-forming dynamics in association football. Journal of systems science and complexity, 26(1), 73-84.

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-related changes in 1/f neural electrophysiological noise. Journal of Neuroscience, 35(38), 13257-13265.

Vuori, J. (2014). Student engagement: buzzword of fuzzword?. Journal of Higher Education Policy and Management, 36(5), 509-519.

Wallot, S., & Kelty-Stephen, D. G. (2018). Interaction-dominant causation in mind and brain, and its implication for questions of generalization and replication. Minds and Machines, 28(2), 353-374.

Ward, K. (2021). Breaking down experience—Heidegger's methodological use of breakdown in Being and Time. European Journal of Philosophy, 29(4), 712-730.

Webber, K. L., Krylow, R. B., & Zhang, Q. (2013). Does involvement really matter? Indicators of college student success and satisfaction. Journal of College Student Development, 54(6), 591-611.

West, G. B., & Brown, J. H. (2005). The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization. Journal of experimental biology, 208(9), 1575-1592.

West, B. J., Geneston, E. L., & Grigolini, P. (2008). Maximizing information exchange between complex networks. Physics Reports, 468(1-3), 1-99.

West, G. (2018). Scale: The universal laws of life, growth, and death in organisms, cities, and companies. Penguin.

Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. Computers in Human Behavior, 32, 123-132.

Wohltjen, S., & Wheatley, T. (2021). Eye contact marks the rise and fall of shared attention in conversation. Proceedings of the National Academy of Sciences, 118(37).

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. Educational and psychological measurement, 73(6), 913-934.

Wrathall, M. A. (2005). How to read Heidegger. Granta Books.

Xanthopoulou, D., Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2009). Reciprocal relationships between job resources, personal resources, and work engagement. Journal of Vocational behavior, 74(3), 235-244.

Zhao, Z., Salesse, R. N., Marin, L., Gueugnon, M., & Bardy, B. G. (2017). Likability's effect on interpersonal motor coordination: Exploring natural gaze direction. Frontiers in psychology, 8, 1864.

Zhou, Y., Calder, B. J., Malthouse, E. C., & Hessary, Y. K. (2021). Not all clicks are equal: detecting engagement with digital content. Journal of Media Business Studies, 1-18.

Appendix

Analysis of Windowed Complexity Matching

Although complexity matching in the first third was not predictive of future improvement, we feel that this might have been because the time windows were too large. As we said previously, several of our best improvers played very little during the beginning of the game, which is why they improved so much. Instead, we took a more fine-grained approach. We split the data into window sizes of 1000 timepoints, with a 50% overlap. We used 1000 because that is the minimum number of data points recommended for use with DFA by Ihlen (2012). We recognize that others have slightly different recommendations, but we found that around 1000 points seemed to be a relatively common minimum. In addition, we tested this on an individual player and found that the results of the DFA was relatively stable using windows of this size. For example, Fig 24 shows the sliding complexity values for the best player. You can see that the line is nearly straight, indicating very little differences between windows, indicating stability of DFA at that window size in this task since play for this player is relatively constant. We did not try any other windows smaller than this.

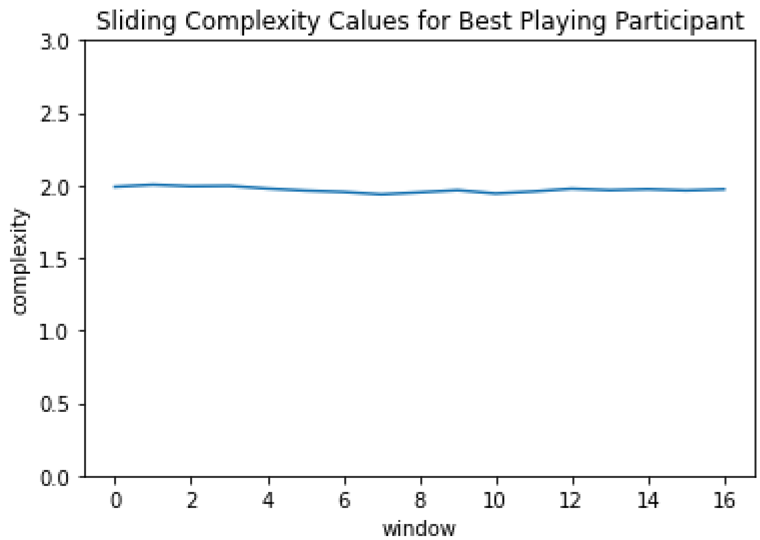

Figure 24. The complexity values in different windows for a player with 100% accuracy. The stability of the line indicates that the measurement of the hurst parameter is relatively stable at this window size.

Doing this we were able to compute complexity matching for each participant with the golden player inside each sliding window. We call this s*liding complexity matching*. We can then compute the change in complexity matching from each window to the next which we termed the *delta sliding complexity matching*. We also computed the accuracy scored in those same windows and the change in accuracy, which is a measure of improvement. We then regressed delta_complexitymatching onto timelagged 1 delta_accuracy (improvement) plotted in Fig 25. A relationship between change in complexity matching and time lagged improvement would indicate a relationship in which complexity matching could be predictive of future improvement. This is what we might expect in our model of engagement that engaged players might be playing well in terms of the movements they are performing without scoring well. Plotted in Fig 25 represents the no pooling model for improvement regressed on change in complexity matching: B = 0.0008, SE B = 0.0003, and the R^2 = < .01 which is very small, p <= .000.

Given that this data is truly nested with 14 data points per person, we also ran a multilevel model which did not converge. We ran a second multilevel model where we removed participants whose slope of improvement was less than |.1| under the belief that those players have small changes in complexity matching and small changes in accuracy and are thereby clustering in the middle and potentially harming homogeneity of variance and causing the model not to converge. This model did converge but still had a small beta of .003 and a standard error of .001 and was significant ($p < .01$). However, as the R^2 above is quite small we find these results to be somewhat weak. However, we do find the graphical evidence to be somewhat compelling and we think this could be an issue with having too short of time series to model this in an appropriate way. The question of interest here is one of time series synchronization. However, the number of samples per participant is too low to do reliable time series modeling (only 14 time points per participant). This is modeled naively as a simple regression of complexity matching on lagged improvement. In future it would be interesting to test this idea on longer time series where we could test a nested model of time series synchronization to determine if complexity matching actually drives improvement.

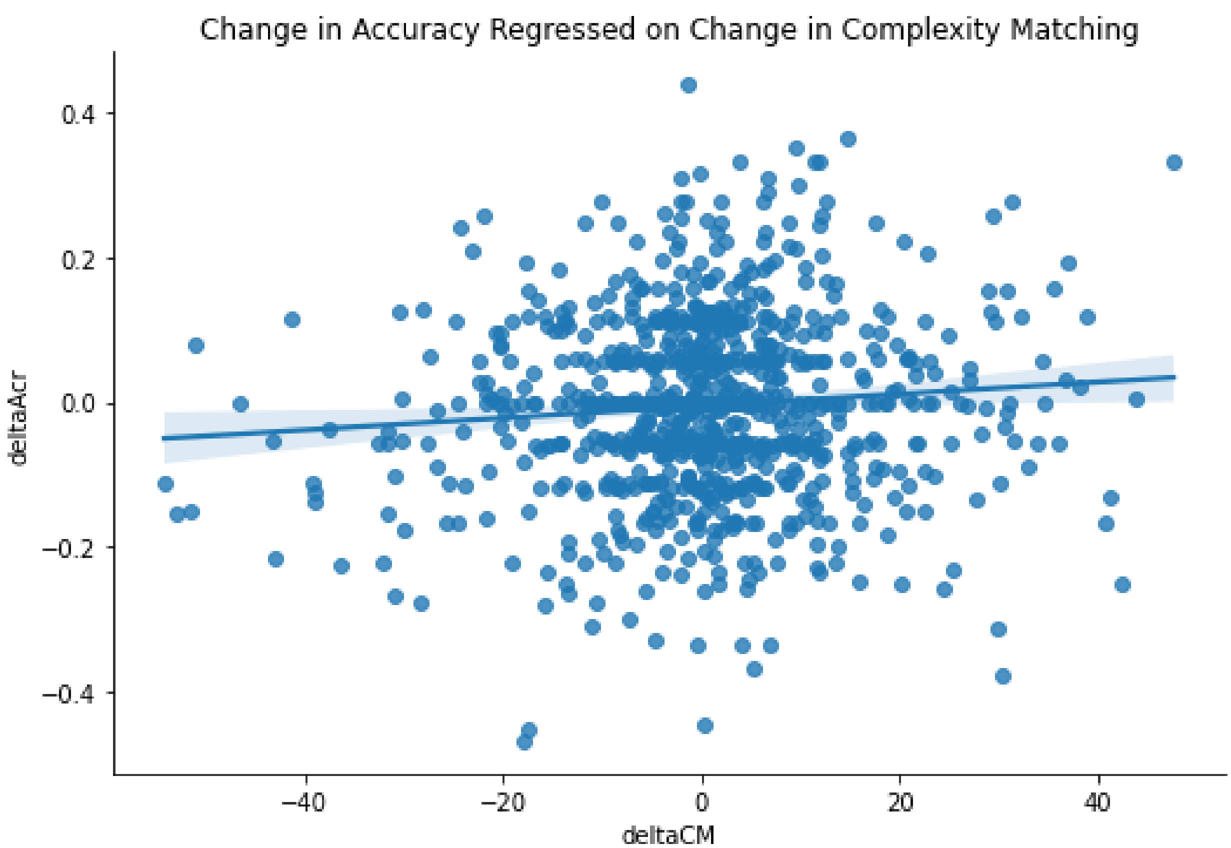

Figure 25. The no pooling model of change in complexity matching at lag 1 predicting improvement. The B is BLANK and significant p = BLANK which indicates that complexity matching is related to improvement in the future.