Using Galvanic Skin Response (GSR) to Measure Trust and Cognitive Load in the Text-Chat Environment

Ahmad Khawaji
National ICT Australia (NICTA)
The University of NSW
Sydney, Australia
ahmad.khawaji@nicta.com.au

Fang Chen
National ICT Australia (NICTA)
The University of NSW
Sydney, Australia
fang.chen@nicta.com.au

Jianlong Zhou
National ICT Australia (NICTA)
Sydney, Australia
jianlong.zhou@nicta.com.au

Nadine Marcus
The University of NSW
Sydney, Australia
nadinem@cse.unsw.edu.au

CHI'15 Extended Abstracts, Apr 18-23, 2015, Seoul, Republic of Korea
ACM 978-1-4503-3146-3/15/04.
http://dx.doi.org/10.1145/2702613.2732766

Abstract

Exchanging text messages via software on smart phones and computers has recently become one of the most popular ways for people to communicate and accomplish their tasks. However, there are negative aspects to using this kind of software, for example, it has been found that people communicating in the text-chat environment may experience a lack of trust and may face different levels of cognitive load [1, 11]. This study examines a novel way to measure interpersonal trust and cognitive load when they overlap with each other in the text-chat environment. We used Galvanic Skin Response (GSR), a physiological measurement, to collect data from twenty-eight subjects at four gradients and overlapping conditions between trust and cognitive load. The findings show that the GSR signals were significantly affected by both trust and cognitive load and provide promising evidence that GSR can be used as a tool for measuring interpersonal trust when cognitive load is low and also for measuring cognitive load when trust is high.

Author Keywords

Trust; Cognitive Load; GSR; Text-Based CMC

ACM Classification Keywords

H5.3. Information interfaces and presentation: Group and Organization Interfaces - collaborative computing, computer-supported cooperative work.

Introduction

Galvanic Skin Response (GSR) is a physiological signal captured easily and cost effectively via the skin [15]. These signals reflect changes in the skin's ability to conduct electricity and are used to indicate the extent of nerve responses [7]. People cannot control signals generated from the GSR device because they are autonomic signals that are extracted from the level of sweat in the skin [12], thus, GSR is considered as a credible physiological measure for the level of sweat in the skin.

The text-chat environment is a form of computer-mediated communication which is very low cost compared with face-to-face communication. A study showed that in an organization, employees commonly used the text-chat medium to communicate more frequently than using face-to-face communication [8]. Although the text-chat environment is commonly used, it has been found that there is a lack of interpersonal trust between the communicators compared to other computer-mediated communication forms such as video [1] and also communicators may face different levels of cognitive load [11]. These findings raise questions about identifying ways to measure the factors which negatively affect communicators for the purpose of providing support to them (such as using linguistic features to measure the level of trust between communicators[9]). However, our research focuses on analyzing the physiological data of communicators in the text-chat environment to see how these data can be affected by certain levels of trust and cognitive load.

Specifically, in this study, we analyze GSR signals (physiological signals) in different and overlapping levels of trust and cognitive load to examine how the signals emitted from the skin of the subjects' fingers can be used as an indicator for each situation to which the subjects were exposed.

Background Literature

Previous research has demonstrated that the extent of trust between communicators in the text-chat environment can be measured. For instance, communicators who trusted each other more used significantly more assent and positive emotion words [4]. Also, cognitive load can be measured via language. People used more varieties of words when they were exposed to a low cognitive load [3].

In relation to using physiological signals to measure trust, researchers found that when they examined eye gaze during web page browsing, people maintained more continuous focus on the pages that they trusted [5]. Also, the ability to measure cognitive load via GSR values is explored. Previous research demonstrated that people's GSR values increased when they were exposed to a high cognitive load which is related to stress [10]. However, stress isn't only associated with cognitive load but also with trust, as demonstrated by a study which showed that people whose trust level was high had low stress, while on the contrary, people whose trust level was low had high stress [2]. Therefore, we also expect in this study to find that stress may result from a lack of trust and this may increase GSR values. However, to the best of our knowledge, no existing study examines GSR signals with the overlapping conditions of trust and cognitive load nor is there a study which examines GSR with interpersonal trust alone. This paper examines GSR signals with four gradients and overlapping conditions of interpersonal trust and cognitive load to verify how two factors, trust and cognitive load, influence the GSR signals and find a way to measure these overlapping factors. Drawing from the literature, we hypothesize that when the participants' trust is low and they are exposed to a high cognitive load, the GSR values will be at their highest level (H1). The GSR values also will be at their lowest level when the participants' trust is high and they are exposed to a low cognitive load (H2).

Method

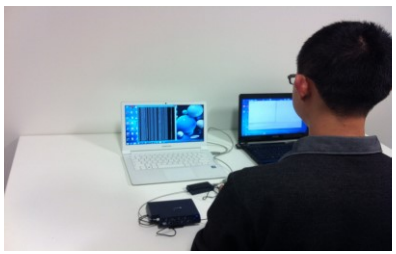

Figure 1: Setup of the experiment (GSR device: ProComp Infiniti System from Thought Technology Ltd).

Twenty-eight students from NSW University and NICTA organization were recruited for this study, their ages ranging from 18 to 40 years (18 males and 10 females). Each participant was assigned randomly to one partner during the experiment, and in total, there were fourteen pairs. The pairs were divided into two groups to manipulate trust: for one group, we enhanced the level of trust between the two participants and for the other group, we decreased the level of trust between the two participants. In each group, the participants were exposed to two levels of cognitive load, low and high. The experiment was designed as follows: 2 trust levels (low/high) x 2 cognitive load levels (low/high) in a mixed design (four conditions): Low Trust-High Cognitive Load (*LTHCL*), Low Trust-Low Cognitive Load (*LTLCL*), High Trust-High Cognitive Load (*HTHCL*) and High Trust-Low Cognitive Load (*HTLCL*).

Interpersonal trust was manipulated between the participants before starting the task. To increase the interpersonal trust between the two participants, we followed two procedures. Firstly, we asked the participants to meet their partner face-to-face for ten minutes and talk, which has been shown to increase trust in the text-chat environment [16]. These participants were given three brainstorming tasks (e.g., if people have an extra thumb on each hand) [13] and were required to discuss this with each other and write three advantages and three disadvantages for each task. Secondly, as the participants will play an investment game, described below, the following paragraph was included as a preface to the instructions of the game to influence increased trust [14]: *"Trust is an essential relationship between people and on this basis, the tasks entrusted to them can be completed successfully. Trust usually leads to the sharing of thoughts and open and honest discussion. However, it is well known that to secure the trust of others, a serious attempt must be made and trust must be exchanged as a starting point. As you can see in the procedure of this game, trust in others is important to obtain high and satisfying profits".* However, to decrease trust, we didn't allow the participants to meet or see each other [16] and also a different paragraph to encourage distrust between the partners was included as a preface to the instructions to the game [14]. After assigning each participant to their partner and manipulating the level of trust, the partners were separated from each other by a partition during the tasks for both the high and low trust conditions.

The GSR data (Figure 1) was collected during the DayTrader task [1, 9], where the partners were allowed to chat with each other using instant messaging and play the investment game in the same window. The rules of this investment game were taken from the Prisoner's Dilemma task where the payoff resulting from the investment game was used to measure trust (when an increase in the group payoff indicates an increase in trust and vice versa). Therefore, this task was chosen to check the level of interpersonal trust between the participants. The participants chatted for four five-minute sessions about their investment in the market. After each chat session, each participant invested with their partner in five separate rounds. Both partners were given $60 each in each investment round and they could invest any amount of money. Each participant was given a payoff after each round. The payoff was as follows: the amount of money which was invested by each participant was summed and tripled and divided between them equally while the money which was not invested was doubled separately and given to each participant. Therefore, in each investment round, the participants who invested less received more money.

During the four chat sessions, all participants were exposed to two levels of cognitive load (high/low). At the same time, each pair of participants was exposed to a high cognitive load task for two chat sessions and to a low cognitive load task for the other two chat sessions. In both cognitive load conditions, during each five-minute chat session, the participants summed twelve random numbers in their heads without using a calculator or pen. These random numbers were shown in pop-up boxes in the game window when the participants chatted. The participants could close these boxes themselves or they would be closed automatically after fifteen seconds. We asked the participants to sum large numbers (between 100 and 300) for the high cognitive load task and smaller numbers (one or two) for the low cognitive load task. The participants completed a cognitive load questionnaire (e.g., How accurately do you think you summed the numbers?) [6] to determine whether the participants were exposed to different cognitive load levels. The participants were paid between $10 and $18 based on their performance of earning money from the investment game and summing the random numbers.

Preliminary Analysis and Results

The results of our approach in relation to manipulating trust and cognitive load levels were as expected. The subjects' trust, as measured by the group payoff, was highest when they met each other face-to-face before they started chatting and after having read a paragraph encouraging trust in the instructions of the task and also the subjects' cognitive load was highest when they summed large random numbers. In relation to the trust rate, the two-tailed t-test showed there were significant differences where the trust rates in both the high trust conditions, *HTLCL* ($M=\$1617.9$, $SD=\$448.5$) and *HTHCL* ($M=\$1816.3$, $SD=\$765.1$), were significantly higher than the trust rate in both the low trust conditions, *LTHCL* ($M=\$1059.7$, $SD=\$611.1$, ($t=3.74$, $p=0.004$ (.vs *HTLCL*)), ($t=4.15$, $p=0.002$ (.vs *HTHCL*))) and *LTLCL* ($M=\$1208.1$, $SD=\$1023.7$, ($t=5.73$, $p<0.00$ (.vs *HTLCL*)), ($t=7.79$, $p<0.00$ (.vs *HTHCL*))). For cognitive load level, the two-tailed t-test showed there were also significant differences where the cognitive load levels in both the high cognitive load conditions, *LTHCL* ($M=29$, $SD=8.3$) and *HTHCL* ($M=31.8$, $SD=5.9$), were significantly higher than the cognitive load level in both the low cognitive conditions, *HTLCL* ($M=23.1$, $SD=6.6$, ($t=2.11$, $p=0.04$ (.vs *LTHCL*)), ($t=4.19$, $p<0.00$ (.vs *HTHCL*))) and *LTLCL* ($M=19.2$, $SD=9.1$, ($t=3.66$, $p=0.003$ (.vs *LTHCL*)), ($t=4.33$, $p<0.00$ (.vs *HTHCL*))).

In this study, we extracted the GSR data of participants only when they were chatting with each other and we examined two GSR features: the average of GSR values and the average of peaks of GSR values.

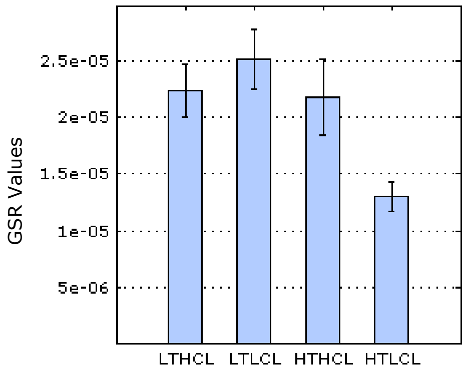

Figure 2: The average of the GSR values.

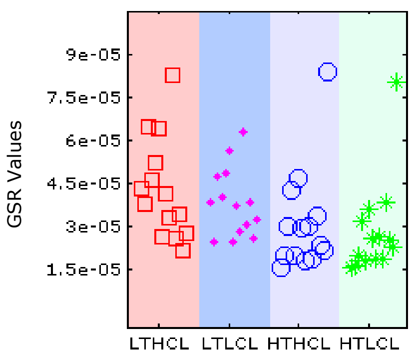

Figure 3: The average of the peaks for each participant (each symbol represents one participant).

Figure 2 shows the average of the GSR values is decreased in the *HTLCL* condition more than the other conditions, and the averages of the four conditions are as follows, from largest to smallest: *LTLCL (M*=2.5E-5, *SD*=9.8E-6*), LTHCL (M*=2.24E-5, *SD*=8.6E-6*), HTHCL (M*=2.2E-5, *SD*=1.3E-5*)* and *HTLCL (M*=1.3E-5, *SD*=5.2E-6*)*. A two-tailed two-way ANOVA was conducted to examine the interaction effect between trust and cognitive load on the averages of GSR and the results showed significant interaction between them (F=5.3, p=0.02). A post-hoc two-tailed t-test was also used to compare the averages of GSR. Six comparisons of these averages were made and three of the six comparisons showed significance differences in the averages of GSR. The results showed that the *HTLCL* averages were significantly lower than the averages of *HTHCL* (t=2.4, p=0.02), *LTHCL* (t=3.5, p=0.001) and *LTLCL* (t=4.1, p<0.00). The other three comparisons between *HTHCL* and *LTHCL*, *HTHCL* and *LTLCL* and *LTHCL* and *LTLCL* didn't show significant differences, the results being as follows (t=0.1, p=0.9), (t=0.8, p=0.4) and (t=0.8, p=0.4), respectively.

Figure 3 illustrates the average of the peaks for each participant. The results of the averages of peaks are as follows, from largest to smallest: *LTLCL (M*=2.52E-5, *SD*=9.83E-6*), LTHCL (M*=2.3E-5, *SD*=8.5E-6*), HTHCL (M*=2.2E-5, *SD*=1.3E-5*)* and *HTLCL (M*=1.4E-5, *SD*=5.2E-6*).* A two-tailed two-way ANOVA was conducted to examine the interaction effect between trust and cognitive load on the averages of peaks and the results showed significant interaction between them (F=5.4, p=0.02). A post-hoc two-tailed t-test was also used to compare the averages of peaks. Six comparisons were made and, similar to the results of averages, three of the six comparisons between the averages of peaks showed significant differences. The results showed that the averages of the *HTLCL* peaks were significantly lower than the averages of the peaks of *HTHCL* (t=2.4, p=0.02), *LTHCL* (t=3.5, p=0.002) and *LTLCL* (t=4.12, p<0.00). The other three comparisons between *HTHCL* and *LTHCL*, *HTHCL* and *LTLCL* and *LTHCL* and *LTLCL* didn't show significant differences, the results being as follows (t=0.1, p=0.8), (t=0.7, p=0.4) and (t=0.7, p=0.4), respectively.

Summary and Ongoing Work

This study examined two GSR features, the average of GSR values and the average of peaks of GSR values and showed that the GSR values were significantly lower when both trust is high and cognitive load is low. Moreover, this study has provided promising findings in relation to indicators for determining the level of trust and cognitive load. Specifically, the results of GSR values show that in a low cognitive load situation, that is in *LTLCL and HTLCL* conditions, GSR can be used to measure the level of interpersonal trust, while in a high trust situation, that is in *HTLCL and HTHCL* conditions, GSR can be used to measure the level of cognitive load.

The results of this study show that hypothesis H1, that is, when the participants' trust is low and they are exposed to a high cognitive load, the GSR values will be at their highest level, isn't supported and hypothesis H2, that is, the GSR values will be at their lowest level when the participants' trust is high and they are exposed to a low cognitive load, is supported. A possible reason for this is that only one negative factor, either high cognitive load or low interpersonal trust, may be enough to increase stress, and consequently, results in increased GSR values as shown in three conditions *LTHCL*, *LTLCL* and *HTHCL*. In the *HTLCL* condition, when participants were exposed to a low cognitive load and their interpersonal trust in their partners was high when they were chatting, we believe the participants were more comfortable and weren't subjected to pressure, which may be the reason why the GSR values were at their lowest level.

The results of this study have implications which can be used. For instance, through the development of a keyboard and mouse which is capable of capturing physiological signals from the fingers of the communicators when they chat, an intelligent system can be built to measure the level of trust and cognitive load to which they were exposed in real time and provide them with suitable assistance.

Our ongoing work will focus on distinguishing between the overlapping levels of trust and cognitive load but via using linguistic and grammatical features to examine the chat contents which were collected during this experiment. Also, we intend to combine the linguistic and grammatical features with GSR features and examine them with several machine learning algorithms, in an attempt to obtain high accuracy in the classification of these overlapping conditions.

Acknowledgements
This work is supported by AOARD under grant No. FA2386-14-1-0007 AOARD 134144 and FA2386-14-1-0022 AOARD 134131.

References
[1] Bos, N., Olson, J., Gergle, D., Olson, G. and Wright, Z. Effects of four computer-mediated communications channels on trust development. In Proc. CHI 2002, ACM Press (2002), 135-140.
[2] Costa, A., Roe, R., and Taillieu, T. Trust within teams: The relation with performance effectiveness. European Journal of Work and Organizational Psychology, 10, 3 (2001), 225-244.
[3] Khawaja, M. A., Chen, F., Marcus, N. Using Language Complexity to Measure Cognitive Load for Adaptive Interaction Design. In Proc. IUI 2010, ACM Press (2010), 333-336.
[4] Khawaji, A., Chen, F., Marcus, N., and Zhou, J. Trust and Cooperation in Text-Based Computer-Mediated Communication. In Proc. OzCHI 2013, 37-40.
[5] Leichtenstern, K., Bee, N., André, E., Berkmüller, U., and Wagner, J. Physiological Measurement of Trust-Related Behavior in Trust-Neutral and Trust-Critical Situations. Trust Management, 358, (2011), 165-172.
[6] NASA. Nasa Task Load Index, Ver. 1.0, 1986.
[7] Peuscher, J. Galvanic Skin Response, Ver. 3, 2012.
[8] Quan-Haase, A., Cothrel, J., and Wellman, B. Instant messaging for collaboration: A case study of a high-tech firm. Computer-Mediated Communication, 10, 4 (2005).
[9] Scissors, L., Gill, A., Geraghty, K., and Gergle, D. In CMC we trust: the role of similarity. In Proc. CHI 2009, 527-536.
[10] Shi, Y., Ruiz, N., Taib, R., Choi, E., and Chen, F. Galvanic skin response (GSR) as an index of cognitive load. In Proc. CHI 2007, 2651-2656.
[11] Thirunarayanan, M, Ryan, C. and Perez-Prado, A. An Exploratory Study of Cognitive Load in Instructional Chat Rooms. In Proc. ICCE 2002.
[12] Westerink, J., Broek, E., Schut, M., Herk, J., and Tuinenbreijer, K. Computing Emotion Awareness Through Galvanic Skin Response and Facial Electromyography. Probing Experience, 8, (2008).
[13] Wang, H-C., Fussell, S. and Cosley, D. From Diversity to Creativity: Stimulating Group Brainstorming with Cultural Differences and Conversationally-Retrieved Pictures. In Proc. CSCW 2011, 265-274.
[14] Zand, D. Trust and Managerial Problem Solving. Administrative Science Quarterly 17, 2 (1972).
[15] Zhou, J., Sun, J., Chen, F., Wang, Y., Taib, R., Khawaji, A., and Li, Z. Measurable Decision Making with GSR and Pupillary Analysis for Intelligent User Interface. ACM Transactions on Computer-Human Interaction (ToCHI), 2014. In press.
[16] Zheng, J., Veinott, E., Olson, J. and Olson, G. Trust without touch jumpstarting long-distance trust with initial social activities. In Proc. CHI 2002, 141-146.