Metadata of the chapter that will be visualized in SpringerLink

Book Title	Augmented Cognition. Human Cognition and Behavior	
Series Title		
Chapter Title	Probing for Psycho-Physiological Correlates of Cognitive Interaction with Cybersecurity Events	
Copyright Year	2020	
Copyright HolderName	Springer Nature Switzerland AG	
Corresponding Author	Family Name	**Mogire**
	Particle	
	Given Name	**Nancy**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Information and Computer Sciences, University of Hawaii at Manoa
	Address	Post 317 1680, East-West Road, Honolulu, HI, 96822, USA
	Email	nmogire@hawaii.edu
Author	Family Name	**Minas**
	Particle	
	Given Name	**Randall**
	Prefix	
	Suffix	
	Role	
	Division	Shidler College of Business
	Organization	University of Hawaii at Manoa
	Address	2404 Maile Way Suite E601f, Honolulu, HI, 96822, USA
	Email	rminas@hawaii.edu
Author	Family Name	**Crosby**
	Particle	
	Given Name	**Martha**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Information and Computer Sciences, University of Hawaii at Manoa
	Address	Post 317 1680, East-West Road, Honolulu, HI, 96822, USA
	Email	crosby@hawaii.edu
Abstract	Changes in psychophysiological signals of the human body are highly revealing of cognitive and emotional responses to stimuli, capturing even subtle and transient events. Based on these properties, changes in recorded psychophysiological signal could reflect changes occurring in the computing network while the human factor is a part of it. This paper outlines a methodology for exploring the psychophysiological correlates of cognitive interaction with cybersecurity events. This continues the discussion on a dissertation research project exploring what neurological and physiological signal changes might reveal in the context of digital interactions from a cybersecurity standpoint.	
Keywords	Human factor - Human-Computer interaction artifacts - Tokens of interaction - Digital evidence - Digital incidents - Digital events - Psycho-physiological signals - Signal changes - Cognition - Cognitive responsiveness - Cybersecurity applications	

Probing for Psycho-Physiological Correlates of Cognitive Interaction with Cybersecurity Events

Nancy Mogire[1](✉), Randall Minas[2], and Martha Crosby[1]

[1] Information and Computer Sciences, University of Hawaii at Manoa,
Post 317 1680, East-West Road, Honolulu, HI 96822, USA
{nmogire,crosby}@hawaii.edu
[2] Shidler College of Business, University of Hawaii at Manoa,
2404 Maile Way Suite E601f, Honolulu, HI 96822, USA
rminas@hawaii.edu

Abstract. Changes in psychophysiological signals of the human body are highly revealing of cognitive and emotional responses to stimuli, capturing even subtle and transient events. Based on these properties, changes in recorded psychophysiological signal could reflect changes occurring in the computing network while the human factor is a part of it. This paper outlines a methodology for exploring the psychophysiological correlates of cognitive interaction with cybersecurity events. This continues the discussion on a dissertation research project exploring what neurological and physiological signal changes might reveal in the context of digital interactions from a cybersecurity standpoint.

Keywords: Human factor · Human-Computer interaction artifacts · Tokens of interaction · Digital evidence · Digital incidents · Digital events · Psycho-physiological signals · Signal changes · Cognition · Cognitive responsiveness · Cybersecurity applications

1 Introduction

1.1 Research Overview

The proliferation of devices that measure and record psycho-physiological signal devices in user space provides an opportunity to harness human cognitive functioning for potential cybersecurity applications.

This research investigates how the electrical signals generated from the functioning of the body, respond to human interaction with digital incidents [1]. If we can find that response-related signal changes are consistently notable, and we can retrieve these changes from recorded signal with an accuracy that is greater than chance, then we can claim that psycho-physiological signals contain markers of digital incidents.

Potential applications of these markers include: in digital investigations for triangulation of other evidence, in cybersafety management as input to tools for regulating immersive digital experiences of locked-in individuals.

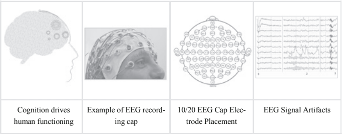
D. D. Schmorrow and C. M. Fidopiastis (Eds.): HCII 2020, LNAI 12197, pp. 1–11, 2020.
https://doi.org/10.1007/978-3-030-50439-7_28

1.2 Cognition and Psychophysiological Artifacts

Cognition processes fall under one of two systems: conscious cognition—also known as system 1 cognition—and automatic (system 2) cognition [2].

During a digital task, conscious cognition is mainly dedicated to the task while automatic cognition attends to a broader scope of elements of human functioning [2, 3] including the task processing, evaluation of its presentation features and assessment of other components of the general environment. Cognition may include appraisal activities such as fetching or forming various heuristics and making of non-conscious judgements.

Cognition is a necessary part of the human functioning that is involved in completing digital tasks [4] (Fig. 1).



Cognition drives human functioning

Example of EEG recording cap

10/20 EEG Cap Electrode Placement

EEG Signal Artifacts

Fig. 1. Cognition, EEG correlates and measurement

Cognitive functioning activates various body systems such as the brain, facial brow muscles, heart and electrodermal systems. These systems generate electrical signals during their functioning [5].

As such, cognition can be said to create psycho-physiological signals artifacts. Examples of such signals include: Electroencephalograms (EEG), Electromyograms (EMG), Electrocardiograms(ECG) and Electrodermograms (EDR). These signals have known structural forms and follow predictable change patterns [6–9].

Various relationships between cognition and psycho-physiological signal change have been studied and documented. For example, some signals have been found to reflect such cognitive experiences as variation in mental workload [10], shift in attentive focus, and experiences of emotional affect such as disgust [11].

1.3 Cybersecurity Context and Thresholds

We assume that cybersecurity events inherently feature, some response-evoking characteristics e.g. salience, aversiveness. What counts as an event is dependent on context, and as such security-relatedness is a wide and varying spectrum.

We differentiate events from incidents by taking events to be those that have potential security implications, while incidents are the events that have links to actual breaches.

Events: can potentially cause security breaches. Incidents: events where occurrence can be directly linked to a breach that occured (Fig. 2).

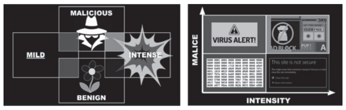


Fig. 2. Security-relatedness is a wide spectrum of variability [1]

1.4 Research Approach Overview

This study involves presentation of cybersecurity related events as stimuli, to participants within a lab setting while they undertake generalized computing tasks that mimic regular device usage.

Throughout each lab session, as the participant works on the tasks and experiences the stimuli, we record psychophysiological signals (EEG, EMG, EDA, ECG) using various body sensors together with relevant body signal acquisition software. We assume that response-evoking properties e.g. salience and aversiveness, are inherent in cybersecurity events.

The study takes a non-blind followed by blind study methodology, and utilizes relevant signal analysis methods including event related potential analysis and wavelet transformation.

In the non-blind phase we match signal changes to event timings using a separately tracked record of event stimuli display timings. In the blind phase, event timings are initially unknown to the primary researcher but held by another researcher. We attempt to identify event timings via signal analysis and then verify against the actual timings. If we can retrieve event timings from signal with accuracy greater than chance, then we can make the claim that psycho-physiological signals contain markers of digital events and incidents [1].

2 Research Design and Process

2.1 Stimuli Design

To be able to probe psychophysiological signals for markers of interaction with cybersecurity events, it was necessary to create conditions mimicking real scenarios that occur in relation to cybersecurity breaches. Towards this goal, we designed a study so that we could present security related material such as pop-ups, page crash reports and error warnings to the study participants while recording their psychophysiological signal data was recorded.

The key assumption is that these security related events would inherently have response evoking properties. Such properties may include dubiousness, aversiveness, salience or other characteristics that elicit cognitive responses [3, 12–14]. These response-provoking elements would be the basis for selection of stimuli events to be displayed in the experimental study setting.

The presentation of these events as stimuli needs be indirect rather than as though such events are the primary computing tasks in order to mirror how such events might occur within a regular computing environment. In this case the study is designed so that they can be presented within the workflow of a computer-based activity serving as a distractor. The activity is packaged in a web-based application that ties together the study introductory and closing material, the assignments a given participants needs to complete for their lab session, and the events and triggers serving as the cybersecurity stimuli.

In summary, the stimuli setup components are as follows:

Distractor Exercise: In this case the exercise involves a set of research tasks that lead to creation of computer workstation specifications for a specified group of professionals. The activity workflow takes participants through a path allowing for the presentation of intended stimuli.

Security Events: These are web elements and content intended as the stimuli to be interacted with in the study. They are selected because they feature various properties associated with security breach scenarios. The properties are initially assumed to be response-evoking and then tested later via a survey during the stimuli validation stage. There are two key selection criteria used to include stimuli events in the setup:

i. **Security-relatedness** i.e. elements should have properties that cause the element to be judged as having security implications e.g. pop-up dialog that prompts a decision on an unsolicited file downloading. Non-security-related events will be utilized as controls.
ii. **Intensity of material** (e.g. visual noise, decision or no decision). Both the security event set and the control event set of elements are further considered on their visual and cognitive demand intensity. Hence the final selections are intended to fit in either into either the intense or the mild category. These selections are later tested via a survey to ensure they elicit the meanings we intended. The purpose of this criteria is to help in the evaluation of the role played by an event's intensity in determining responses evoked in the psychophysiological signal.

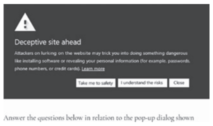
Control Events: These are elements selected because they appear to be decidedly non-security related. The goal is to utilize them as control stimuli for comparative study of responses evoked during interactions with them against those evoked during interactions with security-related stimuli. It is supposed that the responses elicited in the two stimuli groups would be distinct from each other, with the former being less aversive in nature than the latter.

2.2 Stimuli Validation Surveys

To ensure that stimuli is perceived to have the implications we intend, we conducted evaluation surveys to ask a subset of people from the sample population – from which the participant pool for the lab study will be drawn – to rate the elements on various metrics to let us know how these elements are perceived with respect to potential effect on security of a device.

The surveys focused on two goals as discussed next,

Survey 1: Evaluation of User Perceived Significance and Risk Level of Material: The perception of risk may influence the level of responsiveness of a person interacting with digital material. In this survey the goal was to ascertain that the material selected for presentation had a high chance of being interpreted as security-related and potential to elicit responses mimicking those typical in security events (Figs. 3 and 4).



Deceptive site ahead

Attackers on lurking on the website may trick you into doing something dangerous like installing software or revealing your personal information (for example, passwords, phone numbers, or credit cards). Learn more

Answer the questions below in relation to the pop-up dialog shown

Fig. 3. Example of material presented for evaluation: a warning web page

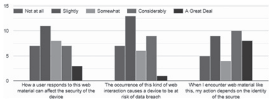


Fig. 4. Sample of summarized results for a subset of the evaluation criteria

Based on the outcomes of these surveys we narrowed down on material that was sound to use as stimuli in as far as perception of security-relatedness was intended. Similarly we identified sound control stimuli against the security-related property.

The selection was limited to materials with evaluation that was not more than 2 SD from the mean, hence abstracting the security relatedness property down to a narrow field where assumptions on general perception could be reasonably made.

Survey 2: Categorization of Material by Security-Relatedness and Intensity: The second survey asked participants to rate material on a likert scale from low to high perceived intensity based on various criteria including visual noisiness and cognitive demand presented by a decision. Similarly, the survey asked participants to rate the material on perceived relevance to security.

Based on the survey, we grouped the elements into four categories based on security relatedness and intensity for use in the analysis phase. for comparative analysis of response signals. The categories used were: Malicious-Intense, Malicious-mild, Benign-Intense, Benign-Mild.

2.3 Other Data Validation Manipulations

- **Single clock setup for timing synchronization:** The same cpu is used to run the stimuli display application and the signal acquisition process. This is one way to achieve millisecond accuracy in the matching of stimuli presentation timing with the timing of response-related changes in recorded signal.
- **Repeated measurement:** Stimuli is presented multiple times with minor variations, to allow for verification of observations by comparing responses to similar stimuli.
- **Handedness questionnaire:** An evaluation of hand dominance versus preference is conducted to account for hand-preference-related variance in speed of response to stimuli.

2.4 Data Collection

Data collection will be conducted in a specialized lab setting with body sensors and data acquisition software that will make the measurement of psychophysiological signals possible.

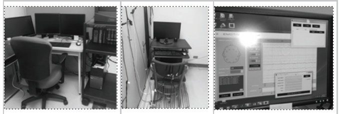
The setup will include a participant area walled into a semi-private space within the lab, where the stimuli display computer will be available for the participant to undertake the computer-based tasks. The researcher will have a workspace away from the participant's work area from where they will manage the data acquisition and storage during the study sessions (Fig. 5).

In the lab, participants will undertake the computer-based exercise hence allowing us to present the security event stimuli. Throughout the session, we will record the participant's body signals using relevant sensors, and corresponding signal acquisition software.

2.5 Data Cleaning and Processing

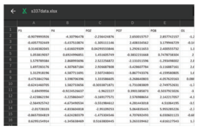
During data collection, the signal files will be initially output in the default formats e.g. EDF for the EEG. To process the data, the files will first be converted into spreadsheets with numerical data.

Once the data format is changed, data is manually inspected for integrity. For example in EEG data collection, the baseline is a fixed default gyroscope setting e.g. 4000 mv and therefore none of the recorded values should not be below this (Fig. 6).



Researchers area | Participant's workspace | Researcher's view

Fig. 5. Data collection setting



s337data.xlsx

	A	B	C	D	E	
1	P3	P4	PO2	PO7	PO8	O1
2	-0.9079993926	-4.30796478	-2.236424876	2.650015757	2.857742157	-1.2
3	-0.4057702449	-3.437510874	-1.305111166	2.408334562	3.179946729	-0.14
4	0.3146382045	-1.616029309	0.04293533846	1.292611653	2.400555732	1.1
5	1.053819037	0.6924996851	1.451605749	-0.3832231668	0.7478738304	2
6	1.579789384	2.868993696	2.522256872	-2.131011596	-1.293498002	2.8
7	1.697283176	4.307687184	2.926687808	-3.428607784	-3.116887161	2.6
8	1.312918196	4.587711691	2.507248041	-3.867743376	-4.159580805	1.6
9	0.4753842766	3.598706396	1.333586605	-3.268640803	-4.057620163	0.088
10	-0.6241460705	1.582733656	-0.3033871871	-1.731083809	-2.749752631	-1.7
11	-1.69499956	-0.9224526637	-1.9622227	0.3901385873	-0.5037923026	-3
12	-2.423862196	-3.225860447	-3.169177571	2.576988654	2.142217057	-4.1
13	-2.564925742	-4.673490534	-3.551984612	4.281443018	4.51064195	-3.5
14	-2.01728103	-4.833604818	-2.95109253	5.064835445	5.955285226	-2.7
15	-0.8657004819	-3.624280179	-1.475304546	4.707692493	6.030601123	-0.66
16	0.6295154914	-1.345638469	0.5163800445	3.263339462	4.616127543	1.7

Fig. 6. Section of spreadsheet depicting data collected with emotiv EEG headset

After inspection, the next steps include: trimming out extraneous signal recordings i.e. before task start and after task end, labelling sensor fields in the data as needed, preparing data for analysis by formatting into matrices and storing as variables. Further processing includes removing baselines, extracting epochs and rejecting extraneous epochs, leaving the cleaned, formatted and epoched data ready for analysis.

2.6 Data Analysis

Analysis will be done using two techniques as follows,

Event Related Potential (ERP) Analysis: This will entail running an Independent Component Analysis algorithm on the EEG signal to isolate the ERP event-related potential(ERP) components contained within it. These components reveal the neurological activity underlying various signal changes.

For this analysis, we will be searching for specific ERP components as per the study hypotheses drawn from theory. If anticipated ERP components are present and corresponding to stimuli presentations, they will be utilized in evaluating the study hypotheses (Fig. 7).

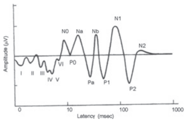


Fig. 7. Graph depicting some of the major ERP Components [16]

Wavelet Transforms: These will be used to analyze and classify the signal change patterns. This process will involve removal of noise from the signal i.e. signal denoising [15], wavelet transformation, feature extraction and classification, determining thresholds and finally locating any structural break points for matching with stimuli display timings to see if there is any correspondence (Fig. 8).

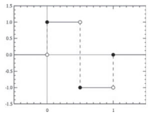


Fig. 8. Example wavelet transform plot: Haar wavelet [17]

2.7 Evaluating Outcomes

The outcomes will be evaluated in two phases:

Non-blind Phase: In this phase, changes in signal will be matched to corresponding stimuli displays if any, by referencing the record of stimuli event display timings. This will facilitate recognition and evaluation of changes in the signal, that relate to security events in particular. Findings from this stage will be used for guidance in the blind phase of the analysis (Fig. 9).

Blind Phase: The blind evaluation will run opposite of the one in the non-blind phase. Starting from psychophysiological signals, we will attempt to locate changes that suggest responses to security events. Based on any changes located, we will hypothesize a timeline of when security-event stimuli was presented (Fig. 10).

This process will not include any attempt to specify the event that was displayed, but rather that a security event was interacted with.

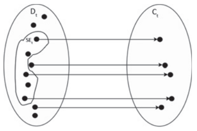


Fig. 9. Finding event markers in signal by mapping security event display timings to changes noted in recorded signal

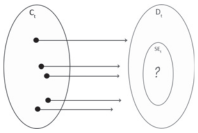


Fig. 10. Changes noted in signal mapped to hypotheses of timings that the corresponding event stimuli was displayed

Afterwards, the hypothesized timings will be verified against the actual event stimuli display timings on record. If we find that we can retrieve the event stimuli timings via the signal, with an accuracy that is greater than chance, then we can claim that the signals contain markers of interaction with those events.

Finally we will work on determining other relevant measures such as error rates and thresholds, which would mark the end of the study.

Fig. 11. Human Computer Interaction (HCI) Outputs = Inputs to Cybersecurity Applications

3 Conclusion

The potential impact of this project is that it highlights and models outputs of human-computer interaction as inputs to cybersecurity applications. If we can harness human cognition in the computing environment for any markers of appraisal, such markers can be useful for various applications (Fig. 11).

For example in digital investigations, they can be used to triangulate other evidence and hence support the justice function. Another envisioned application would be in immersive digital experiences of locked-in individuals, where such markers can be used as input to cybersafety management tools to regulate the experiences of those users. This would contribute towards making cybersafety accessible.

References

1. Mogire, N.: Tokens of interaction: psycho-physiological signals, a potential source of evidence of digital incidents. In: HotSoS 2020, 7–8 April 2020, Lawrence, KS, USA. ACM (2020). ISBN 978-1-4503-7561-0/20/04, https://doi.org/10.1145/3384217.3384226
2. Kahneman, D.: Thinking. Fast and Slow. Macmillan, Basingstoke (2011)
3. Posner, M.I., Snyder, C.R.R.: Attention and cognitive control. In: Solso, R.L. (ed.) Information Processing and Cognition: The Loyola Symposium, pp. 55–85. NJ: Lawrence Erlbaum Associates, Hillsdale (1975). [Ref list]
4. Karray, F., Alemzadeh, M., Saleh, J.A., Arab, M.N.: Human-computer interaction: overview on state of the art (2008)
5. Waller, A.D.: Eight Lectures on the Signs of Life from their Electrical Aspect. EP Dutton, New York City (1903)
6. Loeb, G.E., Loeb, G., Gans, C.: Electromyography for Experimentalists. University of Chicago Press, Chicago (1986)
7. Camm, A.J., et al.: Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task force of the European society of cardiology and the North American society of pacing and electrophysiology. Circulation **93**(5), 1043–1065 (1996)
8. Boucseinm, W.: Electrodermal Activity. Springer, Heidelberg (2012). https://doi.org/10.1007/978-1-4614-1126-0
9. Luck, S.J., Kappenman, E.S.: Electroencephalography and event-related brain potentials. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (eds.) Handbook of Psychophysiology, pp. 74–100, 4th edn. Cambridge University Press, Cambridge (2016). (Cambridge Handbooks in Psychology)
10. Porges, S.W., Campbell, B.A., Hayne, H., Richardson, R.: Attention and information processing in infants and adults. In: Autonomic Regulation and Attention. Lawrence Erlbaum Associates, Hillsdale, pp. 201–223 (1992)
11. Kreibig, S.D., Samson, A.C., Gross, J.J.: The psychophysiology of mixed emotional states. Psychophysiology **50**(8), 799–811 (2013)
12. Lang, P., Bradley, M.M.: The international affective picture system (IAPS) in the study of emotion and attention. In: Handbook of Emotion Elicitation and Assessment, 19–29 April 2007
13. Porges, S.W.: Autonomic regulation and attention. Attention and information processing in infants and adults, pp. 201–223 (1992)
14. Sawaki, R., Luck, S.J.: Capture versus suppression of attention by salient singletons: electrophysiological evidence for an automatic attend-to-me signal. Attention Percept. Psychophys. **72**(6), 1455–1470 (2010)

15. Denoising and Compression (n.d.). https://www.mathworks.com/help/wavelet/denoising-and-compression.html
16. Kuok, A.: Brainstem potentials, 7 June 2007. https://zh.wikipedia.org/wiki/File:Brainstem.PNG. Accessed 2 March 2020
17. Haar wavelet 3 June 2006. https://commons.wikimedia.org/wiki/File:Haar_wavelet.svg. Accessed 2 March 2020

Author Queries

Chapter 28

Query Refs.	Details Required	Author's response
AQ1	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	
AQ2	Please check and confirm if the inserted citations of Figs. 1–11 are correct. If not, please suggest an alternate citations.	