EVALUATING COGNITIVE PROCESSES DURING DECISION-MAKING TASKS USING MOBILE PHONES, COMPUTER MICE, AND KEYBOARDS: THEORY, METHOD, AND APPLICATION

by

David Kim

A Dissertation Submitted to the Faculty of the

ELLER COLLEGE OF MANAGEMENT

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2023

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by: *David Kim*, titled: Evaluating Cognitive Processes during Decision-Making Tasks using Mobile Phones, Computer Mice, and Keyboards: Theory, Method, and Application and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Joseph S. Valacich Date: 5/17/23

Jeffrey L. Jenkins Date: 5/19/23

Laura Brandimarte Date: 5/17/23

Wei Chen Date: 5/19/23

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Joseph S. Valacich Date: 5/17/23
Dissertation Committee Chair
Management Information Systems

2

ACKNOWLEDGEMENTS

First and foremost, my journey, from the initiation to the finish, would have been impossible without the unwavering support of Joe Valacich, my advisor. Over the past five years, his guidance has been instrumental in my academic growth and has shaped my life in countless ways. I am deeply grateful for your constant presence and unyielding guidance throughout the Ph.D. program. Thank you for having me as your student.

I am equally indebted to Jeffrey Jenkins, who was generous enough to illuminate the path toward conducting impactful research using Trace Data. As a fresh first-year Ph.D. student without a programming background, his advice and guidance have been pivotal in developing the critical analytical capabilities central to my research. I sincerely thank him for the invaluable guidance and support he has given.

Also, I am fortunate to have Michael Byrd, Manasvi Kumar, Paul Weisgarber, and Dave Wilson as my colleagues. Thank you for everything. It is a great pleasure to work with you all.

Finally, I would like to thank the faculties, friends, and former students I had from the Management Information Systems Department at the University of Arizona. It was a long but fun journey.

DEDICATION

To Dad, for always reminding me that dissertations are less daunting than assembling IKEA furniture.

To Mom, for the coffees that fueled this dissertation.

To my sister, for always being my emergency contact.

To my brother-in-law, for enduring my sister… and me.

To my niece, for your expert advice on unicorns.

To my nephew, for believing I was writing the world's longest bedtime story.

To all my friends, for pretending to understand what I was talking about.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Trace Data, characterized by fine-grained metadata collected at millisecond precision, has become increasingly popular in Human-Computer Interaction studies to investigate various cognitive states of individuals. However, a significant challenge these studies face is the considerable variation in their theoretical foundation, leading to a diverse array of disconnected empirical findings. This dissertation proposes a unified theoretical framework, Decision Evaluation with Cognitive Observations using Device Engagement (DECODE), grounded in Dual Process Theory, a widely accepted theory that posits that human's thinking is an interplay between intuitive and deliberate processes and Cognitive Control. This widely accepted construct encompasses attention regulation, decision-making, and goal-directed behavior. The DECODE framework proposes that the Cognitive Task Requirements (CTR), which signifies the cognitive demand imposed on an individual by a particular task, and task-specific expertise, influence Cognitive Control (CC) and individuals' engagement in intuitive and deliberate thinking.

This dissertation empirically validates that the differences in behaviors proposed by the DECODE framework can be captured and analyzed with trace data from various Human-Computer Interaction devices, including computer mice, keyboards, and mobile devices. 'Essay One' examines how CTR and individual task-specific expertise influence mouse movement speed. 'Essay Two' further investigates the role of CTR and an individual's task-specific expertise in typing fluency using the keyboard. 'Essay Three' extends previous studies' findings by demonstrating how removing CTR speeds up decision-making using mobile devices. The overall results from all three essays suggest that increasing CTR significantly slows the device usage speed while developing task-specific expertise increases the device usage speed.

CHAPTER I: INTRODUCTION

Human-Computer Interaction (HCI) devices such as computer mice, keyboards, and touch screens are ubiquitous. Modern HCI devices are equipped with high-fidelity sensors that aid in navigation, interaction, and data entry. A growing body of research is demonstrating that metadata about how interaction occurs and how information is entered can be used to understand various cognitive and emotional reactions. Specifically, metadata measuring how such devices are used – i.e., keypress duration, mouse movement speed, or touch pressure – can be used to measure physiological responses indicative of underlying cognitive processes (Freeman and Ambady 2010). This fine-grained metadata is collected at millisecond precision and is increasingly being referred to as trace data (Weinmann et al. in press). Methods for understanding users' cognitive and emotional reactions using computer systems have historically been collected with various neurological technologies, including EEG caps, eye tracking, and fMRI (Gidlöf et al. 2013; Zheng et al. 2014; Zwicker et al. 2011). A drawback of neurological technologies for collecting cognitive and emotional signals is low ecological validity (Holleman et al. 2020) – i.e., these methods are highly controlled and artificial, especially when compared to users' interaction with a website or application in a natural setting. To mirror the interaction methods used in natural settings and overcome ecological validity concerns, trace data from HCI devices is a rapidly emerging research method for understanding users' cognitive and emotional reactions in a variety of contexts.

While there are an increasing number of studies using trace data for a range of contexts (e.g., fraud, usability, emotion, policy compliance, insider threats, etc.), these studies have also leveraged different theoretical frameworks confined to the specific system usage behavior and context (Hibbeln et al. 2017; Jenkins. et al. 2019; Vance et al. 2018). Further, applying these theoretical underpinnings to various types of HCI devices remains largely unexplored. As such, there is a need for a more encompassing theoretical lens for not only integrating the finding of the prior work but also for motivating and guiding future studies, as there is a diverse range of disconnected empirical findings (Refer to review in Chapter 2).

This dissertation draws from Dual-Process Theory (DPT), which asserts that human thought processes consist of intuitive (i.e., System 1) or conscious (i.e., System 2) types of thinking, leveraging Cognitive Control (CC) – a construct representing the mental process that allows an individual to control and allocate mental resources – to demonstrate that device usage speed is an indicator of users' cognitive effort. Integrating empirical findings from the relevant literature, a new theoretical framework, the Decision Evaluation with Cognitive Observations using Device Engagement (DECODE), is proposed, which can be applied to various contexts and HCI devices. To build DECODE, DPT is utilized to explain how individuals process external information in decision-making tasks and leverage varying levels of System 1 and System 2 types of thinking. Further, CC illustrates how increased task-specific expertise on decision-making tasks leads to a shift towards reliance on System 1 types of thinking, which influences the individual's relative dependency on System 1 or System 2 types of thinking.

The main research question addressed in this dissertation is:

> "How do attention-demanding cues (e.g., visual elements) and individual task-specific expertise (e.g., context familiarity) influence the speed of HCI device usage across different contexts and devices?"

A fundamental construct, Cognitive Task Requirements (CTR), is introduced earlier in the discussion to answer this question. CTR signifies the cognitive demands imposed on an individual by a particular task. This concept encompasses attention-demanding cues, such as visual elements and task difficulties, which can either increase or decrease CTR, subsequently influencing individuals' cognitive control. By incorporating the CTR construct, three research essays are developed to empirically demonstrate how the cognitive demands associated with decision-making and individuals' competencies influence computer mice, keyboards, and mobile device usage speed.

'Essay One' examines how CTR and individual task-specific expertise influence mouse movement speed. The study employs a survey with a manipulation condition that encourages individuals to be more attentive during decision-making, thus engaging a higher degree of cognitive control. The analysis reveals that increased CTR slows the mouse cursor movement speed. In contrast, the individuals' task-specific expertise speeds up the decision-making process and, thereby, the mouse cursor movements.

'Essay Two' further examines the role of CTR and an individual's task-specific expertise in typing fluency (i.e., keyboard) context. The essay includes three experiments, with a different focus on either CTR or individual task-specific expertise. The first experiment described in the essay focuses on the impact of CTR on typing fluency by presenting participants with different difficulty levels (e.g., typing their identity-related information vs. typing unfamiliar text). The second experiment focuses on the influence of task-specific expertise by requiring participants to type their name, an unfamiliar international name, and an unfamiliar English name multiple times to examine how context familiarity and competence development through repetition influence typing fluency. The last experiment examines the combined effects of CTR and task-specific expertise on typing fluency in an applied context (e.g., tying familiar and unfamiliar identity information into an online form). The overall results from four experiments indicate that an increase in CTR decreases typing fluency. However, task-specific expertise acts as a moderator, such that as individuals gain task-specific expertise through repetition, the negative impact of the increase in CTR on typing fluency diminishes.

'Essay Three' extends previous studies' findings by demonstrating how removing CTR speeds up decision-making when using mobile devices. To elaborate, this essay differs from the previous studies by focusing on the removal – rather than the addition – of CTR. In this observational study, participants played 20 rounds of a card-draw game where they drew two cards (i.e., attention-demanding cues that increase CTR) each round and were instructed to touch and slide an icon to the "lose" choice when a joker card appeared and to the "win" choice when a joker did not appear. The research design allowed users to voluntarily decide to cheat (i.e., to maximize reward), which also makes them put less emphasis on the evaluation of the cards once a decision to cheat repeatedly has been made.

These essays, combined, provide a better understanding of how CTR and task-specific expertise influence HCI device usage speed. In the remainder of this dissertation, the literature review and theory section will detail the DECODE conceptual framework and foundational schema. This is followed by describing how DECODE applies to each study, context and background, research design, and results.

CHAPTER II: LITERATURE REVIEW & THEORY

Introduction

A literature review of various fields, including Information Systems, Psychology, and Neuropsychology, on trace data, Dual-Process Theory (DPT), and Cognitive Control (CC) was conducted. The insight gained from this review aided in creating a conceptual framework, Decision Evaluation with Cognitive Observations using Device Engagement (DECODE), to guide this research. The review combines the existing theories identified in the literature and explains how they provide a holistic view of how online decisions are made. Specifically, DECODE presents an in-depth, theory-based framework that elucidates the handling of Cognitive Task Requirement (CTR) during decision-making, as well as the moderation of the relationship by task-specific expertise. The users' device usage patterns change depending on task difficulty and users' familiarity with the task.

Overview of the Trace Data

Digital transformation has arrived, and the adoption of digital innovations is growing exponentially in virtually every facet of society. This transformation is streamlining countless online activities including retail (Alzoubi et al. 2022), remote work (Felstead 2022), education (Granić 2022), entertainment (Mulla 2022), social engagements (Tsao et al. 2021), healthcare (Crameri et al. 2022; Wiederhold 2022), and financial services (Khan et al. 2021) to name just a few. Further, just as the number and variety of online services expand, so do the volume and sophistication of the user data collected when interacting with online providers. In terms of total data volume within society, for example, Statista estimated 64.2 zettabytes[1] were "created,

---

[1] A zettabyte is a unit of digital information storage used to denote the size of data. It is equivalent to 1,024 exabytes or 1,000,000,000,000,000,000,000 bytes.

captured, copied, and consumed" in 2020, forecasting a 300% increase to 181 zettabytes by 2024 (Statista 2022). So, while we are literally drowning in data, not all data is equally valuable or actionable (Clark 2021).

As an example, organizations are increasingly capturing and analyzing online consumer data in unprecedented ways, and not all consumer data is the same. For instance, consumer data can be generally categorized as first-party (e.g., collected directly by an organization from users interacting with their website), second-party (e.g., the sharing of one organizations' first-party data with another organization), and third-party (e.g., aggregated data about customer segments obtained from a data service provider – e.g., TransUnion) (Mayer and Mitchell 2012; Schneider et al. 2017). There are also different methods for collecting this tsunami of data, and therefore different types of data, including social media behavior (e.g., posts, likes, and shares), transaction tracking (e.g., purchase behavior), online reviews, website session recording (e.g., assessing usability), cross-site tracking (e.g., third-party cookies), within-site tracking (e.g., trace data), and so on (Cahn et al. 2016; Fernandez et al. 2011; Freelon 2014; Heinonen 2011; Peña-García et al. 2020; Picazo-Vela et al. 2010). While there are different uses for and values of these different types of data, first-party data – that is, data an organization collects about their own customers on their own website – is arguably the most valuable in many use contexts. Further, with highly scalable and low-cost computing, storage, and transit, fine-grained within-site interaction data – i.e., trace data – captured at millisecond granularity can be acted upon in near real-time (Weinmann et al. in press).

A rapidly emerging and novel source of trace data is fine-grained consumer first-party data while interacting with computer applications and online websites. While trace data provides new and novel insights into online consumer behavior, understanding the cognitive mechanisms that drive certain behaviors in an application or online space is a rapidly emerging research area, with many challenges to accurately infer and act upon user behavior outside highly artificial environments (Nunamaker et al. 2011; Vance et al. 2018; Zwicker et al. 2011).

Assessing Trace Data Using Human-Computer Interaction Devices: An In-depth Review

Rooted on various cognitive and social science theories, prior Information Systems (IS) literature in the HCI domain aims to provide objective metrics that can be used to assess multiple online users' behaviors, including habituation (Vance et al. 2018), concealing information (Jenkins. et al. 2019), and fraud (Hibbeln et al. 2014; Weinmann et al. in press). However, previous literature mainly focused on specific situations where humans would experience heightened cognitive load. Thus, the leveraged theories provide the support that applies to a confined set of situations. For instance, Hibbeln and colleagues (2017) drew from Attention Control Theory (Eysenck et al. 2007) to explain how negative emotions heighten cognitive load and thus influence mouse cursor movements (e.g., feeling negative emotions result in slower mouse cursor speeds). Weinmann and colleagues (forthcoming) leveraged Cognitive Load Theory (CLT) and the Response Activation Model (RAM) to demonstrate that more extensive fraud (i.e., more cognitively demanding) increased mouse movement deviations and decreased movement speed (Sweller 2011; Welsh and Elliott 2004). Table 2.1 summarizes the prior literature that assessed trace data using HCI devices. This table provides a comprehensive overview of the relevant studies and outlines the research context, device used, theoretical foundation, and manipulation conditions.

Table 2.1. List of Prior literature and its Theoretical foundation				
Reference	**Research Context**	**Device Used**	**Theoretical Foundation**	**Manipulation Condition**
Hibbeln et al. 2017	Negative emotion	Mouse	Attentional Control Theory	Inducing negative emotions and thereby limiting available cognitive resources
Vance et al. 2018	Habituation	fMRI, Eye tracking	Habituation (Dual-Process) Theory	Presenting polymorphic (i.e., constantly changing) cybersecurity warning
Valacich et al. 2020	Credibility	Mobile Device	Orienting Reflex	Asking risk-relevant questions about the deviant behavior (i.e., cheating) that the user performed
Jenkins. et al. 2019	Insider Threat	Mouse	Response Activation Model (RAM)	Asking risk-relevant questions about the deviant behavior (i.e., cheating) that the user performed
Jenkins et al. 2021	Noncompliance	Mouse	Cognitive Dissonance Theory, Response Activation Model	Providing an option to cheat on an aptitude test and asking follow-up questions about cheating behaviors
Kumar et al. 2021	Response Bias and Data Quality	Mouse	Rational Decision-Making model	Asking risk-relevant questions about the deceptive behavior (i.e., lying about their proficiency in fake software) that the user performed
(Weinmann et al. in press)	Fraud	Mouse	Cognitive Dissonance, Response Activation Model	Incentivizing participants to commit fraud (i.e., higher cognitive load than acting truthfully)

Although these prior studies suggest that device usage efficiency decreases as the increase in cognitive load acts to degrade fine motor control (i.e., controlling your hand movements), there are several limitations. First, heightened cognitive load is a temporal state that does not persist for a prolonged period. Prior literature provides little explanation of the user behaviors under a low cognitive load state. Second, the foundational theories used for these prior studies do not apply to various instances of system usage as they are confined within specific contexts (e.g., feeling negative emotions). For example, users may experience high cognitive load when paying particular attention to online advertisements (Bang and Wojdynski 2016), exploring an unfamiliar website

(Sénécal et al. 2015), and engaging in online learning (Mayer 2019) – which prior theories have little applicability. In summary, the theories primarily focusing on cognitive loads provide little explanation of the underlying mechanisms (i.e., the interplay of low and high cognitive load states) of how humans intake and process information during live system usage. Thus, there is a need for a more encompassing conceptual framework that can be applied to a broader range of situations (i.e., under low or high cognitive load).

In summary, prior literature provides a limited understanding of user behaviors when the users interact with a system in natural settings where they would likely not have a continuous cognitive burden. Further, because users are more likely to habituate to a limited number of websites and applications for daily use, the cognitive demand for daily system usage will also decrease over time. Thus, studying the differences in how individuals behave under familiar and unfamiliar situations provides a more holistic view of ongoing user behaviors in an online environment.

Frameworks / Theories

This section will review theories and frameworks that explain human behaviors and decision-making processes from various disciplines. First, DPT-based frameworks, widely accepted in Psychology, are reviewed. Then, frameworks that focus on CC from the Neuropsychology field are presented. In sum, this section provides an in-depth overview of an individual's decision-making processes by examining DPT-based frameworks from Psychology and CC-focused frameworks from Neuropsychology to integrate insights from multiple disciplines.

Cognitive Miser and Dual-Process Theory Based Framework

"The social world is intrinsically complicated, and our minds are limited, so we take shortcuts," explains Fiske, the psychologist who coined the term Cognitive Miser (BBVA 2020). As Fiske said, human minds have limited cognitive resources (Buschman et al. 2011). To conserve and utilize limited resources, humans rely on intuitive thinking processes rather than engaging in more cognitively demanding, deliberate thinking (Pennington 2000; Stanovich 2018, 2021). Such mental shortcuts are well explained in Dual Process Theory (DPT), which is a widely accepted model in various fields, including psychology, philosophy, economics, and decision-making to explain human reasoning at a primitive level (De Neys and Pennycook 2019; Evans and Stanovich 2013; Kahneman 2011; Neys 2006). It posits that human thinking consists of two distinct types: System 1 (also referred to as intuitive and heuristic) and System 2 (deliberative and analytic). Though there are many different extensions of DPT, the main idea that two distinct thinking processes can characterize human thought processes is a central concept consistent across different DPT variations (De Neys 2017a, 2017b; Evans 2018; Evans and Stanovich 2013; Greene 2023).

System 1 comprises a set of sub-systems that operate with autonomy, and the behavioral output from System 1 is instinctive, immediate, effortless, and fast. When given a reasoning task, the user may use System 1 by default to produce an output, depending on their relative familiarity with it (Miles et al. 2019). In an online environment, for instance, when presented with a routine task (e.g., answering innocuous demographic questions in a survey), a respondent would decide on a response without requiring extensive cognitive resources (i.e., exerting low cognitive load). As System 1 processes are based on heuristics, the decision outcome that relies more on System 1 would be fast, effortless, and autonomous.

However, when given a reasoning task requiring sustained attention and concentration, users may use System 2 thinking to complete the associated reasoning and evaluative activities that entail high cognitive load (Bago and De Neys 2017, 2019). Some examples may include filling out a complex online form (e.g., annual tax form) that requires careful attention to detail, reading and analyzing a dense or highly technical article, or engaging in a strategic game. When an individual is engaged in System 2, the decision outcomes would be slower, more effortful, and deliberate.

Cognitive Control Based Framework

Cognitive Control (CC) is a process that supports flexible, adaptive responses and coordinates complex goal-directed thoughts and behaviors (Nature 2023). Though there is no single, definitive theory that provides a full explanation of CC, it has been conceptualized as a construct that encompasses, individual's ability to regulate various cognitive processes including attention and working memory (Braver 2012; Miller 2000; Ochsner and Gross 2005; Posner et al. 2004).

Given a task, CC allows individuals to inhibit automatic responses and engage in careful decision-making processes to produce an adequate response (Boag et al. 2021; Botvinick and Braver 2015; Braver et al. 2007; Courtney 2004; Engle and Kane 2004; Kool et al. 2017; Straub et al. 2020). As individuals exhibit a high degree of CC when inhibiting automatic responses, a high degree of CC is associated with more reliance on System 2. Thus, CC provides a complementary explanation to DPT by providing an explanation of how individuals are more likely to utilize System 1 or System 2 when completing a task.

Posner and Synder's Dual Process View of Cognitive Control

Given a strong contrast between autonomous and effortful thought processes, the Dual Process View of Cognitive Control suggests that CC plays a crucial role in facilitating a transition from System 2 types of thinking to System 1 types of thinking (Christensen et al. 2016; Posner et al. 2004). The notion begins from the widely accepted view that CC is intrinsically effortful (Eisenberg 2017; Kool and Botvinick 2013, 2013; Kuipers et al. 2017; Nigg 2017). When individuals learn a new task, they initially exert more cognitive effort along with a relatively higher degree of CC. As the learning occurs, individuals become more competent at executing the task, gradually becoming less effortful and more automatic as well as requiring less CC as competency increases. A visual representation of Posner and Synder's view is depicted in Figure 2.1.

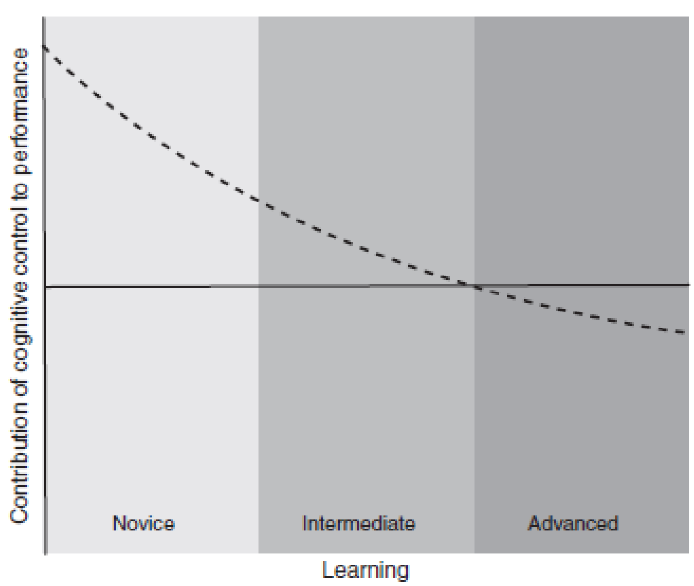


Figure 2.1 Cognitive Control's contribution to performance (Dual Process View – from Christensen et al. 2016)

Hybrid: Autonomous and Mesh View of Cognitive Control

An alternative view to the Dual Process View of Cognitive Control is the Hybrid view, which suggests that CC continues to contribute to performing a task even when an individual has attained a high level of competence (Christensen et al. 2016). The Hybrid view is further broken down into (1) Autonomous, which extends the Dual Process view, and (2) Mesh. A visual representation of the Hybrid view is depicted in Figure 2.2.

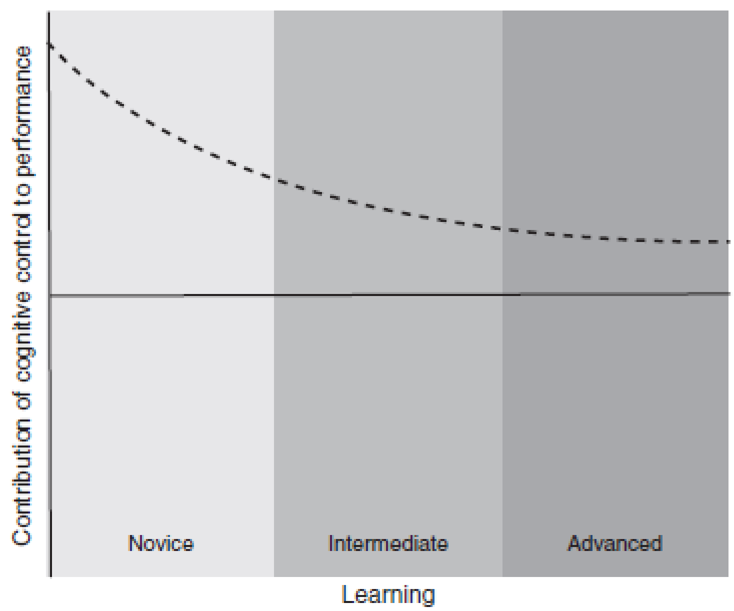


Figure 2.2 Cognitive Control's contribution to performance (Hybrid View - from Christensen et al. 2016)

The Autonomous view suggests that CC plays a vital role in the initial strategic choice of task execution, while the automatic processes guide the action. For instance, when reading a multiple-choice question, the CC would be responsible for reading and comprehending the question. The automatic process would handle moving the mouse cursor toward a chosen answer.

On the other hand, the Mesh view suggests that CC participates in action execution with more influence on task control's strategic aspects. Like the Autonomous view, Mesh would interpret CC as responsible for reading and comprehending the question. However, the Mesh view suggests that CC still contributes while the user is moving the mouse cursor to choose an answer, rather than the process being fully automated.

Table 2.2 compares the role of CC in three different models discussed in Chapter 2: DPT-based, Autonomous, and Mesh models. Overall, DPT and Autonomous views posit that CC contributes little to the execution of motor tasks, which are automated as individuals develop task-specific expertise. On the other hand, Mesh's view asserts that CC makes a direct contribution, even after an individual develops task-specific expertise. Lastly, the discussion of how external cues, such as situational context, distraction, or difficulty, influence individuals' behavior is limited in DPT and Autonomous view. On the other hand, Mesh considers external cues' influence on CC and task execution.

DPT offers a broad, high-level view of thought processes and decision-making. However, DPT is limited to providing a detailed explanation of actions involved in a decision process. First, it is difficult to find the factors that induce individuals to rely more on either System 1 or System 2. Thus, it is not easy to establish clear boundaries for how different activities involve more of System 1 or System 2. Further, no existing literature has examined how the use of System 1 or System 2 manifests as differences in trace data and device usage behaviors.

Table 2.2 Comparing the Role of Cognitive Control in DPT, Autonomous Hybrid, and Mesh Hybrid Models

Phenomenon of Interest		DPT	Hybrid: Autonomous	Hybrid: Mesh
CC's influence on developing a task-specific expertise		CC initially contributes to learning. However, CC's contribution diminishes as motor skill is practiced and refined.	CC's contribution is similar to DPT based model.	CC continues to engage throughout skill development and execution allows for ongoing improvement and adaptation, leading to increased task-specific expertise.
CC's contribution to motor execution related to the task		Once the task is automated, CC makes no contribution to task performance and motor execution.	Indirect contribution to task performance and motor execution.	Direct contribution to Task performance and motor execution.
External Cue's Influence on CC	**Situational Context**	Does not account for the role of situational context.	Considers, but does not emphasize the role of situational context.	Situation awareness directly influences task performance and motor execution.
	Distraction	No influence on the task performance and motor execution.	No influence on the task performance and motor execution	Negatively influence task performance and motor execution.
	Difficulty	Once actions are automated, complexity or difficulty of the tasks have no influence.	No discussion on how complexity and difficulty of the task influences performance and motor execution.	Negatively influence task performance and motor execution.

Limitations of DPT and Hybrid View of Cognitive Control

Likewise, a Hybrid (Autonomous) view of CC partially addresses the limitations of DPT by offering complementary explanations of how one's behaviors may become intuitive and automatic as task-specific competency increases over time. For instance, Posner and Synder's Autonomous view adequately addresses how individuals may attain automaticity over a particular task and become more efficient in executing relevant motor movements. However, the Autonomous view assumes that CC makes no positive contribution to task execution once automaticity is attained. While useful, the aspect that CC plays no contribution when executing highly skilled tasks has been heavily criticized, as there are cases when an individual needs to make strategic decisions (i.e., not automatic) even when executing a highly practiced task (Christensen et al. 2016).

The Hybrid (Mesh) view addresses limitations that Posner and Synder's framework proposes by offering explanations of how the motor skills involved in task executions are acquired, developed, and executed. However, the explanations provided by Mesh are challenging to measure and empirically validate. Specifically, empirical validation of Mesh requires consideration of the types of motor movement involved, individual differences of the collected sample (e.g., expertise), and social and cultural factors. Thus, formulating testable hypotheses for the Mesh framework remains a challenge.

Proposing a Framework and Propositions for Online Decision-Making

Although both DPT and Hybrid views offer slightly different aspects of CC's role in task execution, the concept of developing expertise and automated behaviors remains consistent across all prior conceptual frameworks. Thus, based on the notion that individuals' behaviors become more automatic as they develop task-specific expertise, a framework is proposed that addresses the limitations that were identified in prior DPT and CC-based theoretical frameworks.

First, the proposed framework builds upon the foundational principles of DPT, despite its limitations in fully explaining certain phenomena. Over time, the concerns raised against Posner and Synder's view have increased, with an abundance of empirical research demonstrating that task execution and decision-making involve an interplay between both System 1 and System 2 processes (De Neys 2017a; De Neys and Pennycook 2019; Evans 2008; Kahneman 2011; Miles et al. 2019). This perspective is consistent with Mesh's assertion that CC influences behavioral outcomes even as executions become automated.

Second, the proposed framework leverages CC as a construct that facilitates an interplay between System 1 and System 2 at decision-making is performed. Numerous studies have demonstrated that CC is intrinsically effortful and is necessary when engaged in goal-directed behavior (Chevalier 2018; Chevalier et al. 2018; Westbrook et al. 2019). In online decision-making, incorporating CC into a DPT-based framework provides a thorough explanation of how the interplay of System 1 and System 2 may manifest as differences in device usage data (i.e., trace data generated by commonly used HCI devices).

Lastly, the proposed approach integrates Hybrid frameworks (both Mesh and DPT based) to present a novel concept called Cognitive Task Requirements (CTR), signifying the cognitive demands of a task. The proposed framework then explains how the presence of CTR and the task-specific expertise of individual users influence task performance. To be precise, in online decision-making tasks, numerous external cues, such as visual elements and task difficulties, can increase or decrease CTR and thus influence individuals' cognitive control. This directly influences task performance. The individuals' task-specific expertise on the given task also influences the degree of cognitive effort exerted by CC.

The proposed framework, the Decision Evaluation with Cognitive Observations using Device Engagement (DECODE), is divided into three distinct sections (see Figure 2.3):

(1) Observing: perception of external cues with associated CTR levels.

(2) Deciding: facilitation of System 1 and System 2 thinking through CC.

(3) Behaving: execution of behaviors that can be captured using various HCI devices.

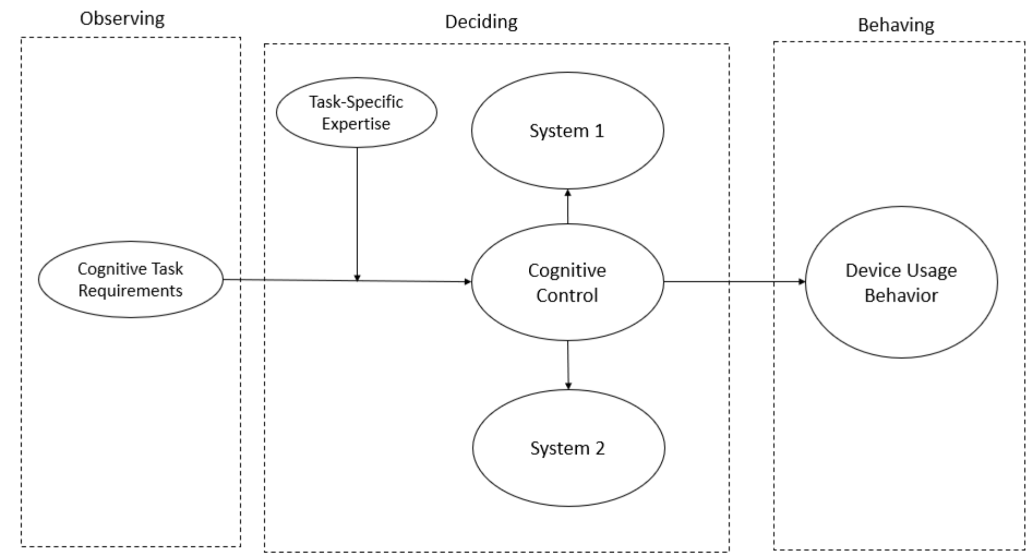


Figure 2.3 Decision Evaluation with Cognitive Observations using Device Engagement (DECODE)

Further, given this backdrop, two propositions that align with the proposed framework are presented:

Proposition 1: Attention-demanding external cues will increase the CTR, resulting in high cognitive control (i.e., System-2 type of thinking) as reflected by slower device usage speed. As CTR increases, individuals need to allocate more attentional resources to process complex information, leading to increased cognitive control and slower responses as they engage in more deliberate thinking.

Proposition 2: Individuals with high task-specific expertise will exhibit lower cognitive control (i.e., System-1 type of thinking) as reflected by faster device usage speed. As experienced individuals have developed more automatic responses and efficient mental shortcuts that allow them to process task-specific information quickly, the need for CC is reduced and thus results in faster device usage.

Research Questions

To empirically validate the DECODE framework, this dissertation explores the influence of CTR and task-specific expertise on decision-making processes during online activities. The investigation is centered around DPT as the underlying theoretical backbone, which posits that human decision-making operates through two distinct yet interrelated systems – the intuitive (System 1) and the analytical (System 2). With CC as a higher-order construct encompassing an individual's ability to regulate cognitive processes such as attention, working memory, and problem-solving, I pose the following as research questions:

> RQ1: How does CTR impact online decision-making within the context of Cognitive Control and Dual Process Theory?

> RQ2: How does an individual's task-specific expertise influence online decision-making within the context of Cognitive Control and Dual Process Theory?

Each research essay proposes a series of hypotheses to test the research questions in a particular device context. As different kinds of HCI devices are used as measurement tools, a consistent measure, such as task completion speed, is needed to indicate the degree of CC exhibited during task execution. Speed, which entails response times and motor action components, has served as an important indicator of cognitive ability during task executions (Ebaid et al. 2017, 2017; Fitts 1966; Hsieh et al. 2008; Kyllonen and Zu 2016; Specka et al. 2000). Further, because of the inherent differences when in using different HCI devices for completing online tasks, the device usage behavior construct is measured using devices specific metrics that align with trace data that can be easily captured while tasks are completed.

CHAPTER 3: ESSAY 1 – AN EMPIRICAL EXAMINATION OF THE DECODE FRAMEWORK USING COMPUTER MICE

Introduction and Related Work

Rooted on various cognitive and social science theories, IS literature in the Human-Computer Interaction (HCI) domain aims to provide objective metrics that can be used to assess multiple online users' behaviors, including habituation to security warnings (Vance et al. 2018), concealing information (Jenkins. et al. 2019), and fraud (Hibbeln et al. 2014; Weinmann et al. in press). Though prior literature explains how users' cognitive states influence device usage behaviors, the underlying theories behind what drives specific user behaviors vary significantly. For instance, Hibbeln and colleagues (2017) drew from Attention Control Theory (Eysenck et al. 2007) to explain how negative emotions influence mouse cursor movements (e.g., feeling negative emotions resulted in slower mouse cursor speeds). On the other hand, Weinmann and colleagues leveraged Cognitive Load Theory (CLT) and Response Activation Model (RAM) to demonstrate that more extensive fraud increased mouse movement deviations and decreased movement speed (Sweller 2011; Welsh and Elliott 2004). Although both studies suggested that mouse movement speed slows down because of increased cognitive load, the instances where online users experience increased cognitive load are not confined to fraud and negative emotional contexts. For example, users' cognitive load also increases when they are focusing on an advertisement (Bang and Wojdynski 2016), exploring an unfamiliar website (Sénécal et al. 2015), or engaging in online learning (Mayer 2019). These interactions cannot be well explained using existing methods and need to be studied under a more encompassing theoretical framework.

This essay specifically aims to theoretically explain how the underlying cognitive processes manifest as the difference in mouse movement speeds using the Decision Evaluation with Cognitive Observations using Device Engagement (DECODE) framework. While prior literature had examined this relationship, the guiding theories across these studies have varied significantly. For instance, Jenkins et al. (2019) paired Response Activation Model with mouse-cursor movements to identify movement characteristics of individuals that are likely to be concealing information. Further, Hibbeln et al. (2017) drew from Attention Control Theory to explain how negative emotions influence mouse cursor movements. Although the theoretical foundations across the studies varied, the results, suggesting that mouse cursor speed and other metrics have been essential indicators of users' cognitive states, remained consistent.

Table 3.1 presents studies that investigated the relationship between mousing speed and various cognitive states, with their respective findings. The result from every study suggests that various cognitive states and factors, such as cognitive load, work stress, negative emotions, and fraud, influence mouse cursor speed. Although confined to computer mice, these findings suggest that analyzing device usage speed can potentially serve as a valuable indicator of users' underlying cognitive states and behaviors.

Table 3.1 Studies Investigating the Relationship between Mouse Speed and Various Cognitive States		
Reference	**Research Context**	**Findings**
Grimes and Valacich 2015	Cognitive Load	Higher cognitive load reduces mouse cursor speed
Banholzer et al. 2021	Work Stress	Work stress exhibited a speed-accuracy trade-off, wherein individuals either increased speed with reduced accuracy or decreased speed while maintaining high accuracy.
Hibbeln et al. 2017	Negative emotion	Negative emotions decreased mouse cursor speed.
Kowatsch et al. 2017	Work Stress	Mouse speed was correlated to valence and arousal.
Weinmann et al. in press	Fraud	Thus, fraudulent responses will exhibit slower mouse cursor movement than honest responses.
Jenkins. et al. 2019	Concealing Information	Mouse cursor speed is slower when individuals try to conceal information.
Macaulay 2004	User Engagement (attention, usefulness, perceived task duration)	Mouse cursor speed was a significant predictor of attention and usefulness

Hypotheses

Concerning RQ1, the presence of attention-demanding cues will increase the Cognitive Task Requirement (CTR) and exhibit a high degree of Cognitive Control (CC); in the case of computer mice, an exhibition of a high degree of CC will slow down the mouse usage speed. To empirically validate the proposed relationship, I impose a manipulation condition that encourages participants to provide socially desirable responses to survey questions. Providing socially desirable responses involves a high degree of CC, as users must inhibit the automatic response (Paulhus 1991; Tourangeau and Yan 2007). Figure 3.1 illustrates the predicted effect of inducing social desirability (i.e., increase in CTR), on mouse cursor speed. The increase in CTR is expected to increase cognitive control and result in slower mouse usage speed.

Given these considerations, Hypothesis 1 can be formulated as follows:

H1: Increase in CTR (e.g., inducing social desirability) will slow down mouse cursor speed.

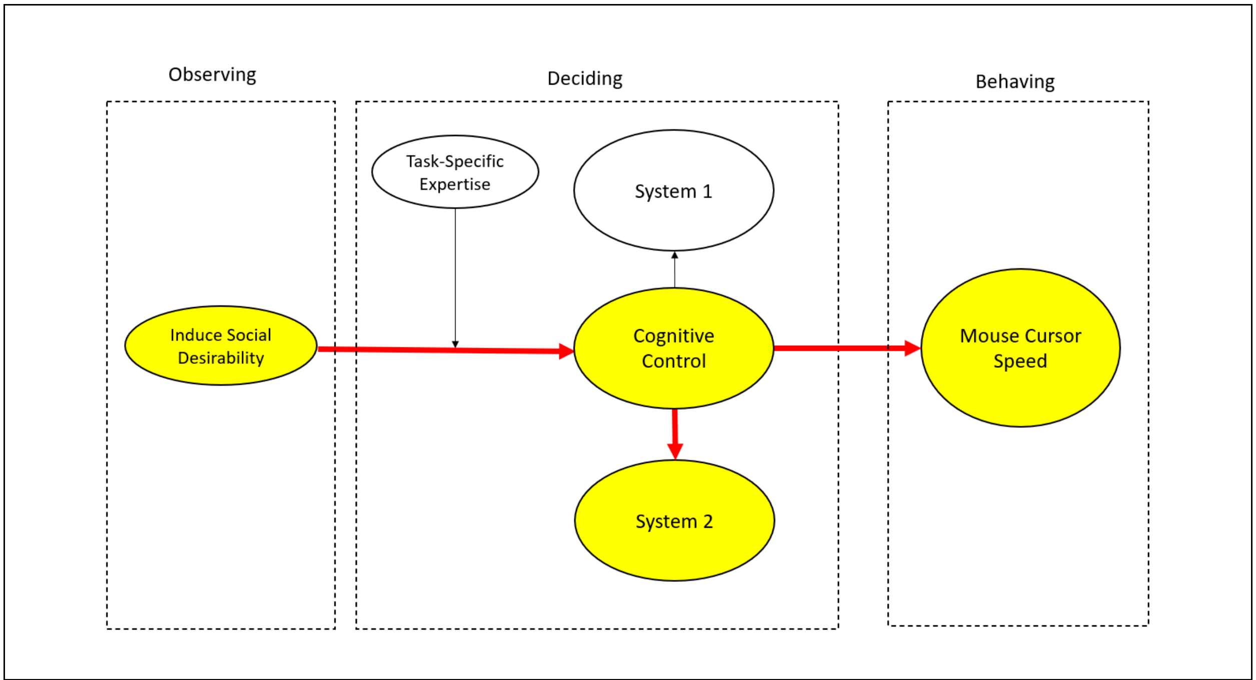


Figure 3.1 Application of DECODE framework to examine how inducing social desirability influences Cognitive Control and mouse cursor speed

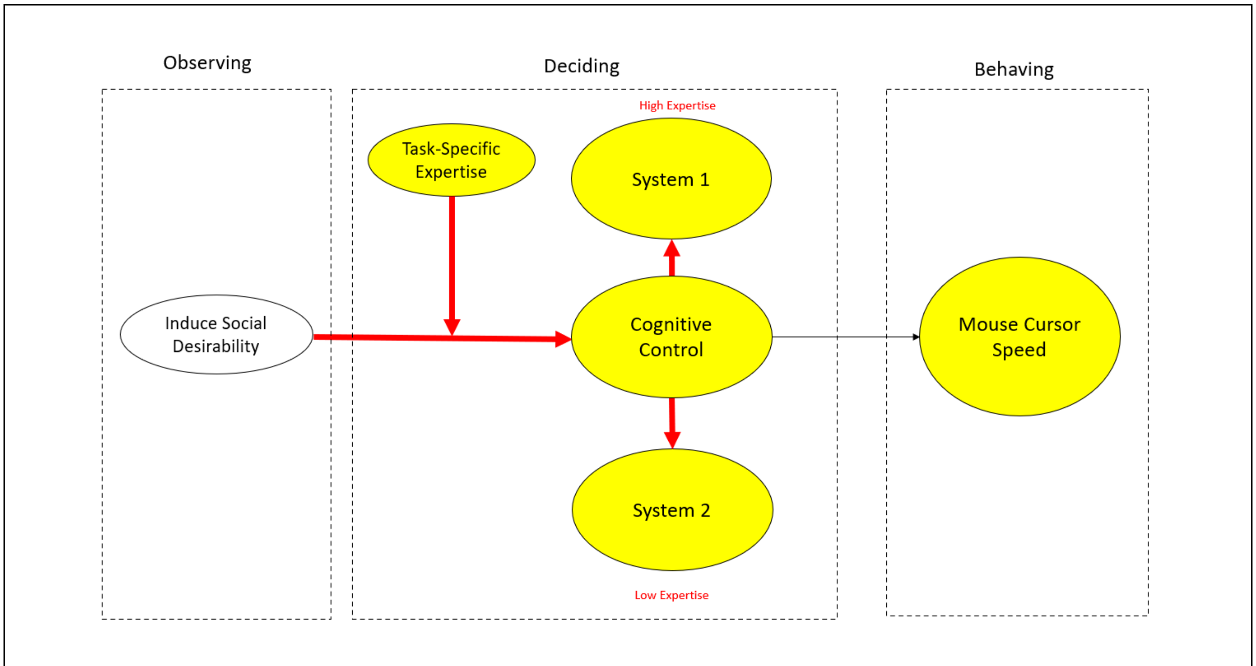


Figure 3.2 Application of DECODE framework to examine how task-specific expertise influences Cognitive Control and mouse cursor speed

Given a task, individuals will have varying degrees of context familiarity (i.e., the extent to which an individual is familiar with the task context). Concerning RQ 2, how context familiarity influences CC and the mouse usage speed can be empirically validated by providing tasks for individuals with high and low context familiarity, as illustrated in Figure 3.2. This figure illustrates the predicted relationship between context familiarity and mouse cursor speed. Low context familiarity, as experienced when answering less familiar questions, is expected to result in slower mouse usage speed due to increased cognitive control. In the case of survey questions, questions with higher context familiarity are also more likely to entail higher task-specific expertise than when answering unfamiliar questions.

In summary, individuals may have high context familiarity and task-specific expertise with answering demographic questions as such questions are more commonly asked in various circumstances (e.g., what is your age or educational level). On the other hand, questions specific to a study (e.g., Big-5 personality inventory questions) will entail relatively lower context familiarity and task-specific expertise when compared to the demographic questions. Given these considerations, Hypothesis 2 can be formulated as follows:

H2: Low context familiarity will slow down mouse cursor speed.

Methodology

Measure: Mouse Movement Speed. The JavaScript code embedded in the study captures the x-coordinates and y-coordinates of the mouse movement at millisecond precision. The captured mouse coordinates are then utilized to map both actual trajectories to calculate the cursor distance per each target object that users interacted with.

The distance between two points $P_t\ (x_t, y_t)$ and $P_{t-1}(x_{t-1}, y_{t-1})$can be calculated using the following formula:

$$d(P_t, P_{t-1}) = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}$$

Here, t represents the timestamp in milliseconds. Similarly, the average speed between data points is derived using:

$$\bar{s} = \frac{\Delta d}{\Delta t}$$

In this equation, Δd denotes the change in distance and Δt denotes the change in time. To ensure that we only capture the mouse movements, additional JavaScript code was implemented to exclude mobile devices from participating in the study. Mouse speed variables were segmented by target and subsequently averaged for analysis. To account for individual differences in mouse usage patterns, all speed measures for each participant were normalized using min-max normalization, which scales the data within a specific range, typically between 0 and 1, while preserving the original distribution of the data. Precisely, the average speed for each question was first calculated. Then min-max normalization was performed at the user level to derive relative speeds per question, thus accounting for individual differences in device usage patterns.

Experimental Design. To empirically validate the hypotheses, a web-based survey was conducted. The survey is a popular and convenient method to gather self-reported information from participants at scale. In this study, the survey was designed to assess participants' task-specific expertise and manipulate the CTR to alter the degree of CC involved in their decision-making process. The survey was divided into three parts:

1. *High Task-Specific Expertise Questions (Part 1).* All participants were asked to complete a demographic questionnaire in the first part of the survey. Demographic questions are commonly used in research to gather information about participants' backgrounds. Not only does such information helps researchers better understand the characteristics of the sample, but it consists of questions that individuals must have seen and answered multiple times in their lifetime (e.g., what is your gender?). Thus, individuals will have High Task-Specific Expertise when answering demographic questions.
2. *Low Task-Specific Expertise Questions (Part 2).* The Big Five Inventory questionnaires consisting of 39 questions that measure the Big Five personality traits (e.g., conscientiousness, extraversion, neuroticism, agreeableness, openness to experience) were subsequently presented. The Big Five personality traits questions were chosen due to their widespread validation in numerous studies, ensuring the reliability and validity of the measures used (McCrae and Costa Jr. 2008). An attention check question was also included to confirm that participants were paying attention throughout the survey. Compared to demographic questions, individuals are expected to have lower task-specific expertise when answering Big Five personality traits questions, as these items require introspection and self-assessment rather than recalling familiar factual information.
3. *Manipulation: Promoting Socially Desirable Responses (Part 3 – Denoted as SDR questions).* The manipulation condition is designed to promote socially desirable responses and thus increase the degree of CC involvement in decision-making. First, participants were primed to think that their answers to the third part of the study were important. Participants were allowed to voluntarily participate in the second part of the survey for an additional payout (e.g., a total of $1.00). The instruction for the third part stated, "The third part of this survey

is an application for a follow-on study that will pay $10 for 10 minutes of work. We will only select a few people to do this follow-on study. If we select you, you can refuse to participate; there is no obligation to do it. We will pay you an extra $0.50 to complete the application, regardless of whether we select you or whether you agree to do the follow-on study if selected." The instruction further stated that "the follow-on study will have you use Excel to do some data analysis, so we are looking for people who have experience with Excel. Experience with Excel's math and statistical functions is a plus but is not required." The follow-up survey comprised multiple subparts related to the user's background and experience. The participants were asked to answer questions about their work hours and occupations (e.g., Are you currently employed in a full-time job?).

Second, when participants answered computer skills questions, they were asked to rate their skills on a non-existent Excel Plugin (i.e., StatView). This question, by design, requires further involvement of CC as users need to either lie about their expertise or respond honestly. All the experience-related questions ranged from 0 (e.g., beginner) to 10 (e.g., expert).

Participants

534 subjects were recruited from Amazon Mechanical Turk. Participants' ages were diverse. About 56% of the participants belonged to a younger crowd between 18-34 years old, with 42% reported being 35 years or older. Of the recruited participants, 41% were female. Of those participants, three hundred and sixty-two subjects completed the third part of the survey. An attention check question was given to all participants before proceeding to the third part.

As this research is primarily interested in how people answer cognitively demanding questions, all samples were retained regardless of whether a person completed part 3 of the survey. This resulted in the final sample size of five hundred and thirty participants with raw mouse movement data of 2.7 million observations.

Results and Discussion

Hypotheses H1 and H2 are evaluated using a one-way Analysis of Variance (ANOVA), followed by Tukey's HSD (honestly significant difference) test across different questions. Specifically, H1 is examined by comparing the average speed across questions in Part 1 and Part 2 with questions in Part 3. The comparison of average speed between Part 1 and Part 2 questions was conducted to test H2. All analyses were conducted using the R statistical software package (R Core Team 2022).

Table 3.2 provides the ANOVA results examining the statistical differences in mouse cursor speeds across the question types. The significance of ANOVA results suggests significant differences in the means of mouse cursor speed across different question types.

Table 3.2 Analysis of Variance Results

Variables	DF	Sum Sq	Mean Sq	F Value	P Value	Significance
Question Types	2	6.996	3.498	150.1	0.0000	****
Residuals	1027	23.934	0.023			
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '						

There was a significant difference in average mouse cursor speed across the question types. Thus, Tukey's HSD was used to compare the average mouse cursor speed. Table 3.3 provides the result of Tukey's HSD, where all evaluated pairs exhibited significantly different. Figure 3.1 and Figure 3.2 present the speed distribution across question types, allowing for visual inspection of speed differences between the question types. In particular, Figure 3.1 displays the distribution of mouse cursor speed for each question type using boxplots with median, quartile, and potential outliers. Figure 3.2 shows the distribution of mouse cursor speed using density plots and visualizes the overall spread and distribution of the data.

Table 3.3. Tukey's HSD Test Results

Pairs Being Tested	Diff	Lower	Upper	P-Value adjusted	Significance
Low – High	-0.0998	-0.1256	-0.0741	0.0000	****
Low – SDR	0.1122	0.0833	0.1412	0.0000	****
High – SDR	0.2121	0.1833	0.2410	0.0000	****
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '					

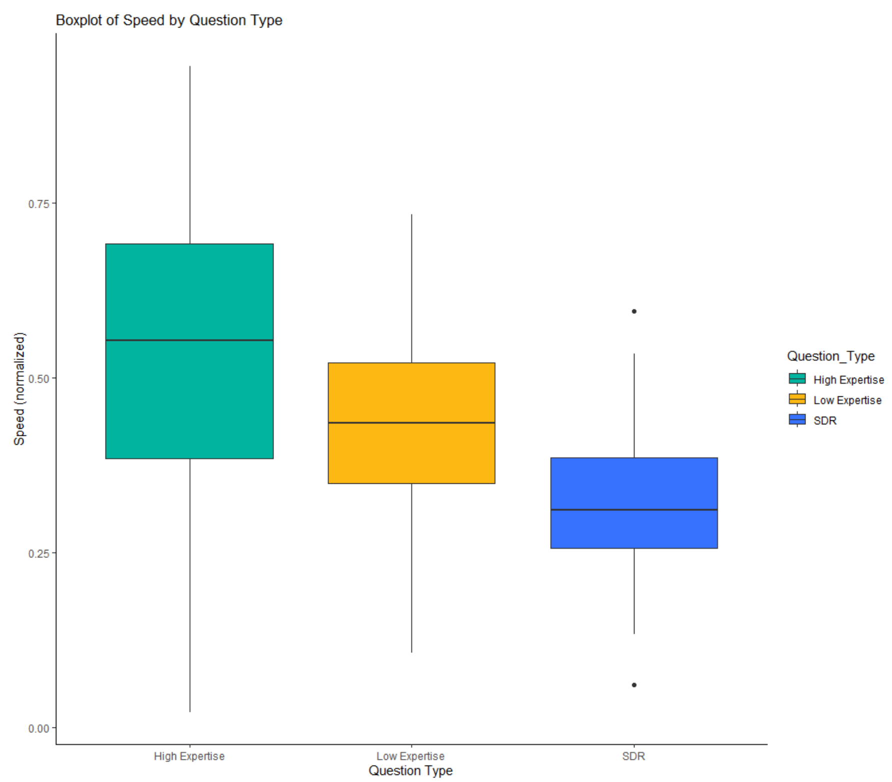


Figure 3.3 Boxplot of normalized speed by question type

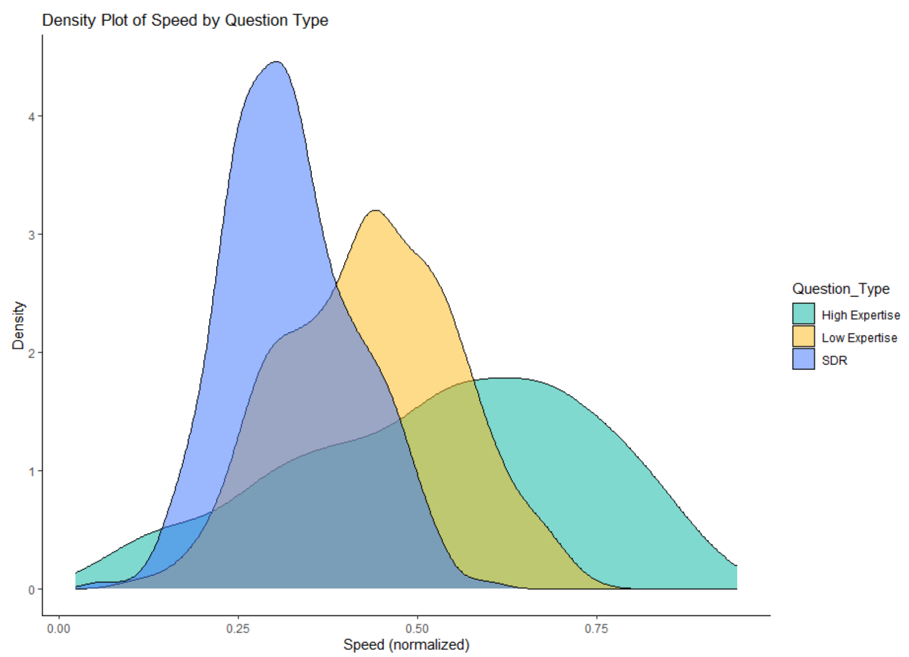


Figure 3.4 Density plot of normalized speed by question type

The ANOVA and Tukey's HSD test suggested that task-specific expertise and CTR influence the mouse cursor speeds. In particular, demographics questions had the highest mouse cursor speeds, followed by Big Five personality Inventory questions and then the questions displayed after the manipulation condition to encourage socially desirable responses. As a result, it can be concluded that H1 and H2 are supported.

Essay One Summary

In this Essay, the DECODE framework was employed to examine how the inducement of social desirability and task expertise influences the mouse cursor speed. The Essay hypothesized that the inducement of social desirability would increase CTR and CC and thus decrease mouse cursor speed, while having low task-specific expertise will also result in slower mouse cursor speed.

A web-based survey, divided into three parts, was designed to test the hypotheses. The survey presented high task-specific expertise questions (demographics), low task-specific expertise questions (Big Five personality traits), and experience-related questions after the manipulation to promote socially desirable responses.

The results, analyzed through ANOVA and Tukey's HSD test, revealed significant differences in mouse cursor speed across different question types. Thus, it was concluded that both task-specific expertise and CTR influence mouse cursor speed.

CHAPTER 4: ESSAY 2 – EMPIRICAL EXAMINATION OF EXAMINATION OF DECODE FRAMEWORK USING COMPUTER KEYBOARD

Introduction and Related Work

'Essay Two' applies the Decision Evaluation with Cognitive Observations using Device Engagement (DECODE) framework to study typing behaviors. Although prior literature has extensively explored the relationship between users' mouse movement and their underlying cognitive states, there is a limited number of studies examining such a relationship in typing behaviors. The limited research in this area presents an opportunity to expand the knowledge on how typing behaviors can be used to understand users' cognitive processes and offer new insights in various application domains (See Table 4.1 for notable studies that investigated such a relationship concerning typing speed). A key takeaway from Table 4.1 is that key transition time has been found to be the most effective measure for studying a variety of cognitive contexts.

Table 4.1. Studies Investigating the Relationship Between Cognitive States and Typing Speed

Reference	Research Context	Findings with respect to speed
Vizer et al. 2009	Stress	Participants under stress exhibited distinctly different typing patterns, producing features that can be used for classification models (75% accuracy in classifying cognitively stressed state).
Epp et al. 2011	Emotion	Key-to-key transition time (i.e., an indicator of speed) was an important feature in identifying various emotional states.
Khanna and Mukundan 2010	Emotion	Positive, Negative, and Neutral emotional states had a significant influence on key-to-key transition time.
Monaro et al. 2018	Deception Detection	Writing time (i.e., speed) and response latency contributed significantly to the identification of liars.

To establish a connection amongst Dual Process Theory (DPT), Cognitive Control (CC), and typing behaviors, focus is placed on motor movements and the explanation of how repeated execution of motor movements results in motor learning (Gliner 1985; Pollock and Lee 1992). Motor learning theories explain how humans develop proficiency in motor movements—ranging from skilled compound movements in athletes (Verburgh et al. 2016) to fine motor movements in musicians (Griffith 2013). As movements are repeatedly executed, they gradually become automatic, requiring less cognitive involvement and increasing consistency (Edwards 2010; Gliner 1985).

Motor learning theories have been used extensively in the study of HCI (Cockburn et al. 2014), including the impact of motor learning on movements of pointing devices, such as mouse movements (Lane and Ziviani 2010). Relevant to this dissertation, motor learning has been demonstrated extensively in the context of keyboard typing movements. Touch typing, for example, is a deeply automatized motor behavior (Sperl and Cañal-Bruland 2020a, 2020b). Computer users type certain pieces of information regularly, developing learned motor patterns that serve as the basis for keyboard-based authentication in which users' unique and consistent typing patterns form a "typing signature" (Joyce and Gupta 1990) against which future patterns can be compared to identify unauthorized use (Karnan et al. 2011; Monrose and Rubin 2000). Although this use of keystroke dynamics does not serve the first-time user context well, there is precedent for classifying first-time users based on typing patterns, such as whether a familiar password is being (re)used during account creation (Jenkins et al. 2014). Following this logic, motor learning is incorporated as a kernel theory that supports the use of keystroke dynamics in determining whether users are engaged in typing tasks with high task-specific expertise (e.g., familiar and well-practiced tasks).

In summary, this essay presents four studies that 1) establish a connection between DPT, CC, and typing behaviors, 2) provide insights into how task-specific expertise and Cognitive Task Requirement (CTR) influence typing behaviors, and 3) demonstrate ecological validity, specifically in the identity fraud space. These studies demonstrate DECODE frameworks' effectiveness in assessing real-time user behaviors in a typing context. By applying the DECODE framework to typing this behavior, this essay seeks to contribute to the existing body of knowledge in HCI as well as validate DECODE in another HCI device context.

Hypotheses

Attention-demanding cues are hypothesized to increase CTR, leading to a high degree of CC. In the context of a computer keyboard, a high degree of CC is expected to result in slower typing speed. While the underlying mechanism behind this proposed relationship is similar to those observed with computer mice, the role of CC in typing behaviors (e.g., entering a sequence of keys) is suspected to differ considerably due to the inherent complexity of typing compared to simple mouse movements.

Further, concerning the DECODE framework, attention-demanding cues (e.g., objective task difficulty) and task-specific expertise can be closely intertwined. Thus, the complete isolation of these constructs when examining their influence on typing speed in a single study is very difficult. For instance, as motor learning happens over time, typing well-practiced words can be performed with less conscious effort. Thus, over time, a person forms a habitual typing behavior that is unique and consistent, whether on a traditional keyboard or a mobile touch keyboard. Because typing is a task where people increase their fluency (i.e., typing speed and accuracy) through practice, motor memory increases as a person repetitively types the exact words over time.

Four experiments were conducted to address the identified issues and assess how CTR and task-specific expertise influence typing speed. ‘Experiment One’ examines the influence of objective task difficulty on typing speed to reveal that increased CTR slows down typing.

‘Experiment two’ focuses on exploring how low task-specific expertise may result in slower typing speed and that the development of expertise may lead to an increase in typing speed. ‘Experiment Three’s’ objective is to apply the findings from the first two experiments to a real-world setting and construct predictive models. Finally, ‘Experiment Four’ extends the previous findings while proposing design improvements and new measures to improve the accuracy of the predictive model.

The specifics of each experiment, including the manipulation, control level, purpose, research questions, and the hypotheses being examined, are outlined in Table 4.2. Figure 4.1 presents the application of the DECODE framework to a typing context, showing the interrelationship between the various factors influencing the typing speed.

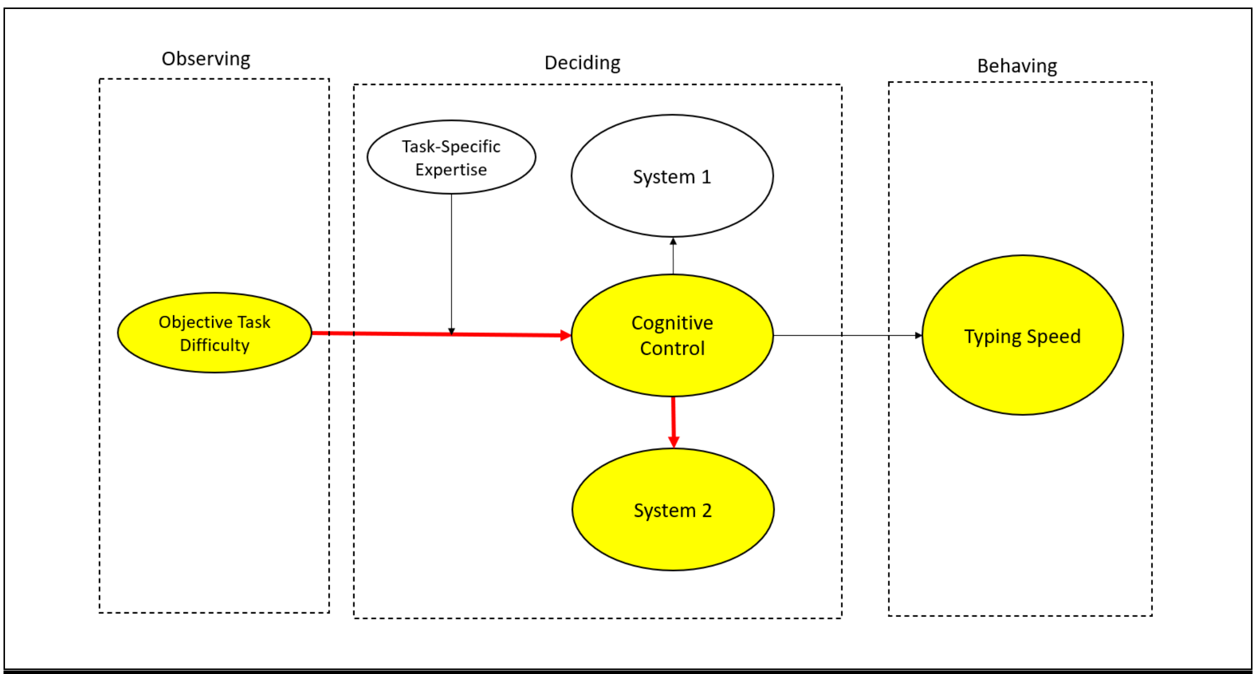


Figure 4.1 Application of DECODE framework to examine how objective task difficulty influences typing speed

Table 4.2. Overview of Experiments in Essay 2: Manipulation, Control, Purpose, Research Questions, and Hypotheses

Experiment	Manipulation	Level of Control	Purpose	RQ being examined	Hypotheses
Experiment 1	Objective task difficulty (Low, Medium, High Difficulty)	High	Examine the influence of objective task difficulty in typing speed.	RQ 1	H1: Increase in CTR (e.g., higher objective task difficulty) will slow down typing speed.
Experiment 2	Objective task difficulty and task-specific expertise	High	Examine the influence of developing a task-specific expertise in typing speed.	RQ 1, 2	H1: Increase in CTR (e.g., higher objective task difficulty) will slow down typing speed. H2: Low task-specific expertise will result in slower typing speed. H3: The development of task-specific expertise will increase typing speed.
Experiment 3	Objective task difficulty	Low	Examine how the findings from Study 1 and Study 2 apply in a real-world setting and propose a predictive model.	RQ 1	H4: Increase in CTR (e.g., higher objective task difficulty) will slow down typing speed on a multi-field form.
Experiment 4	Objective task difficulty	Low	Extend the findings from Study 1,2,3 while proposing a design and measure that can improve the predictive model.	RQ 1	H4: Increase in CTR (e.g., higher objective task difficulty) will slow down typing speed on a multi-field form. H5: Increase in CTR (e.g., presenting unexpected questions) will slow down typing speed on a multi-field form. H6: Increase in CTR (e.g., presenting unexpected questions) will slow progression speed on a multi-field form.

Methodology

Measure: Keystroke Capture. The behavior capture method starts with a script (written in JavaScript) that listens for interaction events during a web interaction. Typing events are logged with timestamps, from which two keystroke dynamics metrics—dwell time and transition time—are calculated. Dwell time refers to how long a given key is pressed down. Transition time refers to the time between the previous key going up and the next key going down (Kim et al. 2023, working paper under review). These metrics are important indicators in studying typing behaviors and can provide insight into users' cognitive states.

When typing a series of characters, the time required to locate the correct key (transition time) and the duration of holding down a key (dwell time) tend to form distinct patterns. It is common for individuals who type efficiently to have smaller or overlapping transitions (Teh et al. 2013), and a user's typing patterns can be used as identifying characteristics. The field of keystroke biometrics exploits these patterns to create a behavioral profile that can be used to authenticate a given user and, conversely, to identify typing patterns that do not match the expected profile for that user (Monrose and Rubin 2000).

Experimental Studies

Experiment 1. 'Experiment One' was conducted to assess RQ1 by examining how the additional CTR imposed by attention-demanding cues caused differences in typing behaviors (See Chapter 2 for RQ1). Considering that individuals are highly familiar with their personal information (e.g., first name, last name) in real-world forms, the study was designed to manipulate the objective task difficulty while holding the rest of the task designs constant.

The procedure used a within-subjects design, manipulating information familiarity across three conditions. To maximize experimental control, participants were asked to type responses into two text input fields: first (given) and last (family) names. The research team designed a web form with an embedded behavior-capturing script to covertly capture and process keystroke behaviors. Three conditions were presented in random order to each participant. Participants entered their own given and family names in the low objective task difficulty condition. In the medium objective task difficulty condition, participants entered the first and last names of a randomly generated common name (e.g., John Smith). The names for this condition were generated from a pool of the most common names from the 1990 US Census—thus, they should be somewhat familiar to most participants. Lastly, the high objective difficulty condition required participants to enter two short (i.e., name-length) Latin words. These Latin words include letter sequences that are far less common than those found in a typical North American name and were thus used as a strong manipulation of unfamiliar information.

Participants. The sample included 64 participants recruited from a large US university. After removing the participants who provided duplicate or incomplete responses, 45 observations remained. Keystroke events were leveraged to derive keystroke metrics, including dwell and transition time.

Results and Discussion. Two commonly used keystroke dynamics measures—dwell time and transition time—were analyzed across the three conditions specified in the previous section. Welch's t-tests were used to assess how the measures of interest statistically differed between the conditions. No significant differences were found when the dwell time was compared across the conditions, suggesting that dwell time may not be sensitive to the variations in the objective task difficulty. Transition time, on the other hand, showed significant differences between medium and high objective task difficulty conditions ($p = 0.0128$) and between the low and high objective task difficulty conditions ($p = 0.0003$). The results indicate that as the objective task difficulty increases, so does the transition time. Table 4.3 summarizes the results that compare each pair of conditions for both dwell and transition time.

Several limitations that informed the refined experimental design for 'Experiment Two' were identified. Specifically, situations in which a user types overly complex text (e.g., Latin letters) are uncommon and may not generalize well to real-world scenarios. Although the analysis results showed that tasks with varying levels of objective difficulty influence typing speeds, further experimentation is needed to ensure the generalizability of the results.

Table 4.3 Summary of Results (Welch's T-tests between each pair of condition)

Measure	Difficulty Condition 1	Difficulty Condition 2	Mean 1	Mean 2	P-value	Significance
Dwell	Low	Medium	123.2422	118.5483	0.4941	
Dwell	Medium	High	118.5483	123.9182	0.4015	
Dwell	Low	High	123.2422	123.9182	0.9213	
Transition time	Low	Medium	95.2490	116.5282	0.2172	
Transition time	Medium	High	116.5282	165.9041	0.0128	*
Transition time	Low	High	95.2490	165.9041	0.0003	***
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '						

Experiment 2. In 'Experiment Two', the aim is to evaluate RQ1 and RQ2 by examining the influence of task-specific expertise on typing speed across tasks with varying levels of objective task difficulty. A within-subjects design is employed, involving three task conditions: (1) typing

their own first and last name (low difficulty), (2) typing a randomly generated English name (medium difficulty), and (3) typing a randomly generated non-English name (high difficulty). To assess the impact of task-specific expertise on typing speed, participants complete each task condition five rounds, allowing for the development of familiarity with the given task. This approach ensures a comprehensive understanding of how task-specific expertise affects typing speed across different levels of task difficulty.

Participants. The experiment involved 203 participants recruited from a large US university. After excluding the participants who submitted duplicate or incomplete responses, 196 observations remained.

Results and Discussion. To evaluate the differences across the 5 rounds of data for each condition, repeated measures Analysis of Variance (ANOVA) was conducted to compare the differences. Repeated measures ANOVA is appropriate for within-subjects design as the participants executed similar tasks across multiple rounds. The repeated measures ANOVA results (See Table 4.4) show significant effects of task type, round order, and their interaction on typing speed. This suggests that the task's objective difficulty and progression influenced participants' typing speed.

Table 4.4. Repeated Measures Analysis of Variance Results

Effect	DF Numerator	DF Denominator	F	P value	Significance	Ges
Type	2	390	194.19311	0.0000	****	0.1923
Order	4	780	65.34577	0.0000	****	0.0390
Type:Order	8	1560	14.95915	0.0000	****	0.0151
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*' 0.1 < ' '						

Similar to the homogeneity of variances assumptions applicable for between-subjects ANOVA, repeated measures ANOVA assumes that the variance of the differences between the related groups is equal (Field 2013). Mauchly's Test for Sphericity was used to assess whether this assumption was met. The results (See Table 4.5) indicate that the variance between all possible pairs of related groups is unequal. This suggests that a correction is needed for the F-test in ANOVA to obtain accurate p-values.

Table 4.5. Mauchly’s Test for Sphericity Results			
Effect	**W**	**P**	**Significance Level**
Type	0.8852	0.0000	****
Order	0.8559	0.0000	****
Type:Order	0.7438	0.0000	****
Significance Code: 0 < ‘****’, 0.001 < ‘***’, 0.01 < ‘**’, 0.05 < ‘*’, 0.1 < ‘ ’			

Since Mauchly's Test for Sphericity results (Table 4.5) indicated that the variance between all possible pairs of related groups is not equal, a correction is needed to adjust the degrees of freedom for the F-test in ANOVA (Girden 1992). To address this issue, Sphericity Corrections were applied to provide adjusted p-values that account for the observed violation of sphericity (See Table 4.6).

Table 4.6. Sphericity Correction Results						
Effect	**GGe**	**P-value (GG)**	**Significance Lev–l - GG**	**HFe**	**P-value (HF)**	**Significance Lev–l - HF**
Type	0.8970	0.0000	****	0.9048	0.0000	****
Order	0.8559	0.0000	****	0.8731	0.0000	****
Type:Order	0.7438	0.0000	****	0.7698	0.0000	****
Significance Code: 0 < ‘****’, 0.001 < ‘***’, 0.01 < ‘**’, 0.05 < ‘*’, 0.1 < ‘ ’						

The significant outcomes of the Sphericity Corrections, demonstrated by the preservation of statistical significance after accounting for sphericity violations using both Greenhouse-Geisser (GG) and Huynh-Feldt (HF) methods, in conjunction with the insights from repeated measures ANOVA and Mauchly's Test, collectively reinforce the validity of these findings. The results indicate a substantial impact of both task types and round order, as well as the interaction between Type and Order on the dependent variable. As a result, post hoc analysis for pairwise comparisons was conducted using Tukey’s HSD.

Table 4.7 includes Tukey's HSD results for round-by-round comparison between task types, with each task type only compared to its matching round counterpart. For instance, the first instance of non-English name typing is only compared to the instance of own name typing or English typing. The findings reveal significant differences between non-English, English, and own name typing tasks in each matching round combination. All pairwise comparisons demonstrated statistical significance with p-values less than 0.0001, suggesting substantial differences exist in typing speed between the tasks in every matching round.

Table 4.7. Tukey's HSD Results (round by round comparisons between types)

Variable 1	V1 Round	Variable 2	V2 Round	Estimate	SE	Df	t-ratio	p-value	Significance
Non-English	1	Own	1	141.546	6.74	1122	21	0.0000	****
English	1	Non-English	1	-65.257	6.74	1122	-9.682	0.0000	****
English	1	Own	1	76.289	6.74	1122	11.319	0.0000	****
Non-English	2	Own	2	104.202	6.74	1122	15.46	0.0000	****
English	2	Non-English	2	-45.263	6.74	1122	-6.715	0.0000	****
English	2	Own	2	58.939	6.74	1122	8.744	0.0000	****
Non-English	3	Own	3	91.829	6.74	1122	13.624	0.0000	****
English	3	Non-English	3	-46.996	6.74	1122	-6.973	0.0000	****
English	3	Own	3	44.833	6.74	1122	6.652	0.0000	****
Non-English	4	Own	4	82.683	6.74	1122	12.267	0.0000	****
English	4	Non-English	4	-39.878	6.74	1122	-5.916	0.0000	****
English	4	Own	4	42.805	6.74	1122	6.351	0.0000	****
Non-English	5	Own	5	69.261	6.74	1122	10.276	0.0000	****
English	5	Non-English	5	-35.133	6.74	1122	-5.212	0.0000	****
English	5	Own	5	34.128	6.74	1122	5.063	0.0000	****
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '									

For the second part of the analysis, a within-task comparison was conducted to examine the differences across the rounds for each type of task. Table 4.8 includes a pairwise comparison across every possible pair of rounds for typing one's own name. In every pair, the results showed no significant differences between the rounds, with p-values ranging from 0.8557 to 1. This suggests that the performance in typing one's own name remains consistent across all five rounds.

Table 4.8. Tukey's HSD Results (typing own name – within-task comparison)

Round # (First)	Round # (Second)	Estimate	SE	Df	t-ratio	p-value	Significance
1	2	6.661	5.26	2321	1.265	0.9953	
1	3	4.753	5.26	2321	0.903	0.9999	
1	4	9.669	5.26	2321	1.837	0.8829	
2	3	-1.908	5.26	2321	-0.362	1	
4	2	-3.008	5.26	2321	-0.572	1	
4	3	-4.916	5.26	2321	-0.934	0.9998	
5	1	-9.976	5.26	2321	-1.895	0.8557	
5	2	-3.315	5.26	2321	-0.63	1	
5	3	-5.223	5.26	2321	-0.992	0.9997	
5	4	-0.307	5.26	2321	-0.058	1	
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '							

Table 4.9 focuses on with-in comparison results for typing a randomly generated English name. In contrast to 4.8, the results indicate that there are several significant differences between the rounds. The differences between the rounds were most significant between the first and the latter rounds with the magnitude of the difference in estimate being greatest between the first and the fifth round. Such differences were not observed in between rounds 2 and 3, 3 and 4, 3 and 5, and 4 and 5. Overall, the results suggest that there was a gradual decline in key transitions, which indicates that the average typing speed gradually increased as the rounds progressed.

Table 4.10 highlights the with-in comparison results for typing a randomly generated non-English name. Similar to the results for the English name typing condition, the differences between the rounds were most significant between the first and the latter rounds, with the magnitude of the difference in estimate being greatest between the first and the fifth round. Such differences were not observed between rounds 2 and 3, 3 and 4, and 4 and 5. One notable difference from the English name typing condition is that the difference between the third and the fifth round was highly significant, suggesting the possibility that the development of task-specific competency has not yet likely occurred until the fourth round of typing tasks.

Table 4.9. Tukey's HSD Results (typing English Name: within-task comparison)

Round # (First)	Round # (Second)	Estimate	SE	Df	t-ratio	p-value	Significance
1	2	24.011	5.26	2321	4.562	0.0005	***
1	3	36.209	5.26	2321	6.88	0.0000	****
1	4	43.153	5.26	2321	8.199	0.0000	****
2	3	12.198	5.26	2321	2.318	0.5773	
4	2	-19.142	5.26	2321	-3.637	0.0224	*
4	3	-6.944	5.26	2321	-1.319	0.9929	
5	1	-52.137	5.26	2321	-9.906	0.0000	****
5	2	-28.126	5.26	2321	-5.344	0.0000	****
5	3	-15.928	5.26	2321	-3.026	0.144	
5	4	-8.983	5.26	2321	-1.707	0.9313	
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '							

Table 4.10. Tukey's HSD Results (typing non-English Name – within comparison)

Round # (First)	Round # (Second)	Estimate	SE	Df	t-ratio	p-value	Significance
1	2	44.004	5.26	2321	8.361	0.0000	****
1	3	54.469	5.26	2321	10.349	0.0000	****
1	4	68.532	5.26	2321	13.021	0.0000	****
2	3	10.465	5.26	2321	1.988	0.8055	
4	2	-24.528	5.26	2321	-4.66	0.0003	***
4	3	-14.063	5.26	2321	-2.672	0.3229	
5	1	-82.261	5.26	2321	-15.629	0.0000	****
5	2	-38.256	5.26	2321	-7.269	0.0000	****
5	3	-27.791	5.26	2321	-5.28	0.0000	****
5	4	-13.729	5.26	2321	-2.608	0.3645	
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '							

The overall results support H1, H2, and H3. Regarding H1, a noticeable difference exists across all conditions. The evident difference between the English Name (Medium difficulty) and Non-English Name (High difficulty) conditions suggests that greater objective task difficulty slows down typing speed.

H2 and H3 are examined simultaneously by inspecting round-by-round differences for each condition (Table 4.8, 4.9, and 4.10). Specifically, H2 is assessed using pairwise comparisons across rounds within each condition. The results indicate no significant differences across rounds for the own name typing condition. However, differences across rounds are observed for the non-English name and English name typing conditions, where participants are expected to have lower task-specific expertise.

The development of task-specific expertise is demonstrated in the results presented in Table 4.9 and Table 4.10. As the rounds progress for English and Non-English Name conditions, typing speeds gradually increase. The boxplot of typing speeds (i.e., normalized key transition) in Figure 4.2 supports these findings. From the visual inspection of the graph, there are clear differences across each typing condition, which aligns with the results from the round-by-round comparison presented in Table 4.7. Further, the gradual decline in key transition time (i.e., increase in typing speeds) indicates that the participants initially had low task-specific expertise, which gradually developed over time.

In summary, the results support H1, H2, and H3, demonstrating the effects of task difficulty and task-specific expertise on typing speeds.

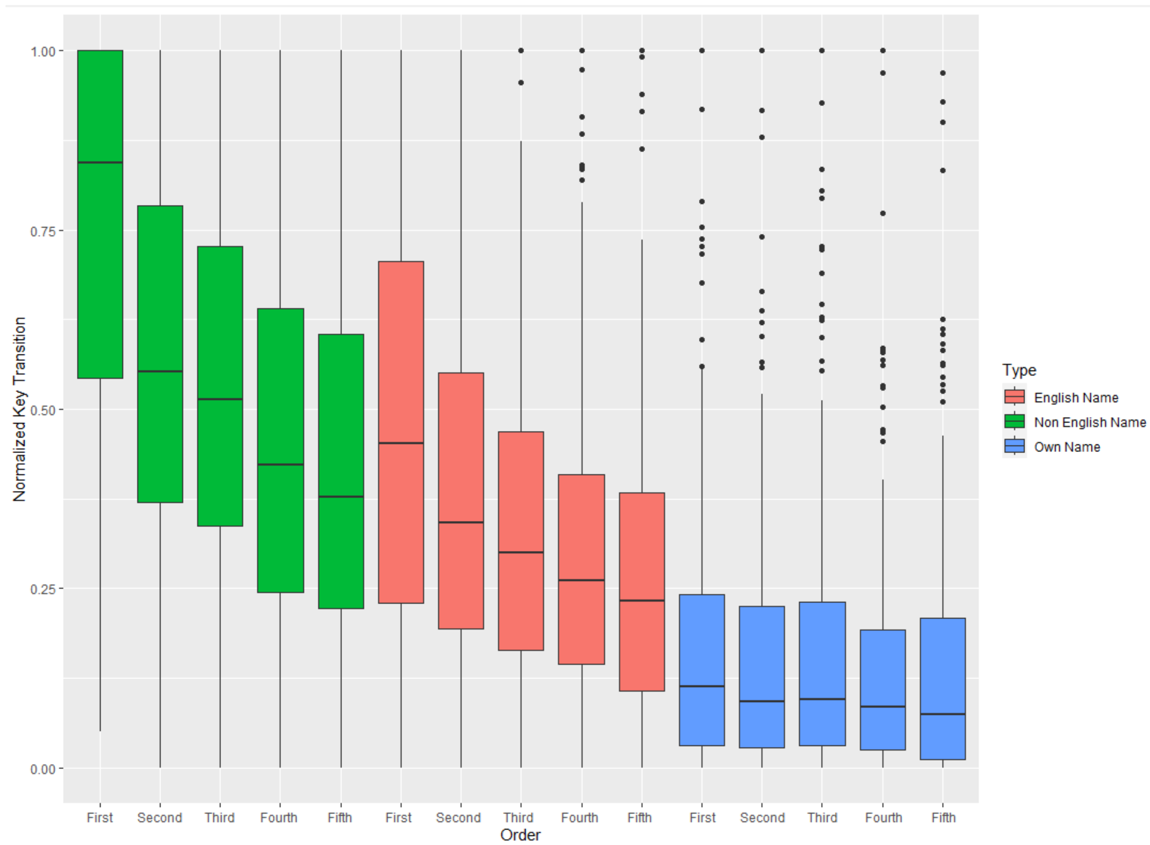


Figure 4.2 Boxplot of typing speed across task types and round order

Experiment 3. ‘Experiment Three’ focuses on extending the first and second experiments’ findings to a real-world setting. Specifically, the second experiment’s findings suggest that individuals maintain consistent typing speeds when executing tasks with high task-specific expertise, such as entering their names. In contrast, typing tasks with low task-specific expertise initially exhibit slower typing speed, gradually becoming faster as individuals develop expertise through repetition.

These observations suggest the potential for designing a privacy-preserving keystroke capture tool with the ability to differentiate fraudulent from genuine identity information. For instance, a typical identity fraud scenario involves a fraudster using stolen or synthetic identity information to submit a loan application or account creation form. The fraudster is likely to have low task-specific expertise when entering stolen or synthetic identity information as they are less familiar with the stolen or synthetic identity details compared to genuine users who regularly enter their information. This difference in familiarity and expertise may manifest as distinctive typing patterns.

First, as the participants were provided with a set of fraudulent identity information in the low task-specific expertise condition, it was necessary to isolate the manipulation's effects on actual typing behaviors from those due to other entry methods. More specifically, because a real-world form was used for data collection, many modern browsers could recognize the common structure of a personal information form and offer to automatically fill in the information for the user (i.e., browser autofill). Additionally, participants might be tempted to simply copy and paste the supplied information into the form fields rather than typing manually. Although either of these behaviors may serve as strong signals, the focus for 'Experiment Three' was on the typing behaviors themselves (However, these and other non-typing behaviors were incorporated into the design and analysis for 'Experiment Four'). A second consideration was that if a fraudster might be aware that their typing behavior was being monitored, he or she may intentionally type fraudulent information as quickly as possible to appear familiar with the information. Thus, a time-pressure element was incorporated as a variation in the low task-specific expertise condition to examine whether this would affect the system's ability to differentiate fraudulent from genuine behaviors.

The experiment employed a within-subjects design, wherein participants completed two tasks using the same form: one requiring high expertise and the other requiring low expertise. The form was designed to mimic a real-world loan application form. Participants were instructed not to use any autofill functionality. JavaScript was embedded to prevent the use of copy and paste.

Participants completed the form twice, once with their personal information (high task-specific expertise) and once with a synthetic identity (low task-specific expertise). Additionally, half of the participants were given a time-pressure manipulation as part of the low task-expertise condition, in which they were asked to type as quickly and accurately as possible. The inclusion of time pressure manipulation allows the examination of how a potential countermeasure technique, such as typing at varying speeds) may influence the predictive models. Analyzing the effects of time pressure on information entry can reveal distinctive indicators or patterns unique to individuals employing countermeasures.

The experimental design ensured that the identity information used in the low task-expertise condition was realistic and minimized the variance due to external factors that may influence our results. Supplied identity information was generated using API calls to retrieve the first and last names from the 1990 US Census Bureau database. The first and last names were randomly paired and included both male and female names. For the address information, a state was randomly chosen from the list: Arizona, Missouri, Alaska, Indiana, Texas, Tennessee, Hawaii, Maine, and Nevada. Subsequently, an API call was made to pick a random city within the chosen state. Lastly, a random number generator generated zip codes and mobile numbers. These pieces of identity information were randomly assembled for each participant, thus randomly distributing any potential confounding effects of a given participant's expertise with any part of the supplied identity (e.g., if the participant happened to share the name or home state with the supplied identity).

This approximates the real-world fraud scenario in which a fraudster might occasionally be more familiar with a given synthetic or stolen identity.

Participants. Participants were recruited from two large universities in the western US. Of the 283 total participants, 35 were excluded due to missing data (24 participants) or because they failed to complete all parts of the study (11 participants). These exclusions left 248 participants in the analysis set.

Results and Discussion. After accounting for differences between our two samples (see Table 4.11), two one-way ANOVAs were conducted to compare the typing patterns across the fields within the form, familiarity conditions, and speed conditions (see Table 4.12). The results suggested significant differences in typing patterns within the form across the fields and between the familiarity conditions. However, there were no differences in the speed condition between the forms. Consequently, the datasets collected across the two universities were combined, and a control variable was added for the speed condition. Since there were only two familiarity conditions, no subsequent analysis was performed.

Table 4.11 T-test Results Across Sample and Time Pressure

Field	P-value (Sample)	Significance	P-value (Time Pressure)	Significance
First	0.372		0.152	
Last	0.890		0.006	***
City	0.664		0.660	
State	0.109		0.823	
E-mail	0.926		0.06	
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '				

Table 4.12 ANOVA Results for Transition Time (fields and familiarity conditions)

Effect	Df	Sum Sq	Mean Sq	F-Value	P-value	Significance
Fields	6	30737071	5122845	320.6	0.0000	****
Familiarity Condition	1	5575036	5575036	348.9	0.0000	****
Speed Condition	1	6665	6665	1.539	0.216	
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '						

Tukey's HSD test was conducted post hoc to examine the pairwise comparison of the fields. The results of Tukey's HSD test, presented in Table 4.13, demonstrate that all pairs, except last name and first name, had significantly different transition times. Lastly, although there were no significant differences in transition time for the speed condition at the form level, the field-level analysis indicated differences in the last name and the email field, and a control variable for the speed variable was added for the subsequent analyses.

Table 4.13 Tukey's HSD Results (Fields and Familiarity Conditions)

Variable 1	Variable 2	Difference	Lower CI	Upper CI	P-adjusted	Significance
Last Name	First Name	15.871	-7.8079	39.55	0.4293	
City	First Name	42.5789	18.8999	66.2579	0.0000	****
State	First Name	51.3708	27.6917	75.0498	0.0000	****
Zip code	First Name	263.0243	239.3453	286.7033	0.0000	****
Mobile Last Four	First Name	209.8311	186.1515	233.5101	0.0000	****
E-mail	First Name	69.7777	46.0987	93.4568	0.0000	****
City	Last Name	26.7078	3.0288	50.3869	0.0154	*
State	Last Name	35.4997	11.8207	59.1787	0.0002	***
Zip code	Last Name	247.1532	223.4742	270.8323	0.0000	****
Mobile Last Four	Last Name	193.96	170.281	217.639	0.0000	****
E-mail	Last Name	53.9067	30.2277	77.5877	0.0000	****
State	City	8.7918	-14.8871	32.4708	0.9266	
Zip code	City	220.4454	196.7663	244.1244	0.0000	****
Mobile Last Four	City	167.2521	143.5731	190.9311	0.0000	****
E-mail	City	27.1988	3.5198	50.8778	0.0125	*
Zip code	State	211.6535	187.9745	235.3325	0.0000	****
Mobile Last Four	State	158.4603	134.7812	182.13931	0.0000	****
E-mail	State	18.4069	-5.272	42.086	0.2474	
Mobile Last Four	Zip code	-53.1932	-76.8722	-29.51421	0.0000	****
E-mail	Zip code	-193.2465	-216.9255	-169.56752	0.0000	****
E-mail	Mobile Last Four	-140.0533	-163.7323	-116.3743	0.0000	****
Significance Code: 0 < '****', 0.001 < '***', 0.01 < '**', 0.05 < '*', 0.1 < ' '						

The ANOVA and Tukey's HSD test suggested that keystroke dynamics in each field contain markers of synthetic information entry. Classification models predicting the fraudulent vs. genuine form (i.e., familiarity condition) were created using the transition time with control variables for the speed condition. The familiarity condition was used as the label for the classification model (i.e., 0 = high familiarity). Since there were two entry events for the 248 participants, 496 observations were available for the classification models. The dataset consists of 50% synthetic and 50% personal information, so the base accuracy rate for random guessing is 50%. A logistic regression model was selected as the baseline model to gauge performance gains. Other popular algorithms, such as support vector machine (SVM), random forest (bagging), and XGBoost (boosting), were also used for comparison.

All models were trained and evaluated using repeated 10-fold cross-validation, with the data reshuffled in each of the 5 repeats. The standard deviation of the evaluation metrics during such a process can be used to evaluate the model's overall stability. Standard evaluation measures, including the average area under the curve (AUC) of the receiver operating characteristics (ROC) curve, average accuracy, the standard deviation of the accuracy, sensitivity, and specificity, are reported in Table 4.14. Overall, all models surpassed the calculated base accuracy rate of 50% and achieved around 80% accuracy, suggesting that the selected features were consistent across the various algorithms. The random forest model delivered the highest performance on the test dataset, with 83.06% accuracy (averaged across repeats). The AUC of the ROC curve was also the highest for the random forest model.

Table 4.14 Model Results (Field-to-Field Measures)

Model	AUC	Accuracy (Test)	Accuracy SD (Test)	Sensitivity	Specificity
Logistic Regression	0.815	81.45%	4.39%	0.8361	0.7937
SVM (Linear)	0.823	82.26%	4.92%	0.8525	0.7936
Random Forest	0.831	83.06%	3.62%	0.8197	0.8413

The results of 'Experiment Three' demonstrate strong differentiation between a fraudulent and genuine form entry from a set of keystroke dynamics features. These results were obtained using keystroke metrics derived using our privacy-preserving architecture, limited to the metadata describing how the information was entered into the fields. In addition, the results were obtained covertly in a standard, real-world form like those commonly used in a customer onboarding scenario.

Despite these encouraging findings, there are several limitations and learnings that informed the changes implemented in 'Experiment Four'. First, field-specific factors created significant differences in keystroke behaviors (see Table 4.9). For example, the number of characters entered into a field can vary significantly (e.g., Arizona vs. AZ). Second, fields requiring numeric entries added noise due to different keyboard layouts (e.g., a number pad vs. number keys along the top of the keyboard) or perhaps due to the relative infrequency with which numbers are typed in day-to-day use as compared to words. Lastly, some fields require multiple key presses (e.g., holding a shift to type the "@" symbol for an email address). These factors likely added noise to the dataset and reduced the predictive accuracy of the classification models.

In addition, the accuracies of the evaluated classification models were lower than other reported accuracies in related keystroke dynamics research in the identity fraud context (Monaro, Galante, et al. 2018; Monaro, Gamberini, et al. 2018), which averaged around 90–95%. This difference in accuracy is partially due to the additional constraints incorporated into the study design to achieve system design goals, namely, (1) using a privacy-preserving architecture that avoids storing or transmitting any personal information and (2) exclusively using short-entry form fields that match those typically used in real-world scenarios. Nevertheless, the overall goal was to meet or exceed the performance of current state-of-the-art procedures while still meeting our system design requirements. As a result, a final iteration of the evaluation design was performed for 'Experiment Four' to increase the predictive power of the classification models while maintaining alignment with system design requirements.

Experiment 4. To obtain stronger differentiation between fraudulent and genuine form-filling behaviors, prior deception literature was reviewed to find additional mechanisms that could be incorporated into the experimental design. One important mechanism that surfaced was a technique for inducing additional cognitive load during deception using "interventions" (Vrij 2008). These interventions typically take the form of unanticipated requests for further detail or elaboration on a portion of the deceiver's story, placing additional cognitive demands on the deceiver (Lancaster et al. 2013). The utility of unexpected questions in reliably revealing deception has been demonstrated in a wide variety of studies (for reviews, see Mac Giolla and Luke 2021, Vrij et al. 2017). Moreover, these unexpected questions have also been adapted to the identity fraud context, revealing stronger differentiation in both mouse dynamics (Monaro et al. 2017a, 2017b) as well as keystroke dynamics data (Monaro, Galante, et al. 2018; Monaro, Gamberini, et al. 2018). Specifically, Monaro, Galante, et al. 2018 captured information about participants' birthdate and birthplace during an early phase of their experiment, and then used that information to derive unexpected questions (e.g., "how old are you" or "in which province were you born"). The typing behaviors for these unexpected questions produced the most important features used to obtain classification accuracies of approximately 90–95%.

Although the technique used by Monaro, Galante, et al. 2018 provides a candidate strategy for increasing the predictive accuracy of our models, some drawbacks of their approach were identified concerning system design goals. Namely, deriving unexpected questions involved the transmission and storage of sensitive information provided during the early phase of the experiment, which works in opposition to our goal of maintaining users' privacy while assessing identity fraud risk. Thus, 'Experiment Four' was designed to address several limitations we identified from 'Experiment Three' while incorporating additional mechanisms to increase the predictive accuracy of our models without compromising the overall system design goals to maintain users' privacy. To accomplish these objectives, Monaro, Galante, et al.'s 2018 work was adapted with extensions and improvements to align with the system goals.

First, the border control scenario used by Monaro, Galante, et al. 2018 required participants in the deception condition to memorize the personal details of a fake identity. As fraudsters in a digital identity fraud context typically work from a set of stolen or synthetic identity information (e.g., in a spreadsheet), memorization is less applicable and induces unnecessary cognitive load. Second, unexpected questions were derived from two data points provided by the user (birthdate and zip code) but the client-side processing architecture was followed to produce these calculations without storing or transmitting their values. As a result, the design employed a simple age calculation based on the birthdate and a city/state lookup based on the provided zip code. The answer metadata—describing how those unexpected questions were answered rather than what answer was provided—was then used in our subsequent analyses.

A final innovation we added for 'Experiment Four' comprised a new set of digital behavioral biometrics different from traditional keystroke dynamics measures to add further predictive power to the classification models. These features were designed to capture additional information that distinguishes an identity fraudster from a genuine user and included information about autofill and copy/paste behavior as well as other metadata about how the participant transitioned from field to field while entering the information. Accordingly, and in contrast to

‘Experiment Three,’ ‘Experiment Four’ participants were allowed to enter the information naturally and without constraining any autofill or copy/paste behavior to better approximate the behavior a real-world fraudster might use to enter stolen or synthetic identity information on a form.

The design used in ‘Experiment Four’ was adapted for ‘Experiment Three’ with several refinements. The number of fields was reduced to four to simplify the form and data analysis, given the variability across the form fields observed in ‘Experiment Three’. The form fields used included two text fields (first name and last name) and two numeric fields (date of birth and zip code). ‘Experiment Four’ again used a within-subjects design in which all participants completed both high and low familiarity in random order. The speed manipulation—in which participants were instructed to enter information as quickly as possible—contributed little to the models in ‘Experiment Three’ and was removed for ‘Experiment Four’. For the high task-specific expertise condition, participants again entered their own personal identity information. For the low task-specific expertise condition, participants were instructed to navigate to an external link where they could download a spreadsheet containing the imposter identity—adding additional realism to the identity fraud scenario in which fraudsters would have access to stolen or synthetic identity information.

After providing the identity information, participants were presented with three unexpected follow-up questions. These were produced using client-side JavaScript to preserve privacy with a simple age calculation and an external API lookup using zip code. The follow-up questions could be answered by searching the web if needed and were thus designed to motivate behavioral signals that differentiate based on the users’ familiarity with the required information.

When individuals committing identity fraud encounter such unexpected follow-up questions in a real-world setting, they would likely need to search the web for the correct answer. This behavior should deviate significantly from the behavior of the identity owner (as expected in the high familiarity condition), who should simply retrieve the identity information from memory and quickly respond with their hometown, state, and age. Importantly, however, these follow-up questions were also chosen because of the ease with which they can be dynamically calculated using client-side, in-browser functionality without the need to store identity information.

Participants. The subjects were recruited from a large public university in the US. Participants who completed the experiment but did not enter the correct information for the low familiarity condition were removed. As the task was designed to take about 10 minutes, subjects who took an abnormally long time (e.g., 30 minutes) or had large breaks in activity time were also removed from the analysis pool. Lastly, we removed participants who used a mouse to copy and paste all the information in the low-familiarity condition. After data cleaning, 157 subjects remained, each with two trials, for a total of 314 observations.

Additional Digital Behavioral Biometrics. In addition to the keystroke dynamics features used in 'Experiment Three', four new features in 'Experiment Four' (see Table 4.15) were designed to capture the strong behavioral signals produced in a typical identity fraud scenario in which a fraudster's copies were stolen or synthetic identity information, as summarized below. Consistent with the privacy-preserving design goal, the calculations required for features were implemented as part of the client-side JavaScript library. The resulting features comprised anonymized metadata describing how information was entered or how the user traversed the form fields on the page.

Table 4.15. Summary of New Digital Behavior Biometrics	
Measure	**Description**
Autofill	If browser autofill is used, this indicates the user owns that information. Autofill is identified as a series of key presses that populate multiple fields simultaneously.
Paste	Paste indicates that the information does not belong to the user. Paste events occur when the user pastes any information into a field using one of several methods.
Tab	The tab key is implicated in most task-switching keyboard shortcuts such as switching to the most recently used application). We derived a binary flag to indicate whether a tab key was used on a given page.
Field Transition Time	Field-to-field transition times are based on timestamps from field "focus" (i.e., field entry) and field interaction data. This feature indicates how efficiently users moved through the form and detects pausing or hesitating while answering a question.

Results and Discussion. Multiple classification models were constructed to validate the performance of the measures, focusing primarily on evaluating the predictive accuracy gains provided by the new study procedures and the new features added in 'Experiment Four'. The baseline model (Equation 1) includes keystroke dynamics measures (e.g., transition and dwell time) from the main identity form as a replication of the model from 'Experiment Three'. Model 1 (Equation 2) adds field transition measures from the main identity form. Model 2 (Equation 3) includes keystroke dynamics and field transition measures for the follow-up form containing unexpected questions. Model 3 (Equation 4) includes all measures from the main form only, and Model 4 (Equation 5) includes all measures for both the main identity form and the follow-up form. Given the relatively small dataset, all models were evaluated using 10-fold cross-validation with 5 repeats.

Baseline: $y = \beta_0 + \beta_1 MF\ Transition + \beta_2 MF\ Dwell + \epsilon$

Model 1: $y = \beta_0 + \beta_1 MF\ Transition + \beta_2 MF\ Dwell + \beta_3\ MF\ Field\ Transition + \epsilon$

Model 2: $y = \beta_0 + \beta_1 MF\ Transition + \beta_2 MF\ Dwell + \beta_3\ MF\ Field\ Transition + \beta_4 FF\ Transition + \beta_5 FF\ Dwell + \beta_6 FF\ Field\ Transition + \epsilon$

Model 3: $y = \beta_0 + \beta_1 MF\ Transition + \beta_2 MF\ Dwell + \beta_3\ MF\ Field\ Transition + \beta_4 Paste + \beta_5 Autofill + \beta_6\ Tab + \epsilon$

Model 4: $y = \beta_0 + \beta_1 MF\ Transition + \beta_2 MF\ Dwell + \beta_3\ MF\ Field\ Transition + \beta_4 FF\ Transition + \beta_5 FF\ Dwell + \beta_6 FF\ Field\ Transition + \beta_7 Paste + \beta_8 Autofill + \beta_9\ Tab\ Press + \epsilon$

The model comparison in Table 4.16 demonstrates the utility of the study and system design changes implemented for ‘Experiment Four’. Specifically, the field transition features (Model 1), data from unexpected questions (Model 2), data from the main form and additional behavioral features (Model 3), and unexpected questions and additional behavioral features (Model 4) progressively improved model performance, culminating in a final model that effectively predicts identity fraud. Notably, the final model performance meets or exceeds that achieved by the current state-of-the-art methods from recent literature (Monaro, Galante, et al. 2018) while providing the added benefits of our privacy-preserving design and depending solely on the short-entry fields typically found on a customer onboarding form.

Table 4.16. Model Performance Results (Averaged Over Five Repeats)

Model	Accuracy	AUC	Sensitivity	Specificity	Precision	Recall	F
Baseline	69.91%	0.6918	0.6041	0.7901	0.7412	0.6041	0.6572
Model 1	86.11%	0.8356	0.8283	0.8929	0.8896	0.8283	0.8523
Model 2	90.62%	0.8735	0.8783	0.8784	0.9281	0.8784	0.9006
Model 3	93.00%	0.9056	0.9183	0.9405	0.9358	0.9183	0.9254
Model 4 – Logistic regression	94.61%	0.8754	0.9385	0.9533	0.9526	0.9385	0.9443
Model 4 – SVM (Linear)	93.64%	0.9025	0.9197	0.9521	0.9512	0.919	0.9337
Model 4 – Random Forest	94.71%	0.8729	0.9319	0.9616	0.9601	0.9319	0.9444
Model 4 – XGBoost	94.90%	0.9214	0.9267	0.9692	0.9660	0.9267	0.9452

One important limitation of the changes we implemented and evaluated in this study is the fact that, in adding the unexpected questions to improve predictive accuracy without compromising user privacy, we partially violated another design requirement regarding not adding additional friction or affecting the customer experience. Although we adapted the unexpected questions approach to minimize the impact of those questions in terms of friction (i.e., adding just three follow-up questions with short responses), we acknowledge that the added performance accuracy obtained in exchange for that added friction may be a tradeoff that some digital platform owners would not be willing to make.

Essay Two Summary

In this Essay, the DECODE framework was employed to examine how the objective task difficulty and task expertise influence typing speed. The Essay hypothesized that increasing objective task difficulty would increase CTR and CC and thus decrease typing speed, while having low task-specific expertise will also result in slower typing speed. Further, the Essay hypothesized that developing task-specific expertise will increase typing speed. The hypotheses were tested on a single field and form entry events.

'Experiment One' examined how objective task difficulty influenced typing behaviors, specifically regarding dwell and transition times. Participants were asked to type their name, a common name, or two Latin words representing low, medium, and high objective task difficulty, respectively. No significant differences were found in dwell time across conditions, but transition time increased with objective task difficulty.

'Experiment Two' examined the influence of task-specific expertise and typing speed across varying levels of objective task difficulty. The experiment aimed to empirically demonstrate that developing task-specific expertise increases the device usage speed (i.e., typing speed). A within-subject experimental design prompting participants into three conditions were employed:

(1) name typing condition (low difficulty), (2) English name (medium difficulty), and (3) non-English name (high difficulty) conditions. Participants entered identity information five times for each condition.

The comparison of typing speeds between the conditions was performed to examine the influence of objective task difficulty. To analyze the influence of developing task-specific expertise on typing speeds, round-by-round pairwise comparisons for each condition were performed. The overall results demonstrated that greater objective task difficulty significantly slowed the typing speed while developing task-specific expertise gradually improved the typing speed.

‘Experiment Three’ sought to apply findings from the first and second experiments to a real-world setting by differentiating fraudulent and genuine identity information through typing patterns. The experiment used a within-subjects design, with participants completing a loan application form using their personal information (high expertise) and a synthetic identity (low expertise). Time-pressure manipulation was introduced to examine the effects of potential countermeasure techniques on predictive models.

‘Experiment Three's’ results demonstrated a strong differentiation between a fraudulent and genuine form. Further, no significant differences were observed between the participants imposed with time-pressure manipulation and the baseline. The predictive model, built using the key transition times from various fields, achieved over 80% accuracy in correctly classifying the identity entry instances (own identity vs. generated).

‘Experiment Four’ aimed to build upon the findings from 'Experiment Three' by identifying additional features that can improve the predictive model's performance while reducing the number of identity fields to those most commonly requested in real-world situations (e.g., first name, last name, date of birth, and zip code).

After providing the identity information, reduced to four fields, participants were presented with three unexpected follow-up questions. These were produced using client-side JavaScript to preserve privacy with a simple age calculation and an external API lookup using zip code. Similar to the prior experiments, the keystroke dynamics features were captured while the participants entered the identity information. Moreover, new sets of Digital behavioral biometric measures, including Autofill, Paste, Tab usage, and Field-to-Field transition time, were derived. The final model that included all the new features resulted in 94.90% accuracy in correctly classifying the identity entry instances.

Overall, the four experiments conducted in this study contribute to a deeper understanding of how typing speed is affected by objective task difficulty and task-specific expertise. Specifically, the application of the DECODE framework provides a theoretical explanation of how the increase in objective task difficulty also increases CTR and thus manifests as slower typing speed. Further, by evaluating the findings in experiments that closely resemble the real world, the research lays the foundation for future advancements in typing pattern analysis and its application to various cybersecurity and authentication contexts.

CHAPTER 5: ESSAY 3 – AN EMPIRICAL EXAMINATION OF DECODE FRAMEWORK USING MOBILE DEVICES

Introduction and Related Work

‘Essay Three’ builds upon the findings of ‘Essay One’ and ‘Essay Two’ to explore the influence of Cognitive Control (CC) on online task execution using mobile devices and the Decision Evaluation with Cognitive Observations using Device Engagement (DECODE) framework. While previous essays have identified detectable behavioral differences between activities that engage or disengage CC, the generalizability of these findings to mobile devices remains uncertain due to inherent differences in usage behavior compared to computer mice and keyboards. Mobile devices equipped with sensors such as accelerometers and gyroscopes can capture motion data with high precision. As users can move and rotate the devices in various ways, the complexity of movement behaviors is significantly different from those observed with computer mice and keyboards.

Despite the challenges and complexity of movement behaviors in analyzing trace data, mobile device data holds immense potential for unveiling novel indicators of users’ underlying cognitive states. Realizing its potential, numerous studies have concentrated on trace data measured using sensors on mobile devices to analyze user behaviors and cognitive states. As shown in Table 5.1, research findings indicate that mobile device sensors can predict various cognitive states. In particular, the mobile device sensors were found to be a helpful indicator in predicting self-reported valence, depressive symptoms, level of cognitive load in typical tasks, and substance use disorder. Similar to the prior literature discussed in ‘Essay One’ and ‘Essay Two’, these studies examining the influence of users’ cognitive states on mobile device usage draw from established psychological theories and relevant literature to explore the relationships between device usage behavior and cognitive states. A list of studies that employed such an approach can be found in Table 5.1.

Table 5.1 Studies Investigating the Relationship Between Mobile Device Usage Behavior and Various Cognitive States

Reference	Research Context	Findings
Ruensuk et al. 2020	Emotion	Built-in sensors on smartphones can be used to predict self-reported valence with high accuracy (more than 92% accuracy)
Wahle et al. 2016	Depression	Smartphone sensor data, including wifi, accelerometer, GPS, and phone use data can be used to detect depressive symptoms.
Cabañero et al. 2020	Cognitive load	Analyzing user interactions with mobile devices can help determine the cognitive load required for different typical tasks on smartphones.
Bach et al. 2021	Voting Behavior	Digital trace data do not accurately predict undecided voters but achieve slightly better results for self-reported voting.
Zech et al. 2022	Substance Use Disorder	Smartphone-based tasks offer compliance, feasibility, and reliability in Ecological Momentary Assessment (EMA) research, playing a crucial role in understanding and predicting substance use disorder through digital trace data.

Hypotheses

Drawing from the prior essays, it is evident that repeated measurement experiments effectively unveil learning progression and changes in device usage behaviors as participants acquire task-specific expertise. Furthermore, experimental designs that closely resemble real-world scenarios (e.g., Experiments three and four from ‘Essay Two’) demonstrate the generalizability of the results obtained from the highly controlled experiments (e.g., Experiments one and two from ‘Essay Two’).

In relation to RQ 1, these findings suggest that observational studies might be better suited for establishing the ecological validity of the results. In particular, there will be instances where users may voluntarily engage or disengage with attention-demanding cues. Examples of engagement include deciding to tap on an advertisement or providing fraudulent responses to sensitive questions. On the other hand, examples of disengagement involve answering survey questions without careful evaluation or ignoring a specific advertisement on a webpage.

Figure 5.1 illustrates the application of the DECODE framework to examine how Cognitive Task Requirement (CTR) influences Cognitive Control and mobile device usage speed. Within the DECODE framework, CTR encompasses not only the cognitive demands of a task but also the changes in the external situation or environment, which can affect users' cognitive states (Christensen et al. 2016; Endsley 1995). By considering the effects of changing external situations on CTR, the DECODE framework allows for exploring how device usage behavior would change when the need to evaluate attention-demanding cues is removed (i.e., lower CTR). While the DECODE framework has been leveraged in scenarios where CTR increases, this new consideration allows the framework to examine what happens when the CTR decreases.

Concerning RQ 2, the development of task-specific expertise can be examined by employing repeated measures study design. Figure 5.2 demonstrates the application of the DECODE framework to examine how task-specific expertise influences Cognitive Control and mobile device usage speed. This figure illustrates the predicted relationship between task-specific expertise and mobile device usage speed. Low task-specific expertise, as experienced when executing unfamiliar tasks, is expected to result in slower mobile device usage, which is expected to improve as the task-specific expertise is developed gradually.

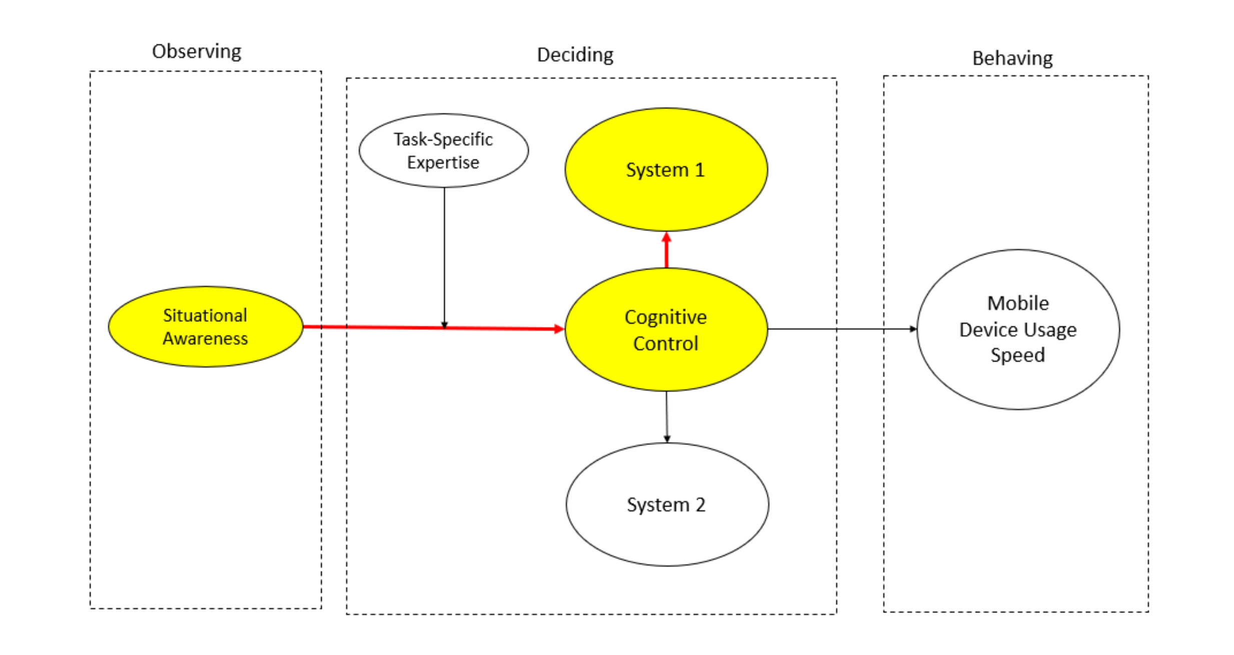


Figure 5.1 Application of DECODE framework to examine how situational awareness influences Cognitive Control and mobile device usage speed

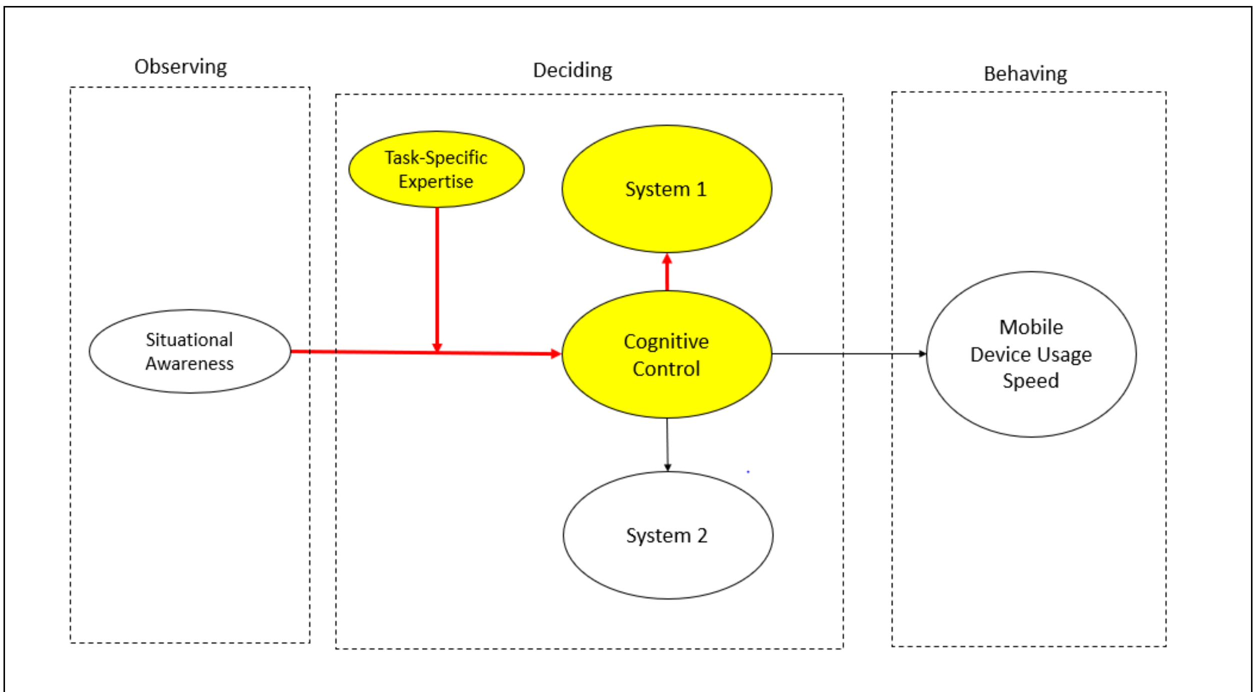


Figure 5.2 Application of DECODE framework to examine how task-specific influences Cognitive Control and mobile device usage speed

Given these backdrops, the following hypotheses are proposed:

H1: Removing the need to evaluate attention-demanding cues (e.g., lower CTR) will alter mobile device usage speed. Users may exhibit different usage patterns without the requirement to process these cues, potentially allowing for faster interactions as cognitive control demands decrease.

H2: Developing task-specific expertise (e.g., over the repeated execution of the task) will gradually improve the mobile device usage speed.

Methodology

Measure: Touch duration, speed, and time taken between rounds. As 'Essay Three' focuses on examining the effectiveness of the DECODE framework using mobile devices, only the measures directly associated with the speed of the task executions are examined. The JavaScript code embedded in the study captures the x-coordinates and y-coordinates of the touch positions at millisecond precision. The captured coordinates are then utilized to map both trajectories to calculate the touch distance per each target object (e.g., slider) with which users interacted.

Similar to the mouse movement distance, the distance between two points can be calculated using the following formula:

$$d(P_t, P_{t-1}) = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}$$

Here, 't' represents the timestamp in milliseconds. Similarly, the average speed between data points is derived using the following:

$$\bar{s} = \frac{\Delta d}{\Delta t}$$

Lastly, the time between the tasks is derived by calculating the differences between the last touch end event and the subsequent touch start events.

Study Design

Based on the learning from the prior essay and the review of the related work, repeated measures observational study was conducted, in which participants played a 20-round card drawing game. Each round consisted of drawing two cards, and there could be a maximum of one Joker card per round. On rounds where a Joker card was displayed, participants were instructed to self-report the game's outcome by sliding their answer to "Lose." In other cases, participants were instructed to slide their answer to "Win" (e.g., rounds without Joker cards). Each win is associated with a gain bonus of $0.05, while a loss results in a bonus deduction of $0.05. All participants had 10 rounds with a Joker card and 10 rounds without a Joker card. Hence, if participants honestly reported each round’s outcome, they would receive a bonus of $0.00 (aside from their base payment for participating in the experiment). There were six pre-selected decks, and each participant was randomly assigned to a deck.

The study design has several implications. First, the participants could choose to cheat by reporting wins when a Joker card was displayed. Realizing this possibility would require them to re-evaluate their approach in subsequent rounds, reflecting changes in situational awareness. Second, as the participants played 20 rounds of the game under the same rule, they were expected to develop task-specific expertise.

Participants

Participants were recruited from Amazon Mechanical Turk to complete the experiment. The base pay rate was set at $0.50 for an approximately 5-minutes long task, which equates to roughly $6 per hour. Bonus payments were also offered based on gameplay results, allowing the maximum payout for each participant to reach $1.50. On average, participants received an additional $0.24 in bonus payment, resulting in an approximate total hourly rate of $9 when combined with the base payment. To ensure data quality, only participants who had completed at least 100 tasks were recruited. Duplicate IP addresses were controlled for, and data collection was limited to mobile devices. The study included 503 valid participants, each of whom was randomly assigned to one of six separate card decks.

Results and Analysis.

Flagging Cheaters. Direct inspection of round outcomes overlooks the experiment's repetitive nature, providing limited insights into how CC acquires and applies sets of rules within the DECODE framework. A subject-level analysis is performed to explore how participants decide to report truthfully or decide to cheat. Participants are categorized as cheaters or truthful based on the number of rounds with mistakes and the ratio of cheating vs. remaining rounds. To better evaluate how CC activity manifests as changes in the device usage pattern, changes in the behavioral patterns of each group are reported as the rounds progress.

During the pre-processing stage, several steps are taken to refine the participant pool and flag the cheaters. The detailed steps for segmenting the groups and flagging the cheaters are as follows:

1. The total number of rounds, cheating count, mistake count, cheating rate, mistake rate, and the first round in which a participant cheated are calculated for each user.
2. For the variable indicating the first round in which a participant cheated, the value is replaced with 'NA,' which means that the participant did not cheat.
3. The cheating score is calculated by subtracting half of the 'mistake rate' from the 'cheating rate.' This allows us to consider instances in which participants may have made mistakes in figuring out ways to cheat (i.e., situational awareness).
4. A threshold for the cheating score is set (in this case, 0.275), and participants with a cheating score above this threshold are flagged as cheaters. Note that the threshold value can be changed to different values based on domain knowledge or empirical analysis.

The derived cheating score considers cheating and mistake rates, striking an appropriate balance. By incorporating only 50% of the mistake rate in the calculation, the possibility of errors in participants' decision-making processes is acknowledged. This approach recognizes that participants might only sometimes execute cheating strategies perfectly and accounts for both intentional and unintentional aspects of cheating and mistakes.

Furthermore, the cheating and mistake rates are evaluated using the threshold for the cheating score. Since there are 10 rounds with Joker cards and 10 rounds without, the maximum number of rounds for cheating or making mistakes is 10. Consequently, cheating and mistake rates are capped at 0.50 (i.e., 50% of the 20 rounds). As only 50% of the mistake rates are considered, the cheating score can be capped at 0.250 if participants played the game in the opposite direction (e.g., sliding to win when Joker appears and sliding to lose otherwise). By setting the cheating score threshold at 0.275, only participants who cheated more than the number of mistakes they made were flagged as cheaters. Of the 503 valid participants, 25 participants were identified as cheaters.

Examining H1 and H2 using Linear Mixed Effects Model. Linear mixed-effects models (LMMs) were used to examine H1 and H2. LMM was chosen because it accounts for within-subject correlations prevalent for within-subject, repeated-measure data (Magezi 2015). Further, LMM allows for the inclusion of both fixed and random effects and thus accounts for random variability (e.g., device-specific differences) in the data.

Before constructing the model, additional pre-process measures were taken to minimize the impact of external factors (e.g., device differences, age groups, external environment) that may influence the measures of interest. A simple min-max normalization was performed across the 20 rounds of card draw games for each participant. As a result, a round with the highest measure of interest (e.g., speed) will have a value of 1, while the round with the lowest measure of interest will have a value of 0.

H1 and H2 were examined using a single LMM model. This is appropriate as H1 and H2 can be assessed by examining the same dependent variable (e.g., Normalized Touch Speed, Normalized Transition Speed, and Normalized Touch Duration). The assessment of H1 involved creating an independent variable named "Cheated," which signals if the person was flagged as a cheater based on the derived cheating score. Since being aware of the cheating option removes the need to evaluate the two cards displayed each round, a cheater is anticipated to become faster in every aspect relating to H1.

Leveraging these two variables, the models to examine H1 and H2 were constructed:

$$Speed = \beta_0 + \beta_1 Cheated + \beta_2 Round\ Number + \beta_3\ Cheated * Round\ Number\ + u_{userId} + \epsilon$$

$$Transition = \beta_0 + \beta_1 Cheated + \beta_2 Round\ Number + \beta_3\ Cheated * Round\ Number + u_{userId} + \epsilon$$

$$Duration = \beta_0 + \beta_1 Cheated + \beta_2 Round\ Number + \beta_3\ Cheated * Round\ Number\ + u_{userId} + \epsilon$$

The interaction term between "Cheated" and "Round Number" allows for examining the impact of situational awareness on mobile device usage speed. A significant interaction effect indicates that changes in situational awareness (e.g., being aware of the cheating option) lead to differences in device usage speed. For H2, the main effect of Rounds Since Cheating captures the influence of task-specific expertise on mobile device usage speed.

The results for LMM (Tables 5.2, 5.3, and 5.4) demonstrated significant round variables for touch speed, touch duration, and transition time. This suggests that as the rounds progressed, both cheaters and non-cheaters experienced changes in their device usage speed. Consequently, H2 was supported. Plots of the average Speed, Transition Time, and Touch Duration over 20 rounds are available in Figures 5.3, 5.4, and 5.5. Specifically, the significance of the round variable in all three instances demonstrates that non-cheaters experienced (1) increased touch speeds, (2) decreased touch durations, and (3) reduced transition times between rounds.

Furthermore, the significant interaction effect between the Cheated and Round variables in the LMM results for speed, transition time, and touch duration indicates that changes in situational awareness (i.e., recognizing the possibility of cheating) impact mobile device usage speed for both cheaters and non-cheaters. Thus, H1 was supported. In particular, the findings revealed that compared to non-cheaters, cheaters experienced (1) a more rapid increase in touch speeds, (2) a more rapid decrease in touch durations, and (3) a more rapid reduction in transition times between rounds.

Lastly, the visual inspection of the cheater vs. non-cheater groups in Figures 5.2, 5.3, and 5.4 suggests clear differences between the cheater and the non-cheater group.

Table 5.2 LMM Result (Speed)						
Factor	**Estimate**	**Std. Error**	**Df**	**t-value**	**P-value**	**Significance**
Cheating (Yes)	-0.0577	0.0327	1541	-1.765	0.0778	
Round	0.0108	0.0004	9358	23.957	0.0000	****
Cheating (Yes):Round	0.0009	0.0021	9353	4.803	0.0000	****
Significance Code: 0 < ‘****’, 0.001 < ‘***’, 0.01 < ‘**’, 0.05 < ‘*’ 0.1 < ‘ ’						

Table 5.3 LMM Result (Transition)						
Factor	**Estimate**	**Std. Error**	**Df**	**t-value**	**P-value**	**Significance**
Cheating (Yes)	0.0177	0.0335	988.8	0.530	0.596	
Round	-0.0127	0.0004	9008	-35.151	0.0000	****
Cheating (Yes):Round	-0.0079	0.0016	9004	-4.827	0.0000	****
Significance Code: 0 < ‘****’, 0.001 < ‘***’, 0.01 < ‘**’, 0.05 < ‘*’ 0.1 < ‘ ’						

Table 5.4 LMM Result (Touch Duration)						
Factor	**Estimate**	**Std. Error**	**Df**	**t-value**	**P-value**	**Significance**
Cheating (Yes)	0.0363	0.0353	1090	1.029	0.3038	
Round	-0.0090	0.0004	9509	-21.340	0.0000	****
Cheating (Yes):Round	-0.0072	0.0019	9505	-3.754	0.0000	****
Significance Code: 0 < ‘****’, 0.001 < ‘***’, 0.01 < ‘**’, 0.05 < ‘*’ 0.1 < ‘ ’						

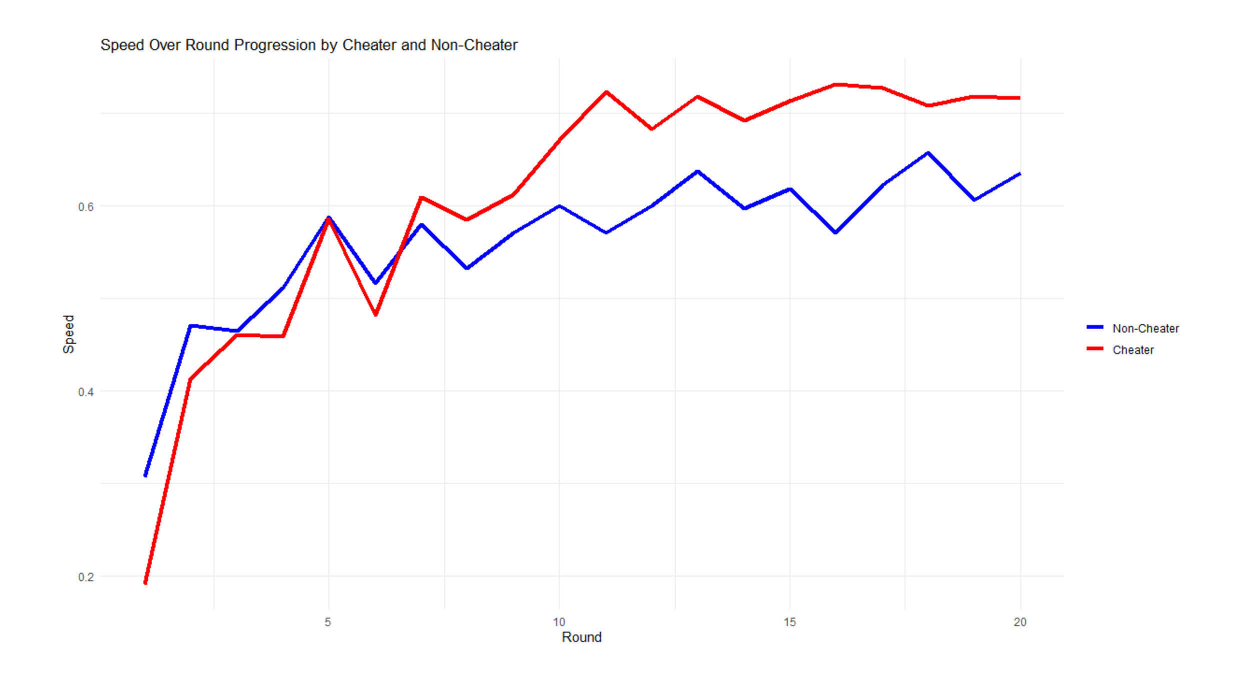


Figure 5.3 Average speed over 20 rounds for Cheater vs. Non-Cheaters

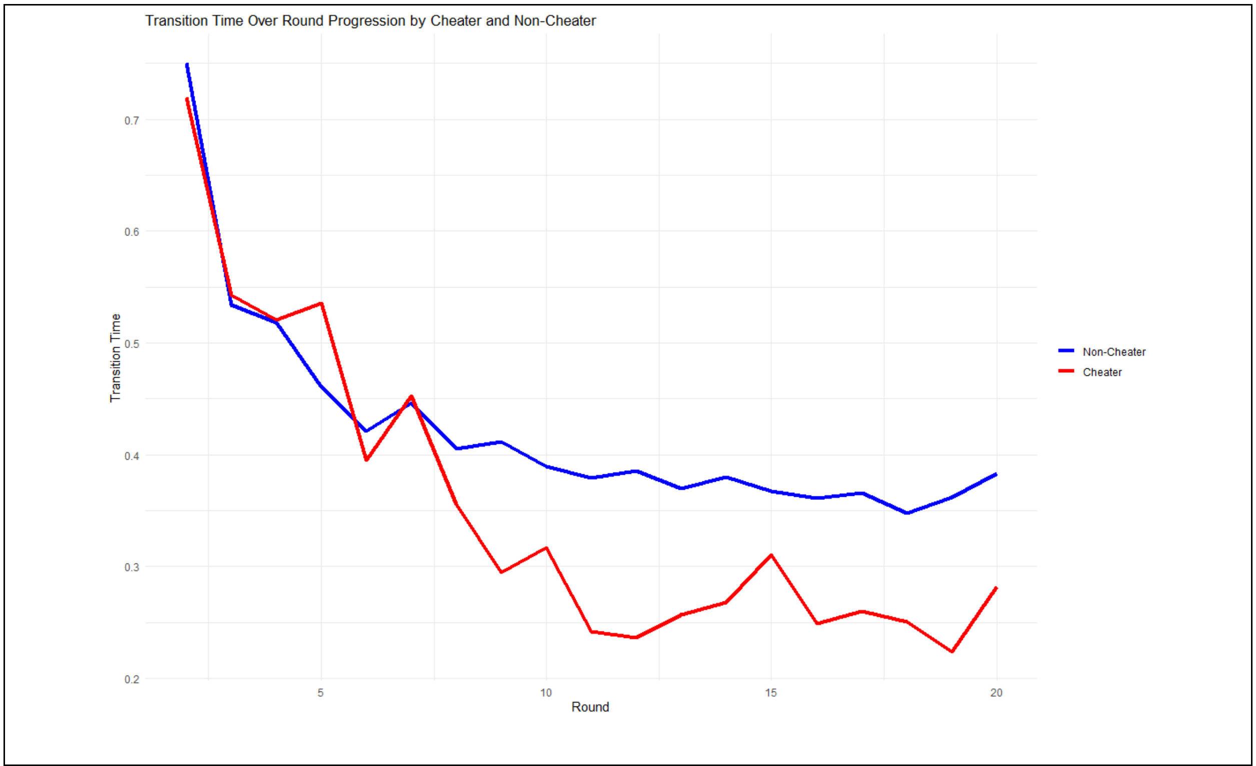


Figure 5.4 Average transition time over 20 rounds for Cheater vs. Non-Cheaters

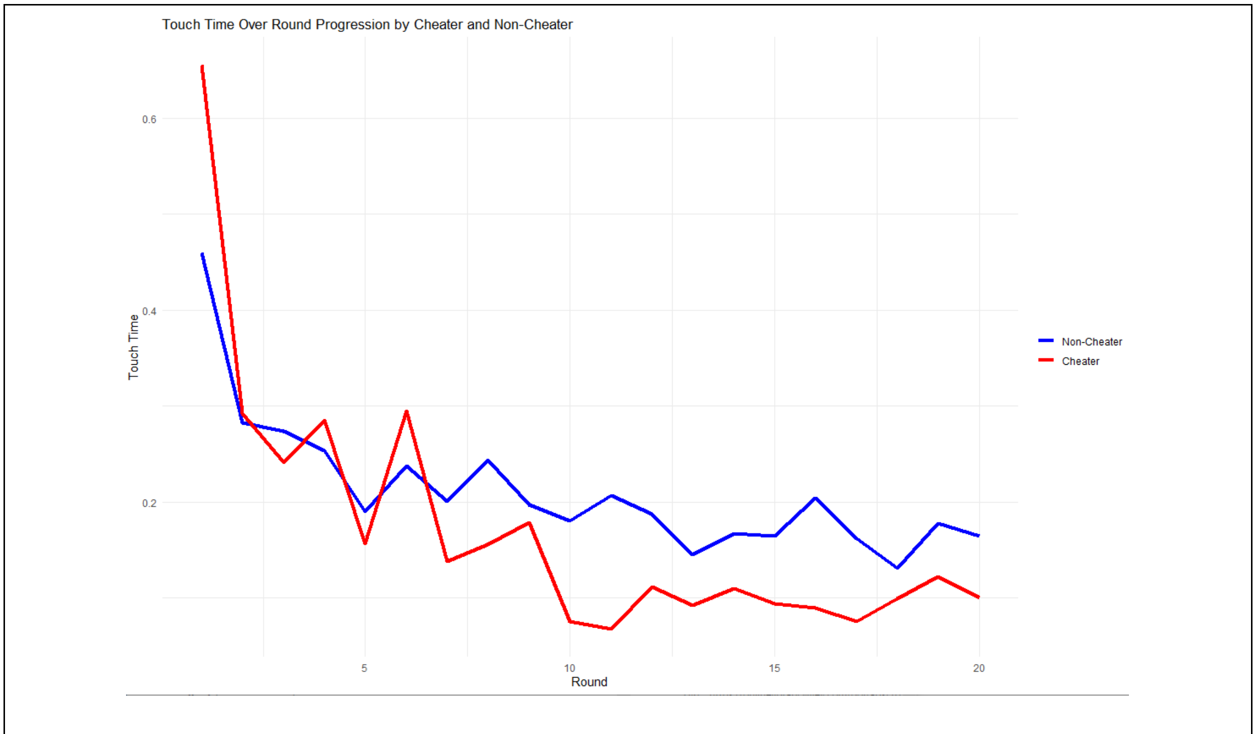


Figure 5.5 Average touch time over 20 rounds for Cheater vs. Non-Cheaters

Essay Three Summary

In 'Essay Three,' the DECODE framework is employed to examine how changes in external situations alter the CTR for the task and how such changes manifest as differences in mobile device usage speed. The study proposes two hypotheses: (1) removing the need to evaluate attention-demanding cues will alter mobile device usage speed, and (2) developing task-specific expertise will gradually improve mobile device usage speed. A repeated measures observational study is conducted in which participants play a 20-round card drawing game, allowing them to cheat by reporting wins when a Joker card is displayed. Based on how the participants played the game, they were classified as cheaters or non-cheaters.

Using a Linear Mixed Effects Model (LMM), the study found that as rounds progressed, both cheaters and non-cheaters experienced a significant increase in device usage speed, supporting Hypothesis 2. Moreover, the significant interaction effect between the Cheated and Round variables in the LMM results indicates that changes in the external situation (i.e., removing the need to evaluate the cards displayed) result in significantly faster device usage speed. This finding suggests that lowering the CTR increases the mobile device usage speed, supporting Hypothesis 1.

CHAPTER VI: DISCUSSION AND CONCLUSION

Overview of Findings and Contributions

The newly proposed Decision Evaluation with Cognitive Observations using the Device (DECODE) framework, a comprehensive and versatile model, has been thoroughly examined through three meticulously designed essays and tested across three different devices. In particular, the three essays aimed to investigate the impact of Cognitive Task Requirement (CTR) and task-specific expertise on device usage speed as proposed in the DECODE framework. By applying the DECODE framework and the usage behaviors across multiple devices, the essays provide support for the proposed propositions (in Chapter 2), research questions (in Chapter 2), and hypotheses (in Chapters 3,4, and 5). The overall results indicate that the increase in CTR slows down the device usage speed while obtaining task-specific expertise is exhibited by faster device usage speed.

The study makes substantial contributions to research. First, the findings from each essay extend the Dual Process Theory (DPT) and Hybrid views of Cognitive Control (CC) into an online space by revealing that high levels of CTR result in higher CC and thus increase the involvement of System-2 types of thinking. This results in slower device usage speed. Second, the inclusion of high task-specific expertise as a construct of the DECODE framework allows the examination of how automaticity over the task is developed and leads to faster device usage speed.

By proposing a theoretically sound framework and empirically validating it in multiple studies and devices, the researchers have access to new models and measurement tools that can effectively and objectively capture the users' underlying cognitive states. The DECODE framework can potentially be applied to various research disciplines. Specifically, the proposed measures for computer mice, keyboards, and mobile devices are helpful for various studies that examine individuals' underlying cognitive processes. For instance, the Information Systems (IS) studies examining the relationships between cognitive load and key user behaviors (e.g., Perceived Ease of Use, User Satisfaction, Reuse Intention) can use the proposed framework and measures to extract new insights from phenomena that have been previously investigated using self-reported measures.

Practical Implications

The DECODE framework also offers valuable practical implications for various industries. First, the real-time assessment of user behavior is a substantial challenge as user behaviors are influenced by various cognitive processes, which are difficult to capture, infer, and act upon outside highly artificial environments (e.g., data collection using fMRI, EEG caps, eye tracking, and so on). When the same techniques are used in a more natural environment, the constructed model can generate strange insights that may lead to poor management decisions. As the DECODE framework involves using devices widely adapted as measurement tools, they are adequate for capturing user behaviors (i.e., typing) in a more natural setting.

Second, by carefully examining the users' device usage speed, practitioners can pinpoint areas where users may exhibit high levels of CC and experience high cognitive load. This lays the groundwork for more effective design and user experience strategies. For instance, the potential application area of DECODE's framework is an assessment of mouse cursor movements as an indicator of cognitive load. This empowers practitioners to design real-time systems to gauge the level of CC and take immediate action:

1. Computer mice example: In e-commerce, mouse cursor speed (e.g., cursor slows down) can evaluate users' interest (i.e., engaging in attention-demanding cues) in specific products. Practitioners can drive additional sales by taking suitable action to engage the users.
2. Computer Keyboard example: Entering information on a form is a typical online task executed regularly. Some examples include but are not limited to instances where users sign up for a website, apply for a loan application, and fill out the identity information for public institutions. By looking into the typing dynamics, practitioners can identify different points within the form where the user may be having difficulties (i.e., putting more reliance on the Type 2 process) while filling out a form online. Experiments three and four of 'Essay Two' show that such insights can be used to identify fraudsters.
3. Mobile device example: As shown in 'Essay Three', touch input patterns can be employed to assess users' interest (i.e., engaging in attention-demanding cues) in various app features or content. By engaging users appropriately, practitioners can enhance user satisfaction and increase app usage. For example, when a user is scrolling down on a website on their mobile phone to find relevant items or products, a slow scrolling speed or prolonged taps on certain sections may indicate increased cognitive load and focused attention. Leveraging this information, practitioners can personalize the content, recommendations, or promotions to better cater to the users' needs and preferences.

Limitations and Future Research

Like all research, this work has limitations. The robustness of DECODE has been demonstrated using three commonly used devices; however, it is still uncertain if the results would generalize to other commonly used devices (e.g., tablet devices). Additionally, users may multi-task while on the internet, potentially influencing measures such as field transitions and generating biased results.

'Essay One' only examined mouse cursor speeds in the context of an online survey. As these surveys consist of simple, goal-oriented tasks, it is necessary to explore whether the results would generalize in broader contexts (e.g., web browsing, system navigation, online shopping). Second, while inducing social desirability is a treatment mechanism that directly influences CC, many other factors can influence CC. Hence, it is essential to investigate these potential influences in detail to see if the obtained results can be replicated.

For 'Essay Two', while the methods could be implemented in any instance where a user enters information on a form, the results might not generalize well to typing instances where the information is less structured (e.g., academic writing). Lastly, given the limited experimental material (e.g., using only one form and examining the DECODE framework with a single device), further research should investigate whether the results concerning design-specific factors (e.g., form designs) are robust. To address these concerns, future research should explore how DECODE can be applied to a broader range of contexts (i.e., device types and users' browsing behaviors).

For 'Essay Three', the dataset is highly imbalanced, as participants were not forced to cheat, and strict rules were applied to classify someone as a cheater. While the estimates represent the behavioral differences between cheaters and honest participants, constructing a machine-learning model without tuning is unsuitable due to the low number of cheaters. Moreover, the results of predictive models are not reported, as the research primarily focuses on applying the DECODE framework to study how task-specific expertise and changes in CTR influence device usage behaviors. Although the experimental design is similar to previous studies examining CC using fMRI machines, fMRI was not employed to ensure participants executed tasks or decided to cheat in natural settings. As a result, the findings should be interpreted cautiously, as they may not generalize well beyond the repeated measures design. In summary, it is crucial to design future studies investigating how changes in CC manifest as differences in device usage behaviors across various circumstances.

Lastly, in addition to the limitations of CC and DPT-based frameworks, there are inherent limitations to the theories themselves. CC and DPT are still actively studied theories and constructs. For example, the Hybrid Dual-Process model, a recent development in the field of psychology, extends the DI model and suggests that a response generated under the influence of System 1 undergoes two distinct intuitive reasoning processes, resulting in (1) a traditional heuristic intuitive response, and (2) a logical, intuitive response (Bago and De Neys 2017, 2020). As DPT continues to be studied, future research should investigate how new developments in this area will impact the DECODE framework. Regarding CC, although the latest developments in the theoretical framework (e.g., Mesh) offer a more sophisticated view of the task execution process, empirical validation remains a challenge. Consequently, there is ongoing research in neuropsychology primarily based on empirical findings from fMRI data. Generalizing these results using the DECODE framework remains a challenge. Overall, future research should continue to evolve and refine the DECODE framework in response to any significant findings that may substantially influence the current theoretical model.

Conclusion

This pioneering research endeavor aimed to explore how integrating the well-established Dual-Process Theory (DPT) and the widely recognized Construct of Cognitive Complexity (CC) can result in a versatile framework adaptable to various online contexts, wherein CTR and task-specific expertise play crucial roles. By consistently demonstrating robust results across a diverse range of devices, this dissertation makes a substantial and noteworthy contribution to the field of Information Systems, setting new standards for future research in this domain. The DECODE framework offers a solid foundation for further exploration and refinement, paving the way for an improved understanding of online behaviors and interactions in the ever-evolving digital landscape.

APPENDIX A: ESSAY 1 INSTRUMENTS

This section presents the complete survey instrument for 'Essay One.'

Thank you for participating in our research. We are interested in learning about how people with different backgrounds answer personality assessment questions. No matter how you answer them, you will still receive the worker reward for completing the survey.

→

What is your current age?

- Under 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 +

What gender do you identify most with?

- Male
- Female

What is the highest degree or level of education you have completed?

- Less than high school diploma
- High school diploma or equivalent degree
- Some college but no degree
- Bachelor’s degree
- Master’s degree
- Doctorate degree

What is your marital status?

- Single, never married
- Married or domestic partnership
- Widowed
- Divorced
- Separated

What is your first language?

- English
- Spanish
- Chinese
- Arabic
- Hindi
- Other

Here are a number of characteristics that may or may not apply to you.

Please answer each question to indicate the extent to which you agree or disagree with that statement.

I am someone who pays attention to this study. Please select 'Agree a little' as the answer to this question.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who enjoys hearing new ideas.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who makes plans and stick to them.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who feels comfortable around people.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who dislikes myself.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who pays attention to details.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who finds it difficult to get down to work.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who believes in the importance of art.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who does not enjoy going to art museums.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who does just enough work to get by.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who is often down in the dumps.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who knows how to captivate people.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who believes that others have good intentions.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who is not easily bothered by things.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who often feels blue.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who cut others to pieces.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who carries out my plans.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who doesn't like to draw attention to myself.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who makes friends easily.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who has a sharp tongue.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who has a good word for everyone.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who panics easily.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who is skilled in handling social situations.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who carries the conversation to a higher level.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who shirks my duties.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who get chores done right away.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who rarely gets irritated.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who has a little to say.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who is the life of the party.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who has a vivid imagination.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who has frequent mood swings.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who respects others.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who suspects hidden motives in others.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who tends to vote for liberal political candidates.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who avoid philosophical discussions.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who is always prepared.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who accepts people as they are.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

I am someone who makes people feel at ease.

Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree Strongly

You have finished the first part of the survey. If you would like to stop now, please select **"End the Survey"** and receive $0.50.

If you would like to continue and do the second part of the survey for a total of $1.00, please select **"Continue with the Second Part."** You will receive a completion code at the end of the second part that will pay the full $1.00. The second part of this survey is an application for a follow-on study that will pay $10 for 10 minutes of work. We will only select a few people to do this follow-on study. If we select you, you can refuse to participate; there is no obligation to do it. We will pay you an extra $0.50 to complete the application, regardless of whether we select you or whether you agree to do the follow-on study if selected. The follow-on study will have you use Excel to do some data analysis, so we are looking for people who have experience with Excel. Experience with Excel's math and statistical functions is a plus but is not required.

End the Survey	Continue with the Second Part

→

Please answer the following questions about your background:

On average, how many hours do you work a week?

- Less than 35 hours per week
- Greater than 35 hours per week
- I am currently not employed
- Prefer not to answer

Have you ever been terminated, laid off, or asked to leave a position?

- Yes
- No

Have you ever been convicted of a crime?

- Yes
- No

Have you ever filed for bankruptcy?

Yes

No

Is your credit rating good?

Yes

No

How would you describe yourself?

Rate your proficiency with computers.

Beginner					Intermediate					Expert
0	1	2	3	4	5	6	7	8	9	10

What is your level of communication skills -- oral and written?

Beginner					Intermediate					Expert
0	1	2	3	4	5	6	7	8	9	10

Do you normally develop plans and follow them?

Never					Sometimes					Always
0	1	2	3	4	5	6	7	8	9	10

It is better to follow the rules and take longer than to save time by breaking the rules.

Never					Sometimes					Always
0	1	2	3	4	5	6	7	8	9	10

I am comfortable working under tight time deadlines.

Never					Sometimes					Always
0	1	2	3	4	5	6	7	8	9	10

I prefer to work by myself rather than work in teams.

Never					Sometimes					Always
0	1	2	3	4	5	6	7	8	9	10

Is it better to be the center of attention than to be overlooked.

Never					Sometimes					Always
0	1	2	3	4	5	6	7	8	9	10

How easy is it to get along with people where you currently work?

Very Easy					Possible					Impossible
0	1	2	3	4	5	6	7	8	9	10

Have you ever had difficulty working with a manager?

Never					Sometimes					Always
0	1	2	3	4	5	6	7	8	9	10

Team harmony is more important than my personal success.

Never					Sometimes					Always
0	1	2	3	4	5	6	7	8	9	10

What would your current or most recent supervisor say about you?

Rate your experience with Microsoft Excel.

Beginner					Intermediate					Expert
0	1	2	3	4	5	6	7	8	9	10

Rate your level of experience using the StatView plugin for Excel?

Beginner					Intermediate					Expert
0	1	2	3	4	5	6	7	8	9	10

Describe a specific example when you used Microsoft Excel to create a report. What type of report did you create and why was it needed? This example can include reports used for classroom, work, or personal reasons.

APPENDIX B: ESSAY 2 INSTRUMENTS

This section presents the experimental instrument for 'Essay Two.'

Experiment 1:

Please type **your first and last name** in the text box below.

→

Please type the following name in the text box below: **David Russell**

→

Please type the following text in the text box below: **ceterus paribus georginorium**

→

Experiment 2:

What gender do you identify most with?

- ○ Male
- ○ Female
- ○ Non-binary / third gender
- ○ Prefer not to say

What is your current age?

- ○ Under 18
- ○ 18 - 24
- ○ 25 - 34
- ○ 35 - 44
- ○ 45 - 54
- ○ 55 +
- ○ Prefer not to answer

What is your first language?

- ○ English
- ○ Spanish
- ○ Chinese
- ○ Arabic
- ○ Hindi
- ○ Other
- ○ Prefer not to answer

Next, you will be asked to type a **Non-English name** multiple times.

<u>You must complete the entire study in one session or you will NOT receive credit.</u>

Note: Please do **<u>NOT</u>** copy and paste**.** You will receive 0 credit.

Please type **Guda Hakim** in the text box below:

→

Eller

Next, you will be asked to enter an **English name** multiple times.

<u>You must complete the entire study in one session or you will NOT receive credit.</u>

→

Eller

Note: Please do **<u>NOT</u>** copy and paste**.** You will receive 0 credit.

Please type **Lynnette Anderson** in the text box below:

→

Next, you will be asked to type **your name** multiple times.

<u>You must complete the study in one session or you will NOT receive credit.</u>

→

Note: Please do **<u>NOT</u>** use autofill. You will get 0 credit.

Please type **your full name**:

→

Experiment 3:

Use your personal information to complete the form below:

Please do **NOT** use autofill.

First Name:

Last Name:

Home City:

State (Spell out):

Zip Code:

Phone (last 4 digits):

Primary Email Address:

Use this information to complete the form below:

First Name: John
Last Name: Wallace
Home city: Augusta
State (spell out): Nevada
Zip Code: 86789
Mobile Number: (899)-749-2134
Email address: john170@yahoo.com

First Name:	John
Last Name:	Wallace
Home City:	Augusta
State (Spell out):	Nevada
Zip Code:	86789
Phone (last 4 digits):	2134
Primary Email Address:	john170@yahoo.com

Experiment 4:

During this study, you will be asked to enter information on a standard web form over several pages. Please follow the directions carefully.

This task should take less than 10 minutes to complete. **You must complete all parts of the study in one session.**

After you have complete all portions of the study, you will be given a field where you can enter your Net ID.

Next, you will enter some **personal (permanent) identity** information.

The entered information **will NOT be saved** and will only be kept long enough to verify the data entry accuracy.

You must complete all parts of the study in one session.

→

Enter your personal information in the fields below.

First Name:

Last Name:

Date of Birth (MM/DD/YYYY):

MM/DD/YYYY

Zip code (Permanent Address):

→

To earn **full credit**, you **must correctly answer** the following questions:

Which **city** is associated with the following zipcode:

Which **state** is associated with the following zipcode:

(Please type the full name of the state.)

What is the age of a person born on ?

→

Next, you will enter some **imposter identity** information.

You must complete all parts of the study in one session.

→

Eller

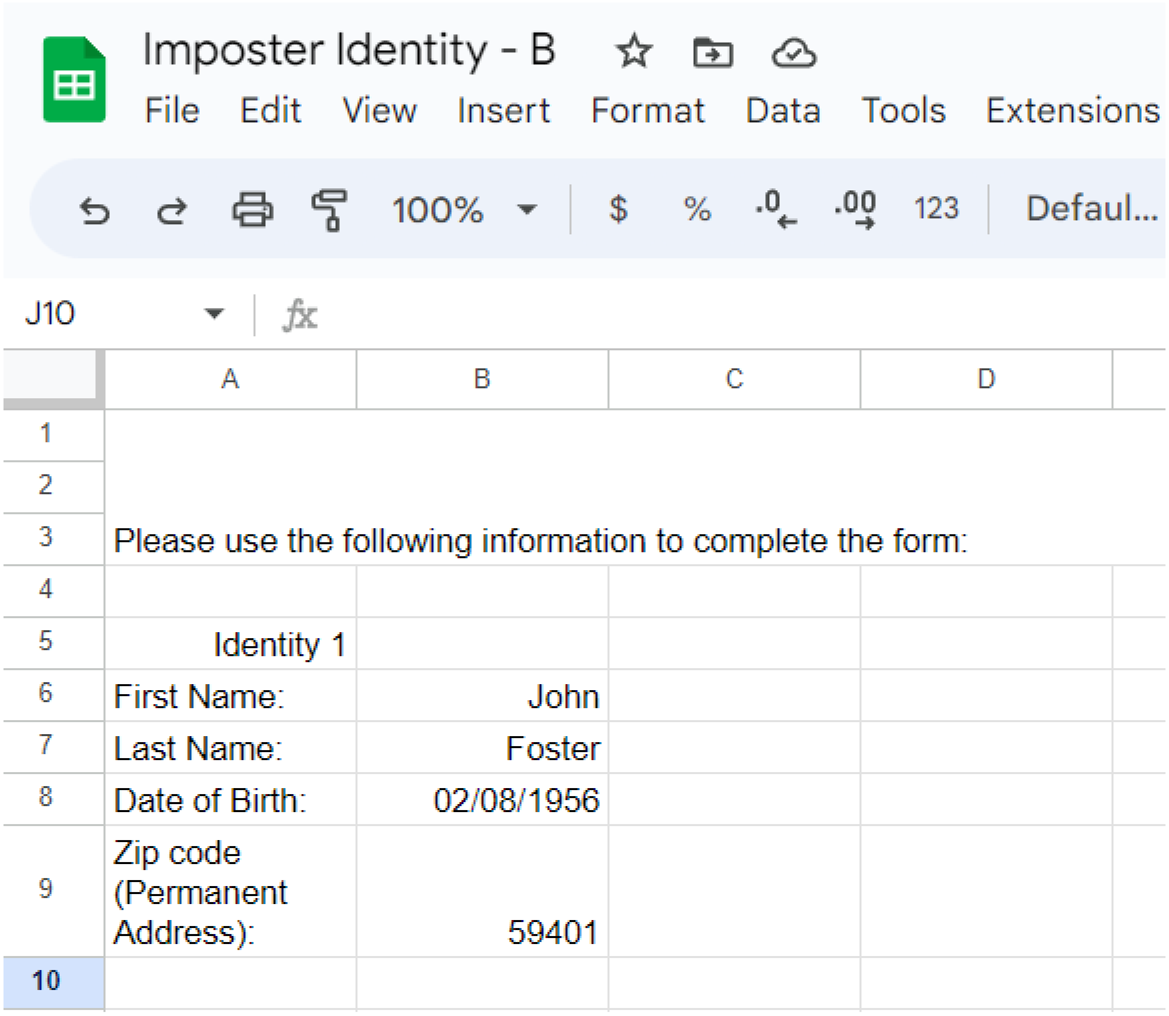
Enter imposter information in the fields below.

Click the following link to find the imposter information.

Link to Imposter Identity Info

First Name:	
Last Name:	
Date of Birth (MM/DD/YYYY):	MM/DD/YYYY
Zip code (Permanent Address):	

→



Imposter Identity - B

File Edit View Insert Format Data Tools Extensions

100% $ % .0 .00 123 Defaul...

J10

	A	B	C	D
1				
2				
3	Please use the following information to complete the form:			
4				
5	Identity 1			
6	First Name:	John		
7	Last Name:	Foster		
8	Date of Birth:	02/08/1956		
9	Zip code (Permanent Address):	59401		
10				

APPENDIX C: ESSAY 3 INSTRUMENTS

This section presents the complete instrument for an observational study in 'Essay Three.'

Thank you for participating in this study!
This study must be taken from a mobile device.
Please click the link directly from your mobile device
Thanks!

→

You are using Safari or Mobile Safari. As mentioned in the Task HIT, this task must be performed in the Chrome browser. If you want to continue with this task, please reopen the link on your mobile device in the Chrome browser.

Please click the → button where you will be redirected to the end of the survey.

Thank you.

You are using a Windows/Microsoft phone or the Edge browser. As mentioned in the Task HIT, this task must be performed on a non-Windows/Microsoft phone in the Chrome browser. If you want to continue with this task, please reopen the link on a non-Windows/Microsoft mobile device in the Chrome browser.

Please click the → button where you will be redirected to the end of the survey.

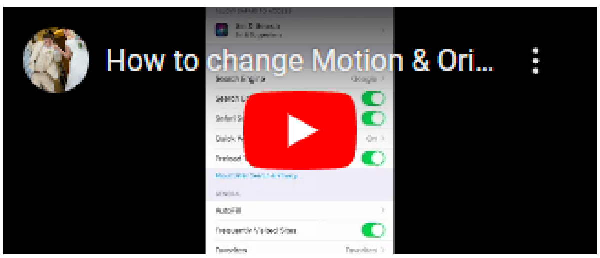
Thank you.

We are collecting device motion and orientation data for this study.

It appears that your device does not provide this data. You are ineligible for this study. You will now be redirected to the end of the survey.

Thank you.

We are collecting device motion and orientation data for this study.
<u>Please make sure that motion and orientation access is turned on for your device</u>.
You can navigate to Settings → Safari to turn on the access. Once the motion and orientation access is turned on, please reload this page and continue forward.



Please click here to view the consent form for this study. When you are done, please close the tab/window to return to this page and click Next (>) to continue.

Acknowledging the consent form:
By taking part in this study, you are allowing your responses to be used for research purposes.

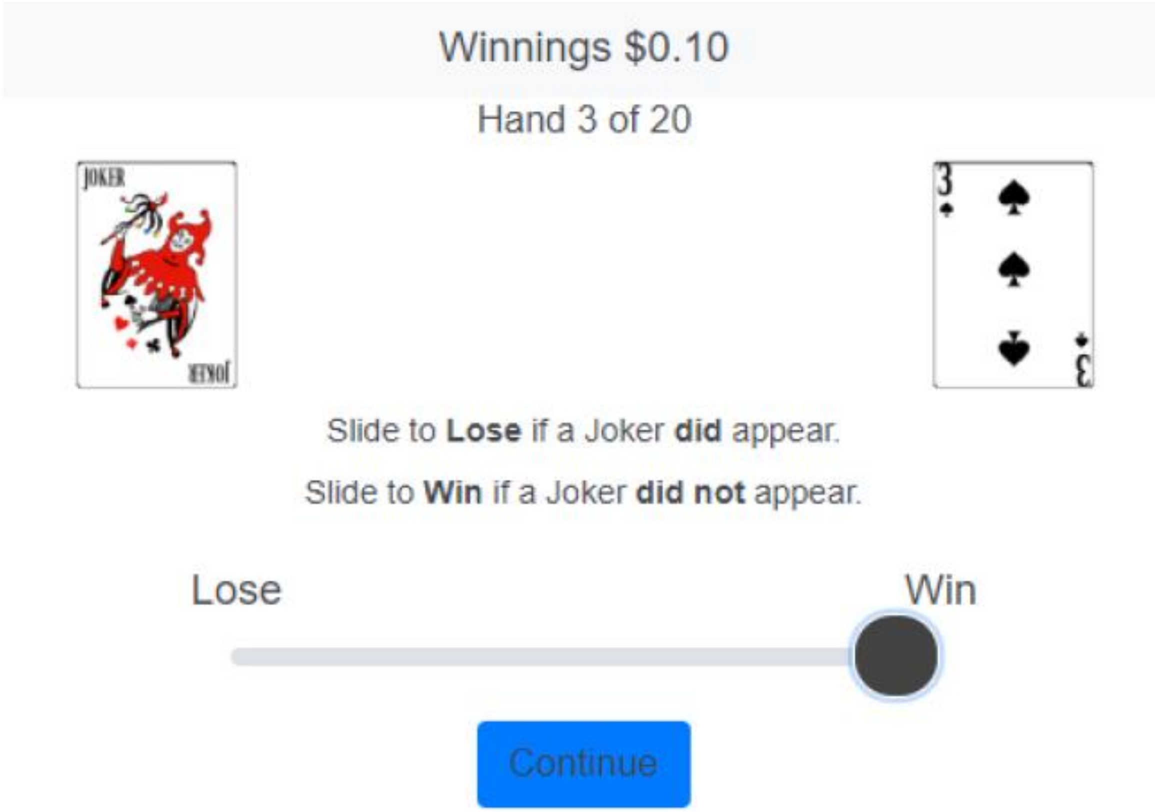
Thank you for participating in our research. We are interested in learning about how people with different backgrounds like playing a game called "Joker." No matter how well you play, you will still receive the worker reward for completing the survey. **Depending upon how many rounds you win, you will receive additional payout as a bonus!**

Joker uses a deck of 120 playing cards. Cards can be Joker, or be one of four possible suits (Hearts, Diamonds, Spades or Clubs) with a value that ranges from 1 (Ace) to 9 (numbers) or be one of 3 face cards (Jack, Queen or King).

Rules: You begin the game with $0.00 and will play 20 rounds; each round is worth $0.05. When you win, **$0.05 will be added** to your winning balance. When you lose, **you lose $0.05**. Remember, by participating, you will get at least the Worker Reward for completing the survey.

To Play: For each round, two cards are randomly selected and simultaneously displayed. If a Joker appears, select "Lose." If a Joker doesn't appear, select "Win".

Good luck!


Winnings $0.10
Hand 3 of 20
JOKER
3
Slide to Lose if a Joker did appear.
Slide to Win if a Joker did not appear.
Lose
Win
Continue

REFERNCES

Alzoubi, H., Alshurideh, M., Kurdi, B., Alhyasat, K., and Ghazal, T. 2022. "The Effect of E-Payment and Online Shopping on Sales Growth: Evidence from Banking Industry," *International Journal of Data and Network Science* (6:4), pp. 1369–1380.

Bach, R. L., Kern, C., Amaya, A., Keusch, F., Kreuter, F., Hecht, J., and Heinemann, J. 2021. "Predicting Voting Behavior Using Digital Trace Data," *Social Science Computer Review* (39:5), SAGE Publications Inc, pp. 862–883. (https://doi.org/10.1177/0894439319882896).

Bago, B., and De Neys, W. 2017. "Fast Logic?: Examining the Time Course Assumption of Dual Process Theory," *Cognition* (158), Elsevier, pp. 90–109.

Bago, B., and De Neys, W. 2019. "The Intuitive Greater Good: Testing the Corrective Dual Process Model of Moral Cognition.," *Journal of Experimental Psychology: General* (148:10), American Psychological Association, p. 1782.

Bago, B., and De Neys, W. 2020. "Advancing the Specification of Dual Process Models of Higher Cognition: A Critical Test of the Hybrid Model View," *Thinking & Reasoning* (26:1), Taylor & Francis, pp. 1–30.

Bang, H., and Wojdynski, B. W. 2016. "Tracking Users' Visual Attention and Responses to Personalized Advertising Based on Task Cognitive Demand," *Computers in Human Behavior* (55), Elsevier, pp. 867–876.

Banholzer, N., Feuerriegel, S., Fleisch, E., Bauer, G. F., and Kowatsch, T. 2021. "Computer Mouse Movements as an Indicator of Work Stress: Longitudinal Observational Field Study," *Journal of Medical Internet Research* (23:4), JMIR Publications Toronto, Canada, p. e27121.

BBVA. 2020. "Susan Fiske and Shelley Taylor Recognized with a Frontiers of Knowledge Award for Revealing the Role of Cognitive Bias in Social Relations," *NEWS BBVA*, , April 16. (https://www.bbva.com/en/susan-fiske-and-shelley-taylor-recognized-with-a-frontiers-of-knowledge-award-for-revealing-the-role-of-cognitive-bias-in-social-relations/, accessed April 18, 2023).

Boag, R. J., Stevenson, N., van Dooren, R., Trutti, A. C., Sjoerds, Z., and Forstmann, B. U. 2021. "Cognitive Control of Working Memory: A Model-Based Approach," *Brain Sciences* (11:6), MDPI, p. 721.

Botvinick, M., and Braver, T. 2015. "Motivation and Cognitive Control: From Behavior to Neural Mechanism," *Annual Review of Psychology* (66), Annual Reviews, pp. 83–113.

Braver, T. S. 2012. "The Variable Nature of Cognitive Control: A Dual Mechanisms Framework," *Trends in Cognitive Sciences* (16:2), Elsevier, pp. 106–113.

Braver, T. S., Gray, J. R., Burgess, G. C., and others. 2007. "Explaining the Many Varieties of Working Memory Variation: Dual Mechanisms of Cognitive Control," *Variation in Working Memory* (75), p. 106.

Buschman, T. J., Siegel, M., Roy, J. E., and Miller, E. K. 2011. "Neural Substrates of Cognitive Capacity Limitations," *Proceedings of the National Academy of Sciences* (108:27), National Acad Sciences, pp. 11252–11255.

Cabañero, L., Hervás, R., González, I., Fontecha, J., Mondéjar, T., and Bravo, J. 2020. "Characterisation of Mobile-Device Tasks by Their Associated Cognitive Load through EEG Data Processing," *Future Generation Computer Systems* (113), pp. 380–390. (https://doi.org/10.1016/j.future.2020.07.013).

Cahn, A., Alfeld, S., Barford, P., and Muthukrishnan, S. 2016. "An Empirical Study of Web Cookies," in *Proceedings of the 25th International Conference on World Wide Web*, pp. 891–901.

Chevalier, N. 2018. "Willing to Think Hard? The Subjective Value of Cognitive Effort in Children," *Child Development* (89:4), Wiley Online Library, pp. 1283–1295.

Chevalier, N., Dauvier, B., and Blaye, A. 2018. "From Prioritizing Objects to Prioritizing Cues: A Developmental Shift for Cognitive Control," *Developmental Science* (21:2), Wiley Online Library, p. e12534.

Christensen, W., Sutton, J., and McIlwain, D. J. F. 2016. "Cognition in Skilled Action: Meshed Control and the Varieties of Skill Experience: Cognition in Skilled Action," *Mind & Language* (31:1), pp. 37–66. (https://doi.org/10.1111/mila.12094).

Clark, S. 2021. "Drowning in Data? How To Ensure Your Data Strategy Isn't Hurting Your Brand?," *CMSWire.Com*, , March 15. (https://www.cmswire.com/digital-experience/drowning-in-data-is-your-data-strategy-helping-or-hurting-your-brand/, accessed April 18, 2023).

Cockburn, A., Gutwin, C., Scarr, J., and Malacria, S. 2014. "Supporting Novice to Expert Transitions in User Interfaces," *ACM Computing Surveys (CSUR)* (47:2), ACM New York, NY, USA, pp. 1–36.

Courtney, S. M. 2004. "Attention and Cognitive Control as Emergent Properties of Information Representation in Working Memory," *Cognitive, Affective, & Behavioral Neuroscience* (4:4), Springer, pp. 501–516.

Crameri, K.-A., Maher, L., Van Dam, P., and Prior, S. 2022. "Personal Electronic Healthcare Records: What Influences Consumers to Engage with Their Clinical Data Online? A Literature Review," *Health Information Management Journal* (51:1), SAGE Publications Sage UK: London, England, pp. 3–12.

De Neys, W. 2017a. *Dual Process Theory 2.0*, Routledge.

De Neys, W. 2017b. “Bias, Conflict, and Fast Logic: Towards a Hybrid Dual Process Future?,” in *Dual Process Theory 2.0*, Routledge, pp. 47–65.

De Neys, W., and Pennycook, G. 2019. “Logic, Fast and Slow: Advances in Dual-Process Theorizing,” *Current Directions in Psychological Science* (28:5), Sage Publications Sage CA: Los Angeles, CA, pp. 503–509.

Ebaid, D., Crewther, S. G., MacCalman, K., Brown, A., and Crewther, D. P. 2017. “Cognitive Processing Speed across the Lifespan: Beyond the Influence of Motor Speed,” *Frontiers in Aging Neuroscience* (9), Frontiers Media SA, p. 62.

Edwards, W. H. 2010. *Motor Learning and Control: From Theory to Practice*, Cengage Learning.

Eisenberg, N. 2017. “Commentary: What’s in a Word (or Words)–on the Relations among Self-Regulation, Self-Control, Executive Functioning, Effortful Control, Cognitive Control, Impulsivity, Risk-Taking, and Inhibition for Developmental Psychopathology–Reflections on Nigg (2017),” *Journal of Child Psychology and Psychiatry* (58:4), Wiley Online Library, pp. 384–386.

Endsley, M. R. 1995. “Toward a Theory of Situation Awareness in Dynamic Systems,” *Human Factors* (37:1), SAGE Publications Sage CA: Los Angeles, CA, pp. 32–64.

Engle, R. W., and Kane, M. J. 2004. *Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control.*, Elsevier Science.

Epp, C., Lippold, M., and Mandryk, R. L. 2011. “Identifying Emotional States Using Keystroke Dynamics,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Vancouver BC Canada: ACM, May 7, pp. 715–724. (https://doi.org/10.1145/1978942.1979046).

Evans, J. S. B. 2018. *Dual Process Theory: Perspectives and Problems.*, Routledge/Taylor & Francis Group.

Evans, J. S. B., and Stanovich, K. E. 2013. “Dual-Process Theories of Higher Cognition: Advancing the Debate,” *Perspectives on Psychological Science* (8:3), Sage Publications Sage CA: Los Angeles, CA, pp. 223–241.

Evans, J. S. B. T. 2008. “Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition.,” *Annual Review of Psychology* (59), United States, pp. 255–278. (https://doi.org/10.1146/annurev.psych.59.103006.093629).

Eysenck, M. W., Derakshan, N., Santos, R., and Calvo, M. G. 2007. “Anxiety and Cognitive Performance: Attentional Control Theory.,” *Emotion* (7:2), American Psychological Association, p. 336.

Felstead, A. 2022. *Remote Working: A Research Overview*, Routledge.

Fernandez, A., Insfran, E., and Abrahão, S. 2011. "Usability Evaluation Methods for the Web: A Systematic Mapping Study," *Information and Software Technology* (53:8), Elsevier, pp. 789–817.

Field, A. 2013. *Discovering Statistics Using IBM SPSS Statistics*, sage.

Fitts, P. M. 1966. "Cognitive Aspects of Information Processing: III. Set for Speed versus Accuracy.," *Journal of Experimental Psychology* (71:6), American Psychological Association, p. 849.

Freelon, D. 2014. "On the Interpretation of Digital Trace Data in Communication and Social Computing Research," *Journal of Broadcasting & Electronic Media* (58:1), Taylor & Francis, pp. 59–75.

Freeman, J. B., and Ambady, N. 2010. "MouseTracker: Software for Studying Real-Time Mental Processing Using a Computer Mouse-Tracking Method," *Behavior Research Methods* (42:1), pp. 226–241. (https://doi.org/10.3758/BRM.42.1.226).

Gidlöf, K., Wallin, A., Dewhurst, R., and Holmqvist, K. 2013. "Using Eye Tracking to Trace a Cognitive Process: Gaze Behaviour During Decision Making in a Natural Environment," *Journal of Eye Movement Research* (6:1). (https://doi.org/10.16910/jemr.6.1.3).

Girden, E. R. 1992. *ANOVA: Repeated Measures*, sage.

Gliner, J. A. 1985. "Purposeful Activity in Motor Learning Theory: An Event Approach to Motor Skill Acquisition," *The American Journal of Occupational Therapy* (39:1), The American Occupational Therapy Association, Inc., pp. 28–34.

Granić, A. 2022. "Educational Technology Adoption: A Systematic Review," *Education and Information Technologies* (27:7), Springer, pp. 9725–9744.

Greene, J. D. 2023. "The Dual-Process Theory of Moral Judgment Does Not Deny That People Can Make Compromise Judgments," *Proceedings of the National Academy of Sciences* (120:6), National Acad Sciences, p. e2220396120.

Griffith, S. 2013. *The Jazz Practice Kinesthetic: Motor Learning Theory Applied to Jazz Brass Practice*, Sam Griffith.

Grimes, G. M., and Valacich, J. S. 2015. "Mind Over Mouse: The Effect of Cognitive Load on Mouse Movement Behavior," in *Proceedings of the International Conference on Information Systems - Exploring the Information Frontier, ICIS 2015, Fort Worth, Texas, USA, December 13-16, 2015*, T. A. Carte, A. Heinzl, and C. Urquhart (eds.), Association for Information Systems. (http://aisel.aisnet.org/icis2015/proceedings/HCI/9).

Heinonen, K. 2011. "Consumer Activity in Social Media: Managerial Approaches to Consumers' Social Media Behavior," *Journal of Consumer Behaviour* (10:6), Wiley Online Library, pp. 356–364.

Hibbeln, M., Jenkins, J. L., Schneider, C., Valacich, J. S., and Weinmann, M. 2014. "Investigating the Effect of Insurance Fraud on Mouse Usage in Human-Computer Interactions," in *35th International Conference on Information Systems (ICIS 2014): Building a Better World Through Information Systems*, Association for Information Systems.

Hibbeln, M., Jenkins, J. L., Schneider, C., Valacich, J. S., and Weinmann, M. 2017. "How Is Your User Feeling? Inferring Emotion Through Human-Computer Interaction Devices," *MIS Quarterly* (41:1), pp. 1–21. (https://doi.org/10.25300/MISQ/2017/41.1.01).

Holleman, G. A., Hooge, I. T., Kemner, C., and Hessels, R. S. 2020. "The 'Real-World Approach'and Its Problems: A Critique of the Term Ecological Validity," *Frontiers in Psychology* (11), Frontiers Media SA, p. 721.

Hsieh, Y.-H., Chen, K.-J., Wang, C.-C., and Lai, C.-L. 2008. "Cognitive and Motor Components of Response Speed in the Stroop Test in Parkinson's Disease Patients," *The Kaohsiung Journal of Medical Sciences* (24:4), Wiley Online Library, pp. 197–203.

Jenkins, J. L., Grimes, M., Proudfoot, J. G., and Lowry, P. B. 2014. "Improving Password Cybersecurity through Inexpensive and Minimally Invasive Means: Detecting and Deterring Password Reuse through Keystroke-Dynamics Monitoring and Just-in-Time Fear Appeals," *Information Technology for Development* (20:2), Taylor & Francis, pp. 196–213.

Jenkins., J. L., Proudfoot, J. G., Valacich, J. S., Grimes, G. M., and Nunamaker, J. F. Jr. 2019. "Sleight of Hand: Identifying Concealed Information by Monitoring Mouse-Cursor Movements," *Journal of the Association for Information Systems* (20), pp. 1–32. (https://doi.org/10.17705/1jais.00527).

Jenkins, J. L., Valacich, J. S., Zimbelman, A. F., and Zimbelman, M. F. 2021. "Detecting Noncompliant Behavior in Organizations: How Online Survey Responses and Behaviors Reveal Risk," *Journal of Management Information Systems* (38:3), Routledge, pp. 704–731. (https://doi.org/10.1080/07421222.2021.1962600).

Joyce, R., and Gupta, G. 1990. "Identity Authentication Based on Keystroke Latencies," *Communications of the ACM* (33:2), ACM New York, NY, USA, pp. 168–176.

Kahneman, D. 2011. *Thinking, Fast and Slow*, macmillan.

Karnan, M., Akila, M., and Krishnaraj, N. 2011. "Biometric Personal Authentication Using Keystroke Dynamics: A Review," *Applied Soft Computing* (11:2), Elsevier, pp. 1565–1573.

Khan, M. Z., Khan, Z. U., Hameed, A., and Zada, S. S. 2021. "On the Upside or Flipside: Where Is Venture Capital Positioned in the Era of Digital Disruptions?," *Technology in Society* (65), Elsevier, p. 101555.

Khanna, P., and Mukundan, S. 2010. “Recognising Emotions from Keyboard Stroke Pattern,” *International Journal of Computer Applications* (11). (https://doi.org/10.5120/1614-2170).

Kim, D., Valacich, J. S., Jenkins, J. L., Kumar, M., and Wilson, D. W. 2023. *Are You, You? Covertly Fighting Identity Fraud While Preserving Privacy with Keystroke Dynamics*.

Kool, W., and Botvinick, M. 2013. “The Intrinsic Cost of Cognitive Control,” *Behavioral and Brain Sciences* (36:6), Cambridge University Press, pp. 697–698.

Kool, W., Shenhav, A., and Botvinick, M. M. 2017. “Cognitive Control as Cost-Benefit Decision Making,” *The Wiley Handbook of Cognitive Control*, Wiley Online Library, pp. 167–189.

Kowatsch, T., Wangenheim, F., and Filler, A. 2017. *Design And Lab Experiment Of A Stress Detection Service Based On Mouse Movements*.

Kuipers, M., Richter, M., Scheepers, D., Immink, M. A., Sjak-Shie, E., and van Steenbergen, H. 2017. “How Effortful Is Cognitive Control? Insights from a Novel Method Measuring Single-Trial Evoked Beta-Adrenergic Cardiac Reactivity,” *International Journal of Psychophysiology* (119), Elsevier, pp. 87–92.

Kumar, M., Kim, D., Valacich, J. S., Jenkins, J. L., and Dennis, A. R. 2021. “Improving the Quality of Survey Data: Using Answering Behavior as an Alternative Method for Detecting Biased Respondents,” in *Proceedings of the Twentieth Annual Pre-ICIS Workshop on HCI Research in MIS*, , December 12.

Kyllonen, P. C., and Zu, J. 2016. “Use of Response Time for Measuring Cognitive Ability,” *Journal of Intelligence* (4:4), MDPI, p. 14.

Lancaster, G. L., Vrij, A., Hope, L., and Waller, B. 2013. “Sorting the Liars from the Truth Tellers: The Benefits of Asking Unanticipated Questions on Lie Detection,” *Applied Cognitive Psychology* (27:1), Wiley Online Library, pp. 107–114.

Lane, A. E., and Ziviani, J. M. 2010. “Factors Influencing Skilled Use of the Computer Mouse by School-Aged Children,” *Computers & Education* (55:3), Elsevier, pp. 1112–1122.

Mac Giolla, E., and Luke, T. J. 2021. “Does the Cognitive Approach to Lie Detection Improve the Accuracy of Human Observers?,” *Applied Cognitive Psychology* (35:2), Wiley Online Library, pp. 385–392.

Macaulay, M. 2004. “The Speed of Mouse-Click as a Measure of Anxiety during Human-Computer Interaction,” *Behaviour & Information Technology* (23:6), pp. 427–433. (https://doi.org/10.1080/01449290412331294651).

Magezi, D. A. 2015. “Linear Mixed-Effects Models for within-Participant Psychology Experiments: An Introductory Tutorial and Free, Graphical User Interface (LMMgui),” *Frontiers in Psychology* (6). (https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00002).

Mayer, J. R., and Mitchell, J. C. 2012. “Third-Party Web Tracking: Policy and Technology,” in *2012 IEEE Symposium on Security and Privacy*, IEEE, pp. 413–427.

Mayer, R. E. 2019. “Computer Games in Education,” *Annual Review of Psychology* (70), Annual Reviews, pp. 531–549.

McCrae, R. R., and Costa Jr., P. T. 2008. “The Five-Factor Theory of Personality.,” in *Handbook of Personality: Theory and Research, 3rd Ed.*, New York, NY, US: The Guilford Press, pp. 159–181.

Miles, A., Charron-Chénier, R., and Schleifer, C. 2019. “Measuring Automatic Cognition: Advancing Dual-Process Research in Sociology,” *American Sociological Review* (84:2), SAGE Publications Sage CA: Los Angeles, CA, pp. 308–333.

Miller, E. K. 2000. “The Prefontral Cortex and Cognitive Control,” *Nature Reviews Neuroscience* (1:1), Nature Publishing Group UK London, pp. 59–65.

Monaro, M., Galante, C., Spolaor, R., Li, Q. Q., Gamberini, L., Conti, M., and Sartori, G. 2018. “Covert Lie Detection Using Keyboard Dynamics,” *Scientific Reports* (8:1), Nature Publishing Group, p. 1976. (https://doi.org/10.1038/s41598-018-20462-6).

Monaro, M., Gamberini, L., and Sartori, G. 2017a. “Identity Verification Using a Kinematic Memory Detection Technique,” in *Advances in Neuroergonomics and Cognitive Engineering: Proceedings of the AHFE 2016 International Conference on Neuroergonomics and Cognitive Engineering, July 27-31, 2016, Walt Disney World®, Florida, USA*, Springer, pp. 123–132.

Monaro, M., Gamberini, L., and Sartori, G. 2017b. “The Detection of Faked Identity Using Unexpected Questions and Mouse Dynamics,” *PloS One* (12:5), Public Library of Science San Francisco, CA USA, p. e0177851.

Monaro, M., Gamberini, L., Zecchinato, F., and Sartori, G. 2018. “False Identity Detection Using Complex Sentences,” *Frontiers in Psychology* (9), Frontiers Media SA, p. 283.

Monrose, F., and Rubin, A. D. 2000. “Keystroke Dynamics as a Biometric for Authentication,” *Future Generation Computer Systems* (16:4), Elsevier, pp. 351–359.

Mulla, T. 2022. “Assessing the Factors Influencing the Adoption of Over-the-Top Streaming Platforms: A Literature Review from 2007 to 2021,” *Telematics and Informatics*, Elsevier, p. 101797.

Nature. 2023. “Cognitive Control - Latest Research and News | Nature.” (https://www.nature.com/subjects/cognitive-control, accessed April 18, 2023).

Neys, W. D. 2006. “Dual Processing in Reasoning: Two Systems but One Reasoner,” *Psychological Science* (17:5), SAGE Publications Sage CA: Los Angeles, CA, pp. 428–433.

Nigg, J. T. 2017. “Annual Research Review: On the Relations among Self-Regulation, Self-Control, Executive Functioning, Effortful Control, Cognitive Control, Impulsivity, Risk-Taking, and Inhibition for Developmental Psychopathology,” *Journal of Child Psychology and Psychiatry* (58:4), Wiley Online Library, pp. 361–383.

Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., and Patton, M. W. 2011. “Embodied Conversational Agent-Based Kiosk for Automated Interviewing,” *Journal of Management Information Systems* (28:1), Taylor & Francis, pp. 17–48.

Ochsner, K. N., and Gross, J. J. 2005. “The Cognitive Control of Emotion,” *Trends in Cognitive Sciences* (9:5), Elsevier, pp. 242–249.

Paulhus, D. L. 1991. “Measurement and Control of Response Bias.,” in *Measures of Personality and Social Psychological Attitudes.*, Measures of Social Psychological Attitudes, Vol. 1., San Diego, CA, US: Academic Press, pp. 17–59. (https://doi.org/10.1016/B978-0-12-590241-0.50006-X).

Peña-García, N., Gil-Saura, I., Rodríguez-Orejuela, A., and Siqueira-Junior, J. R. 2020. “Purchase Intention and Purchase Behavior Online: A Cross-Cultural Approach,” *Heliyon* (6:6), Elsevier, p. e04284.

Pennington, D. C. 2000. *Social Cognition*, Psychology Press.

Picazo-Vela, S., Chou, S. Y., Melcher, A. J., and Pearson, J. M. 2010. “Why Provide an Online Review? An Extended Theory of Planned Behavior and the Role of Big-Five Personality Traits,” *Computers in Human Behavior* (26:4), Elsevier, pp. 685–696.

Pollock, B. J., and Lee, T. D. 1992. “Effects of the Model’s Skill Level on Observational Motor Learning,” *Research Quarterly for Exercise and Sport* (63:1), Taylor & Francis, pp. 25–29.

Posner, M. I., Snyder, C. R., and Solso, R. 2004. “Attention and Cognitive Control,” *Cognitive Psychology: Key Readings* (205), Psychology Press New York, NY, pp. 55–85.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. (https://www.R-project.org/).

Ruensuk, M., Cheon, E., Hong, H., and Oakley, I. 2020. “How Do You Feel Online: Exploiting Smartphone Sensors to Detect Transitory Emotions during Social Media Use,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (4:4), ACM New York, NY, USA, pp. 1–32.

Schneider, M. J., Jagpal, S., Gupta, S., Li, S., and Yu, Y. 2017. “Protecting Customer Privacy When Marketing with Second-Party Data,” *International Journal of Research in Marketing* (34:3), Elsevier, pp. 593–603.

Sénécal, S., Fredette, M., Léger, P.-M., Courtemanche, F., and Riedl, R. 2015. "Consumers' Cognitive Lock-in on Websites: Evidence from a Neurophysiological Study," *Journal of Internet Commerce* (14:3), Taylor & Francis, pp. 277–293.

Specka, M., Finkbeiner, T., Lodemann, E., Leifert, K., Kluwig, J., and Gastpar, M. 2000. "Cognitive-Motor Performance of Methadone-Maintained Patients," *European Addiction Research* (6:1), Karger Publishers, pp. 8–19.

Sperl, L., and Cañal-Bruland, R. 2020a. "Interindividual Differences in the Capability to Change Automatized Movement Patterns," *Acta Psychologica* (204), Elsevier, p. 103027.

Sperl, L., and Cañal-Bruland, R. 2020b. "Reducing Proactive Interference in Motor Tasks," *Journal of Motor Behavior* (52:3), Taylor & Francis, pp. 372–381.

Stanovich, K. E. 2018. "Miserliness in Human Cognition: The Interaction of Detection, Override and Mindware," *Thinking & Reasoning* (24:4), Taylor & Francis, pp. 423–444.

Stanovich, K. E. 2021. "Why Humans Are Cognitive Misers and What It Means for the Great Rationality Debate," *Routledge Handbook of Bounded Rationality*, Routledge, pp. 196–206.

Statista. 2022. "Total Data Volume Worldwide 2010-2025," *Statista*. (https://www.statista.com/statistics/871513/worldwide-data-created/, accessed April 18, 2023).

Straub, E., Kiesel, A., and Dignath, D. 2020. "Cognitive Control of Emotional Distraction–Valence-Specific or General?," *Cognition and Emotion* (34:4), Taylor & Francis, pp. 807–821.

Sweller, J. 2011. "Cognitive Load Theory," in *Psychology of Learning and Motivation* (Vol. 55), Elsevier, pp. 37–76.

Teh, P. S., Teoh, A. B. J., and Yue, S. 2013. "A Survey of Keystroke Dynamics Biometrics.," *TheScientificWorldJournal* (2013), United States, p. 408280. (https://doi.org/10.1155/2013/408280).

Tourangeau, R., and Yan, T. 2007. "Sensitive Questions in Surveys.," *Psychological Bulletin* (133:5), United States, pp. 859–883. (https://doi.org/10.1037/0033-2909.133.5.859).

Tsao, S.-F., Chen, H., Tisseverasinghe, T., Yang, Y., Li, L., and Butt, Z. A. 2021. "What Social Media Told Us in the Time of COVID-19: A Scoping Review," *The Lancet Digital Health* (3:3), Elsevier, pp. e175–e194.

Valacich, J. S., Jenkins, J. L., Byrd, M., Kim, D., and Kumar, M. 2020. "Using Accelerometer and Gyroscope Data in Common Mobile Devices to Assess Credibility," in *AMCIS 2020 Proceedings. 25.*, Association for Information Systems.

Vance, A., Jenkins, J. L., Anderson, B. B., Bjornn, D. K., and Kirwan, C. B. 2018. "Tuning Out Security Warnings: A Longitudinal Examination of Habituation Through FMRI, Eye Tracking, and Field Experiments," *MIS Quarterly*.

Verburgh, L., Scherder, E., Van Lange, P., and Oosterlaan, J. 2016. "The Key to Success in Elite Athletes? Explicit and Implicit Motor Learning in Youth Elite and Non-Elite Soccer Players," *Journal of Sports Sciences* (34:18), Taylor & Francis, pp. 1782–1790.

Vizer, L. M., Zhou, L., and Sears, A. 2009. "Automated Stress Detection Using Keystroke and Linguistic Features: An Exploratory Study.," *International Journal of Human-Computer Studies* (67), Netherlands: Elsevier Science, pp. 870–886. (https://doi.org/10.1016/j.ijhcs.2009.07.005).

Vrij, A. 2008. *Detecting Lies and Deceit: Pitfalls and Opportunities*, John Wiley & Sons.

Vrij, A., Fisher, R. P., and Blank, H. 2017. "A Cognitive Approach to Lie Detection: A Meta-Analysis," *Legal and Criminological Psychology* (22:1), Wiley Online Library, pp. 1–21.

Wahle, F., Kowatsch, T., Fleisch, E., Rufer, M., Weidt, S., and others. 2016. "Mobile Sensing and Support for People with Depression: A Pilot Trial in the Wild," *JMIR MHealth and UHealth* (4:3), JMIR Publications Inc., Toronto, Canada, p. e5960.

Weinmann, M., Valacich, J., Schneider, C., Jenkins, J. L., and Hibbeln, M. T. in press. *The Path of the Righteous: Using Trace Data to Understand Fraud Decisions in Real Time*, SSRN Scholarly Paper, Rochester, NY. (https://papers.ssrn.com/abstract=4008223).

Welsh, T. N., and Elliott, D. 2004. "Movement Trajectories in the Presence of a Distracting Stimulus: Evidence for a Response Activation Model of Selective Reaching," *The Quarterly Journal of Experimental Psychology Section A* (57:6), Taylor & Francis, pp. 1031–1057.

Westbrook, A., Lamichhane, B., and Braver, T. 2019. "The Subjective Value of Cognitive Effort Is Encoded by a Domain-General Valuation Network," *Journal of Neuroscience* (39:20), Soc Neuroscience, pp. 3934–3947.

Wiederhold, B. K. 2022. "Metaverse Games: Game Changer for Healthcare?," *Cyberpsychology, Behavior, and Social Networking*, Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New …, pp. 267–269.

Zech, H. G., Reichert, M., Ebner-Priemer, U. W., Tost, H., Rapp, M. A., Heinz, A., Dolan, R. J., Smolka, M. N., and Deserno, L. 2022. "Mobile Data Collection of Cognitive-Behavioral Tasks in Substance Use Disorders: Where Are We Now?," *Neuropsychobiology* (81:5), Karger Publishers, pp. 438–450. (https://doi.org/10.1159/000523697).

Zheng, W.-L., Dong, B.-N., and Lu, B.-L. 2014. "Multimodal Emotion Recognition Using EEG and Eye Tracking Data," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, , August, pp. 5040–5043. (https://doi.org/10.1109/EMBC.2014.6944757).

Zwicker, J. G., Missiuna, C., Harris, S. R., and Boyd, L. A. 2011. "Brain Activation Associated with Motor Skill Practice in Children with Developmental Coordination Disorder: An FMRI Study," *International Journal of Developmental Neuroscience* (29:2), pp. 145–152. (https://doi.org/10.1016/j.ijdevneu.2010.12.002).