2012 45th Hawaii International Conference on System Sciences

Measuring Users' Mental Strain when Performing Technology Based Surgical Tasks on a Surgical Simulator Using Thermal Imaging Technology

Jon R. Pluyter
Tilburg University
J.R.Pluyter@uvt.nl

Anne –F. Rutkowski
Tilburg University
A.Rutkowski@uvt.nl

Jack J. Jakimowicz
Delft University of Technology
Department of Research & Education, CZE Hospital Eindhoven
J.J.Jakimowicz@tudelft.nl

Carol S. Saunders
University of Central Florida
csaunders@bus.ucf.edu

Abstract

Information Systems (IS) researchers rely heavily on self-report measures, especially when studying the use, impact and adoption of Information Technology (IT). Recently psychophysiological and neurophysiological tools have been presented as an efficient way to gather measures and improve IS theory. The study presents thermal imaging as a technology that can be used to uncover different dimensions of IT constructs such as mental strain in high-tech contexts. Mental strain is an important factor in understanding the effects of using a new IT. Such complementary measures to self-reported scales can be used to triangulate results and reduce common method variance. In this paper we describe the use of thermal imagery to assess the effectiveness of curriculum protocols in handling stress when learning technology based surgery. Practical applications of the use of thermal imagery are then presented and discussed.

1. Introduction

An interesting paradox in the Information Systems (IS) field is that while concerned with technology, IS researchers principally use self-report measures to investigate Information Technology (IT) use, adoption, and intent to adopt. At least 57% of the empirical studies conducted on diffusion, impact and adoption of IT primarily rely on self-reports [1]. However, self-reports are subject to a range of errors and biases that might threaten the reliability and validity of empirical studies [2, 3]. Indeed such reports reflect the respondents' subjective experience and interpretation of a researched construct and therefore might lack objectivity [4]. The main advantage is that self-reports measures are relatively easy to obtain and are often the only feasible way to measure variables and constructs. Variables and constructs are particularly meaningful in building, developing, and testing Explicative and Predictive IS theories [5]. Increasing emphasis in the IS discipline has been placed on improving instrument quality and IS theory rigor. Still, Straub, Boudreau and Gefen [6] report that "IS positivist researchers still have major barriers to overcome in instrument, statistical and other forms of validation" (p. 5).

Recently, IS research went one step further in applying neuroscience theories and methods to advance our understanding of adoption theories and IT artifact design. The term neuroIS has been coined to describe that new idea (see [7]). Dimoka and Davis [8] used these advances in the medical technological world to test one of the best known Explicative and Predicative theories in the IS field: the Technology Adoption Model (TAM) originally developed by Davis, Bagozzi, & Warshaw [9]. They recommend using psychophysiological as well as brain imaging tools to inform IS theory.

In this paper we report our experience in using a psychophysiological tool to measure mental strain while using Augmented Reality (AR) surgical simulators to train psychomotoric skills required for technology based surgery in a controlled training setting [10]. This technology is used to assess the effectiveness of curriculum protocols in handling stress when learning technology based surgery as a substitute to self reported scales. More particularly we report data collected and analyzed to validate the use of unobtrusive thermal imaging technology to gather information on the potential mental strain of users learning how to manipulate surgical instruments. Indeed, research has demonstrated that excessive levels of mental strain have been associated with impaired performance and low ease of use in past research [11]. The curriculum protocols that we are testing can reduce the mental strain of the surgeon

978-0-7695-4525-7/12 $26.00 © 2012 IEEE
DOI 10.1109/HICSS.2012.31

and reduce overload created by IT in the operating room [12]. Studies employing the thermal imaging camera will lead then to increase surgeon performance and impact the diffusion of Augmented Reality training devices in the healthcare curriculum [13, 14].

The structure of the paper is as follows. The background section informs the reader on the (1) psychophysiological tools and (2) importance of efficiently measuring mental strain in sensitive contexts such as the operating room. The study, results, and possible applications are presented in the next three sections. We conclude that thermal imaging technology is a non-invasive, safe, comfortable, and a relatively cheap way to study the antecedents and consequences of technology adoption. It is also a complementary measure of mental strain of the IT user. Moreover the results of the data use easily-interpretable, explicit colors, and do not require knowing a lot about the human brain functioning.

2. Psychophysiological tools and mental strain

Motivated to better understand and support IS theories, IS researchers recently turned to neurophysiological tools [7]. Two main categories of neurophysiological tools exist: psychophysiological and brain imaging (see [7] for details). *Psychophysiological* tools measure changes in the physiology of the human body such as cardiac activities and skin conductance spawned by eye tracking. *Brain imaging* tools measure changes in brain neural activities. Recently, Dimoka and Davis [8] used functional Magnetic Resonance Imaging (fMRI) techniques to match TAM constructs of perceived complexity and ease of use to a certain level of brain activity when dealing with the visual stimulus of a web offer. Such tools enable the measurement of users' responses to IT with data drawn directly from the human body (see [7] for an exhaustive list of such technologies).

Our focus is on the use of psychophysiological tools in medical settings. Using medical methods in healthcare research is relatively common. For example physiological parameters such as systolic blood pressure, urinary epinephrine or cortisol awakening responses have been used to establish stress in relationship between perceived work overload and chronic worrying of paid participants [15] or woman nurses [16].

Psychophysiological tools enable the measurement of users' responses to the training technology with physiological data that are not susceptible to social desirability bias [7]. That is important since response bias is particularly likely to occur when the construct of interest is sensitive in nature [3]. Mental strain is often so sensitive that it is not discussed in healthcare settings. It is also not part of the surgeon culture to accept stress as an inevitable part of practice [17]. However, it is important to deal with mental strain effectively. One way of doing so is to increase the adoption of AR technologies to improve surgeon training.

Mental strain is referred to in thermal imaging literature as cognitive workload [18], mental workload and mental demand [19], mental effort [11], frustration [20], and stress [21]. In fact these represent cognitive and emotional manifestations of overload [12]. Previous studies demonstrated that overload negatively impacts performance [22] and the adoption and diffusion of technology [23].

Mental strain is associated with temperature changes of specific areas of the human face due to activity alteration of the autonomic nervous system [19]. It triggers subtle increases in the temperature of the skin directly surrounding the periorbital [11, 18]. Excessive levels of mental strain measured by the thermal imaging camera are associated with impaired performance and low perceived ease of use of the technology under study [11].

3. Study

3.1 Thermal imaging technology

Facial temperature can be measured unobtrusively with a thermal imaging camera (see Figure 1). It measures infrared radiation naturally emitted by any object. The amount of radiation emitted increases with the temperature of the object [21]. The thermal imaging camera accordingly constructs a thermographic image of the object such as the human face (see Figure 2). Thermal imaging of the face is used in other contexts such as mental strain in car drivers [19]. It potentially is a valuable tool to objectively and continuously measure facial temperature of IT users including surgeons and surgical residents performing technology based surgery [24].

The Thermoview 8300 infrared thermal imaging camera (Sensor Partners, Drunen, The Netherlands; see Figure 1) was used to measure the participant's facial temperature. This camera allows continuous measurement of infrared radiation in the spectral range of 8-14μm with a thermal sensitivity of 0.08 degrees Celsius. This makes it particularly suitable for detecting relatively subtle temperature differences in the human face. We positioned the camera on a tripod and focused it on participants' faces in a somewhat upward angle to provide an optimal image of the participants' periorbital area. The participants' position and camera angle remained constant throughout the simulator sessions as all exercises were performed using a simulator screen with a fixed position. The study was performed in a temperature-controlled room to minimize the impact of external factors on the skin temperature of the participants.

Figure 1. Thermal imaging technology

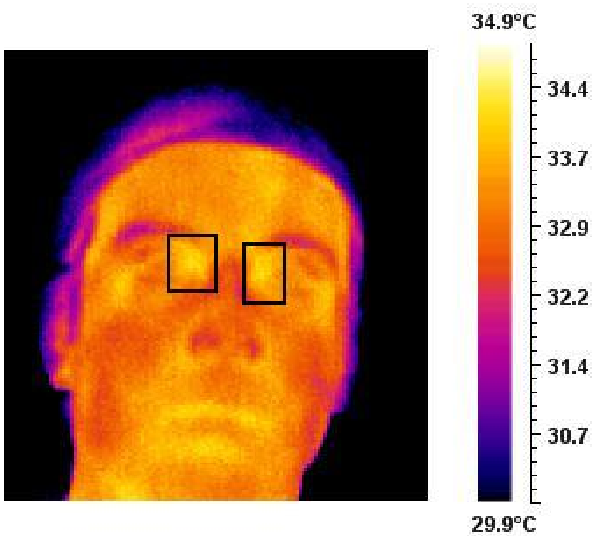

Figure 2. Thermographic image: periorbital area marked with rectangle

3. 2 Surgical technology

The Haptica Promis I Augmented Reality (AR) simulator (Haptica, Dublin, Ireland) is a validated training device that is already used for training surgeons [10, 14]. It was used in this study. Previous research empirically demonstrated that simulated surgery imposes high mental strain on the user [25]. A set of three basic surgical tasks (i.e., aiming and translocating, stretching, and cutting simulated tissue) was selected to match the prior knowledge of the participants about the tasks. These tasks mimic component tasks needed to perform laparoscopic surgery such as gallbladder removal. This kind of simulation is applicable for surgical training curricula.

3.3 Participants and procedures

Twenty-one participants were invited to an intensive training session on the surgical AR simulator. Participants should have no prior experience with AR, nor knowledge of the exact nature of the medical task. If they had, their chunking abilities built through previous experience [12], may have biased the results. The subjects were undergraduate business students who had similar backgrounds regarding IT usage. We excluded one participant from the sample set because he was an online gamer. We did not want to use medical students, in this pilot phase, because they may have been familiar with AR and the medical tasks that were being performed in the study. Participants (mean age 23, 81% male, 19% female) completed surgical tasks in a timeframe of two hours, divided over two runs. During each run the set of three basic surgical tasks was performed once by each of the participants in executing the tasks in the same order. Periorbital skin temperature was measured at baseline, and after the first and second run of technology based surgery on the surgical simulator. Participants were instructed to sit on a chair upon arrival at the temperature controlled training lab. Their baseline temperature was measured 10 minutes after arrival. This period allowed the participants to acclimatize to the training room temperature and to refrain from engaging in complex cognitive activities for a considerable period of time (cf. [19]).

4. Results

Periorbital temperature was measured at baseline and after the first and second run of technology based surgery on the surgical simulator. Table 1 and Figure 3 provide an overview of the periorbital temperature at these three points in time. Analysis using the paired sample t-test revealed that periorbital temperature after the first simulator run was significantly higher than baseline (p=.001). Periorbital temperature after the second simulator run was also significantly higher than baseline (p=.0001).

Table 1. Periorbital temperature in degrees Celsius at baseline and during surgery on the simulator

	Baseline	Surgery on simulator (first run)	Surgery on simulator (second run)
Mean (SD)	34.39 (.56)	34.90* (.49)	34.95* (.45)
Min	33.50	34.20	33.90
Max	35.60	35.90	35.60
* significantly different from baseline at alpha=.05			

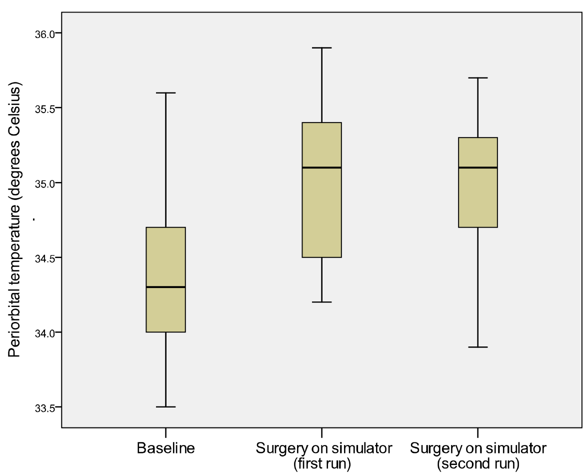

Figure 3. Periorbital temperature during baseline and surgery on the simulator

5. Discussion

The results indicate that facial skin temperature measured by the thermal imaging technology is higher during technology based simulated surgery than during the baseline. Hence, thermal imaging appears to be a viable technology to gather physiological measures of IT users' mental strain when using the AR training device. The technology can be used to measure skin temperature continuously and unobtrusively without inducing anxiety or stress in the study subjects. It can be used to complement and possibly triangulate self-report measures to increase validity of the study and reduce common method variance. It extends the set of psychophysiological measures of IT constructs proposed by [7]. As opposed to other physiological measures such as fMRI, thermal imaging is a passive modality. It depends solely on radiation naturally emitted by the subject [21]. Hence it does not raise additional health concerns for the IT user whose reactions are being measured. Moreover thermal imaging is non-invasive and unobtrusive. It does not interfere with the tasks performed by the IT user, nor does it limit the users' movement while performing the task.

Thermal imaging can be used as a complementary measurement to improve Explicative and Predictive theory in IS. For example, Ahuja & Thatcher [23], when using self-reports in their research, found that overload and associated manifestations of mental strain negatively impact technology diffusion. Thermal imaging can be used to complement and triangulate these results using a physiological measure. Physiological measures are also often more accurately predict behavior than self report measures such as those currently used in, adoption and post-adoption research [7]. Thermal imaging of IT users' mental strain has a wide range of potential applications in diffusion and use of IT. Several applications are discussed both within and beyond the healthcare community.

5.1 Possible applications: diffusion and use of IT

Thermal imaging of mental strain could potentially be used to support or accompany training and assessment of surgeons' and surgical residents' cognitive ability to deal with mentally demanding technology based surgery. It can help identify mental strain and associated stress during technology based surgery and assess the negative impact on performance that surgical staff is typically reluctant to recognize [17]. For example, surgical residents demonstrating excessive mental strain when performing basic surgical tasks could be instructed to follow additional simulation training before to operating on real patients (cf. [24]). Thermal imaging could be used to assess the cutoff point where trainees can proceed from training to practice. One could think of similar applications in other high-tech and high-risk professions such as aviation and aerospace.

Excessive levels of mental strain when using IT resulting in impaired performance may have far reaching consequences. Thermal imaging of mental strain could be used interactively with the IT user to improve performance [21]. For example, thermal imaging could potentially detect excessive levels of mental strain imposed by IT during technology based surgery, office work, or even technology mediated auctions. The IT user can be notified and instructed to take a cool down period in order to avoid cognitive overload that might result in surgical errors, degraded performance levels, or faulty auction bids. Finally, managers could use thermal imaging to identify which potentially interesting technologies could best be adopted and diffused within the organization as a function of the mental strain they impose on the IT user.

5.2 Limitations and future research

This lab-setting study is aimed at testing the appropriateness of thermal imaging as a physiological measure to help identify IT users' mental strain during technology based surgery on a surgical simulator. Therefore we only compared technology based surgery to baseline. It would be interesting to experimentally manipulate mental strain on various levels (e.g. low-medium-high) to determine the sensitivity and specificity of the thermal imaging technology measuring mental strain. Future studies should also attempt to identify correlations between physiological measures of mental strain and self-reports. For example, the camera could be used to study the antecedents of adoption such as ease of use in tests of the TAM model.

Also, it would be interesting to use the technology within a real setting with experienced surgeons, or advanced medical students. Indeed lab settings are often too remote from the real life. We are planning to use the thermal imaging technology to measure mental strain of the surgical team during real technology based surgery in the operating room.

Like any other measure, the use of thermal imaging technology raises challenges that require attention. From a data analysis perspective it is rather difficult to delineate the exact region of interest (e.g. forehead, periorbital, nose). Also, currently there is no consensus about the unit of analysis. Authors use measures such as mean and standard deviation [11], or hottest pixels [21] to represent the temperature of the region of interest. Further research is required to identify the delineation and unit of analysis that best represents the construct of interest. From a data collection perspective, thermal imaging of the face requires a relatively stable position of the subject. Moving subjects can however be traced using a motion tracking system (see [21]). Ethically, subjects have to be informed that they are being monitored and researchers need approval from the Institutional Research Board. Finally, thermal imaging requires a temperature controlled room to control the impact of external sources (e.g. direct sunlight, heating, air-conditioning) on body temperature.

It should be noted that psychophysiological measures should be complementary to excising measures including self-reports and brain imaging tools. Multiple sources of measurement allow triangulation and might shed light on different dimensions of the construct of interest. For example, thermal imaging can extend the assessment of the body's physiological manifestation of mental strain when confronted with technology. Brain imaging tools can shed light on associated brain activity. Self-reports can be used to gather data on the subjective experience and perception of mental strain. All measures can inform the researcher on different aspects of mental strain imposed by IT. They also counterbalance part of each other's methodological weaknesses and reduce the impact of common method variance.

6. Conclusion

IS researchers strongly rely on self-report measures, especially when studying the use, impact and adoption of IT. Self-reports are subject to several errors and biases that might threaten the reliability and validity of the study, including response bias and common method variance. This study empirically demonstrates that thermal imaging of mental strain of the IT user potentially provides a promising complementary physiological measure. Psychophysiological measures can be used to uncover different dimensions of IT constructs such as mental strain when using newly adopted technologies [11]. They can also be used in testing TAM and Theory of Planned Behavior Models when measuring antecedents of adoption such as ease of use to the extent that a tool that is easier to use should create less strain. Finally, they can be used to triangulate results of self report measures and brain imaging tools to reduce common method variance. Thermal imaging additionally has multiple practical applications. These include diffusion and use of simulation technology in healthcare curricula and IT in various social communities.

Finally, we conclude that the thermal camera has three main advantages that make it worthwhile to consider. First of all it does not require a large budget nor extensive knowledge to interpret the data. Ethically speaking, it can be used without disturbing the subjects or causing any potential damage. This could be especially useful when conducting research in a real (vs. lab) setting, such as the operating room. Third, the interpretation of the results is computerized and based on visualization effects. In that respect the technology is rather intuitive. We were among the first to use thermal imaging to efficiently map mental strain in surgical technology use. We think it could be a good tool to investigate and improve models of diffusion and use of IT.

7. Acknowledgements

The authors would like to thank Delft Faculty of Industrial Design (Delft University of Technology, Delft, the Netherlands) for providing the thermal imaging camera and accompanying technical support. The authors also would like to acknowledge the Skillslab of Catharina Hospital (Eindhoven, the Netherlands) for their generous technical assistance. This research was partially supported by the Scientific Fund Catharina Hospital.

8. References

[1] Y. Dwivedi, M. D. Williams, B. Lal, and A. Schwarz, "Profiling adoption, acceptance and diffusion research in the information systems discipline", Paper presented at the 16th European Conference on Information Systems, Galway, Ireland, 2008.

[2] D. T. Campbell and D. W. Fiske, "Convergent and discriminant validation by the multitrait-multimethod matrix", Psychological Bulletin, 56(2), 1959, pp. 81-105.

[3] S. I. Donaldson and E. J. Grant-Vallone, "Understanding self-report bias in organizational behavior research", Journal of Business and Psychology, 17(2), 2002, pp. 245-260.

[4] E. M Hufnagel and C. Conca, "User response data: The potential for errors and biases", Information Systems Research, 5(1), 1994, pp. 48-73.

[5] S. Gregor, "The nature of theory in information systems", MIS Quarterly, 30(3), 2006, pp. 611-642.

[6] D. Straub, M. C. Boudreau, and D. Gefen, "Validation guidelines for IS positivist research", Communications of the Association for Information Systems, 13(1), 2004, pp. 380-427.

[7] A. Dimoka, R. D. Banker, I. Benbasat, F. D. Davis, A. R. Dennis, D. Gefen, A. Gupta, A. Ischebeck, P. Kenning, P. Pavlou, G. Müller-Putz, R. Riedl, J. vom Brocke, and B. Weber, "On the Use of Neurophysiological Tools in IS Research: Developing a Research Agenda for NeuroIS", MIS Quarterly forthcoming, 2011.

[8] A. Dimoka and F. D. Davis, "Where does TAM reside in the brain? The Neural Mechanisms underlying Technology Adoption", Proceedings of the 29th International Conference on Information Systems, Paris, France, 2008.

[9] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: a comparison of two theoretical models", Management Science, 35(8), 1989, pp. 982-1003.

[10] M. Schijven and J. Jakimowicz, "Virtual reality surgical laparoscopic simulators", Surgical endoscopy, 17(12), 2003, pp. 1943-1950.

[11] M. L. Reyes, J. D. Lee, Y. Liang, J. D. Hoffman, and R. W. Huang, "Capturing driver response to in-vehicle human-machine interface technologies using facial thermography", Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Big Sky, Montana, 2009.

[12] A. –F. Rutkowski and C. S. Saunders, "Growing pains with information overload", Computer, 43(6), 2010, pp. 94-96.

[13] A. G. Gallagher, E. M. Ritter, H. Champion, G. Higgins, M. P. Fried, G. Moses, C. D. Smith, and R. M. Satava, "Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training", Annals of Surgery, 241(2), 2005, pp. 364.

[14] J. Jakimowicz and A. Cuschieri, "Time for evidence-based minimal access surgery training—simulate or sink", Surgical Endoscopy, 19(12), 2005, pp. 1521-1522.

[15] W. Schlotz, J. Hellhammer, P. Schulz, and A. A. Stone, "Perceived work overload and chronic worrying predict weekend–weekday differences in the cortisol awakening response", Psychosomatic Medicine, 66(2), 2004, pp. 207-214.

[16] I. B. Goldstein, D. Shapiro, A. Chicz-DeMet, and D. Guthrie, "Ambulatory blood pressure, heart rate, and neuroendocrine responses in women nurses during work and off work days", Psychosomatic Medicine, 61(3), 1999, pp. 387-396.

[17] S. Arora, L. Hull, N. Sevdalis, T. Tierney, D. Nestel, M. Woloshynowych, A. Darzi, and R. Kneebone, "Factors compromising safety in surgery: stressful events in the operating room", The American Journal of Surgery, 199(1), 2010, pp. 60-65.

[18] J. Stemberger, R. S. Allison, and T. Schnell, "Thermal Imaging as a Way to Classify Cognitive Workload", Canadian Conference Computer and Robot Vision, Ottawa, Ontario, Canada, 2010.

[19] C. K. L. Or and V. G. Duffy, "Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement", Occupational Ergonomics, 7(2), 2007, pp. 83-94.

[20] C. Puri, L. Olson, I. Pavlidis, J. Levine, and J. Starren, "StressCam: non-contact measurement of users' emotional states through thermal imaging", Conference on Human Factors in Computing Systems, Portland, OR, USA, 2005.

[21] I. Pavlidis, J. Dowdall, N. Sun, C. Puri, J. Fei, and M. Garbey, "Interacting with human physiology", Computer Vision and Image Understanding, 108(1-2), 2007, pp. 150-170.

[22] C. O'Reilly, "Individuals and information overload in organizations: is more necessarily better?", Academy of Management Journal, 23(4), 1980, pp. 684-696.

[23] M. K. Ahuja and J. B. Thatcher, "Moving beyond intentions and toward the theory of trying: effects of work environment and gender on post-adoption information technology use", MIS Quarterly, 29(3), 2005, pp. 427-459.

[24] B. Bass, E. Silverman, P. Dowdall, T. Bourlai, I. Pavlidis, and B. Dunkin, "High definition thermal and visual facial monitoring: A valuable metric for surgical skill acquisition", Sages Annual Conference, Paris, 2008.

[25] R. Berguer, W. D. Smith, and Y. H. Chung, "Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery", Surgical endoscopy, 15(10), 2001, pp. 1204-1207.