Using Wearable Devices for Non-invasive, Inexpensive Physiological Data Collection

James Gaskin
Brigham Young University, Utah
james.gaskin@byu.edu

Thomas Meservy
Brigham Young University, Utah
tmeservy@byu.edu

Jeffrey Jenkins
Brigham Young University, Utah
jeffrey_jenkins@byu.edu

Jacob Steffen
Brigham Young University, Utah
steffen.h.jacob@gmail.com

Katherine Payne
Brigham Young University, Utah
kcpayne@byu.edu

Abstract

Using sensors to gather physiological data about users can provide valuable insights for Information Systems (IS) research that are not availed through traditional measures. While useful in many laboratory settings, many of these physiological sensors (e.g., fMRI, EEG, EKG, etc.) are impractical and severely limited in other scenarios due to (1) prohibitive cost, (2) small sample size, (3) invasiveness, and (4) the difficulty to match <u>psycho</u>logical traits to <u>physio</u>logical measures. In this study, we demonstrate how inexpensive consumer-grade wearable technologies overcome these first three limitations while we extend existing research on exploring the fourth limitation.

1. Introduction

Direct measurement of users' physiology can provide valuable information about users' cognitive and emotional states [6]. Whereas IS researchers have historically relied on external measures of cognition and emotion such as survey responses or observed behavior, physiological data about users can provide "a richer account of user cognition than that obtained from any other source, including the user himself" [30, p. 73]. These physiological measures include cardiovascular (heart rate, heart rate variability, blood volume pulse), electrodermal (galvanic skin response), electro-encephalography, electromyography, and respiratory data, to name a few [22, 26]. The insights garnered from this physiological data can be used to design more effective systems, create adaptive systems, and help users make more informed decisions [23, 34].

However, until recently, devices used to capture physiological measures have been bulky, expensive, and invasive—potentially confounding experimental data in some settings, and often resulting in small sample sizes [11]. Beyond this, it is often very difficult to match this physiological data with psychometric or behavioral traits. While the development of consumer wearable technologies does not resolve this last issue, new possibilities have opened up for capturing relevant physiological measures in an unobtrusive and inexpensive way using wearable devices [12]. As such, sample sizes are able to increase and correlations will become more apparent with the benefit of stronger statistical power, thus leading to more precise matching between the physiological and psychological measures

Extant literature on wearable devices largely focuses on the potential for healthcare, especially the potential to amass large amounts of personal data over time that can give greater insight into patient progress and potential treatments [7]. However, we highlight the enormous potential of consumer-grade wearable devices for capturing objective physiological measures in a non-invasive and inexpensive way to allow larger sample sizes that are more ecologically valid experiments.

As noted, the instruments in many social science studies that have leveraged physiological data are often invasive and obtrusive and, as such, can confound the experiment. When the participants in such studies are tethered, probed, strapped, and wired, it becomes more difficult to ignore the artificiality of the experiment environment and may directly interfere with participants' affective states and interaction with a system. The experimental control this affords can impact the face-validity of the experiment and creates a sterile environment that is often not reflective of real-

HĬCSS

world settings, as it is difficult to implement similar designs in practice.

Current wearable technologies offer an encouraging solution to many of the aforementioned limitations of past methods. Intrusiveness is minimized as the devices are small, comfortable, wireless, and can be worn while performing a wide range of activities. Relatively speaking, the devices are also inexpensive. Thus, if the measures captured by wearable devices are consistent, accurate, and reliable, they open up many doors to more valid and practical research. Further, there is a propitious opportunity to link physiological correlates to non-biased measurements, corroborating, or even used in place of, self-report instruments such as the Likert scale. In this study, we explore whether consumer-grade wearable devices can capture data with sufficient granularity, precision, and unobtrusiveness to demonstrate their worth as a method for gathering data for research in science and in practice in the context of cognitive and affective state changes.

2. Studies Leveraging Physiological Data

Over the past decade, the IS discipline has become increasingly interested applying physiological measures to IS related constructs. The popularized term for this line of inquiry is NeuroIS. A growing community of IS scholars seeks to use neurophysiological measures to more deeply understand IS phenomenon including technostress [40], cognitive load [13], trust [16], emotions [1], HCI [5, 9, 35], social media [29], online search and filtering of information [19, 28], technology acceptance and adoption [17], and more [34].

Most studies that specifically discuss or make use of wearable technologies are done from a medical perspective. For example, some studies have focused on the advantages of continuous measurements in combating various health challenges, including sleep apnea and cardiovascular diseases [32]. Others stress the advantage of extending the range of data collection to patients' homes in providing insight for disorders such as Parkinson's disease [7].

A number of past studies have attempted to map out significant physiological correlates for a number of different psychological constructs. Obtaining objective measures for psychological constructs is paramount, as other measures are easily confounded by differences afforded to individual subjects [2]. For certain psychological constructs such as flow or immersion, the very act of self-reporting is contradictory. Because a key component of flow is the loss of self-awareness, capturing self-reported data requires the subject to exit the state of flow or try to recall it, thus decreasing the

validity of the data captured [33]. Captured physiological measures offer a potential remedy to such dilemmas. We next provide examples of how invasive and limited such studies can be.

One of the larger studies involving physiological measures is Léger et al. [22]. They used neurophysiological measures to establish physiological correlates of cognitive absorption as subjects interacted with an ERP training game. While playing the game, each of the subjects was connected to an extensive amount of hardware. These included gelled electrodes on their palms to measure electrodermal activity, electro-encephalographic (EEG) sensors on their scalp and ear lobes to measure brain activity, and sensors placed over the sternum and the seventh intercostal (between ribs) space on both the left and right sides of the chest to measure heart rate. Results showed that inclusion of neurophysiological measures help their predictive model for cognitive absorption. The sample size for the study was 36.

Other studies have measured different constructs and used other physiological measures. For example, Wagner et al. [44] sought to understand the correlates of emotion, specifically measuring joy, anger, pleasure, and sadness. A single subject listened to music as various four-channel biosensors were used to record electromyogram, electrocardiogram, skin conductivity, and respiration change. Mauri et al. [26] tracked 30 subjects as they browsed their personal Facebook pages in order to better understand how flow contributes to social media use. Tracking equipment used included a respiration belt placed on the chest, EMG sensors attached to the muscles near the eyebrows, two electro-encephalogram (EEG) sensors positioned on the left and right frontal lobes, with the respective reference electrodes on the ear lobes (all applied with conductive paste), as well as two skin conductivity adhesive patches placed on the nondominant palm, and a Blood Volume Pulse (BVP) sensor placed on the top of the index finger. Results indicated a correlation between Facebook use and a core flow state of high valence and arousal [26].

De Manzano et al. [14] also investigated flow by measuring physiology while 21 classical pianists performed various pieces. Measures obtained included muscle tone of facial muscles recorded by Ag/AgCl surface electrodes attached to muscle sites on the left side of the face. Participants' skin was prepared with a low-alcohol detergent to minimize impedance. Head movements were registered using a 3-axis accelerometer attached to a headband. Thoracic respiration was measured using a piezo-electric respiratory belt transducer placed around the chest

below the nipple line. Heart period and blood pressure measurements were based on arterial pulse pressure acquired with an IR plethysmograph attached to the right ear.

Similar studies that utilize brain imaging to capture data have comparably small sample sizes and measuring constraints. For example, Benbasat et al. [6] explored interactions with recommendation agents of both genders and various ethnicities, but used fMRI scans to overcome potential reporting biases. The sample size for the study was 24. Venkatraman et al. [42] used fMRI scans in addition to biometrics and eye tracking to predict the persuasiveness of 30-second advertisements. Thirty-three subjects were scanned for the fMRI portion of the study, while 29 subjects participated in the biometric/eye tracking portion wearing various physiology tracking devices for collecting skin conductance, heart rate, and respiration data.

All aforementioned studies strive to make significant contributions to mapping out physiological correlates of various psychological phenomena. However, across all studies, two major weaknesses are apparent: small sample sizes and use of potentially intrusive (confounding) devices. The sample sizes used are an understandable limitation, as many of the tools used are expensive and require lengthy and involved setups for each participant. While the limitations of sample size are understandable, the effect on statistical power derived from such results must also be taken into account when evaluating conclusions drawn from such studies. The intrusive nature of the devices used to capture the data should also be called into question. If the very act of using intrusive devices elicits a reaction that confounds the measurements taken, the validity of any conclusion is significantly weakened. For example, wearing chest straps, cranial electrodes, and palm electrodes may distract participants and prevent them from entering a flow state. The use of modern consumer wearable technologies offers a potential remedy to these problems. If these devices can be shown to be sensitive enough to consistently capture granular data, future studies with increased sample sizes and less intrusive measurement tools can be performed.

3. Theory-based Expectations

In this exploratory paper, we focus primarily on using wearable devices to measure granular changes in affective states, which is relevant to many constructs of interested to the IS discipline, including flow and immersion. There is a large body of literature related to measuring affective states or emotions [8, 15, 43]. A predominant framework [8] suggests that emotions vary along three dimensions: valence, arousal, and dominance. Valence is the positive- or negative-ness of the emotion. For example, joy is positive whereas fear is negative. Arousal refers to the excitability of the emotion. Peacefulness is low in arousal but a thrilling situation would create high arousal. Dominance refers to the extent that you feel like you are in control of your situation and environment. An emotion that is empowering and liberating would be high in dominance and one where the emotion makes you feel small or helpless would be low in dominance.

Several studies have sought to identify specific emotions using heart rate, including amusement, sadness, anger, fear, surprise, and disgust [18, 20, 31, 38]. Research related to measuring arousal using heart rate has largely focused on the change of beats per minute from the presentation of a stimuli [3, 39]. Findings suggest that average heart rates tend to decrease over time for both low-arousal and high-arousal conditions from the onset of the emotional stimuli [39]. However, high-arousal images decelerate more than low-arousal images [39]. It is important to note that other dimensions, such as valence, also impact heart rate measurement. Other studies have found that heart rate was higher for high arousal images compared to low arousal images even when valence and vividness were controlled for [41].

A handful of articles have investigated how to measure emotion using heart rate over a longer period of time (i.e., minutes instead of seconds) [21, 38]. Schwartz et al. [38] had participants imagine vivid situations where participants felt happiness, sadness, fear, anger, and relaxation for two minutes. They compared heart rates after this time period to baselines for each participant. Results indicate that anger and fear increased average heart rate the most (10-11 bpm) followed by happiness and sadness (6-7 bpm). Relaxation was similar to the control condition.

Research on heart rate and emotions has also provided additional insights. For example, negative emotions have a longer lasting impact on heart rate than positive emotions [10]. Interestingly, power spectrum analysis has revealed that certain emotions (e.g., anger, appreciation) can be differentiated by comparing different parts of the power spectrum captured while experiencing those emotions [27]. A similar approach has also been used to understand heart rate changes over an extended period of time—50 minutes—while participants watched comedies or tragedies [37].

Beyond heart rate, or beats per minute, additional measures have been developed and are now commonly used including inter-beat interval and heart rate

variability [24]. Inter-beat interval is the time between individual heart beats and heart rate variability (HRV), though sometimes measured in various ways, is the variation of the interval between heartbeats. HRV has also been used to detect different emotional states. For instance, "a low HRV can indicate a state of relaxation, whereas an increased HRV can indicate a potential state of mental stress or frustration" [21]. **Thus, we expect that high arousal emotions have higher variability than low arousal emotions.** Heart rate variability analysis has also been used as a mechanism to understand when emotions are regulated [4]. We intentionally replicate this known effect in this study. This allows us to hold constant the known effect while evaluating the utility of the data collection technology.

In this study, we focus specifically on emotional arousal. Emotional arousal was selected because it is a construct of interest across a number of lines of inquiry within Information Systems and is included in several past IS related studies. Obtrusive and invasive instruments have been able to capture a measure of emotional arousal—increased heart rate. However, this measure can also be captured and is commonly available in numerous inexpensive wearables thanks to advances in photoplethysmography. Importantly, using arousal allows us to explore the efficacy of wearable devices in measuring changes in physiological responses (e.g., heart rate) in response to constructs of interest to IS.

4. Study Design and Method

To explore whether we can observe granular changes in heart rate (captured via wearable devices) due to changes in arousal, we conducted a within-subject experiment manipulating arousal in two conditions: low arousal vs. high arousal. Arousal was manipulated by having participants interact with low- and high-arousal stimuli on the Samsung Gear VR goggles. Virtual Reality (VR) was chosen to manipulate arousal because it provides life-like immersive interactions that are effective at manipulating emotional arousal [36]. We also chose to use VR as our treatment because of the subtlety of the changes it would induce in the participants' physiology as a response to subtle changes in their affective/cognitive states. If the consumer-grade wearable can detect these subtle changes, then it is likely useful for less subtle scenarios and experiments.

After consenting to the research, participants put on the Microsoft Band (version 2) wearable bracelet. The Microsoft Band (**Figure 1**) is worn on the wrist and provides physiological data about heart rate,

perspiration, and movement, among other information. The Microsoft Band was chosen due to the ease of custom app development that allowed us to extract the physiological measures at custom intervals.[1]

Figure 1. Microsoft Band wearable bracelet

Participants were led to a room where they put on the Samsung Gear headset and interacted with a randomly assigned condition (the low- or high-arousal condition). Prior to seeing the low- or high-arousal stimuli in the condition, participants interacted with a neutral stimulus—a benign 3D office meeting room (Figure 2)—to adjust to the novelty of the VR goggles, to help nullify the effects of prior emotions, and to settle physiological conditions (e.g., heart rate). Participants were allowed to look around the office for approximately two minutes. Afterwards, the low- or high-arousal condition was randomly shown, which also lasted approximately two minutes. Participants then completed a short survey to capture self-report measures of the experience and afterwards started the other condition (e.g., if the participant randomly saw the high-arousal condition first, the participant would see the low-arousal condition next). Prior to starting the second condition, participants again saw the baseline stimulus for two minutes to help nullify the effects of the prior emotion. The second condition lasted approximately two minutes again. Finally, after completing the second condition, participants again completed a survey to capture self-report measures. The two conditions are described below.

---

[1] We purchased and experimented with many other wearable devices, including the Basis Peak, Apple Watch, W/ME, FitBit, and we researched many other devices' features. We settled on the Microsoft Band because it accommodated user-generated custom apps and could be manipulated to report heart rate and other measures at a user-defined interval. Most other devices did not allow custom apps and only reported physiological measures at broad intervals (e.g., every 1 to 10 minutes).

Figure 2. Baseline stimulus prior to each condition

Low-Arousal Condition. In the low-arousal condition, participants interacted with a 360-degree immersive beach scene. The beach scene was selected to induce feelings of low-arousal, such as relaxation and serenity. Snapshots of the beach are shown in Figure **3**.

Figure 3. Scene from low-arousal condition

High-Arousal Condition. In the high-arousal condition, participants interacted with a 360-degree immersive action movie. In the movie, participants are chased by an alien through the subway system in a metropolitan city. The action movie was selected to induce high-arousal emotions such as excitement or thrill. Snapshots of the movie are shown in **Figure 4**.


Figure 4. Scenes from high-arousal condition

Measures. We created a custom app to record heart rate, galvanic skin response (sweat), and movement while participants experienced the baseline condition (office room) and low- and high-arousal conditions. The app allowed us to restrict the data collection to precisely the time period when participants were viewing the actual media. From the heart rate data, we calculated the mean, max, min, standard deviation, and range for each participant. In the post-survey following each condition, we captured self-report measures about the experience. These included a measure of emotional arousal using the self-assessment manikin (SAM) scale—a 9-point pictorial scale of emotion [8].

Participants. Over a two week period (four separate experiment sessions), we collected data from 113 participants. The participants were recruited from various business courses at a large private university. Among the participants, 59.3% were male and 40.7% were female. The average age was 21.83 (stdev: 2.90). We also asked to what extent participants had experience with VR prior to the experiment. Seventy-five percent reported never having used VR prior to the experiment, with 18.6% having used it only once prior to the experiment, 5.6% having used it a few times prior, and only one participant reported having used it many times prior to the experiment. Thus, our participants were largely new to the VR experience, which should work in our favor to increase arousal due to novelty.

Figure 5 visually depicts the flow of our experiment.

Figure 5. Experiment Flow
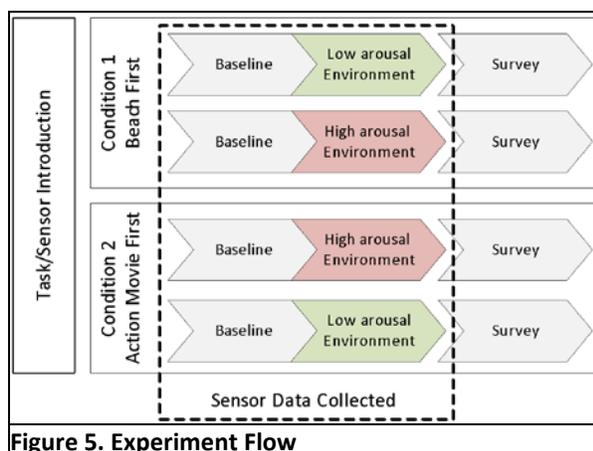
5. Analysis

We first assessed whether our manipulations of emotional arousal were successful. To do this, we conducted a paired t-test comparing the self-reported emotional arousal in the low-arousal condition to the high-arousal condition. We found that participants reported significantly higher arousal in the high-arousal condition (M = 5.440, SD = 2.045) compared to the low-arousal condition (M = 3.484, SD = 1.747); t(90) = 7.549, p <0.001. Thus, our manipulations were deemed successful.

Next, we analyzed whether the arousal treatment influenced users' physiology as measured through the wearable device. For this analysis, we used five statistics of heart rate variability as our dependent variables: heart rate min, max, range, standard deviation, and mean. We specified linear mixed-effects models that used arousal (0 = low arousal treatment, 1 = high arousal treatment) as a fixed effect to predict the heart rate dependent variables. In addition, we included the subject id as a random effect in the model, controlling for heart rate differences inherent to different participants as well as other natural differences among participants that may influence heart rate (e.g., previous experience with VR).. Significant levels of the fixed effect (arousal) was determined through likelihood-ratio tests. Conditional[2] and marginal[3] r-squared estimates were calculated for each linear-mixed effect model to assess effect size.

---

[2] Explained variance from random (subject id) and fixed (arousal) effects

[3] Explained variance from fixed effects only

6. Findings

Table 1 displays the means and standard deviations for each heart rate statistic by condition. **Table 2** displays the results of the linear mixed effects model. With the exception of the mean heart rate, all other heart rate variables were significantly different between the two conditions. Participants in the high arousal condition had a significantly lower min, higher max, greater standard deviation, and greater range in heart rate.

Table 1. Mean and standard deviations for heart rate statics by group

Variable	Low Arousal Condition (Beach)	High Arousal Condition (Movie)
Average Mean	72.83 (8.33)	73.03 (8.76)
Average Max	79.85 (8.62)	81.57 (12.42)
Average Min	66.30 (8.64)	65.35 (8.25)
Average StDev	3.71 (1.40)	4.30 (3.03)
Average Range	13.66 (4.05)	16.22 (10.02)
Mean shown with standard deviation in parentheses		

Table 2. Summary of results for heart rate

Heart rate Variable	Arousal Fixed Effect Estimate	χ^2	R² Cond	R² Marg
Mean	0.212	0.18 (NS)	.838	.000
Max	1.717	2.68*	.560	.006
Min	-0.947	2.92*	.803	.003
StDev	0.584	2.91*	.042	.015
Range	2.561	5.45*	.087	.028
NS (not significant); *p <.05, **p <.01, ***p <.001				

7. Discussion

We began this study by identifying four major limitations to collecting physiological data: (1) prohibitive cost, (2) small sample size, (3) invasiveness, and (4) the difficulty to match psychological traits to physiological measures. This study acts as a proof-of-concept for demonstrating the viability of using consumer-grade wearable technologies as a means of remedying, to some extent, these four major limitations. As far as we are aware, this is the first study of this kind to use consumer grade wearable technologies to collect physiological data for scientific research. In this respect, our approach is novel and contributes to laying a foundation for future studies with similar intent.

Prohibitive Cost. Many of the physiological data collection instruments used in extant literature are exorbitantly expensive to purchase and to use, and are often difficult to access due to these costs. For example,

an fMRI machine costs between $500,000 and $3 million USD. To rent usage of the machine can cost upwards of $600 per hour, accommodating only a couple participants per hour. An EEG monitor costs multiple thousands of dollars, as do hospital-grade heart rate monitors and blood oxygen sensors. The smart bracelet we used in this study currently costs $175 USD, which is comparable to other smart health monitors on the market today. We required only two of these bracelets for the current study.

Small Sample Size. Many prior studies have been severely limited due to sample size restrictions. These restrictions are partly due to the prohibitive costs. The sample sizes in the studies we surveyed in the literature review section were typical of most other studies collecting physiological data. Common sample sizes are a couple to a few dozen. The largest we surveyed included 36 participants [22]. Our preliminary study has included 113—over three times the highest study we surveyed—and we are still collecting data, with a goal of over 200 within a one month (total) experiment period (data collection twice per week). Sample sizes like these can now provide the statistical power required for more generalizable findings and help to minimize type 2 errors (false negatives).

Another reason for small sample sizes in many studies is the time-intensiveness and expertise required for setting up each participant and running the hardware and software. Placing electrodes, chest straps, and positioning participants precisely (e.g., in an fMRI machine) requires special expertise, training, and time. The average time for calibrating and setting up our participants was about ten seconds (just to put on the bracelet) and required no expertise or training. The entire experiment was conducted by undergraduate research assistants without requiring faculty presence beyond a trial run—mainly for pacing and verifying all hardware and software worked as expected.

Invasiveness. Most physiological instruments used in extant studies are highly invasive—as we have described in our literature review. They often require probes, tethers, awkward positions, and constraints, among other inhibiting aspects. These conditions act to confound the very measures they seek to capture. The device we used was as unobtrusive as a bracelet. To verify this, we surveyed the participants (immediately after their first VR session) to discover if the bracelet was distracting, invasive, or considered an obstacle to their movement or interaction. We used a 10-point scale where 1 indicated not distracting, uncomfortable or an obstacle, and 10 indicated the extreme opposite (i.e., very invasive, etc.). On this 10-point scale, 100% responded with a 1, 2, or 3, indicating it was not

distracting. For discomfort, 91% responded with a 1, 2 or 3, and 97% responded similarly to indicate it was not an obstacle.

As further evidence that the participants were able to completely forget they were wearing the device, we collected data on immersion ("I was able to block out most other distractions", "I was absorbed in what I was doing", and "I was immersed in the VR"). On a 5-point Likert-scale of agreement, participants averaged a 4.6 out of 5, indicating that they were highly immersed in the VR experience, and therefore not distracted by the device.

Match Psychological and Physiological Measures. It is difficult to match physiological patterns with psychological traits. This remains difficult. In the current study, we have demonstrated that a simple set of heart rate measures consistently correlates with arousal. While this finding is not new [25], it is novel in that the measurement was done using an inexpensive, unobtrusive, consumer-grade wearable device. We expect more sophisticated measures, such as pattern analysis of measures over time—rather than means, ranges, standard deviations, etc.—would facilitate more precise matching of physiological data with psychological states. However, such sophisticated analyses are beyond the scope of this exploratory proof-of-concept study.

7.1 Limitations and Future Research

As an exploratory proof-of-concept, this study is limited in a few ways. First, we collected data from a single population (students) who are generally young and healthy. Results may differ with a more heterogeneous population. Additionally, we did not control for health, weight, height, general fitness level, general reactivity of the autonomic nervous system to exogenous triggers, or other preexisting conditions which may have affected heart rate, although we did included the subject id as a random effect in the model, controlling for heart rate differences inherent to different participants. Second, we only reported results with regard to heart rate. Consumer-grade wearable devices can collect many other types of measures, such as galvanic skin response (i.e., perspiration levels), skin temperature, movement in 3D space, electro-encephalography (EEG), among others. Heart rate is the best established correlate of arousal, and as such, was a good candidate measure for an exploratory study. However, GSR and EEG, among other measures, should be explored for their potential correlations with psychological measures. Any progress we can make on this front will help to mitigate the weaknesses of self-

report bias. Third, we limited our data to blunt measures, such as mean, min, max, and standard deviation. Precise matching with psychological traits is more likely with more precise measures, such as temporal pattern analyses.

8. Conclusion

Using wearable technologies for scientific data collection has a bright future. These devices will only get more sophisticated, less expensive, more accurate, and more intelligent. Immediately upon the horizon is the Microsoft Hololens which includes all of the sensors contained in existing smart activity/health trackers, as well as a built in EEG. Another device on the horizon is the Embrace monitor, which includes sensors never before included on a wearable device. These sensors monitor your autonomic nervous system and stress levels. As devices improve, scientists will be able to gather more data with greater ease. As sample sizes increase, so too will our ability to identify patterns in the data, just as has been done in the field of genome mapping. As these patterns emerge and are connected to psychological, cognitive, and behavioral traits and states, we will be able to begin moving toward a less biased set of measures for constructs of interest.

9. References

[1] M. T. P. Adam, M. Hagenau, D. Neumann and C. Weinhardt, *Emotions in Electronic Auctions-A Physio-Economic Approach on Information Systems*, *ECIS*, Citeseer, 2008, pp. 315-325.

[2] K. L. Ailawadi, R. P. Dant and D. Grewal, "The difference between perceptual and objective performance measures: An empirical analysis", Marketing Science Institute, 4 (2004), pp. 1-25.

[3] J. Anttonen and V. Surakka, *Emotions and heart rate while sitting on a chair*, *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 2005, pp. 491-499.

[4] B. M. Appelhans and L. J. Luecken, "Heart rate variability as an index of regulated emotional responding", Review of general psychology, 10 (2006), pp. 229.

[5] I. Benbasat, "HCI research: Future challenges and directions", AIS Transactions on Human-Computer Interaction, 2 (2010), pp. 1.

[6] I. Benbasat, A. Dimoka, P. A. Pavlou and L. Qiu, *Incorporating Social Presence in the Design of the Anthropomorphic Interface of Recommendation Agents: Insights from an fMRI Study*, *ICIS*, 2010, pp. 228.

[7] P. F. Binkley, "Predicting the potential of wearable technology", Engineering in Medicine and Biology Magazine, IEEE, 22 (2003), pp. 23-27.

[8] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential", Journal of behavior therapy and experimental psychiatry, 25 (1994), pp. 49-59.

[9] J. V. Brocke, R. Riedl and P.-M. Léger, "Application strategies for neuroscience in information systems design science research", Journal of Computer Information Systems, 53 (2013), pp. 1-13.

[10] J. F. Brosschot and J. F. Thayer, "Heart rate response is longer after negative emotions than after positive emotions", International Journal of Psychophysiology, 50 (2003), pp. 181-187.

[11] K. S. Button, J. P. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson and M. R. Munafò, "Power failure: why small sample size undermines the reliability of neuroscience", Nature Reviews Neuroscience, 14 (2013), pp. 365-376.

[12] D. J. Cook and W. Song, "Ambient intelligence and wearable computing: sensors on the body, in the home, and beyond", Journal of ambient intelligence and smart environments, 1 (2009), pp. 83-86.

[13] A. O. de Guinea, R. Titah and P.-M. Léger, "Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research", Computers in Human Behavior, 29 (2013), pp. 833-844.

[14] Ö. de Manzano, T. Theorell, L. Harmat and F. Ullén, "The psychophysiology of flow during piano playing", Emotion, 10 (2010), pp. 301.

[15] P. Desmet, *Measuring emotion: Development and application of an instrument to measure emotional responses to products*, *Funology*, Springer, 2003, pp. 111-123.

[16] A. Dimoka, "What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study", Mis Quarterly (2010), pp. 373-396.

[17] A. Dimoka and F. D. Davis, "Where does TAM reside in the brain? The neural mechanisms underlying technology adoption", ICIS 2008 Proceedings (2008), pp. 169.

[18] P. Ekman, R. W. Levenson and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions", Science, 221 (1983), pp. 1208-1210.

[19] K. J. Fadel, T. O. Meservy and M. L. Jensen, "Exploring Knowledge Filtering Processes in Electronic Networks of Practice", Journal of Management Information Systems, 31 (2015), pp. 158-181.

[20] J. J. Gross and R. W. Levenson, "Emotion elicitation using films", Cognition & emotion, 9 (1995), pp. 87-108.

[21] A. Haag, S. Goronzy, P. Schaich and J. Williams, *Emotion recognition using bio-sensors: First steps towards an automatic system*, *ADS*, Springer, 2004, pp. 36-48.

[22] P.-M. Léger, F. D. Davis, T. P. Cronan and J. Perret, "Neurophysiological correlates of cognitive absorption in an enactive training context", Computers in Human Behavior, 34 (2014), pp. 273-283.

[23] Q. H. Mach, M. D. Hunter and R. S. Grewal, "Neurophysiological correlates in interface design: An HCI perspective", Computers in Human Behavior, 26 (2010), pp. 371-376.

[24] M. Malik, "Heart rate variability", Annals of Noninvasive Electrocardiology, 1 (1996), pp. 151-181.

[25] E. J. Malmstrom, E. Opton Jr and R. S. Lazarus, "Heart rate measurement and the correlation of indices of arousal", Psychosomatic Medicine, 27 (1965), pp. 546-556.

[26] M. Mauri, P. Cipresso, A. Balgera, M. Villamira and G. Riva, "Why is Facebook so successful? Psychophysiological measures describe a core flow state while using Facebook", Cyberpsychology, Behavior, and Social Networking, 14 (2011), pp. 723-731.

[27] R. McCraty, M. Atkinson, W. A. Tiller, G. Rein and A. D. Watkins, "The effects of emotions on short-term power spectrum analysis of heart rate variability", The American journal of cardiology, 76 (1995), pp. 1089-1093.

[28] T. O. Meservy, M. L. Jensen and K. J. Fadel, "Evaluation of competing candidate solutions in electronic networks of practice", Information Systems Research, 25 (2013), pp. 15-34.

[29] D. Meshi, C. Morawetz and H. R. Heekeren, "Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use", Frontiers in human neuroscience, 7 (2013), pp. 439.

[30] B. S. Minnery and M. S. Fine, "Neuroscience and the future of human-computer interaction", interactions, 16 (2009), pp. 70-75.

[31] F. Nasoz, C. L. Lisetti, K. Alvarez and N. Finkelstein, *Emotion recognition from physiological signals for user modeling of affect*, *Proceedings of the 3rd Workshop on Affective and Attitude User Modelling (Pittsburgh, PA, USA*, 2003.

[32] N. Oliver and F. Flores-Mangas, *HealthGear: a real-time wearable system for monitoring and analyzing physiological signals*, *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, IEEE, 2006, pp. 4 pp.-64.

[33] C. Peifer, *Psychophysiological correlates of flow-experience*, *Advances in flow research*, Springer, 2012, pp. 139-164.

[34] R. Riedl, R. D. Banker, I. Benbasat, F. D. Davis, A. R. Dennis, A. Dimoka, D. Gefen, A. Gupta, A. Ischebeck and P. Kenning, "On the foundations of NeuroIS: reflections on the Gmunden Retreat 2009", Communications of the Association for Information Systems, 27 (2010), pp. 15.

[35] R. Riedl, A. B. Randolph, J. vom Brocke, P.-M. Léger and A. Dimoka, "The potential of neuroscience for human-computer interaction research", SIGHCI 2010 Proceedings (2010).

[36] G. Riva, F. Mantovani, C. S. Capideville, A. Preziosa, F. Morganti, D. Villani, A. Gaggioli, C. Botella and M. Alcañiz, "Affective interactions using virtual reality: the link between presence and emotions", CyberPsychology & Behavior, 10 (2007), pp. 45-56.

[37] S. Sakuragi, Y. Sugiyama and K. Takeuchi, "Effects of laughing and weeping on mood and heart rate variability", Journal of physiological anthropology and applied human science, 21 (2002), pp. 159-165.

[38] G. E. Schwartz, D. A. Weinberger and J. A. Singer, "Cardiovascular Differentiation of Happiness, Sadness, Anger, and Fear Following Imagery and Exercise", Psychosomatic Medicine, 43 (1981), pp. 343-364.

[39] R. F. Simons, B. H. Detenber, T. M. Roedema and J. E. Reiss, "Emotion processing in three systems: The medium and the message", Psychophysiology, 36 (1999), pp. 619-627.

[40] S. Tams, K. Hill, A. O. de Guinea, J. Thatcher and V. Grover, "NeuroIS-alternative or complement to existing methods? Illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research", Journal of the Association for Information Systems, 15 (2014), pp. 723.

[41] C. Vanoyen Witvliet and S. R. Vrana, "Psychophysiological responses as indices of affective dimensions", Psychophysiology, 32 (1995), pp. 436-443.

[42] V. Venkatraman, A. Dimoka, P. A. Pavlou, K. Vo, W. Hampton, B. Bollinger, H. E. Hershfield, M. Ishihara and R. S. Winer, "Predicting advertising success beyond traditional measures: New insights from neurophysiological methods and market response modeling", Journal of Marketing Research, 52 (2015), pp. 436-452.

[43] S. R. Vrana, E. L. Spence and P. J. Lang, "The startle probe response: a new measure of emotion?", Journal of abnormal psychology, 97 (1988), pp. 487.

[44] J. Wagner, J. Kim and E. André, *From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification*, *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, IEEE, 2005, pp. 940-943.