Chapter 30

Web Usability Testing With Concurrent fNIRS and Eye Tracking

Siddharth Bhatt[1], Atahan Agrali[1], Kevin McCarthy[1], Rajneesh Suri[1], Hasan Ayaz[1,2,3]

[1]Drexel University, Philadelphia, PA, United States; [2]University of Pennsylvania, Philadelphia, PA, United States; [3]Children's Hospital of Philadelphia, Philadelphia, PA, United States

INTRODUCTION

Digitization in the financial sector has given rise to a new industry in its own right: fintech. Within financial services, wealth management has been one of the prime beneficiaries of the digitization wave. According to the consulting firm KPMG,[1] global investment in fintech companies increased from US$4 billion in 2012 to US$47 billion in 2015. Today every major player in the wealth management industry offers digital portfolio management services to its clientele. With all major players having digitized their services, the competitive advantage no longer lies in firms merely having digital capabilities; rather, the winners will be firms which can provide the best possible wealth management experience by designing web interfaces that give consumers multiple features on the website along with ease of use. Given customers' need to self-manage their portfolios, the quality of the web interface offered by financial firms is becoming decisive. Thus creating and maintaining web interfaces which allow customers to have a pleasant experience on the firm's website have become a top priority for wealth management firms.

Unfortunately, firms often learn about issues with their websites only after such websites are launched. Poorly designed websites often signal an inferior quality of the brand(s) associated with the website[2]; they make it difficult for customers to navigate with ease and hence make it difficult to process information on the website. There is sufficient evidence to suggest that when people find it difficult to process product information, their evaluation of the brand becomes negative.[3,4] According to an internet performance research firm, 95% of surveyed IT executives felt that poor websites resulted in loss of revenue,[5] hence it is important that web platforms are designed in ways that enable consumers to attain their wealth management goals with the least amount of friction. To that end, it is becoming common to involve end users in both development and testing of such websites.[6] However, the methods employed to elicit users' feedback remain subjective and rudimentary. Most web usability studies employ qualitative and/or quantitative surveys to collect users' assessment of a website, but self-reports and similar techniques are severely limited in their ability to bring forth users' insights objectively. Further, these measures fail to highlight specific areas that might need improvement, thus severely limiting a company's ability to make changes to the website to reflect user feedback.

Such shortcomings could be overcome if the fintech industry took advantage of the new developments in neuroergonomic research. Drawing on the findings from neuroergonomics, this research proposes a novel multimodal approach to assessing how websites could be tested by making end users perform different types of ecologically valid tasks in a systematic and repeated manner while behavioral activity and user physiological measures are logged. The approach used in this research combines the strengths of neurophysiological methods with traditional survey tools to help elicit rich and specific feedback that customary tools fail to provide. Recent developments in neuroergonomics have provided methodologies which can benefit web assessment studies. Neuroergonomics is defined as the study of the human brain in relation to performance while at work and in everyday settings.[7,8] Within this field, the portable and wearable brain imaging technology of functional near-infrared spectroscopy (fNIRS) has emerged as a practical sensor modality to capture cortical oxygenation changes in the brain. fNIRS has been used to assess variety of cognitive functions, such as working memory, attention, and decision-making.[9–11] These sensors have been used in real-world field settings to assess workload of air traffic controllers using a new generation of user interfaces, and in training unmanned vehicle operators with customized flight simulator interfaces.[12] More recently, miniaturized and battery-operated wireless versions of fNIRS have been developed[13] and used to monitor individuals in ambulatory and outdoor settings, such as when participants compare spatial navigation aid interfaces.[14] fNIRS as a noninvasive and portable technology can be used

Neuroergonomics. https://doi.org/10.1016/B978-0-12-811926-6.00030-0

with other data collection methods, and provides cognitive and affective-state-related brain activity data.[12,15–17] In this study we aimed to investigate usability of a new web-based investment tool via simultaneous fNIRS and eye-tracking measurements. The brain and body measures were used in addition to traditional self-reported survey measures and task performance measures on the website to provide a comprehensive assessment of modulated task difficulty and user engagement.

METHODS

This multimodal study combined four different metrics (survey measures, performance measures, neural correlates, and eye-tracking metrics) to assess how users perceived a new wealth management web platform.

The 37 participants in the study (46% female, mean age = 47. 54) were composed of three groups of users: first-time users of the website, the company's current clients, and its employees . Users interacted with the prototype of the new website and performed several investment-related tasks that they would normally undertake on a wealth management platform.

To assess the website an experimental protocol with two levels of task difficulty was devised for all task types: easy tasks were designed so that users could readily find the information needed to complete the task on the same webpage where they started, and more challenging, difficult tasks required users to search deeper for information necessary to complete the tasks. While easy tasks required little navigation and planning, difficult tasks required search, working memory, and planning, thus testing how easy it was to perform the given task on the website. Navigation included going to different webpages on the site, changing the format of the financial information presented, and using other functions of the webpage. Users went through 37 different tasks on the website, and all four metrics were collected for each task. Contrasting the four dimensions of metrics on task difficulty provided a rigorous test of the website and the approach.

RESULTS

Analyses of data across each of the four modalities were conducted using task difficulty (easy versus difficult) and group (first-time users versus clients versus employees) as fixed factors in a mixed-model effect analysis. Together these metrics revealed significant differences between easy and difficult tasks across all task domains on the website.

Survey Measures

The survey measures (Fig. 30.1) comprised a single item assessing participants' perception of the difficulty of all tasks on a nine-point Likert scale (1 = very difficult, 9 = very easy). The linear mixed-effect model revealed a significant effect of difficulty on the measure ($F(1, 351) = 131.05$, $P = .00$), indicating that the participants could differentiate between easy and difficult tasks. There was no main effect of group ($F(2, 32) = .22$, $P = .80$) or interaction between group and task difficulty ($F(2, 351) = .12$, $P = .88$). Participants found it easier to find information for easy tasks ($M_{easy} = 7.68$, S.D. = 1.79) compared to difficult tasks ($M_{difficult} = 6.20$, S.D. = 2.48).

Behavioral Measures

Accuracy on the task was used as the measure of user performance on the website (Fig. 30.2). The user obtained a score of 1 (in contrast to 0) on the task if he/she succeeded in accurately performing an investment-related task on the new platform. The linear mixed-effect model revealed a significant effect of difficulty on accuracy ($F(1, 1070) = 66.20$, $P = .00$), indicating that participants could differentiate between easy and difficult tasks. There was no main effect of group ($F(2, 34.2) = 1.78$, $P = .18$) or interaction between group and task difficulty ($F(2, 1070) = .51$, $P = .59$). Participants had higher accuracy in performing easy tasks (86%) compared to difficult tasks (66.2%).

Eye-Tracking Measures

Fixation count is the most commonly used eye-tracking metric to assess users' visual experience (Fig. 30.3). The number of fixations on screens with the tasks was used as the dependent variable in the mixed model. The linear mixed-effect model revealed a significant effect of difficulty on fixation count ($F(1, 1301.4) = 207.68$, $P = .00$). There was no main effect of group ($F(2, 31.9) = .30$, $P = .73$) or interaction between group and task difficulty ($F(2, 1301.4) = 2.19$, $P = .11$). Participants had fewer fixations for easy tasks (M = 25.27, S.D. = 21.33) compared to difficult tasks (M = 50.63, S.D. = 31.04).

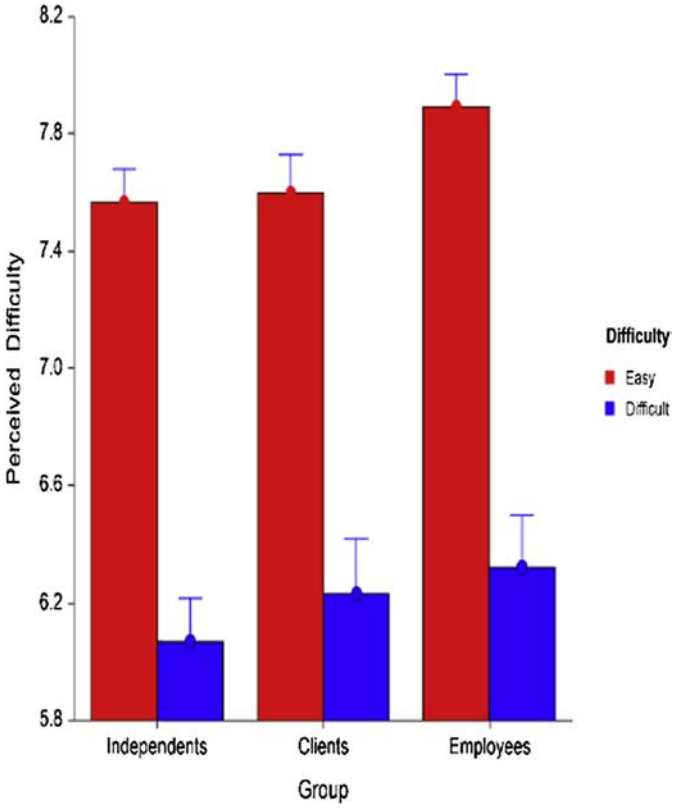

FIGURE 30.1 Survey measures (perceived difficulty) for easy versus difficult tasks.

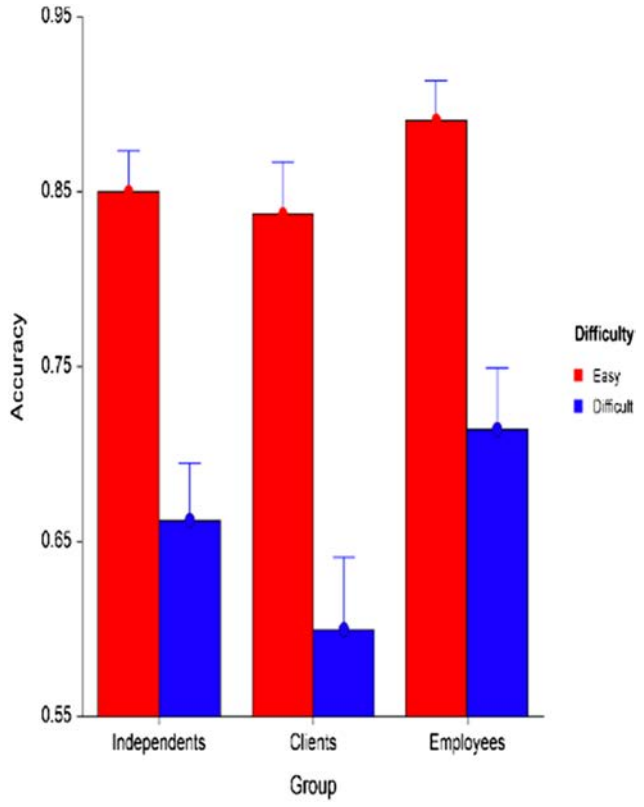

FIGURE 30.2 Performance measures (accuracy) for easy versus difficult tasks.

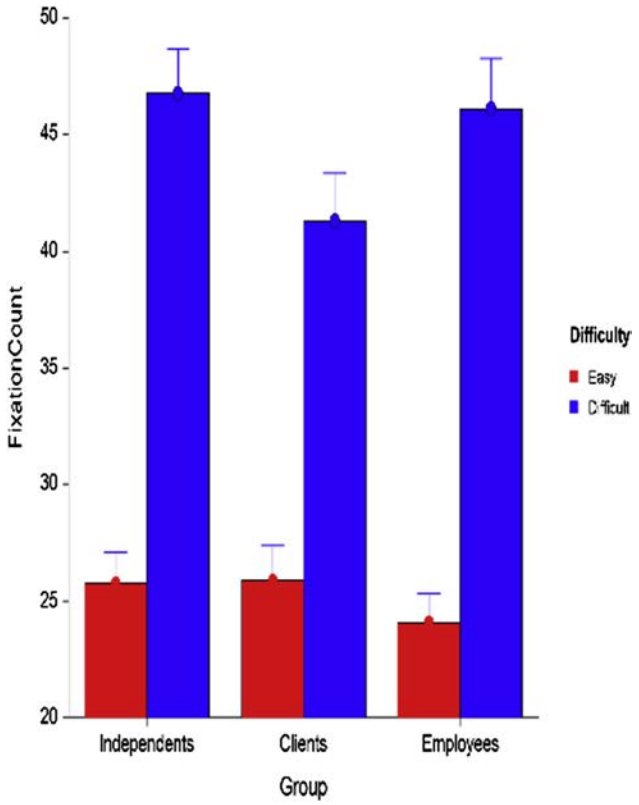

FIGURE 30.3 Eye-tracking measures (fixation count) for easy versus difficult tasks.

Neural Measures

While survey and performance measures helped distinguish between easy and difficult tasks, neural measures revealed additional insights that survey and performance measures could not uncover (Fig. 30.4). An analysis of the neural activity revealed patterns of brain activity that help explain why participants found the difficult tasks more challenging. Across all four quadrants, participants in all three groups displayed much higher mental effort (as measured by oxygenated hemoglobin changes) for difficult tasks compared to easy tasks (Q1: $F(1, 1009.2) = 14.07$, $P = .00$; Q2: $F(1, 1022.6) = 6.32$, $P = .01$; Q3: $F(1, 954.2) = 9.11$, $P = .00$; Q4: $F(1, 992) = 5.15$, $P = .02$). There was no main effect of group or interaction between group and task difficulty.

CONCLUSION

Given the large investment in web platforms, wealth management companies need better methods of testing their platforms before launching them. Traditionally, web usability studies have relied on survey measures as the only method to assess users' perceptions of a website. However, surveys are deficient in uncovering critical insights, and this concern has been echoed in the literature for a long time. In this research we proposed and tested a new paradigm in web usability which incorporates a multimodal approach. Using a systematic and repetitive block-design-based procedure coupled with technologies such as brain imaging and eye tracking, we gained objective measures of engagement and additional insights such as areas of the website which users found difficult to navigate, issues with symbolism, and effectiveness of visualization and delivery of data, among others. Put together, the four metrics lead to a more holistic and rigorous testing of the website which improves the provider's confidence in offering the new website to its clientele.

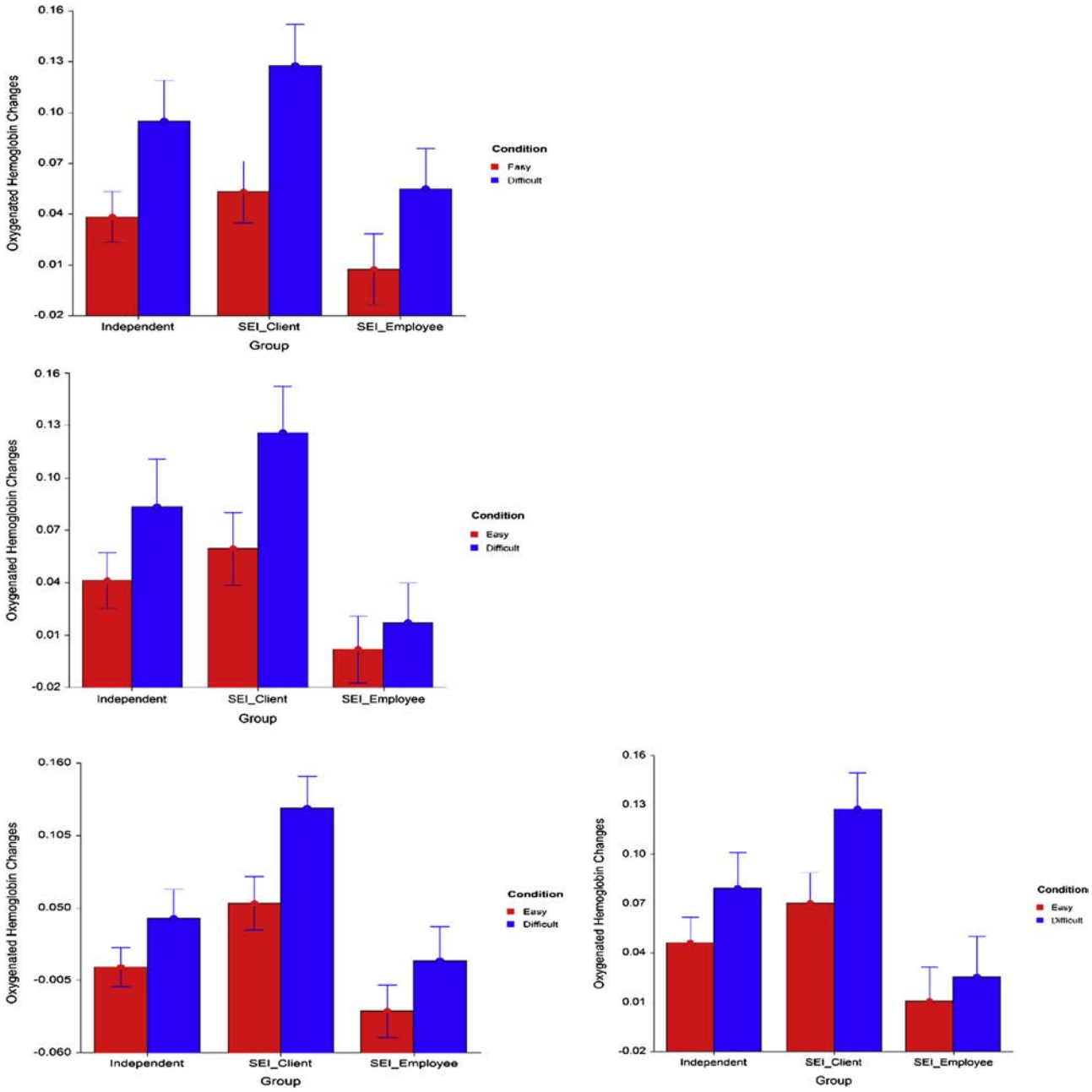

FIGURE 30.4 Neural measures (oxygenated hemoglobin changes in Q1, Q2, Q3, and Q4) for easy versus difficult tasks.

REFERENCES

1. KPMG International Cooperative. *The pulse of fintech Q4 2016*. 2017. Retrieved from: https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2017/02/pulse-of-fintech-q4-2016.pdf.
2. Gregg DG, Walczak S. The relationship between website quality, trust and price premiums at online auctions. *Electronic Commerce Research* 2010;**10**(1):1–25.
3. Reber R, Schwarz N, Winkielman P. Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Personality and Social Psychology Review* 2004;**8**(4):364–82.
4. Song H, Schwarz N. If it's hard to read, it's hard to do: processing fluency affects effort prediction and motivation. *Psychological Science* 2008;**19**(10):986–8.

5. Dyn. *World's it executives Agree: Vital online Sales Hurt by poor website performance up to 50 percent of the time (web log post)*. March 31, 2015. Retrieved October 22, 2017, from: http://dyn.com/blog/worlds-it-executives-agree-vital-online-sales-hurt-by-poor-website-performance-up-to-50-percent-of-the-time/.
6. Kujala S. User involvement: a review of the benefits and challenges. *Behaviour and Information Technology* 2003;**22**(1):1–16.
7. Parasuraman R. Neuroergonomics: research and practice. *Theoretical Issues in Ergonomics Science* 2003;**4**(1–2):5–20.
8. Parasuraman R, Rizzo M, editors. *Neuroergonomics: the brain at work*. Oxford University Press; 2008.
9. Ayaz H, Shewokis PA, Bunce S, Izzetoglu K, Willems B, Onaral B. Optical brain monitoring for operator training and mental workload assessment. *Neuroimage* 2012;**59**(1):36–47.
10. Di Domenico SI, Rodrigo AH, Ayaz H, Fournier MA, Ruocco AC. Decision-making conflict and the neural efficiency hypothesis of intelligence: a functional near-infrared spectroscopy investigation. *NeuroImage* 2015;**109**:307–17.
11. McKendrick R, Ayaz H, Olmstead R, Parasuraman R. Enhancing dual-task performance with verbal and spatial working memory training: continuous monitoring of cerebral hemodynamics with NIRS. *NeuroImage* 2014;**85**:1014–26.
12. Ayaz H, Onaral B, Izzetoglu K, Shewokis PA, McKendrick R, Parasuraman R. Continuous monitoring of brain dynamics with functional near infrared spectroscopy as a tool for neuroergonomic research: empirical examples and a technological development. *Frontiers in Human Neuroscience* 2013;**7**.
13. McKendrick R, Parasuraman R, Ayaz H. Wearable functional near infrared spectroscopy (fNIRS) and transcranial direct current stimulation (tDCS): expanding vistas for neurocognitive augmentation. *Frontiers in Systems Neuroscience* 2015;**9**.
14. McKendrick R, Parasuraman R, Murtza R, Formwalt A, Baccus W, Paczynski M, Ayaz H. Into the wild: neuroergonomic differentiation of hand-held and augmented reality wearable displays during outdoor navigation with functional near infrared spectroscopy. *Frontiers in Human Neuroscience* 2016;**10**.
15. Hill AP, Bohil CJ. Applications of optical neuroimaging in usability research. *Ergonomics in Design* 2016;**24**(2):4–9.
16. Mandrick K, Chua Z, Causse M, Perrey S, Dehais F. Why a comprehensive understanding of mental workload through the measurement of neurovascular coupling is a key issue for Neuroergonomics? *Frontiers in Human Neuroscience* 2016;**10**.
17. Mehta RK, Parasuraman R. Neuroergonomics: a review of applications to physical and cognitive work. *Frontiers in Human Neuroscience* 2013;**7**.

FURTHER READING

1. Dapp T, Slomka L, Hoffmann R. *Fintech—the digital (r) evolution in the financial sector*. Frankfurt am Main: Deutsche Bank Research; 2014.
2. Mackenzie A. The fintech revolution. *London Business School Review* 2015;**26**(3):50–3.