Appreciating the Performance of Neuroscience Mining in NeuroIS Research: A Case Study on Consumer's Product Perceptions in the Two UI Modes—Dark UI vs. Light UI

Francis Joseph Costello
Sungkyunkwan University
joe.costehello@gmail.com

Min Gyeong Kim
Sungkyunkwan University
Mgkim9541@gmail.com

Kun Chang Lee
Sungkyunkwan University
kunchanglee@gmail.com

Abstract

The goal of the current study was to explore the potential of neuroscience mining (NSM) for comprehending NeuroIS paradigms. NSM is an interdisciplinary field that combines neuroscience and business mining, which is the application of big data analytics, computational social science, and other areas to business problems. NSM makes it possible to apply predictive models to NeuroIS datasets, such as machine learning and deep learning, to find intricate patterns hidden by conventional regression-based analysis. First, we predicted 28 individual EEG power spectra separated brainwave data using a Random Forest (RF) model. Next, we used NSM to precisely predict how consumers would perceive a product online, depending on whether a light or dark user interface (UI) mode was used. The model was then used to extract more precise results that could not be obtained using more conventional linear-based analytical models using sensitivity analysis. The benefits of using NSM in NeuroIS research are as follows: (1) it can relieve the burden of the three-horned dilemma described by Runkel and McGrath; (2) it can enable more temporal data to be directly analyzed on the target variables; and (3) sensitivity analysis can be performed on a condition/individual basis, strengthening the rigor of findings by reducing sample bias that can be lost in grand averaging of data when analyzed with methods like GLM.

1. Introduction

In their review of NeuroIS, vom Brocke and colleagues (2020) noted that the field had made significant strides since its founding in 2007. The authors of this review article identified four major challenges that they felt were essential for the field's future advancement within the larger field of information systems (IS). The authors noted potential improvements that NeuroIS might bring to established IS theories, like those relating to the creation and use of IS. However, no mention of improvements to methodological approaches was made. The use of NeuroIS tools, EEG (Electroencephalogram), fNIRS (Functional near-infrared spectroscopy), and Eye Tracking, for example, can generate enormous amounts of data due to the temporal velocity with which sensors can acquire neurophysiological data, as noted by Riedl and Léger in their book on the fundamentals of NeuroIS (2016). Despite this, the majority of the studies cited in three significant reviews of the NeuroIS literature (Fischer et al., 2019; Riedl et al., 2017; Xiong & Zuo, 2020) show that the majority of published articles used conventional analysis techniques (for example, linear models (GLM) and regressions such as ANOVA) to test research questions.

These methods have clear advantages, perform well when testing null hypotheses, and give strong signals about the potential root causes of a research question. However, we found three problems with these tactics. First, they are constrained because they need to be used with smaller subsets of data extracted from a larger pool of data, typically via average pooling into a single scalar value for use with, for instance, GLM. This ignores potential additional causal inference indicators that could be present in the intricate larger temporal data set used for averaging. The models used to analyze the neurophysiological data are built as representations of potential activity during a controlled event and then assumed to represent self-reported behavioral responses. This is the second criticism

URI: https://hdl.handle.net/10125/102754
978-0-9981331-6-4

leveled at these methodologies. Analytical models built on the GLM framework do not directly analyze or forecast neurophysiological data from given behavioral responses. As a result, it is challenging to rigorously establish a causal link between the dependent variable and the participant's neurophysiological data. Finally, Runkel and McGrath's three-horned dilemma applies to NeuroIS researchers because they primarily focus on conducting research in a laboratory setting (Runkel & McGrath, 1972; abbreviated as RMTHD hereafter). All methodological approaches, according to RMTHD, have some control, realism, and/or generality limitations. In addition, researchers encounter a problem that causes them to lose one or two of these three crucial research dimensions when they try to emphasize one of them, claim Chang et al. (2014). For instance, control is frequently stressed in NeuroIS studies at the expense of authenticity and generalizability.

By using a big data lens to analyze the neurophysiological data found in NeuroIS, the current study aims to address these problems. We suggest the use of a neuroscience mining (NSM) approach that incorporates advanced data mining and analytics as a supplementary analytical approach within the field of NeuroIS because the era of big data is opening up new opportunities for conducting high-quality research on social science phenomena (Chang et al., 2014). NSM is an interdisciplinary field that combines neuroscience and business mining, which is the application of big data analytics, computational social science, and other fields to business problems. To find complex patterns hidden by conventional regression-based analysis, NSM enables the application of predictive models such as machine learning (Witten et al., 2016) and deep learning (Lecun et al., 2015) to NeuroIS datasets. The following research question is what we'd like to address as a result:

RQ1: How can NSM help NeuroIS research circumvent RMTHD in their experimentation?

RQ2: Can NSM help overcome the potential loss of generality and realism when control is emphasized in NeuroIS?

This paper emphasizes the significance of using an analytical framework based on NSM that allows for the direct examination of neurophysiological data derived from behavioral data and further investigation of the target problem. We also used sensitivity analysis to delve deeper into the data and draw more immediate conclusions from our predictive models (Cortez & Embrechts, 2013). The findings are a part of a larger, ongoing study that looks at consumer motivations, particularly actions like approach motivation toward products and visual aesthetics (Ye et al., 2020). The results of this study also offer a dynamic picture of how neurophysiological and self-reported data can interact within a single predictive model from a methodological standpoint. Additionally, because several models were run, we could respond to our RQs without any ambiguity in the outcomes, eliminating any one-shot bias that might have been attained when only one predictive model was used.

This essay is structured as follows: after presenting a review of the aesthetics-focused NeuroIS literature, we examine the case study on UI modes. We then go into the methodological methods for data collection and predictive modeling. The empirical results of applying NSM techniques are provided in the following section. After looking into the findings' limitations and potential future directions, we finally discuss their implications.

2. Literature Survey

2.1. NeuroIS and IS Aesthetics

IS is tasked with combining various data sources to create knowledge that can help users make decisions. As a result, the literature that is currently available has concentrated on different facets of IS use and design, such as aesthetics on websites (Riaz et al., 2018; Ye et al., 2020). The use of NeuroIS has recently deepened our understanding of aesthetics. Additionally, neuro-based studies frequently compile a combination of self-reported and neurophysiological data to provide information on IS paradigms, pushing the boundaries of knowledge in behavioral and design-oriented information systems research (vom Brocke et al., 2020).

Additional studies have shown how visual aesthetics may influence the establishment of a consumer's impression. According to the research, expressive aesthetics may increase arousal, while classical aesthetics may decrease it (Ye et al., 2020). Furthermore, it has been shown that gender distinctions can influence how website aesthetics are perceived. For instance, Nissen and Krampe (2021) demonstrated that the PFC's brain activity was increased by visually appealing, practical, and easy-to-use e-commerce websites (prefrontal cortex). The use of NeuroIS in IS research has been shown to continually increase our understanding of the effects of website design and aesthetics, making it a suitable and reliable methodology for this research topic.

The complexity of the brain's various functions is still unknown because it is compared to a "black box." But recent developments in brain imaging technology have opened up new research avenues, particularly for IS paradigms (Dimoka et al., 2011). For example, the measurement of psychophysiological postsynaptic electric potentials on the brain's surface is possible using the electroencephalogram (EEG), a popular NeuroIS technique (Kuan et al., 2014). These electric potentials can identify brainwaves as 100 cycles per second sound wave patterns. Since EEG can record up to 2,000 different measurements per second, it is one of the best tools for determining brain activation after stimulus exposure (Riedl & Léger, 2016).

2.2. Case Study: Consumer's Product Perceptions and UI Modes

It is common knowledge that consumer perceptions are influenced by how products are marketed (Fiore et al., 2000). As an illustration, research has demonstrated that altering certain aspects of product images can significantly affect how consumers interact with products and the brands they represent online (Fiore et al., 2005). According to additional research, simple designs with straightforward search paths promote online shoppers' enjoyable and compelling shopping experiences (Elliot & Fowell, 2000). Aslam also found that in a marketing context, dark colors exude an aura of sophistication, expensiveness, and aesthetic appeal (Aslam, 2006). Because of these factors, perceived attractiveness was considered to be crucial.

Giving customers product images that are the most accurate representations of the intended use is another factor for online retailers to take into account. It has been shown that doing so can increase online shoppers' intent to purchase (Then & Delong, 1999). A need for a product will emerge if its intended use is thought to be necessary by the consumer. Therefore, we concentrated on needs and operationalized them through the usefulness of the product. Accordingly, a product's utility is defined as its capacity to satisfy customer needs (Cooper, 1979; Dahl et al., 1999; Henard & Szymanski, 2001), and this study includes utility as a dependent variable. These two factors were chosen for analysis even though other factors might have been more appropriate for understanding how UI affects consumer perceptions of online products.

These are common knowledge in the world of online retail, but the growth of mobile shopping has created new difficulties. Researchers in human-computer interaction claim that studying user interfaces for mobile devices gives them insight into a very different environment than when people shop online using personal computers (Lee & Benbasat, 2004). Therefore, online retailers have had to alter their websites and user interfaces to accommodate the new wave of mobile commerce customers (Ngai & Gunasekaran, 2007). A new design featuring customization for device UI in the form of a Dark UI mode was introduced in 2019 with the release of Apple iOS 13 and Android 10. Users of these systems and devices have had the option to alternate between the traditional light UI mode and the new dark UI mode ever since it was released. Researching the effects on consumer behavior when faced with product presentation within the dark UI mode in the context of online shopping has become necessary as a result of the introduction of the dark UI mode.

3. Methodology

3.1. Experiment, data collection, and preprocessing

To obtain participants, a notice on the college's homepage was utilized. Twenty-eight South Korean undergraduates were conveniently sampled ($M_{age} = 24.61$; 18 males and 10 females) in the experiment for a fee. We used a homogenous sample of undergraduate students to minimize differences in social class, cultural, and educational factors. The weekly online shopping frequency of participants was about 1.58, and it was found that females shopped about 0.7 times more frequently than male (M_{female}: 2.06, M_{male}: 1.38). All participants were right-handed and had no neurological issues. In addition, approval from the IRB's ethics committee was obtained prior to the commencement of the experiments. In addition, approval from the IRB's ethics committee was obtained prior to the commencement of the experiments. The participants were randomly assigned to dark UI or light UI modes. The stimuli were presented in the form of a tablet on a 29-inch computer screen to simulate shopping on a tablet. We used the software PsychoPy to run the stimuli on a computer, which also recorded the timestamps of the stimuli into the data (Peirce, 2007). Except for the different user interfaces, each condition's stimuli design was identical. Five generic product categories were used for testing: fashion, electronic home appliances, computers, jewelry, and furniture. Each category displayed fifteen images for five seconds each. Randomizing the order of the images ensured that no two participants saw the same order of images during the experiments. Using two questions, behavioral data was gathered. We used a three-point scale to assess perceived attractiveness and utility in response to the statements "this product appears to be attractive to me" and "this product appears to be useful to me" (1 = strongly disagree, neither disagree nor agree, 3 = strongly agree).

This study analyzed EEG data using the Emotive EPOC+ Model 2.0 (EEG) headset with 14 channels and two reference electrodes at the P3 and P4 sites. The electrode signals were amplified after going through a low-pass filter with a cutoff frequency of 83 Hz, a preamplifier, and a high-pass filter with a cutoff frequency of 0.16 Hz. The Fast Fourier Transformation (FFT) (Hamming Window), which breaks down each signal into a function of frequencies, was used to extract the EEG spectral bands (Theta, Alpha, Beta low, Beta high, and Gamma). A vector space was created after the EEG data underwent preprocessing to be used with the NSM predictive models.

The experimental dataset initially consists of timestamp sequences that are unable to forecast the DVs separately. As previously mentioned, we used a specific duration for each stimulus when recording brain waves. A total of 40 timestamps were recorded throughout the procedure, which were then divided using 14 sensors and 5 brain waves. With the exception of the DVs, the resulting dataset contained 2800 features. Additionally, samples with fewer than 40 timestamps were disregarded because we were unable to properly match the vector space for learning (including missing values).

3.2. Neuroscience Mining Models Used

To determine the preliminary results we could obtain, we evaluated our data using a variety of machine learning and deep learning families. We could not, however, include all of the testing results due to space limitations. The results of our testing using tree-based algorithms are presented instead. This is because they were the most successful overall at this task, which is in line with earlier studies that used predictive machine learning to analyze EEG data (see Barral et al., 2017; Buettner et al., 2020). Overall, three tree-based algorithms—Random Forest (RF), XGBoost (XGB), and Decision Tree—were used for analysis (DT).

3.3. Data-sampling Sensitivity Analysis (DSSA)

In this paper, we adapt the data-based sensitivity analysis, or DSSA, in a modified form from Cortez and Embrechts (2013). We must ascertain the sensitivity of these findings in order to aid in the removal of sample bias since the characteristics of the data have a significant impact on data-driven predictive models. Additionally, this method enables participant analysis on an individual basis to help uncover variations in participants' perceptions of each product category. This is accomplished by conducting an analysis of product categories within the subject. The results are then measured across participants using a one-out sampling technique to ensure they are consistent across all participants. Using DSSA allowed for a thorough analysis of the target problem and allowed the chance to respond to the previously proposed RQs.

4. Empirical Results

The dependent variables used in this study (attractiveness, usefulness) each consist of three classes (as highlighted in section 2.1.). Therefore, the prediction

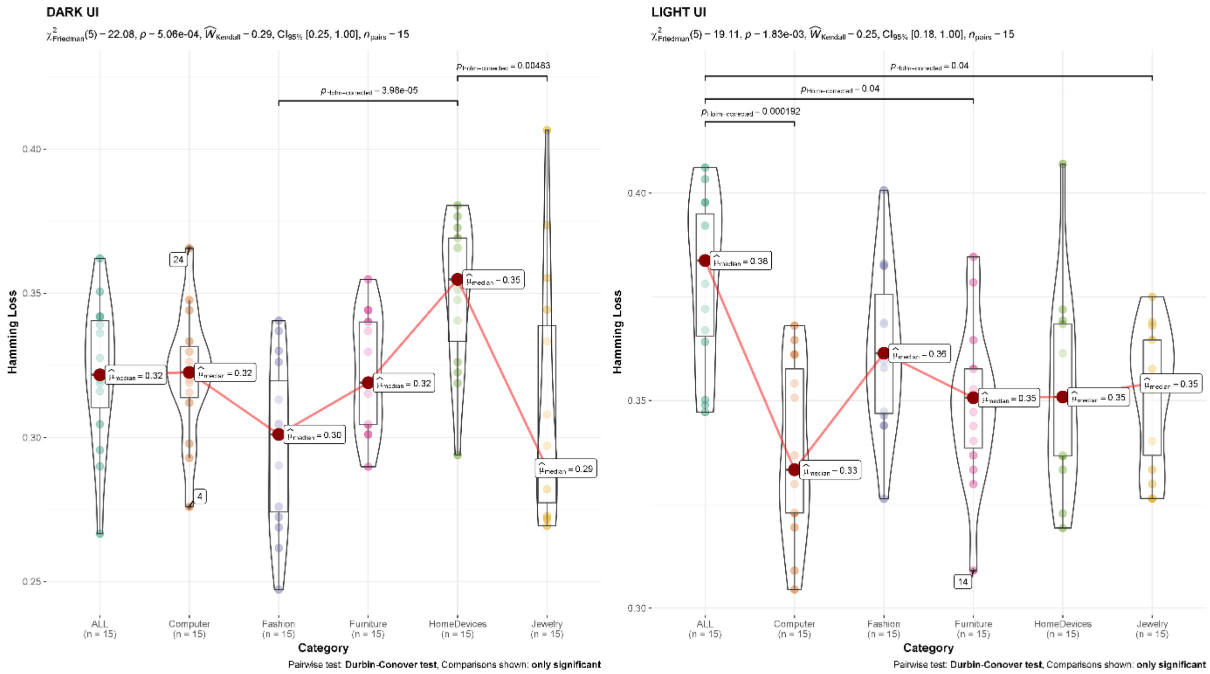

Figure 1. Results from the Random Forest (RF) multi-label classification testing on all the samples and individual samples for each product category.

model requires a multilabel classification approach that includes two dependent variables. To evaluate the performance of the multilabel prediction model, we could not use the standard accuracy metric used in binary classification problems. Instead, we employed Hamming Loss and Exact Match. The specific calculation of Hamming loss is shown in Eq1, where (N: number of total samples / L: number of labels / xor: Exclusive OR). Where $y_{I,j}$ is the real target value, and $\hat{y}_{I,j}$ is the predicted value. The Exclusive OR operator returns zero when the target and identical prediction are employed. Next, we use the Exact Match metric (see Eq2). The Exact match is a strict measure of the multilabel classification performance that compares

accuracy. It reflects only when the model correctly classifies every label that an example has, without any false positives.

$$Hamming\ Loss = \frac{1}{|N|\cdot|L|}\sum_{i=1}^{|N|}\sum_{j=1}^{|L|} \mathrm{xor}(y_{i,j}, \hat{y}_{i,j}) \quad \text{(Eq1)}$$

$$Exact\ Match = \frac{1}{|N|}\sum_{i=1}^{|N|} \mathrm{I}(Y_i = \hat{Y}_i) \quad \text{(Eq2)}$$

4.1. Initial Classifier Prediction Results

We carried out a series of predictive experiments using the data from the model tuning subset to choose the top-performing algorithm from the tree-based family of candidate algorithms. The algorithm that performed the best was RF, which had a Hamming Loss (Exact Match) of .320 (.408) for the dark UI mode sample and .372 (.3187) for the light UI mode sample (see Table 1). As a result, we decided to test the RF multilabel classification model further.

Table 1. Initial Prediction Results.

UI	Model	Hamming Loss	Exact Match	# of Samples
Dark	**RF**	**0.320**	**0.408**	624
	XGB	0.363	0.320	
	DT	0.392	0.288	
Light	**RF**	**0.372**	**0.318**	643
	XGB	0.390	0.256	
	DT	0.473	0.233	

4.2. Main NSM Results

The dark UI mode has a smaller Hamming loss of .032 when compared to the light UI mode for the primary

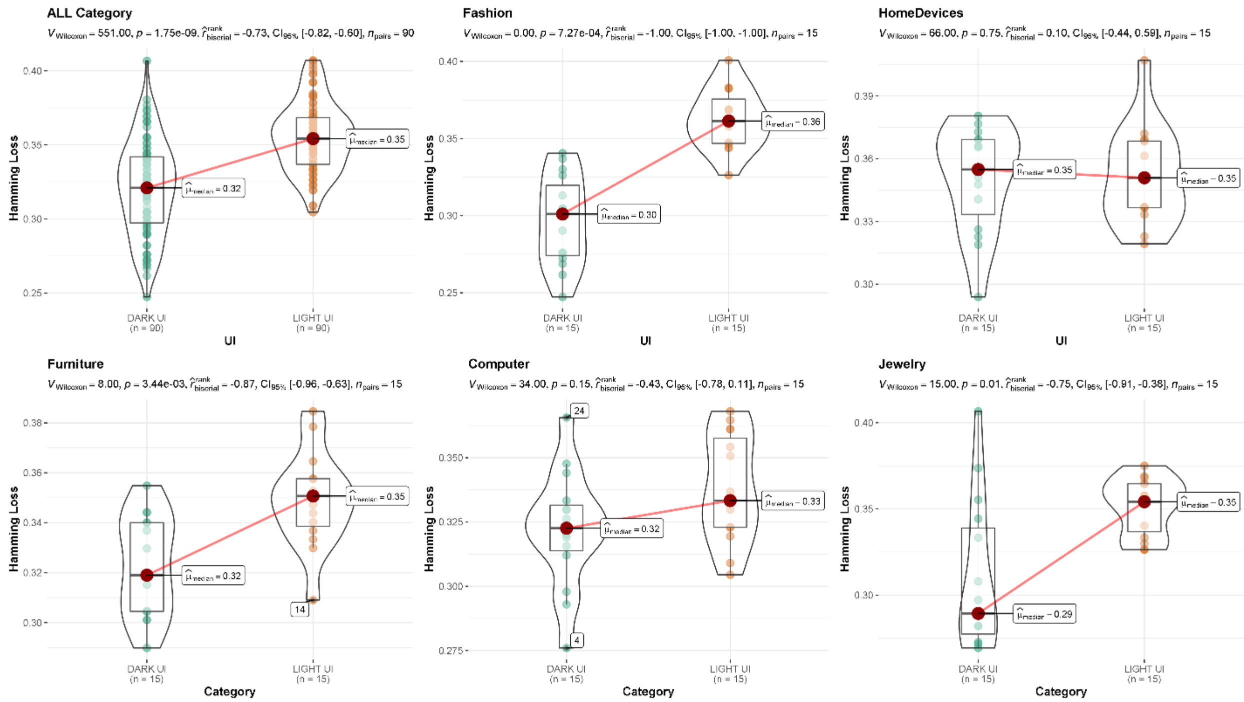

Figure 2. Wilcoxon tests between all the samples for both the dark and light UI mode.

sample, including all features, as shown in Figure 1, while the light UI mode is predictive with a loss of .038. Participants who viewed the dark UI mode were expected to perceive products as more attractive and useful than those who viewed the light UI mode. With the help of NSM's prediction technique, we can clearly distinguish between the two UI modes. Additionally, it was discovered that jewelry, fashion, and furniture showed lower Hamming loss results for the dark UI mode compared to the light UI mode, suggesting that participants thought these categories were more attractive and had greater potential utility. In contrast, there weren't many differences between computers and household appliances.

4.3. Comparison of Product Categories

The significance of each individual's pairwise comparison was evaluated in the analysis that follows. We therefore use the Wilcoxon test for each pairwise analysis, including the main sample and the different categories (see Figure 2). First, we discovered that in the primary comparison, there was a significant difference between the dark UI mode and the light UI mode. Wilcoxon tests confirmed the previous analysis, showing that jewelry, fashion, and furniture all had the same unfavorable, statistically significant outcome. Additionally, when testing computers and home appliances, we discovered no discernible difference between the dark and light UI modes.

We can therefore draw the conclusion that the NSM approach was able to predict a distinct difference between the two UI modes, which was also largely consistent with our initial hypotheses regarding the impact of the dark UI mode on consumer perceptions of the products.

4.4. Across-Participant Analysis

Our most recent analysis validated the previously obtained predictive results using a different sampling technique. In this experiment, we ran the models again after removing one participant at a time. The analysis of individual differences that cannot be seen when testing the entire sample is made possible by using this sampling technique. As can be seen in Figure 3, the result of the light UI mode was insignificant, showing that the outcome was the same for all participants. It's interesting that the dark UI mode didn't produce the same outcome. The humming loss was reduced from .40 in five cases to .30 as a result of this person being excluded from the sample. This suggests

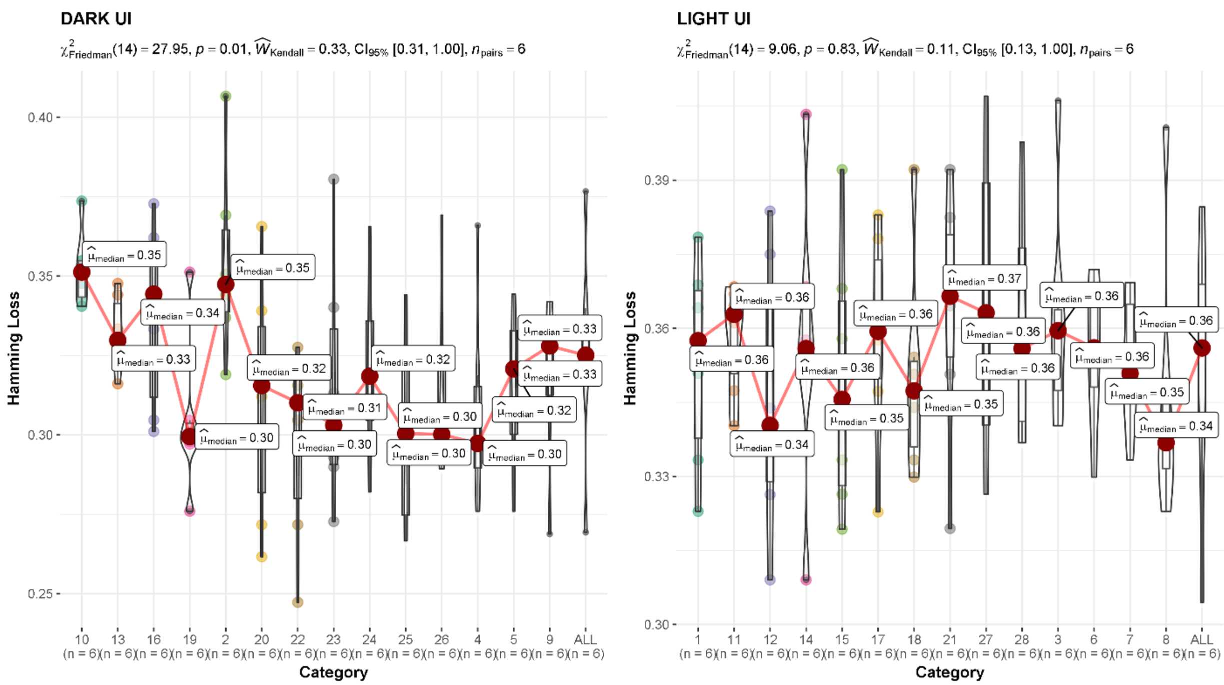

Figure 3. Across-participant sampling analysis: One person is removed, and the model is run again to gauge if the model performance adjust much deeming this person (un)important for this prediction score.

that about a third of the sample preferred the dark UI mode when it came to how attractive and useful they thought online products were. With a humming loss score of .35, which was significantly higher than the sample average, we found that two people in this sample were less affected by the dark UI mode.

The predictive modeling demonstrated that the dark UI mode consistently reduced humming loss compared to the light UI mode, even after taking individual differences into account in this analysis. This provides additional evidence of the dark UI mode's potential impact on participants' opinions of online goods and should be viewed as a reliable reflection of those opinions. The significance of each pairwise comparison was assessed individually in the next analysis. As a result, we employ the Wilcoxon test for each pairwise analysis, taking into account both the main sample and the different categories. In our initial comparison, we found that there was a statistically significant difference between the dark and light UI modes. Wilcoxon tests revealed that jewelry, fashion, and furniture all had the same negative, statistically significant result, consistent with previous analysis, indicating that there is a genuine distinction between these two UI modes in terms of participants' perceptions of online products. Additionally, we found no discernible difference between the dark and light UI modes when testing household appliances and computers. With the help of these tests, we were able to more rigorously statistically confirm our initial findings and offer compelling evidence of the potential effects of both UI modes in an online environment.

5. Discussion and Conclusions

The field of NeuroIS has advanced considerably since its founding in 2007. Numerous studies have been published in esteemed journals and have contributed significant ideas that are applicable to all practitioners of information systems (IS) (see, for example, Dimoka, 2010; Riedl et al., 2010; Vance et al., 2018).

However, a recent review paper examining the present and future of the NeuroIS field found opportunities for the field to advance areas like IS design or the use of IA (vom Brocke et al., 2020). This, however, disregarded the need for methodological and analytical design advancement in the field. Currently, laboratory experiments are frequently used in NeuroIS research to clarify effects within IS phenomena. The RMTHD, also known as Runkel and McGrath's three-horned dilemma (Runkel & McGrath, 1972), contends that generality and realism are lost in laboratory experiments that prioritize control. As a result, this paper suggests a novel technique called neuroscience mining (NSM). This method typically uses deep learning or machine learning models to decipher the complexity of neurophysiological data before attempting to forecast the significance of the data for the dependent variable (s). In this study, we also used sensitivity analysis to conduct a more in-depth analysis. Samples of the data representing various product categories were taken, and each sample was then individually examined. NSM, in contrast to conventional approaches like GLM, can provide a solution to this methodological challenge by analyzing both sequential temporal data within the target problem and enabling direct examination of the DV using the neurophysiological data.

We used the case study of consumers' perceptions of online products in two settings—the less common dark user interface (UI) mode and the more common light UI mode—to conduct NSM analysis. We specifically carried out an Electroencephalogram (EEG) experiment on 28 people in a lab setting, and we used this data for NSM analysis. We found convincing evidence of the effects of the dark UI mode on participants' perceptions of the attractiveness and value of online products, as shown in Figures 2 and 3. We found that the dark UI mode consistently outperformed the more traditional light UI mode in terms of prediction results when using a Random Forest (RF) model for multilabel classification. In particular, participants' EEG data in the dark UI mode for the product categories fashion (.30), jewelry (.29), and furniture (.32) was a strong predictor of their rated perceptions of perceived attractiveness and usefulness using the Hamming Loss metric (the standard metric for multilabel classification problems). The findings also showed no statistically significant distinctions between computers and home appliances. Across all of our tests, we could find no evidence that the light UI mode predicted higher perceived product utility and attractiveness. Overall, employing an NSM approach enabled us to examine the data in a novel manner that enabled us to eke out findings that were not possible to achieve using the common standard metrics used in NeuroIS.

5.1. Implications

Firstly, our study is one of the first to use an NSM methodology to try and predict participants' perceptions of online products. To our knowledge, this is also the first paper to use a multilabel machine learning approach to understand participants' EEG data while shopping online. When using models like GLM, where DVs are not tested in the same analytical models, the proposed method reveals interactions that are not obvious and offers greater analytical flexibility. Additionally, a non-linear RF algorithm was used to predict the outcomes, suggesting that non-linear models can sometimes be used to analyze paradigms in the field of NeuroIS.

Last but not least, on a broader scale, we think that the approach described in this paper for examining consumer product perceptions based on alternative UI designs has succeeded in addressing some of the issues raised in Runkel and McGrath's three-horned dilemma (Runkel & McGrath, 1972). Because controlled laboratory experiments are frequently used in neuroIS research, generalizability and realism are lost. Since more of the temporal data in this study was subjected to NSM (rather than being grand averaged with GLM, for example), the results are more broadly applicable because the brain waves during this time period more accurately reflect the physiological reactions to the stimuli. Therefore, we contend that by utilizing computational social science, NSM can be seen as moving us closer to solving this three-horned conundrum by utilizing big data and data analysis techniques built into the AI paradigm (Chang et al., 2014).

5.2. Limitations and Future Work

We must now discuss the shortcomings of this work and make recommendations for future research that can help to address these shortcomings. First, despite the fact that the multilabel classification model produced respectable Hamming Loss results for the target values, we did not train the model to account for additional potentially important variables like price or surrounding features and how these variables might contribute to the possibility of complexity overload. Future research must make every effort to explore all viable options because running them with NSM is not a difficult task, even though the focus was on method rather than context.

Another limitation of this study is that we did not separate all EEG signal bands when predicting our target variables. As a result, our findings cannot definitively state which EEG spectral band has the greatest influence on the dependent variable. For two reasons, we decided against conducting this analysis in this paper: first, our training data samples were small due to the brief stimulus durations, and second, space constraints prevented us from including all possible analyses. We wanted to show how useful NSM is for understanding IS phenomena. Although we do not think that band separation will affect the overall findings, it should be considered when NSM is used in academic papers submitted to prestigious IS journals.

Then, using a one-sample-out methodology, we carried out a sensitivity analysis based on multiple model predictions. A more well-liked approach, in contrast, has emerged in other paradigms of IS research where variables within models can be taken out and reanalyzed. This machine learning/deep learning method, also referred to as the relative importance analysis method (Loureiro et al., 2018), may be equally successful in addressing issues with NSM RQs and enabling practitioners to think about a variety of fresh perspectives when examining NeuroIS research paradigms.

Finally, our demographics might be seen as a drawback. Since the research was done in the same nation as the participants, they all had a common cultural background and language. As a result, this sampling strategy might have a built-in convenience bias that can only be removed by looking at the same phenomenon in a more diverse sample in terms of race and ethnicity.

6. Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1F1A1074808). This article's correspondence should be sent to kunchanglee@gmail.com.

7. References

Android Authority. (2020). We asked, you told us: Just about everyone uses dark mode.

Aslam, M. M. (2006). Are you selling the right colour? A cross-cultural review of colour as a marketing cue. Journal of Marketing Communications, 12(1), 15–30.

Barral, O., Kosunen, I., & Jacucci, G. (2017). No need to laugh out loud: Predicting humor appraisal of comic strips based on physiological signals in a realistic environment. ACM Transactions on Computer-Human Interaction, 24(6). 1-29

Buchner, A., Mayr, S., & Brandt, M. (2009). The advantage of positive text-background polarity is due to high display luminance. Ergonomics, 52(7), 882–886.

Buettner, R., Beil, D., Scholtz, S., & Djemai, A. (2020). Development of a Machine Learning Based Algorithm To Accurately Detect Schizophrenia based on One-minute EEG Recordings. HICSS-53, 3216–3225

Chang, R. M., Kauffman, R. J., & Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. Decision Support Systems, 63, 67–80.

Choi, W., & Tulu, B. (2017). Effective Use of User Interface and User Experience in an mHealth Application. HICSS-50, 3803–3812

Cortez, P., & Embrechts, M. J. (2013). Using sensitivity analysis and visualization techniques to open black box data mining models. Information Sciences, 225, 1–17.

Dimoka, A., Pavlou, P. A., & Davis, F. D. (2011). NeuroIS: The potential of cognitive neuroscience for information systems research. Information Systems Research, 22(4), 687–702.

Elliot, S., & Fowell, S. (2000). Expectations versus reality: a snapshot of consumer experiences with Internet retailing. International Journal of Information Management, 20(5), 323-337

Fiore, A. M., Kim, J., & Lee, H. H. (2005). Effect of image interactivity technology on consumer responses toward the online retailer. Journal of Interactive Marketing, 19(3), 38–53.

Fiore, A. M., Yah, X., & Yoh, E. (2000). Effects of a product display and environmental fragrancing on approach responses and pleasurable experiences. Psychology and Marketing, 17(1), 27–54.

Fischer, T., Davis, F. D., & Riedl, R. (2019). NeuroIS: A survey on the status of the field. In Lecture Notes in Information Systems and Organisation, 29, 1–10.

Kuan, K., Zhong, Y., & Chau, P. (2014). Informational and normative social influence in group-buying: Evidence from self-reported and EEG data. Journal of Management Information Systems, 30(4), 151–178.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

Lee, Y. E., & Benbasat, I. (2004). A framework for the study of customer interface design for mobile commerce. International Journal of Electronic Commerce, 8(3), 79–102.

Löffler, D., Giron, L., & Hurtienne, J. (2017, September). Night mode, dark thoughts: Background color influences the perceived sentiment of chat messages. In IFIP Conference on Human-Computer Interaction, 184-201

Loureiro, A. L. D., Miguéis, V. L., & da Silva, L. F. M. (2018). Exploring the use of deep neural networks for sales forecasting in fashion retail. Decision Support Systems, 114, 81–93.

Ngai, E. W. T., & Gunasekaran, A. (2007). A review for mobile commerce research and applications. Decision Support Systems, 43(1), 3–15.

Nissen, A., & Krampe, C. (2021). Why he buys it and she doesn't – Exploring self-reported and neural gender differences in the perception of eCommerce websites. Computers in Human Behavior, 121 1, 106809.

Park, J., Lennon, S. J., & Stoel, L. (2005). Online product presentation: Effects on mood, perceived risk, and purchase intention. Psychology & Marketing, 22(9), 695-719.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. Journal of neuroscience methods, 162(1-2), 8-13.

Riaz, A., Gregor, S., Dewan, S., & Xu, Q. (2018). The interplay between emotion, cognition and information recall from websites with relevant and irrelevant images: A Neuro-IS study. Decision Support Systems, 111, 113–123.

Riedl, R., & Léger, P.-M. 2016. Fundamentals of NeuroIS-information systems and the B rain. Cham, Switzerland: Springer Publishing.

Riedl, R., Fischer, T., & Léger, P. M. 2017. A Decade of NeuroIS Research: Status Quo, Challenges, and Future Directions. ICIS 2017 Proceedings, 0–28.

Runkel, P. J., & McGrath, J. E. (1972). *Research on human behavior: A systematic guide to method* (1st ed.). Rinehart and Winston.

Then, N. K., & Delong, M. R. (1999). Apparel shopping on the Web. Journal of Family and Consumer Sciences, 91(3), 65.

vom Brocke, J., Hevner, A., Léger, P. M., Walla, P., & Riedl, R. (2020). Advancing a NeuroIS research agenda with four areas of societal contributions. European Journal of Information Systems, 29(1), 9–24.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. (2016). *Data Mining: Practical Machine Learning Tools and Techniques* (4th eds). Morgan Kaufmann Publishers Inc.

Xiong, J., & Zuo, M. (2020). What does existing NeuroIS research focus on?. Information Systems, 89, 101462.

Ye, X., Peng, X., Wang, X., & Teo, H. H. (2020). Developing and testing a theoretical path model of web page impression formation and its consequence. Information Systems Research, 31(3), 929-949.