

9-30-2021

Rigor, Relevance, and Practical Significance: A Real-life Journey to Organizational Value

Joseph S. Valacich

University of Arizona, valacich@arizona.edu

Jeffrey L. Jenkins

Brigham Young University, jjenkins@byu.edu

Follow this and additional works at: <https://aisel.aisnet.org/thci>

Recommended Citation

Valacich, J. S., & Jenkins, J. L. (2021). Rigor, Relevance, and Practical Significance: A Real-life Journey to Organizational Value. *AIS Transactions on Human-Computer Interaction*, 13(3), 350-368. <https://doi.org/10.17705/1thci.00153>

DOI: 10.17705/1thci.00153

This material is brought to you by the AIS Journals at AIS Electronic Library (AISeL). It has been accepted for inclusion in AIS Transactions on Human-Computer Interaction by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.



Transactions on Human-Computer Interaction

Volume 13

Issue 3

9-2021

Rigor, Relevance, and Practical Significance: A Real-life Journey to Organizational Value

Joseph S. Valacich

Management Information Systems, University of Arizona, valacich@arizona.edu

Jeffrey L. Jenkins

Information Systems, Brigham Young University, jjenkins@byu.edu

Follow this and additional works at: <http://aisel.aisnet.org/thci/>

Recommended Citation

Valacich, J. S. & Jenkins, J. L. (2021). Rigor, relevance and practical significance: A real-life journey to organizational value. *AIS Transactions on Human-Computer Interaction*, 13(3), pp. 350-368.

DOI: 10.17705/1thci.00153

Available at <http://aisel.aisnet.org/thci/vol13/iss3/4>



Transactions on Human-Computer Interaction

Research Commentary

DOI: 10.17705/1thci.00153

ISSN: 1944-3900

Rigor, Relevance, and Practical Significance: A Real-life Journey to Organizational Value

Joseph S. ValacichManagement Information Systems
University of Arizona**Jeffrey L. Jenkins**Information Systems
Brigham Young University

Abstract:

In this essay, we describe a research journey focusing on how to analyze mouse cursor movements, typing fidelity, and data from other human-computer interaction (HCI) devices to better understand the end-user online experience. We begin by defining organizational value and how it relates to other aspects that researchers use to assess academic research quality. We then describe and contrast our research journey by demonstrating key research milestones: from achieving statistical significance to achieving practical significance and, finally, to reaching relevance to practice. We then explain how we crossed the chasm between academic research and technology commercialization (i.e., the last research mile). We conclude by describing the process one can follow to develop an initial prototype—the minimal viable product (MVP)—and how demonstrations with potential customers provides continuous insight and validation for evolving the commercial product capabilities to meet constantly changing and evolving customer and industry needs.

Keywords: Human-computer Interaction (HCI), HCI Dynamics, Commercialization, Organizational Value, Minimal Viable Product (MVP), Lean Startup.

Fiona Nah was the accepting senior editor for this paper.

1 Introduction

In 2011, along with a small group of collaborators from around the world, we began researching how to analyze mouse cursor movements, typing fidelity, and data from other human-computer interaction (HCI) devices to better understand the end-user online experience. After conducting intense research in this area for many years, we invented and enabled the technology that we today call Human Analytics™ and co-founded a company called Neuro-ID (see www.Neuro-ID.com) in 2014 that has since grown into a successful venture capital-backed corporation that has begun to transform the lending, insurance, e-commerce, and other digital worlds. Our patented software monitors, analyzes, and scores HCI devices by evaluating how users enter various types of information into online forms (e.g., a loan application) or interacts with websites (e.g., going through an e-commerce purchasing process) (for details on the patents, see Valacich & Jenkins, 2019, 2020). This rich behavioral data (e.g., about how a person clicks, types, and swipes) can help firms identify possible fraud (e.g., excessive edits, unusual answering behavior, etc.) and friction (e.g., navigation or data entry confusion) and better identify genuine customers.

As of 2021, the systems based on our research process trillions of data points to help hundreds of millions of end users. With Human Analytics, organizations can identify friction on their websites and make guided improvements to the user experience. In addition, the data can help organizations to better partition large digital customer populations into different strata based on their behavior. Thus, for example, organizations can more deeply scrutinize individuals who exhibit fraudulent behavior yet provide a more streamlined process to individuals who behave like genuine customers. The ability to partition large customer populations allows organizations to better utilize their limited resources and, thereby, lower fraud and risk while simultaneously increasing the rate at which they convert good customers.

Just as we have learned much since we began our research journey, so too has the information systems (IS) discipline continued to evolve and mature. Throughout the past several decades, researchers have debated what type of research the IS discipline should value and encourage. The discipline has weighed the value of rigor over relevance and vice versa (Davenport & Markus, 1999; Lee & Hubona, 2009). We have debated the role and importance of the information technology (IT) artifact (Benbasat & Zmud, 2003; Orlikowski & Iacono, 2001; Robey, 2003). More recently, the discipline has discussed big data's influence and emerging role as both the scholars and methods in IS community have diversified (Grover, Lindberg, Benbasat & Lyytinen, 2020) and the value of relevance and practical significance relative to other factors such as rigor and statistical significance (Mohajeri, Mesgari, & Lee, 2020). In this essay, we do not comment on what information systems research should or should not be or how we can improve it; rather, we describe our research journey from conducting rigorous foundational research to studies with practical significance before finally commercializing our research and achieving organizational value (which Nunamaker, Briggs, Derrick, and Schwabe (2015) refer to as the last research mile).

Because organizational value can be multifaceted and complex, we first define the concept and describe several ways in which one can measure academic research's organizational value. We believe that all research has value to someone depending on their own personal journey. We also believe that, to demonstrate organizational value, one must first understand and demonstrate statistical and practical significance and relevance. However, we emphasize that statistical and practical significance and relevance constitute necessary but insufficient conditions to provide organizational value. In describing our research journey, we hope to help other researchers better navigate their own journeys toward more impactful research.

2 Organizational Value

We have found that, like life itself, developing deep insights in a particular research context is a journey (Quote Investigator, n.d.). Also, like life, one cannot hurry the research journey. As one's work evolves, one sees signposts—results and insight—that inform one's direction, pace, and urgency. While one may have a desire to complete a journey as quickly as possible, we have found that critical learning is often serial in nature. In other words, we mean that one cannot learn concept B until one fully understands concept A, and, as one learns, it informs one's direction. We have been fortunate to have had many great companions and collaborators on our research journey over the years. We thank them all for their contributions and friendship, and we look forward to more discoveries ahead!

To help readers understand our perspective and what we mean by organizational value, we briefly define and discuss some important concepts and terms to help readers better understand how we describe our

journey. Most fundamentally, we follow Babbie (1998) in defining research as a systematic inquiry—using both inductive and deductive methods—to describe, explain, predict, and control observed phenomenon. Like many other disciplines, the IS discipline uses numerous methodologies and perspectives to conduct research. We believe that our work resides under the sociotechnical umbrella (Sarker, Chatterjee, Xiao, & Elbanna, 2019) because we do behavioral theory-driven empirical research and design science. Before we introduce organizational value, we first define three important concepts: 1) statistical significance, 2) practical significance, and 3) relevance. In statistical hypothesis testing in quantitative studies, a result has statistical significance when it does not likely occur by chance. While statistical significance represents an important threshold to achieve in our research, it alone cannot deliver organizational value. In addition, regarding the value of statistical significance, there is a growing concern of p-hacking, which refers to misreporting true effect sizes (Head, Holman, Lanfear, Kahn, & Jennions, 2015). In practice, researchers accomplish p-hacking via selectively reporting results to achieve statistical significance. Assuming ethical methods and reporting, statistical significance indicates only that one has sufficient evidence to conclude that an effect exists but not that it necessarily matters from a practical, relevance, or organizational value standpoint.

Therefore, we also need to define practical significance, which refers to whether an acquired research result (e.g., a statistically significant result) is useful in the real world versus in theory. To assess practical significance, researchers use metrics such as effect sizes, accuracy rates, and a finding's coverage. In addition, in the real world, having a large effect or a high accuracy rate may not be enough. For instance, how well do existing solutions already solve this problem relative to a new approach? If a new approach only marginally improves performance over existing approaches or has excessive implementation or operating costs that lower its return on investment (ROI) relative to these existing approaches, it may not provide meaningful value to an organization.

To determine whether an acquired research result (i.e., a statistically significant result) is useful in the real world versus in theory, we must understand the real-world context of interest. Thus, we must compare a new approach to the existing solutions in practice and understand the problem's importance to gain insight into its potential value. In short, not all statistically significant differences represent interesting or meaningful ones. Practical significance, from a real-world perspective, involves more than effect size, more than accuracy rates, and more than coverage. While a result may relate to practice, it may not solve a significant problem, or existing solutions may already solve it. Researchers also use relevance to assess research's value. We define relevance as responsiveness to business and industry needs, which makes it useful and easy to understand or put into practice. In short, relevance focuses on whether practitioners can easily understand and consume. Some key factors that influence this understandability include:

- Whether a research approach and its finding are intuitive,
- Whether the findings are noteworthy,
- Whether practitioners can take action due to the findings, and
- Whether findings can make a “meaningful difference”.

Thus, we argue that relevance—like statistical and practical significance—constitutes a necessary but insufficient condition to provide organizational value.

Thus, we now turn to organizational value, which we define as research that *may have* statistical and practical significance and relevance but *must also solve* a problem that *matters* to organizations. Organizational value involves *much* more than relevance. In the Lean startup language (Ries, 2011), delivering organizational value relates to *crossing the chasm* (i.e., finding the product-market fit). Product-market fit refers to the degree to which a product satisfies a strong market demand (i.e., the last research mile) (Nunamaker et al., 2015). We discuss the last research mile in Section 4.

For research to have organizational value (i.e., to have value that one can build a successful product or company on), several things must come into alignment. For instance, the problem one focuses on solving with one's research-driven solution must be “big enough to matter” to a meaningful part of a market. Furthermore, one's research-driven solution must be better than current approaches; it must explain unique variance. It must also be intuitive and easy to understand. Also, one needs to make the solution easy to consume, which means that companies must find it easy to integrate into their technology stack and into their day-to-day decision making processes. It must not add additional risk from a cybersecurity standpoint or from a data privacy standpoint (e.g., GDPR). In sum, the research-driven solution must not be just a better mouse trap—it must be a significantly better mouse trap. It must be bigger, better, faster, and cheaper.

To help researchers understand whether their research has the potential to gain significant organizational value, they can ask various questions to understand its current or potential value such as:

- Strategic value: how vital is your research for enabling an organization's competitive strategy?
- Problem size: how large is the problem for an organization or industry that your research addresses?
- Alternatives: how well do existing solutions solve the problem that your research addresses?
- Operational value: what is the ROI an organization can achieve by deploying your research?
- Time to value: how quickly can organizations gain measurable value from your research?
- Scale: how easily (e.g., in terms of time, cost, and effort) can one scale the solution that your research proposes?

In sum, in academic research, there are standards for determining the cutoffs or thresholds for “good” research. For instance, p-values can differ in whether they constitute a “meaningful” positive result. In some research domains, 5 out of 100 or 1 out of 1000 may prove meaningful. In organizational contexts, however, 20 or even 30 out of 100 may prove meaningful. Likewise, with practical significance, one may look for a large effect size or a strong area under the curve (AUC) score for a machine learning model, but, in the real world, a company may highly value an AUC of 0.6 depending on the context (e.g., the problem size, alternative solutions, etc.). Furthermore, regarding relevance, researchers need deep knowledge about the context and value relative to other available solutions. Thus, researchers must connect academic research to organizational value; but, to achieve organizational value, they need to ensure their research meets the minimum statistical significance, practical significance, and relevance thresholds. Importantly, however, even if researchers have met all these preliminaries, they have no guarantee that they have or can achieve organizational value. Thus, even our “best” academic research only has the potential to yield organizational value (see Figure 1). In Section 3, we describe our journey to obtain organizational value.

3 The Research Journey

In this section, we review some of our past work. We begin with studies that had relatively low ecological validity (i.e., studies with only sufficient relevance to the test population) before we moved onto studies that had practical significance and, finally, to studies that directly pertained to practice. As such, over the past decade, our research has evolved and matured, but we have built it on a foundation of strong theory and concepts from the neurological and cognitive sciences. We first describe some foundational theories and concepts that have guided our work.

3.1 Brief Theory Overview

Foundational research from scholars in the cognitive and neurological sciences has shown that collecting and analyzing how users interact with a computer system can reveal insights into their emotional and cognitive states. More specifically, this prior research has unequivocally demonstrated that cognitive and emotional changes influence fine motor control (Freeman et al., 2011) and that one can use various HCI devices to collect raw data streams from these devices at millisecond precision. As such, tracking mouse movements, interaction data from touch screens, keystroke behaviors, and the like has emerged into a scientific methodology, which we refer to as human-computer interaction (HCI) dynamics. HCI dynamics refers to the science of collecting raw data streams of HCI data (e.g., timestamps and key presses for typing, screen x- and y-coordinate positions and timestamps for mouse movements, etc.), transforming it into metrics, and interpreting the metrics to provide objective insight about a person's decision-making and other psychological processes (i.e., Human Analytics). After reviewing mouse-tracking studies in their HCI dynamics review, Freeman, Dale, and Farmer (2011, p. 59) made the following suggestion:

Movements of the hand...offer continuous streams of output that can reveal ongoing dynamics of [cognitive] processing, potentially capturing the mind in motion with fine-grained temporal sensitivity.

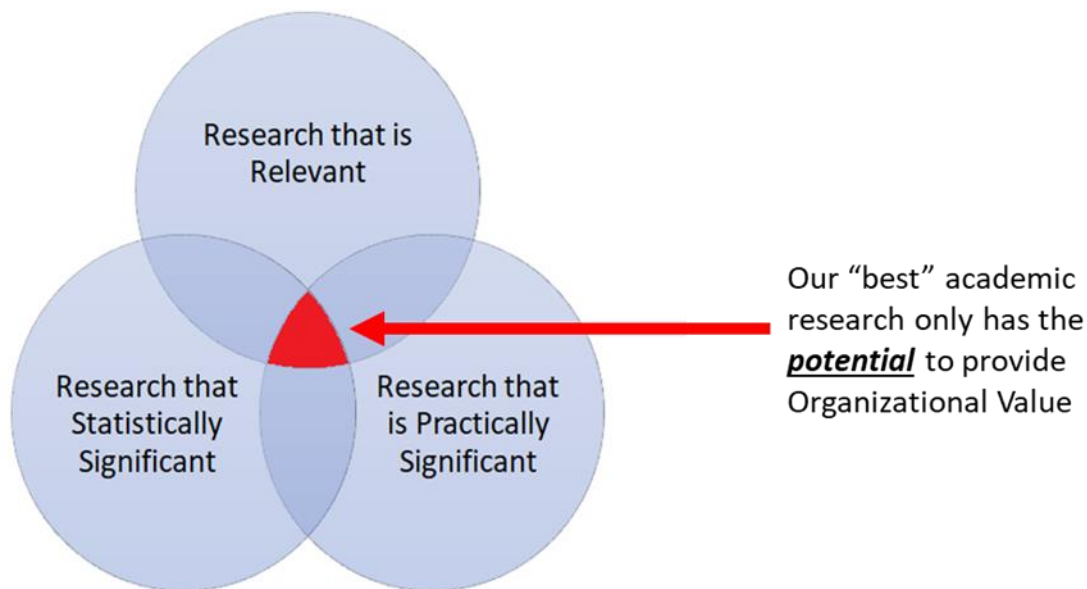


Figure 1. Characteristics of Academic Research that Has the Potential to Provide Organizational Value

Over the past several years, numerous studies from diverse fields have used HCI dynamics as a methodology to study various cognitive and emotional processes. For example, research has shown HCI dynamics to predict decision conflict (McKinstry, Dale, & Spivey, 2008), attitude formation (Jenkins & Valacich, 2015), whether individuals conceal their racial prejudices (Wojnowicz, Ferguson, Dale, & Spivey, 2009), response difficulty (Horwitz, Kreuter, & Conrad, 2017), response certainty (Bodily et al., 2015), dynamic cognitive competition changes (Dale, Kehoe, & Spivey, 2007), perception formation (Cloutier, Freeman, & Ambady, 2014), and emotional reactions (Hibbeln, Jenkins, Schneider, Valacich, & Weinmann, 2017).

Several foundational theories support one's ability to use HCI dynamics to infer changes in cognitive and emotional states. Together, these theories explain how cognitive and emotional changes influence fine motor control and, thereby, user interactions with HCI devices. To learn more about the key theories and concepts that we leveraged in our prior and ongoing academic research, we direct readers to Jenkins et al. (forthcoming) for an up-to-date review and comparison of the most impactful academic research in this area.

3.2 Achieving Statistical Significance

In much of our early work, we focused on gaining insight into the relationship that cognitive and emotional changes have with fine motor control. In one example (Grimes & Valacich, 2015), we reported a method to detect heightened cognitive load using mouse cursor movements. Specifically, we reported the results from a laboratory experiment in which we manipulated cognitive load—low, medium, and high—using a 1x3 within-subjects design while collecting mouse movement data.

We designed the task's interface to resemble the interface that other mouse-tracking studies that researchers in the cognitive- and neuro-science area conducted during the same period used (e.g., Freeman et al., 2011). The interface had low ecological validity and, thus, provided great control for clear-cut hypothesis testing (Dennis & Valacich, 2001). Building on this prior work, we implemented an anchor button in the interface at the bottom center of the screen to increment each study iteration and to give the mouse cursor an anchor point (in terms of both time and location) for each trial, response buttons at upper left and right corners of the screen, and a stimuli presentation area in the middle of the screen (see Figure 2). Each time a participant clicked the anchor button (i.e., next), the stimuli (i.e., a number in this study) appeared for approximately one second and then disappeared. We required participants to click one response button (i.e., yes or no) and then click the anchor button again to advance. Participants completed 100 iterations of each task and then self-reported the task's difficulty.

We designed the first task to require relatively little cognitive effort (i.e., a low cognitive load). We instructed participants to monitor the stream of 100 numbers for the number 5. When they saw the number 5, they had to click the green button at the top right corner of the screen. When they saw any number other than 5, they

had to click the red button at the top left of the screen. As expected, participants performed this task successfully: they marked the correct response 99.6 percent of the time on average (SD = 0.9, min = 95, max = 100).

We designed the second task to elicit higher cognitive load (i.e., a medium cognitive load) by requiring lag-1 number recall and simple comparison of two numbers. As in the first task, we showed participants a stream of numbers. However, this time, we instructed them to click the green button at the top right corner of the screen if a number larger than the previous one appeared on the screen and to click the red button at the top left corner of the screen if a number smaller than the previous one appeared on the screen. Overall, participants performed this task slightly less successfully than the first task: they marked the correct response 95.6 percent of the time on average (SD = 5.5, min = 74.4, max = 100).

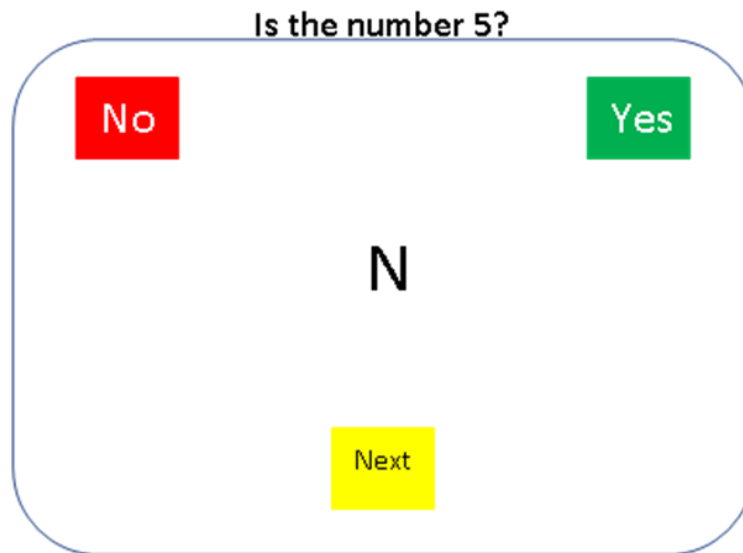


Figure 2. Screen Design with Low Ecological Validity to Examine the Influence of Increasing the Cognitive Load (Adapted from Grimes & Valacich, 2015)

We designed the third task to elicit considerably higher cognitive load. As with the two previous tasks, we presented participants with a stream of 100 numbers. This time, we instructed them to add the two previous numbers in their head to click the green button at the top right corner of the screen if the current number on the screen was larger than the sum of the previous two numbers or the red button at the top left corner of the screen if it was smaller than the sum of the two previous numbers. Participants performed this task less successfully than the other two: they provided correct response 81.3 percent of the time on average (SD = 13.2, min = 46.2, max = 98.7).

A multilevel linear model revealed a significant difference with a medium effect ($d = 0.53$) between the scores for the first and second tasks ($b = 4.01$, $t(104) = 2.71$, and $p < 0.01$) and a significant difference with a large effect ($d = 1.87$) between the scores for the second and third tasks ($b = 14.28$, $t(104) = 9.64$, and $p < 0.001$). These results suggest that the manipulation did elicit three distinct cognitive load levels. We further found that participants exhibited a significantly longer task duration, longer mouse movements, more direction changes, and slower speed when under a high cognitive load. In Figure 3, we show a participant's movements in the first 10 rounds (out of a total of 100 rounds for each task) for the first task compared with participant's movements in the first 10 rounds for the third task. Note the differences in the movement traces and the slower speeds (i.e., 21 versus 38 seconds to complete 10 rounds of the task).

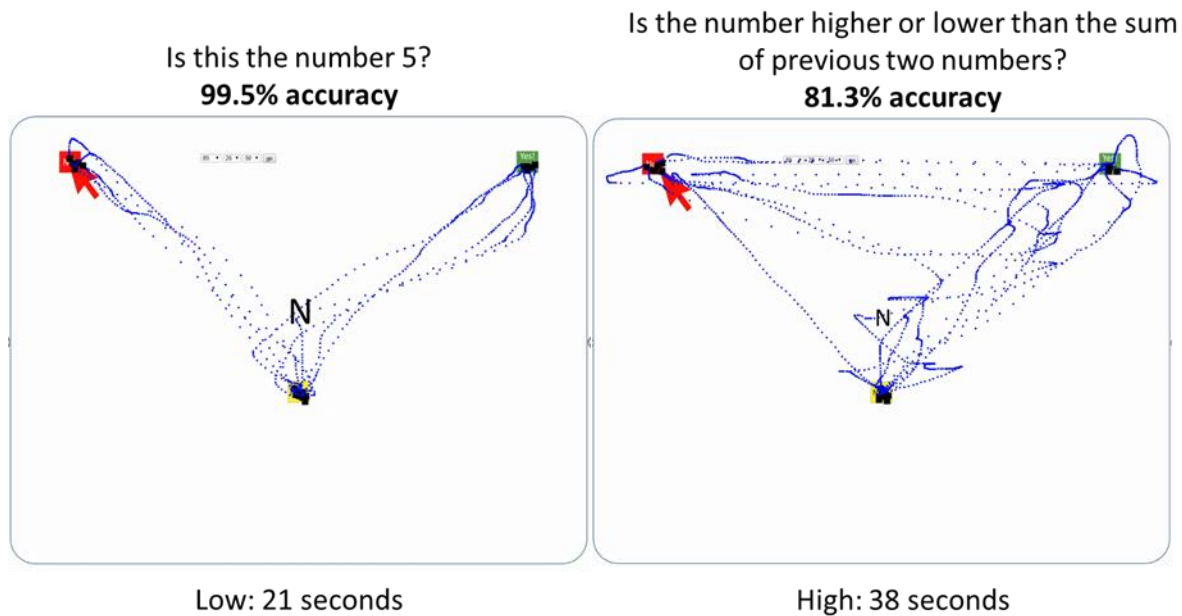


Figure 3. Mouse Cursor Movements for the First Task (Low Cognitive Load) and the Third Task (a High Cognitive Load) after 10 Rounds (from Grimes and Valacich, 2015)

3.3 Achieving Practical Significance

As we continued to work on trying to understand this research area in depth and more broadly, we strove to make our work more practical and relevant while maintaining strong experimental control. Researchers cannot easily achieve both statistical and practical significance in a single study. Thus, in Hibbeln et al. (2017), we reported the results from three related studies to understand how negative emotion influences mouse cursor movements. More specifically, we focused on framing our work with respect to customer interactions in an e-commerce context. To obtain both statistical and practical significance, we conducted three studies: in the second and third studies, we built on the methods and findings from the prior ones to demonstrate the practical value that one can gain from understanding mouse cursor movements in various online contexts.

In the first study, an experiment with participants from MTurk, we randomly manipulated negative emotions and then monitored participants' mouse cursor movements as they completed a number-ordering task. We manipulated participants' emotions by randomly giving them a fair and positive aptitude test versus an unfair and negative aptitude test. In the fair and positive (baseline) conditions, we asked easy test questions and told participants at the end that they were smart and performed well. In the unfair and negative (treatment) conditions, we asked hard questions, timed how long participants took to answer them, and made the user experience hampered and frustrating. At the end of the study, we told participants in the frustration-inducing treatment that they underperformed on the test. We measured frustration using a self-report measure (i.e., the self-assessment manikin—see Figure 4). We found that frustration significantly differed across the baseline and treatment conditions. We collected the mouse cursor movements and stored them in a database for later analysis while all participants dragged and dropped the numbers in ascending order (see Figure 5). Additionally, we found that negative emotions significantly increased the mouse cursor distance and reduced the speed during the number-sorting task.

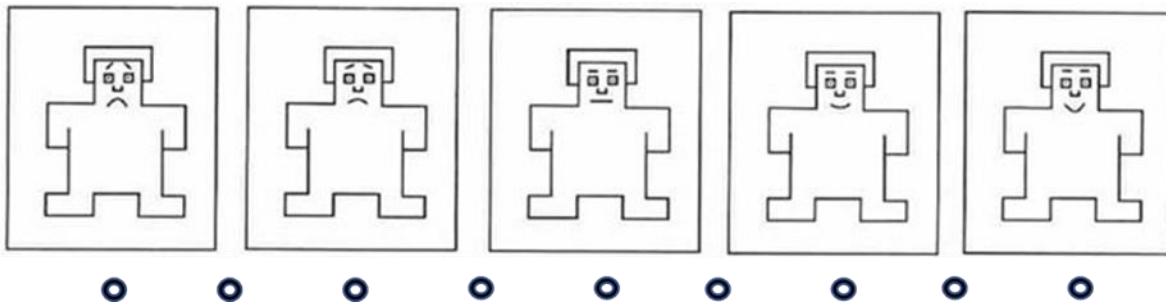


Figure 4. The Self-Assessment Manikin (SAM) Capturing the Participants' Post-Treatment Valence prior to the Number-sorting Task

To receive compensation, please order the following number in ascending order (from lowest to highest) by dragging the numbers from the list (on the left) to the box (on the right).

ITEMS

13424

13485

13421

13410

13429

13599

ORDER NUMBERS HERE

Figure 5. The Number-sorting Task to Capture the Mouse Cursor Distance and Time

In the second study, an experiment with university students as participants, we randomly manipulated negative emotions and then monitored participants' mouse cursor movements while they interacted with a mock e-commerce site. Specifically, we asked participants to complete a goal-directed task on an e-commerce website to pretend to purchase a J. Crew Abingdon Laptop Bag for a 17-inch laptop. The website was designed so that there was only one obvious link on each page that would lead the user closer to goal attainment (finding the correct laptop bag). After purchasing the laptop bag, participants completed a post-experiment survey. While they interacted with the webpages, we recorded their mouse cursor movements and stored them in a database for later analysis as in the first study.

We randomly assigned half of the participants to the negative-emotion (treatment) condition where we manipulated various aspects of the website's usability. For instance, we added a download delay (i.e., the webpage loading speed) and error messages to induce negative emotions (Ceaparu et al., 2004; Galletta et al., 2004). We randomly assigned the other half of the participants to the baseline condition in which the website loaded webpages error free and without delay. In both treatments, the system recorded the mouse cursor movements after loading each webpage and anchoring the mouse cursor to approximately the same location. Hence, the mouse cursor movement analysis had the same beginning and ending points regardless of the condition (and excluded the negative emotion manipulation that only the negative-emotion condition contained). Consistent with the first study, we found longer mouse cursor distances and slower movement speeds for participants in the negative-emotion treatment condition. Additionally, we used participants' mouse cursor distance and time to infer the presence of negative emotion (i.e., participants' treatment condition) with an overall accuracy rate of 81.7 percent.

In the third study, an observational study with participants from universities in Germany and Hong Kong, we used a repeated-measure design to monitor the mouse cursor movements while participants interacted with an online product configurator from either Dell or Volkswagen. The participants reported their level of negative emotion after each step in the configuration process. We found that we could use the mouse cursor

distance and speed to infer the level of negative emotion—as the tasks became more or less frustrating—with an out-of-sample R2 of 0.17.

In addition to traditional metrics to determine practical significance, we believe that researchers need to be able to position research in a real-world context. The results from the above three studies show that we could better understand digital customers' emotions in various online contexts. Specifically, assessing emotional valence and emotional changes during live system use with temporal precision (i.e., in real time) has promising value in many online contexts. At this stage, we remained far away from providing guidance on how to deploy mouse-cursor tracking for such real-time organizational value. Nevertheless, visionary IT leaders could see this emerging method's promise. In fact, CIONET, the global professional organization for IT executives, named our study, which began in 2011, the 2018 research paper of the year. Other visionary practitioners used our work to stimulate several practitioner articles on what the future might hold as our ability to capture and interpret fine-grained HCI data at scale becomes a reality. For instance, imagine a computer that can predict when a user becomes angry (Surendranath, 2015).

3.4 Achieving Relevance

We previously argued that relevance focuses on research's value to organizations and practitioners (e.g., if they can understand, act on, etc.). In other words, academic practice must pass a high bar to achieve meaningful relevance to practice. Given this backdrop, we pursued a research context that pertained highly to practice: identifying fraudulent online entries (Weinmann, Valacich, Schneider, Jenkins, & Hibbeln, forthcoming).

Given digital business's rapid growth rate and prevalence, countless business and governmental processes rely on online customers to complete online forms and enter data. However, when entering such data, customers commonly enter fraudulent information, which results in costly consequences for organizations and society. Furthermore, detecting fraudulent responses in online forms and data entry often involves much difficulty, time, and costs.

While fraud commonly occurs in many online contexts, we focused specifically on insurance fraud in one study we report on in this paper. The Coalition Against Insurance Fraud (2021) has estimated that insurance fraud exceeds US\$80B annually in the United States alone. In addition, applying for insurance or reporting insurance claims has increasingly moved online. Thus, we focused on designing a study that would allow practitioners to easily connect the dots to a specific use case and context—submitting an insurance claim where participants could freely choose to inflate damage to increase the payout from a claim-reporting system.

We had participants enter five randomly ordered accidents into an accident claim reporting system and report damages using an online damage reporting form (see Figure 6). During this data entry, we recorded participants' mouse cursor movements. In each scenario, participants could choose to commit fraud by claiming more damages than presented. Thus, claiming more than the presented damage—committing fraud—resulted in a higher financial payout to the participant.

We found that participants moved their mouse significantly more slowly and with greater deviation when entering fraudulent responses. Further, we found that not only the decision to commit fraud but also the fraud's extent influenced mouse movements in that participants who committed more extensive fraud displayed increased movement deviation and decreased movement speed. These results demonstrate the efficacy of analyzing mouse cursor movements to detect fraud during online transactions in real time, which enables organizations to proactively detect fraud as it happens at Internet scale.

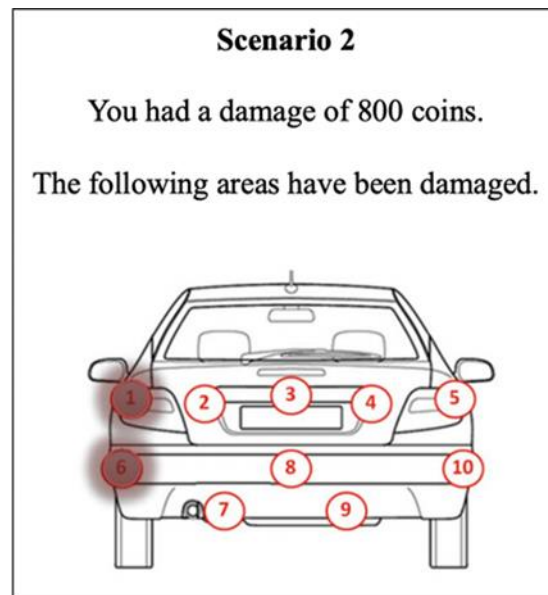


Figure 6. Repair Sample Scenario (Repair Costs: 800; Number of Damages: 2) (from Weinmann et al., forthcoming)

4 Crossing the Research Chasm to Organizational Value

Rogers' (2003) diffusion of innovation theory explains how, why, and at what rate new ideas and technology spread. The theory separates different market segments into five adopter categories: innovators (will take a risk with a new product and buy or try a new technology first), early adopters (will try new products and innovations before the majority), early majority, late majority, and laggards.

In his best-selling book *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*, Moore (1999) built on Rogers' (2003) diffusion theory and focused on various nuances and requirements involved in marketing high technology products during the early start-up period. He argued that innovators and early adopters have very different product expectations than individuals who adopt technology later in the diffusion model. Innovators and early adopters seek innovation and new capabilities. Alternatively, individuals who adopt a technology later have different product expectations and avoid risk more. They demand products to have proven benefits. For a technology company to succeed, it must cross the chasm from selling to early risk-takers, to selling to mainstream markets. We have lived through this process over the past several years. It can be quite painful and frustrating, and many great ideas fail at this point.

We both graduated from the Management Information Systems (MIS) PhD program of the University of Arizona and were influenced by Jay Nunamaker and his never-ending journey to innovate and commercialize his ideas. In 2015, he wrote a paper in which he described how he has gone from "desk-chair science", where most academic research ends, to commercializing some of his research, which he referred to as navigating the "last research mile" (Nunamaker et al., 2015). According to Nunamaker et al. (2015), going the last research mile means using scientific knowledge and methods to address important unsolved problem classes for real people with real stakes in the outcomes. The last research mile proceeds in three stages: proof-of-concept research to demonstrate a solution's functional feasibility, proof-of-value research to investigate how a solution can create value across various conditions, and proof-of-use research to address complex operational feasibility issues. The last research mile ends only when practitioners routinely use a solution in the field.

Over the past several years, along with our academic research partners and Neuro-ID colleagues, we have traveled the last research mile. We have found it to be an endless road. In our terminology, the research mile ends when we can deliver organizational value to a market beyond innovators and early adopters. Human Analytics can benefit many types of online businesses and processes, and, thus, numerous vertical markets that each has some overlapping and unique requirements. Thus, as we move from one vertical

market to another, we must repeatedly cross the chasm to organizational value. However, as one achieves success in one market, one requires less effort, time, and resources to cross the next.

4.1 Building and Evolving the Minimal Viable Product

When we realized that we could commercialize and provide organizational value with our research, we knew we needed a working prototype—what Nunamaker et al. (2015) refer to as a proof of concept. The Lean Startup (Ries, 2011) terminology refers to this prototype as the minimum viable product (MVP). Conceptually, the MVP has enough features to potentially attract innovators and early adopters and can help one gain insight into missing product capabilities and their priority when demonstrating the MVP to more risk-averse customers (i.e., not innovators or early adopters). One can add these “missing” product features to the product roadmap, a seemingly never-ending priority list of product features and capabilities derived from customer demonstrations (Olson, 2015). Because scrum and other agile development methodologies focus on validating and iteratively improving products based on customer feedback, the MVP plays a central role in not only revealing what a product lacks but also the needed capabilities’ order and urgency (i.e., the product roadmap).

We provide a sample visualization from our first MVP in Figure 7. This MVP could capture mouse movements and clicks when users answered a series of yes-no questions on a single page form. The MVP did not score the data in real time but stored it on a Web service for later processing. This figure comes from a page on an early adopter’s website, a payday lender, which helped partition customers into different categories based not only on their final response but also on how confidently the respondent answered important questions. In Figure 7, the undulating continuous line reflects the user’s mouse movements and the dots reflect various clicks. Here, we highlight what we referred to as a “risk-relevant” question compared to “baseline” questions (the other four). The respondent had significantly lower confidence for the highlighted question (both greater movement deviations and slower speeds), and our algorithm flagged this respondent as being “concerning” with respect to this question. Interestingly, the lender still gave the application a loan. However, we later found out that the individual defaulted on the loan. Thus, we found that our technology provided meaningful insights not just in the laboratory but, most importantly, in the real world.

During the MVP building and evolving process, we obtained valuable knowledge that built on our earlier fundamental scientific research. In the real world, one encounters situations (in our case, behaviors on websites) that one cannot anticipate in the laboratory. Based on experience, we adopted this mantra: “If it could happen, it will happen at Internet scale”. In other words, if we had any chance to observe behavior no matter how rare or obscure it might be, we would absolutely see it at Internet scale. The increased control that laboratory experiments afford means researchers often obtain relatively clean data. However, data in the real world can be astronomically huge and messy. Thus, we had an opportunity to investigate a whole new class of research questions: how could we filter out “Internet-scale” noise and enhance a signal? How could we develop algorithms that could handle very large amounts of data? How could we extend our research’s generalizability to obscure and rare cases? What behaviors associated with real outcomes (e.g., fraud or friction) did we not theorize or anticipate when conducting the fundamental research, and why did they predict this outcome? What applicability boundaries did our research have?

We gained immense insights far beyond what we could have gained from conducting our fundamental research alone via exploring these new questions. Yet, we would have lost much if we skipped the fundamental research stage. Due to the fundamental research stage, we understood “why” and the links between online behaviors, psychology, neuroscience, and outcomes. As such, we could better discover scientifically robust and defensible new features and learnings during the MVP stage and could move away from indefensible correlations between data and outcomes or correlations that a confounding factor in the real world influenced. We learned that one requires a scientifically robust and defensible solution to engage and provide real value to various companies that make real-time decisions based on this data.

Did you provide a working phone number in this application?

Yes No

Is the bank account number you provided, an accurate number assigned to your personal bank account?

Yes No

Are you able to afford the scheduled payment per pay period listed on your loan agreement?

Yes No

Did you provide an accurate address, city, state, and zip code in this application?

Yes No

Did you provide at least one personal reference in this application?

Yes No

Next

Figure 7. Demonstration of the First Minimum Viable Product (MVP)

4.2 Reaching Organizational Value

Over the next few years, our MVP evolved greatly: a process that led to organizational value. To reach organizational value—to solve a problem that matters to organizations—we needed to answer new research questions; that is, to achieve proof of use (Nunamaker et al., 2015).

Some questions concerned how we could maximize value for customers, such as:

- What factors influence the data's usability?
- What economic impact does the use of behavioral data to reduce fraud and friction have?
- What findings can extend to other use cases?

Some questions concerned how to make the value more consumable and scalable, such as:

- What factors influence a technology implementation's ease of use?
- What use cases are most scalable?
- How can one design a system to be scalable?

Other questions concerned how to more quickly deliver new value, such as:

- How can organizations effectively manage distributed, cross-functional teams?
- How organizations motivate people to be more productive?
- How can organizations maintain culture amid the remote work that the COVID-19 pandemic forced on organizations?

As an example, we explored finding other markets that our behavioral data (i.e., Human Analytics) could help transform. Initially, we applied our research to help reduce fraud and friction in an online lending context. In such contexts, people typically must fill out an application that may ask for questions about their identity, employment, and finances. We realized that a tangential market, the insurance applications market, could benefit from behavioral data (see Figure 8). Like with online lending, insurance applications typically comprise an online form. As such, several of our research findings from the online lending context naturally extended to an online insurance application context. However, the two contexts also differ in some important ways. Unlike online lending forms, online insurance application forms are typically longer and may also ask about assets (cars, houses, etc.) in scenarios where people might have multiple assets (e.g., multiple cars). Hence, organizations in the two contexts have different expectations and baselines for what constitutes

friction. In addition, insurance fraud behavior often differs from online lending fraud behavior; hence, sometimes, different features distinguish fraudulent from genuine users in an insurance context. We used the opportunity to expand and scale to investigate how our research extended to a new domain and identify additional insights and features that ensured the science applied to it.

While many of the early findings we uncovered still applied to value creation in newer markets, we discovered new knowledge that helped ensure we maximized our research's value for customers in new vertical markets. In every case, we built new research on our previous research. For example, the foundational research provided the theoretical and explanatory power that we used to create robust and rigorously defensible features and models. The previous MVP research in one domain (e.g., online lending) provided a foundation that we could use to move more quickly in a new domain (e.g., insurance). In addition, this previous research provided valuable insight into a particular finding's boundary conditions and provided insight into higher-level patterns that humans find more innate and, thereby, better generalize across markets. Hence, the proof of use journey often begins sequentially and linearly builds on itself. Yet, across each stage of the journey, as one crosses the chasm for each new vertical market, one can gain valuable insights.

The screenshot displays the ACME Insurance quote process, divided into four main sections: Policy Detail, Driver Detail, Vehicle Detail, and Final Review. The current view is the Policy Detail section, which includes the following elements:

- Navigation:** Buttons for "Get Quote", "Policy Info", and "Payment".
- Section Headers:** "Address Information", "Garaging/Parking Address", and "Vehicle Usage".
- Form Fields and Interactions:**
 - Two questions with "Yes/No" radio buttons: "Is your mailing address different from your parking/garage address?" and "Have you moved in the past 6 months?".
 - Text input fields for "Street Address" (4916 Dancing Lights Ave), "City" (Las Vegas), "State" (Nevada), and "Zip Code" (89130).
 - Another "Yes/No" radio button question: "Is this your permanent residence?".
 - A text input field for "Average Miles Driven per Year" (53,000).
 - A final "Yes/No" radio button question: "Is this vehicle used in the course of business other than driving to and from work?".
 - "Back" and "Continue" buttons at the bottom.
- Quote Details Panel (Right):**
 - Quote Number: 6798234
 - Down Payment: \$73.57
 - Future Payments: \$61.13
 - Total Premium: \$678.00
 - > Quote Details
- Help Center Panel (Right):**
 - Call us toll free: 1-800-555-8765
 - Email Us | Chat Now | Call Me
- Footer:** ACME Insurance | ACME Loans, ©2018 Neuro-ID, Inc. All rights reserved.

Hand-drawn lines in red and black trace the user's path through the form, starting from the top navigation, moving through the questions and input fields, and ending at the "Continue" button.

Figure 8. Neuro-ID's Technology for Capturing Complex Behavioral Data as People Interact with a Broad Range of Online Systems (source: Neuro-ID, Inc.)

5 Conclusion

Research is a journey. In this paper, we describe our research journey from basic and then more applied academic research to commercialization. We define research's organizational value (i.e., research that may have statistical and practical significance and relevance but must also solve a problem that matters to organizations). We show examples about how a new technology can provide such value relative to other existing technological solutions (e.g., strategic value, problem size, alternatives, operational value, time-to-value, and scale). We also describe and contrast key research milestones: from achieving statistical significance to achieving practical significance to reaching relevance to practice. Indeed, while important, these milestones do not guarantee the technology will result in successful commercialization. We then explain how we crossed the chasm between academic research and technology commercialization, what Nunamaker et al. (2015) referred to as the last research mile. To successfully navigate this last mile, we described how we used our initial prototype system—the MVP—to gain feedback via demonstrations to potential customers that we used to identify new product features and their relative priority. Ultimately, we developed a successful commercial product (i.e., Human Analytics) and company (i.e., Neuro-ID). In this paper, we share our story so that others who may be considering taking their academic research to commercialization can learn from our experience.

Acknowledgments

We have been fortunate to have many collaborators on various projects related to this topic. We especially want to thank Michael Byrd, Alan Dennis, Mark Grimes, Martin Hibbeln, David Kim, Manasvi Kumar, Jay Nunamaker, Jeff Proudfoot, Christoph Schneider, Markus Weinmann, Parker Williams, Ryan Wright, Aaron Zimbelman, and Mark Zimbelman. We would also like to thank our many Neuro-ID colleagues.

References

- Babbie, E. R. (1998). *The practice of social research*. Boston, MA: Wadsworth.
- Benbasat, I., & Zmud, R. W. (2003). The identity crisis within the IS discipline: Defining and communicating the discipline's core properties. *MIS Quarterly*, 27(2), 183-194.
- Bodily, R., Harris, S., Jenkins, J., Larsen, R., Sandberg, D., Stokes, S., Valacich, J., & Williams, P. (2015). A multi-experimental examination of analyzing mouse cursor trajectories to gauge subject uncertainty. In *Proceedings of the Americas Conference on Information Systems*.
- Cloutier, J., Freeman, J. B., & Ambady, N. (2014). Investigating the early stages of person perception: The asymmetry of social categorization by sex vs. age. *PLoS One*, 9(1), e84677.
- Coalition against Insurance Fraud. (2021). *The impact of fraud*. Retrieved from <https://insurancefraud.org/fraud-stats/>
- Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition*, 35(1), 15-28.
- Davenport, T. H., & Markus, M. L. (1999). Rigor vs. relevance revisited: Response to Benbasat and Zmud. *MIS Quarterly*, 23(1), 19-23.
- Dennis, A. R., & Valacich, J. S. (2001). Conducting experimental research in information systems. *Communications of the Association for Information Systems*, 7, 1-41.
- Freeman, J., Dale, R., & Farmer, T. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*. Retrieved from <https://doi.org/10.3389/fpsyg.2011.00059>
- Grimes, G. M., & 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*.
- Grover, V., Lindberg, A., Benbasat, I., & Lyytinen, K. (2020). The perils and promises of big data research in information systems. *Journal of the Association for Information Systems*, 21(2), 268-291.
- Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biol*, 13(3), e1002106.
- Hibbeln, M., Jenkins, J. L., Schneider, C., Valacich, J. S., & Weinmann, M. (2017). How is your user feeling? Inferring emotion through human-computer interaction devices. *MIS Quarterly*, 41(1), 1-21.
- Horwitz, R., Kreuter, F., & Conrad, F. (2017). Using mouse movements to predict web survey response difficulty. *Social Science Computer Review*, 35(3), 388-405.
- Jenkins, J. L., & Valacich, J. S. (2015). Behaviorally measuring ease-of-use by analyzing users' mouse cursor movements. In *Proceedings of the Fourteenth Annual Workshop on HCI Research in MIS*.
- Jenkins, J. L., Valacich, J. S., Zimbelman, A., & Zimbelman, M. (Forthcoming). Detecting noncompliant behavior in organizations: How responses and behaviors reveal the risk. *Journal of Management Information Systems*.
- Lee, A. S., & Hubona, G. S. (2009). A scientific basis for rigor in information systems research. *MIS Quarterly*, 33(2), 237-262.
- McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science*, 19(1), 22-24.
- Mohajeri, K., Mesgari, M., & Lee, A. S. (2020). When statistical significance is not enough: Investigating relevance, practical significance, and statistical significance. *MIS Quarterly*, 44(2), 525-559.
- Moore, G. A. (1999). *Crossing the chasm, marketing and selling high-tech products to mainstream customer*. New York, NY: HarperCollins.
- Nunamaker Jr., J. F., Briggs, R. O., Derrick, D. C., & Schwabe, G. (2015). The last research mile: Achieving both rigor and relevance in information systems research. *Journal of Management Information Systems*, 32(3), 10-47.

- Olsen, D. (2015). *The lean product playbook: How to innovate with minimum viable products and rapid customer feedback*. Hoboken, NJ: Wiley.
- Orlikowski, W. J., & Iacono, C. S. (2001). Desperately seeking the “IT” in IT research—a call to theorizing the IT artifact. *Information Systems Research*, 12(2), 121-134.
- Quote Investigator (n.d.). *Life is a journey, not a destination*. Retrieved from quoteinvestigator.com/2012/08/31/life-journey/
- Ries, E. (2011). *The lean startup*. New York, NY: Crown Publishing.
- Robey, D. (2003). Identity, legitimacy and the dominant research paradigm: An alternative prescription for the IS discipline: A response to Benbasat and Zmud’s call for returning to the IT artifact. *Journal of the Association for Information Systems*, 4(7), 352-359.
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.
- Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The sociotechnical axis of cohesion for the IS discipline: Its historical legacy and its continued relevance. *MIS Quarterly*, 43(3), 695-720.
- Surendranath, A. (2015). Computers can detect your anger. *Medindia*. Retrieved from <https://www.medindia.net/news/healthinfocus/computers-can-detect-your-anger-156212-1.htm>
- Valacich, J. S., & Jenkins, J. L. (2019). *Fraudulent application detection system and method of use*. U.S. Patent Publication No. 20160328572.
- Valacich, J. S., & Jenkins, J. L. (2020). *Automated detection method for insider threat*. U.S. Patent Publication No. 20200163605.
- Weinmann, M., Valacich, J. S., Schneider, C., Jenkins, J. J., & Hibbeln, M. (Forthcoming). *The path of the righteous: Using trace data to understand fraud decisions in real time*. *MIS Quarterly*.
- Wojnowicz, M. T., Ferguson, M. J., Dale, R., & Spivey, M. J. (2009). The self-organization of explicit attitudes. *Psychological Science*, 20(11), 1428-1435.

About the Authors

Joseph (Joe) S. Valacich is the Gary M. Munsinger Professor of Entrepreneurship and Innovation in the Eller College of Management at the University of Arizona. His research focuses on human-computer interaction (HCI), deception detection, cyber security, and e-business. He has published more than 100 scholarly articles in top journals such as the *Academy of Management Journal*, *The Accounting Review*, *Information Systems Research*, *Journal of Applied Psychology*, *Journal of the Association for Information Systems*, *Journal of Management Information Systems*, *MIS Quarterly*, *Management Science*, and *Organizational Behavior and Human Decision Processes*.

Jeffrey L. Jenkins is an Associate Professor of Information Systems at Brigham Young University. He has developed, extensively validated, and patented algorithms to detect deception, cognitive conflict, uncertainty, and emotional reactions while people are using electronic input devices. His research also explores methodologies for detecting and mitigating insider security threats. His research has been published in the top information systems journals such as the *Journal of Management Information Systems*, *MIS Quarterly*, *Information Systems Research*, and *Journal of the Association of Information Systems*.

Copyright © 2021 by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints or via e-mail from publications@aisnet.org.



Editor-in-Chief

<https://aisel.aisnet.org/thci/>

Fiona Nah, Missouri University of Science and Technology, USA

Advisory Board

Izak Benbasat, University of British Columbia, Canada
John M. Carroll, Penn State University, USA
Phillip Ein-Dor, Tel-Aviv University, Israel
Dennis F. Galletta, University of Pittsburgh, USA
Shirley Gregor, National Australian University, Australia
Elena Karahanna, University of Georgia, USA
Paul Benjamin Lowry, Virginia Tech, USA

Jenny Preece, University of Maryland, USA
Gavriel Salvendy, University of Central Florida, USA
Ben Shneiderman, University of Maryland, USA
Joe Valacich, University of Arizona, USA
Jane Webster, Queen's University, Canada
K.K. Wei, Singapore Institute of Management, Singapore
Ping Zhang, Syracuse University, USA

Senior Editor Board

Torkil Clemmensen, Copenhagen Business School, Denmark
Fred Davis, Texas Tech University, USA
Gert-Jan de Vreede, University of South Florida, USA
Soussan Djamasbi, Worcester Polytechnic Institute, USA
Traci Hess, University of Massachusetts Amherst, USA
Shuk Ying (Susanna) Ho, Australian National University, Australia
Matthew Jensen, University of Oklahoma, USA
Atreyi Kankanhalli, National University of Singapore, Singapore
Jinwoo Kim, Yonsei University, Korea
Eleanor Loiacono, College of William & Mary, USA
Anne Massey, University of Massachusetts Amherst, USA
Gregory D. Moody, University of Nevada Las Vegas, USA

Lorne Olman, Claremont Graduate University, USA
Stacie Petteer, Baylor University, USA
Choon Ling Sia, City University of Hong Kong, Hong Kong SAR
Heshan Sun, University of Oklahoma, USA
Kar Yan Tam, Hong Kong U. of Science & Technology, Hong Kong SAR
Chee-Wee Tan, Copenhagen Business School, Denmark
Dov Te'eni, Tel-Aviv University, Israel
Jason Thatcher, Temple University, USA
Noam Tractinsky, Ben-Gurion University of the Negev, Israel
Viswanath Venkatesh, University of Arkansas, USA
Mun Yi, Korea Advanced Institute of Science & Technology, Korea
Dongsong Zhang, University of North Carolina Charlotte, USA

Editorial Board

Miguel Aguirre-Urreta, Florida International University, USA
Michel Avital, Copenhagen Business School, Denmark
Gaurav Bansal, University of Wisconsin-Green Bay, USA
Ricardo Buettner, Aalen University, Germany
Langtao Chen, Missouri University of Science and Technology, USA
Christy M.K. Cheung, Hong Kong Baptist University, Hong Kong SAR
Tsai-Hsin Chu, National Chiayi University, Taiwan
Cecil Chua, Missouri University of Science and Technology, USA
Constantinos Coursaris, HEC Montreal, Canada
Michael Davern, University of Melbourne, Australia
Carina de Villiers, University of Pretoria, South Africa
Gurpreet Dhillon, University of North Carolina at Greensboro, USA
Alexandra Durcikova, University of Oklahoma, USA
Andreas Eckhardt, University of Innsbruck, Austria
Brenda Eschenbrenner, University of Nebraska at Kearney, USA
Xiaowen Fang, DePaul University, USA
James Gaskin, Brigham Young University, USA
Matt Germonprez, University of Nebraska at Omaha, USA
Jennifer Gerow, Virginia Military Institute, USA
Suparna Goswami, Technische U.München, Germany
Camille Grange, HEC Montreal, Canada
Juho Harami, Tampere University, Finland
Khaled Hassanein, McMaster University, Canada
Milena Head, McMaster University, Canada
Netta Iivari, Oulu University, Finland
Zhenhui Jack Jiang, University of Hong Kong, Hong Kong SAR
Richard Johnson, Washington State University, USA
Weiling Ke, Southern University of Science and Technology, China

Sherrie Korniak, Memorial U. of Newfoundland, Canada
Yi-Cheng Ku, Fu Chen Catholic University, Taiwan
Na Li, Baker College, USA
Yuan Li, University of Tennessee, USA
Ji-Ye Mao, Renmin University, China
Scott McCoy, College of William and Mary, USA
Tom Meservy, Brigham Young University, USA
Stefan Morana, Saarland University, Germany
Robert F. Otondo, Mississippi State University, USA
Lingyun Qiu, Peking University, China
Sheizaf Rafaeli, University of Haifa, Israel
Rene Riedl, Johannes Kepler University Linz, Austria
Lionel Robert, University of Michigan, USA
Khawaja Saeed, Wichita State University, USA
Shu Schiller, Wright State University, USA
Christoph Schneider, IESE Business School, Spain
Theresa Shaft, University of Oklahoma, USA
Stefan Smolnik, University of Hagen, Germany
Jeff Stanton, Syracuse University, USA
Chee-Wee Tan, Copenhagen Business School, Denmark
Horst Treiblmaier, Modul University Vienna, Austria
Ozgun Turetken, Ryerson University, Canada
Wietske van Osch, HEC Montreal, Canada
Wei-quan Wang, City University of Hong Kong, Hong Kong SAR
Dezhi Wu, University of South Carolina, USA
Fahri Yetim, FOM U. of Appl. Sci., Germany
Cheng Zhang, Fudan University, China
Meiyun Zuo, Renmin University, China

Managing Editor

Gregory D. Moody, University of Nevada Las Vegas, USA