2013

IS Cognitive Load: An Examination of Measurement Convergence

Mary Dunaway
University of Arkansas - Fayetteville, MDunaway@walton.uark.edu

Zachary Steelman
University of Arkansas - Fayetteville, zach@zachsteelman.com

IS Cognitive Load: An Examination of Measurement Convergence

Doctoral Student Roundtable

Mary M. Dunaway
University of Arkansas - Fayetteville
mdunaway@walton.uark.edu

Zachary R. Steelman
University of Arkansas - Fayetteville
zach@zachsteelman.com

ABSTRACT

Despite the growth and interest in information processing research, understanding the supporting role of information systems (IS) has been limited. While cognitive processing of information has been examined in learning environments with traditional learning tasks, the investigation of cognitive load within complex simulated IS learning environments has received less attention. Traditional measurement allows for a broad user evaluation of the ISs and actual usage from a holistic perspective; however, detailed synchronous evaluation of cognitive load during the usage of the IS may allow for more accurate assessment of how system features influence cognitive load and subsequent performance outcomes. Therefore, this research attempts to integrate traditional subjective and physiological measurements to examine cognitive load within a dynamic simulated IS learning environment. This research study focuses on how subjective and objective physiological (galvanic skin response (GSR), heart rate variability (HRV), and electroencephalography (EEG) measures of cognitive load compare in simulated IS training environments.

Keywords

Cognitive Load, NeuroIS physiological tools, training, IS learning.

INTRODUCTION

Recent directions in research have led to the emergent evaluation of theories and foundations with objective physiological and neurological measurement (Dimoka, Pavlou & Davis, 2011). This thrust has been an attempt to validate and evaluate methods of collecting data on the physiological changes that occur in an individual during various phenomena. Typically, empirical evidence of theoretical testing has been evaluated with self-report scale-based measurement items that attempt to capture an unobservable behavior. Due to the latency characteristics, post-hoc evaluation, and self-report bias of many of our traditional techniques, the evaluation of measurement error, validity, and reliability can be a significant concern compared to objective measures.

Within the field of IS an emerging research stream of NeuroIS has developed which focuses on "the use of cognitive neuroscience theories, methods, and tools to inform IS research" (Dimoka et al. 2007 pp. 1). The use of neuroscience and physiological measurement techniques can inform and enrich our ability to capture and measure objective data during IS phenomena, compare traditional techniques and methodologies, and examine the differences and commonalities within each method. This ability for cross examination and triangulation of methods via multi-trait multi-method (MTMM) techniques allows researchers to delve deeper into complex phenomena via multiple measure comparison to assess construct validity. An area of interest in both the ISs research as well as the mature stream of psychophysiological assessment is cognitive load.

Cognitive load (CL) can be defined as the cognitive effort made by an individual to understand and perform his/her task (Sweller, 1988). The utilization of ISs is a core support tool assisting individuals with handling large amounts of data in an attempt to reduce the cognitive load. Prior literature has well-grounded the concept and simplified its theoretical association to short-term or working memory (Ayres, 2006; Paas, Renkel, & Sweller, 2003; Sweller, 2006). Cognitive load can typically be extracted through three measurement techniques: subjective measures, performance measures, and neurophysiological measures (Galy, Cariou, & Melan, 2011). Yet, these methods still provide limited detail of the association with cognitive brain processing. Recent advances in cognitive neuroscience and psychophysiological measurement allow for the capture of objective measures during the phenomenenon of interest without relying on post-hoc, self-reported evaluations of cognitive load.

The purpose of this study is to empirically examine the effect of cognitive load on learning and performance outcomes utilizing subjective, psychophysiological, and neurophysiological measures. We conduct controlled lab experiments utilizing ERPsim, an SAP simulation training environment, to examine cognitive load within IS learning environments by HRV, GSR, and EEG in real-time. We cross-validate the psychophysiological and neurophysiological measures with traditional survey-based scales via a multitrait-multimethod (MTMM) technique to examine construct validity for the cognitive load measures. These measures are further investigated to explore their influence on learning and performance outcomes (e.g. task performance, satisfaction, effectiveness).

COGNITIVE LOAD THEORY

Cognitive Load Theory (CLT; Yeung, Jin & Sweller, 1998) focuses on the aspects of mental architectures of learners which influence their performance in learning tasks. One of the foundational assumptions surrounding CLT is that individuals are working with a limited amount of working memory and requirements to complete various mental tasks. When the required cognitive load of a task does not exceed the available working memory of an individual there are adequate mental resources to integrate and absorb the information required.

Germane load is utilized for the schemata construction within an individual's long term memory and is highly effective for the learning process (Kalyuga, 2009). A variety of strategies have been developed around the use of worked examples which provide increased germane load and learning capabilities for individuals (Pass & Van Gog, 2006; Paas & Van Merrienboer, 1994). Therefore, developers of training and learning environments must find ways to optimize the cognitive load of individuals such that the training itself lowers both intrinsic and extraneous load while increasing the germane load presented within the environment. These dimensions of cognitive load are considered to be additive such that their total must remain below available mental resources to enable learning without creating a cognitive overload. Therefore, the development of training and learning environments relies upon the active monitoring and examination of cognitive load components affected to develop effective and efficient learning methods.

COGNITIVE LOAD MEASUREMENT

Traditionally researchers have utilized two types of methods to measure cognitive load levels via techniques such as survey questionnaires and secondary tasks (Palinko et al., 2010; Cierniak, Scheiter, & Gerjets, 2008). These post-hoc cognitive load measures are a proxy for the actual difficulty and mental effort of an individual across the entire task timeline.

SUBJECTIVE SCALE

Most often in IS research, subjective scales capture psychological measures of behaviors such as knowledge, abilities, attitudes, and shared understandings (Galliers & Land, 1987). Much of the research utilizes a survey method where responses are self-reported by individuals. The empirical evidence is used to draw inferences about the validity of theory and measurement (Peter & Churchill, 1986). There is an underlying presumption that subjective assessments should correspond reasonably closely with objective data. However, no subjective scale is without measurement error. The scales are developed to capture measures as accurately as possible, or at least partially represent the constructs of interest.

Largely, respondents' framework of assessment can change in the course of learning due to adaption processes or as a response to motivational and emotional changes to decrease reliability (Schnotz, Wolfgang, & Kurschner, 2007). Yet, there are several advantages of subjective ratings. They are simple and easily applicable; moreover subjective ratings are most often captured in a natural setting, which increases the ecological validity while revealing valuable data. Thus, the varied subjective scale outcomes suggest the need to investigate additional measures to enhance accuracy and explanation for CL nomological and theoretical validity.

OBJECTIVE PERFORMANCE MEASURES

Objective performance measures include both a primary and secondary tasks (Gwizdka, 2010). Primary performance measures include the number of errors, accuracy, task completion time relative to the user population time, and ratio of the actual completion time to a baseline (Wickens, Hellenberg & Xu, 2002). These type of tasks are applicable when the task performance pace is externally controlled. For example, an IT worker receives instructions on how to fix a PC problem. When the individual in this scenario controls the task pace, the applicability to the measure becomes uncertain.

Secondary tasks typically involve some aspects of monitoring external events which can be delivered periodically though the visual or auditory channel. The delivery choice of the secondary task is highly dependent on the primary task. Methods that involve performance on a secondary task are called dual-task techniques (Kim & Rieh, 2005; Cegarra & Chevalier, 2008). CL measures derived from performance on a secondary task are based on limited cognitive resources. The secondary task performance may increase as more resources are required by the primary task. Despite the ability for this technique to be highly sensitive and reliable in detecting CL levels it has received limited utilization (Sweller, 1988; Chandler & Sweller, 1996). This utilization may limit ecological validity when the inflated artificial resources are required for the tasks of interests.

PSYCHOPHYSIOLOGICAL METHODS

The physiological measurement techniques focused on in this study are objective and include EEG, GSR, and HRV. These measurement techniques have been found in prior literature related to CL (Sweller, 2010; Schnotz, Wolfgang, & Kurscher, 2007; Minassian et al., 2004; Mulder, 1992). Sweller's (2010) results provided greater interpretation of CL effects in particular when two participants experienced the same amount of overall CL effects were not easily detectable. In addition to dual-task measures, physiological measurement techniques have the advantage of reflecting dynamic, real-time data collection during a task.

EEG is identified as a physiological index that can serve as an online, continuous measure of cognitive load that has higher sensitivity of CL fluctuations when overall CL measures fail to detect differences in cognitive processing (Antonenko & Niederhauser 2010). In the GSR method, current is passed through the body with the skin resistance measured (active GSR) or the current generated by the body itself (passive GSR) (Schnotz & Kurschner, 2007). The advantages of the GSR method are that it provides a relatively simple method for examining the function of the sympathetic autonomic nervous system and that it is not prone to the introspective skills of the individual. The disadvantage of the method is that it cannot be used in natural settings, which reduces its ecological validity. Finally, amplitudes tend to habituate and vary depending on the experimental conditions. Accordingly, the framework of reference for data interpretation can also change in the course of learning.

The use of heart HRV rate variability for measuring cognitive load is based on the assumption that controlled processing is related to a specific cardiovascular state that manifests itself in the HRV power spectrum band (Mulder, 1992). Cognitive effort is supposed to be directly related to control processing, which in turn causes a change in the power spectrum. Paas et al. (1994) found, that this method was not more useful than subjective ratings. Alternatively, DeRivecourt et al. (2008) findings show that HRV from short data segments provided more insight in intermediate levels of mental effort. Their results support a detection of change in mental workload. Given the inconsistent findings, this study will further provide insights into HRV validity for CL measures.

The following section details our methodology and experiment to compare the similarities and differences across multiple cognitive measurement techniques. We examine these effects within the simulated SAP enterprise system ERPsim which provides a dynamic software environment that mimics real economic activities in accelerated time to provide users with real-world experience with the software. Each participant provides multiple cognitive load measurement responses to allow comparison and validity of the techniques within this unique environment.

METHOD

Participants

Graduate and undergraduate student participants at a large Midwestern university will be recruited for this study. Participants will be presented with detailed consent forms outlining the task, equipment being used, and approximate duration of the study. The experimental protocol duration is approximately one and a half hours. Participants will be compensated $20.00 for their participation. Demographic and descriptive data will be produced for sample.

Research Design

A controlled lab experiment is conducted to examine the changes in cognitive load between and within individuals categorized by varying levels of SAP and ERP technology expertise. For this study, the participants will be randomly assigned into five teams. Each team will play the same scenario against computer players. Within each team, every player is assigned a difficulty level. In the experiment pre- and post- test, individuals will complete an ERP expertise evaluation (Cronan et al., 2010) and the self-report cognitive load survey. Figure 1 depicts the experiment timeline.

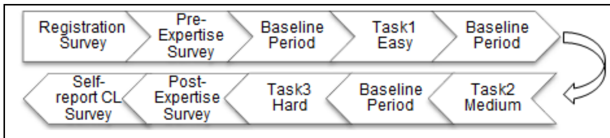

Figure 1. Experimental Timeline

Training System and Task

The utilization of simulated software environments allows for individuals to experience near real-world situations and then applies that knowledge to situations outside of the training environment. Detailed computer logs and click-stream data can be associated continuously with the cognitive load levels. Traditional measurement techniques typically rely upon post-hoc measurement of cognitive load which can limit a researcher's ability to fully measure CL phenomena.

The task for this experiment consists of a series of simulation tasks utilizing the SAP enterprise resource planning (ERP) system implemented through the ERPsim software (Léger et al., 2007). ERPsim is a dynamic simulator of SAP's ERP system which allows for the immersion of individuals into the cash-to-cash business process. The simulator allows individuals to play against each other or the computer in the manufacturing, distribution, and sale of products where transactions are generated by the simulator within a virtual economy in condensed time periods. The simulation experimental design has been utilized in previous studies capturing psychophysiological measurement and has been found to be a viable testing (Léger et. al, 2010; Caya et al., 2011; Léger et al., 2012; Caya et al., 2012) and training environment (Léger et al., 2011; Cronan et al., 2012; Cronan & Douglas, 2013).

Psychophysiological and Neurological Apparatus

For all neurological and psychophysiological measurement, BIOPAC MP150 data acquisition system will utilized. HRV is captured with two electrodes utilizing silver-oxide gel applied by one of the researchers and placed on the left ankle, right ankle, and left forearm. GSR will be collected with a single electrode connected to the middle finger of the left hand. Each electrode and leading wire is held down with surgical tape to reduce potential disconnects and loosening during the procedure.

The B-Alert X10, wireless helmet will capture a matrix of EEG signals in standard neurological locations. This type of apparatus allows for consistent placement of the electrodes to reduce measurement error and increase reliability.

MEASURES

Subjective measurements will be collected from two measures of cognitive load typically utilized in previous literature: the Paas 9-point mental effort scale (Paas, 1992) and the NASA-TLX (Task Load Index) (Hart & Staveland, 1988). The Paas 9-point mental effort scale ranges from (1) very, very low mental effort to (9) very, very high mental effort. The NASA-TLX utilizes six dimensions of task load (Mental, Physical, Temporal, Performance, Effort, and Frustration) ranging on scales from 0 to 100 to assess the required mental workload for each individual. Each measure will be captured at the end of each trial and before the baseline rest period starts.

Objective performance of the simulation task is measured by examining both the total profits generated by the individual as well as the ratio of the individual's profit to the computer's profit to calculate a relative index of performance between individuals and across difficulty levels. Satisfaction with the training simulation is captured with a 4 item scale adapted from Bhattacherjee (2001) which asks participants to evaluate how they feel about their overall system experience using a semantic comparison scale.

Experimental Controls

Our experimental design comprises three distinct levels of task difficulty for each participant and dummy variables for each level as well as an ordinal variable for regression analyses. Additionally, to validate experimental design conditions, we capture a subjective measure of perceived task difficulty (Paas, 1994). Each individual's pre-and post- training expertise will be evaluated using an 18-item ERP knowledge scale (Cronan et al. 2010). These measures capture the objective knowledge of Enterprise System Management, Business Process, and SAP Transaction Skills knowledge. Also, we control for a series of individual level attributes which are believed to have potential influence within our study including age, gender, grade point average, college major, computer self-efficacy (Compeau & Higgins, 1995), and computer anxiety (Thatcher & Perrewe, 2002).

CONTRIBUTIONS

This project will provide three main contributions to the literature on cognitive load, dynamic simulated training environments, and IS design and implementation. First, our utilization of a variety of measurement tools to examine cognitive load via surveys, neurological, and psychophysiological tools provides insights into which type of measurement can be utilized for examining different aspects of cognitive load. Second, our results provide a deeper understanding into the variability of CL as well as how these measures convergence. Third, we contribute to the literature on dynamic simulation training environments by exploring how these measures can provide greater insights into user experience and learning.

ACKNOWLEDGMENTS

We would like to thank Dr. Fred Davis, Dr. Pierre-Majorique Léger, and Dr. Paul Cronan for their endearing support, guidance, and encouragement for our research ideas and projects in the area of NeuroIS.

REFERENCES

1. Antonenko, P. D., & Niederhauser, D. S. (2010). The influence of leads on cognitive load and learning in a hypertext environment. Computers in Human Behavior, 26(2), 140-150.
2. Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. Applied Cognitive Psychology, 20(3), 287-298.
3. Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. Mis Quarterly, 351–370.
4. Caya, O., Léger, P. M., Grebot, T., & Brunelle, E. (2012). Integrating, sharing, and sourcing knowledge in an ERP usage context. Knowledge Management Research & Practice.
5. Caya, O., Brunelle, É., Leger, P. M., & Grebot, T. (2012, January). An Empirical Study on Emotions, Knowledge Management Processes, and Performance within Integrated Business Process Teams. In System Science (HICSS), 2012 45th Hawaii International Conference on System Sciences, 514-524.
6. Cegarra, J., & Chevalier, A. (2008). The use of tholos software for combining measures of mental workload: Toward theoretical and methodological improvements. Behavior Research Methods, 40(4), 988-1000.
7. Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. Applied Cognitive Psychology, 10(2), 151-170.
8. Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? Computers in Human Behavior, 25(2), 315-324.
9. Compeau, D. R., & Higgins, C. A. (1995). JSTOR: MIS Quarterly, Vol. 19, No. 2 (Jun., 1995), pp. 189-211. MIS Quarterly, 19(2).
10. Cronan, T. P., Léger, P.-M., Charland, P., Robert, J., & Babin, G. (2010). Comparing Objective Measures and Perceptions of Cognitive Learning in an ERP Simulation Game: A Research Note. Information Technology Research Institute, ITRI-WP153-0910.
11. Cronan, T. P., Léger, P. M., Robert, J., Babin, G., & Charland, P. (2012). Comparing Objective Measures and Perceptions of Cognitive Learning in an ERP

Simulation Game A Research Note. Simulation & Gaming, 43(4), 461-480.
12. Cronan, T. P., & Douglas, D. E. (2013). Assessing ERP Learning (Management, Business Process, and Skills) and Attitudes. Journal of Organizational and End User Computing (JOEUC), 25(2), 59-74.
13. Dimoka, A., Pavlou, P. A., & Davis F. (2007). Neuro-IS: The potential of cognitive neuroscience for information systems research. Proceedings of the 28th International Conference on Information Systems, 11-14.
14. DeRivecourt, D., Kuperus, M. N., Post, W. J., & Mulder, L. J. M. (2008). Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight, Ergonomics, 51(9), 1295-1319.
15. Galliers, R. D., & Land, F. F. (1987). Choosing Appropriate Information Systems Research Methodologies. Communications of the ACM, 30(11), 900-902.
16. Galy, E., Cariou, M., & Mélan, C. (2011). What is the relationship between mental workload factors and cognitive load types? Journal of the International Organization of Psychophysiology.
17. Gwizdka, J. (2010). Distribution of cognitive load in web search. Journal of the American Society for Information Science and Technology, 61(11), 2167-2187.
18. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Human Mental Workload (Vol. 1, pp. 139-183).
19. Kalyuga, S. (2009). Instructional designs for the development of transferable knowledge and skills: A cognitive load perspective. Computers in Human Behavior, 25(2), 332-338.
20. Kim, Y. M., & Rieh, S. Y. (2005). Dual-task performance as a measure for mental effort in library searching and web searching. Proceedings of the 68th Annual Meeting of the American Society for Information Science and Technology, 42.
21. Léger, P.-Majorique, Davis, F., Perret, J., and Dunaway, M., "Psycholphysiological Measures of Cognitive Absorption" (2010). SIGHCI 2010 Proceeding, Paper 9. http://aisel.aisnet.org/sighci2010/9.
22. Léger, P.-Majorique, Charland, P., Feldstein, H. D., Robert, J., & Babin, G. (2011). Business Simulation Training in Information Technology Education Guidelines for New Approaches in IT Training Trends in Higher Education From Situational Theory to Simulation Games. Journal of Information Technology Education, 10.
23. Léger, P. M., Cronan, P., Charland, P., Pellerin, R., Babin, G., & Robert, J. (2012). Authentic OM problem solving in an ERP context. International Journal of Operations & Production Management, 32(12), 1375-1394.
24. Léger, P.-majorique, Robert, J., Babin, G., Pellerin, R., & Wagner, B. (2007). ERPSim. HEC Montreal, Montreal, Qc.: ERPsim Lab (erpsim.hec.ca).
25. Minassian, A., Granholm, E., Verney, S., & Perry, W. (2004). Pupillary dilation to simple vs. complex tasks and its relationship to thought disturbance in schizophrenia patients. International Journal of Psychophysiology, 52(1), 53-62.
26. Mulder, L. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. Biological Psychology, 34(2), 205-236.
27. Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. Journal of Educational Psychology, 84(4), 429-434.
28. Paas, F., Merrienboer, J.J.G, and Adam, J.J. (1994). Measure of cognitive load in instructional research. Perceptual and Motor Skills, 79, 419-430.
29. Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. Educational Psychologist, 38(1), 1-4.
30. Paas, F., & Van Gog, T. (2006). Optimising worked example instruction: Different ways to increase germane cognitive load. Learning and Instruction, 16(2), 87-91.
31. Paas, F., & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. Journal of Educational Psychology, 86(1), 122-133.
32. Palinko, O., Kun, A. L., Shyrokov, A., and Heeman, P. (March, 2005). Estimating Cognitive Load Using Remote Eye Tracking in a Driving Simulator, Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, 141-144.
33. Peter, J. P., & Churchill Jr, G. A. (1986). Relationships among research design choices and psychometric properties of rating scales: A meta-analysis. Journal of Marketing Research, 23(1), 1-10.
34. Schnotz, W., & Kurschner, C. (2007). A reconsideration of cognitive load theory. Educational Psychology Review, 19(4), 469-508.
35. Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. Educational Psychology Review, 22(2), 123-138.
36. Thatcher, J. B., & Perrewe, P. L. (2002). An Empirical Examination of Individual Traits as Antecedents to Computer Anxiety and Computer Self-Efficacy. MIS Quarterly, 26(4), 381-396.
37. Wickens, C. D., Hellenberg, J., & Xu, X. (2002). Pilot maneuver choice and workload in free flight. Human Factors: The Journal of the Human Factors and Ergonomics Society, 44(2), 171-188.