Using cardiovascular features to classify state changes during cooperation in a simulated bomb defusal task

Mark Dennison[1], Cathy Neubauer[2], Tony Passaro[2], Andre Harrison[1], Stefan Scherer[2], and Pete Khooshabeh[1]

1 U.S. Army Research Laboratory, Adelphi, USA
mdennison@ict.usc.edu, {antony.d.passaro.civ, andre.v.harrison2.ctr, peter.khooshabehadeh2.civ}@mail.mil
2 USC Institute for Creative Technologies, Playa Vista, USA
{neubauer, scherer}@ict.usc.edu

Abstract. Teams of two individuals worked together in a high-intensity simulated bomb diffusing task. Half the teams were given icebreaker social time to increase comfort and familiarity with each other and the remaining half of the teams served as controls and did not meet until the task began. Electrocardiography and impedance cardiography were recorded to examine cardiac changes during task cooperation. Changes in ventricular contractility showed that individuals who had taken part in the icebreaker showed increased task engagement over time whereas controls showed the opposite. Data also trended to show that icebreaker participants were in a challenge state and controls were in a threat state during the final thirty seconds of bomb defusal. Finally, we show that a set of cardiac features can be used to classify participant data as belonging to the icebreaker or control groups with an accuracy as high as 88%.

Keywords: psychophysiology, engagement, team cohesion, state classification, machine learning, cardiovascular analysis

1 Introduction

Performance in a high-intensity environment is often accompanied by changes in psychological and physiological state. These states, such as threat, anxiety, or fatigue, can influence human performance and resilience in the face of a difficult task. Access to information about changes in these states can be used to improve decisions in the real world and to adapt training using human-state aware systems and interfaces. Automatic behavior analysis, tracking, and machine learning offer new ways to assess the dynamic changes in state that are simply not possible using other traditional measures such as interviews or questionnaires.

In this paper we discuss how changes in cardiovascular features can be used to classify state changes during a high-intensity task that required cooperation with a teammate. The data analyzed are from an experiment previously conducted by Neubauer et al. [1]. The primary experimental manipulation was the familiarity of the teammates to examine how team intimacy affected resilience to stress and changes in state. In this paper, we contribute three novel analyses of the dataset from [1]. First, physiological measures, namely electrocardiogram and impedance cardiography, are used to assess changes in task engagement as a function of time [2]. Next, we look at how team intimacy relates to the development of individuals exhibiting a challenge or threat state during moments of high demand and stress [3,4]. Finally, we use machine learning to classify participant data as belonging to either of two experiment groups. The primary assumption behind this experiment is that teammates who become more familiar with one another before a high-stress task may be able to develop different methods of handling stress and may approach the task with a greater desire to succeed.

We propose the following ideas which will serve as our two research questions:

- RQ1: Do participants show cardiovascular changes that are characteristic of a challenge state or a threat state that depend on their prior meeting with their teammate?
- RQ2: Can we classify which experimental group a participant came from using only their physiological data?

2 Methods

2.1 Experimental Design

The study by Neubauer [1] used a 2x2 between-subjects design with two experimental conditions using 20 gender-matched pairs for each condition. The between-subjects factor was manipulation of team intimacy, which was the inclusion of an 'Icebreaker Conversation' (IB) between teammates prior to the experiment or the lack of an icebreaker conversation to serve as a Control (CT) condition. The dependent variables of interest were subjective ratings of engagement before and after the task and physiological response (discussed in further detail below). Figure 1 shows how the experiment was setup and the timeline of the actual task with a two-person team.

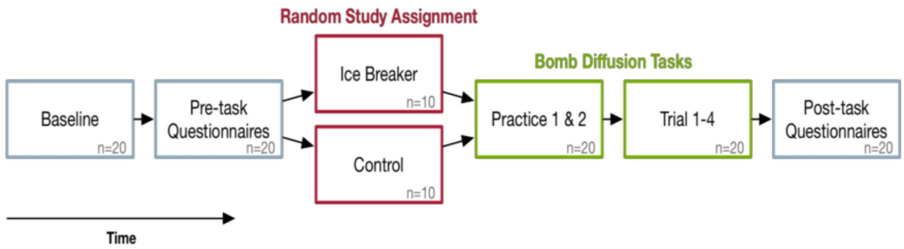

Fig. 1. Timeline of the experimental setup using a two-person team. Baseline physiological measures were taken prior to the pre-task subjective engagement assessment and experimental manipulation. The practice and main tasks occurred after the experimental manipulation followed by the post-task assessment of subjective engagement.

Half of the participants were placed into to an Icebreaker scenario before participating the task. They were then given five minutes to talk with their teammate by create a greater level of intimacy within the team setting. Discussions were guided by a short narrative script. Topics on this script included information about their social life as well as fears and emotional memories experienced. The remaining half of participant teams served as the Control and did not have this conversation prior to starting the tasks.

The task used in this experiment was a simulated "bomb defusal" game [5]. Participants first completed a practice phase of the game to acquaint themselves. After practice, participants then completed two rounds of the game. Each game had a duration of roughly five minutes. During each game, one team member acts as the bomb diffuser and the other serves as the manual reader. The manual reader was given a document with the directions on how to go about defusing the bomb. The manual reader was told that it was their responsibility to provide correct information to allow their teammate to successfully defuse the bomb. After completion of each game, the team members switched roles (i.e., each team member was given the opportunity to be both the diffuser and the manual reader twice during the main task. This resulted in the completion of four games in total.

The simulated bomb has three randomized "modules" that differed in their level of difficulty as well as what steps were needed in order to solve them. The task of the 'diffusor' was to communicate clearly what the modules looked like so that their teammate could provide the correct instructions on how to disarm each individual module in accordance with the manual. Because each participant could only see their own display (i.e., the 'diffuser' could not see the screen with the manual and the 'manual reader' could not see bomb), clear communication was necessary for success. Participants were given five minutes to defuse each bomb. If a mistake was made (i.e., cutting the wrong wire or pressing an incorrect button within a module) three times, the bomb exploded and the game ended.

2.2 Participants

Twenty pairs of gender-matched individuals (20 male, 20 female) were recruited for this study via Craigslist. Age range was from 18- 65 years with an average age of 40.03 years (SD = 13.39). All participants were unfamiliar with each other, had normal or corrected-to-normal vision, no history of epilepsy, were not taking any medication and English was their primary language.

2.3 Measures

Subjective Questionnaire Measures. Participants completed the Dundee Stress State Questionnaire (DSSQ) [6] before and after the task to assess their subjective ratings of engagement. The pre-task DSSQ assesses 11 dimensions of mood, motivation, and cognition in performance settings. Scales are grouped into three higher-order factors associated with task engagement (e.g., energy, task motivation and concentration), distress (e.g., tension, low confidence) and worry (e.g., self-esteem, task-related thoughts) symptoms. For this paper we focus only on the engagement factor.

Cardiovascular Measures. A BIOPAC MP150 [7], was used with a standard lead II electrode configuration to record electrocardiography (ECG), impedance cardiography, and changes in blood pressure. Data were recorded continuously for each participant throughout the game and analyzed offline.

The raw time series for each game were first segmented into thirty second intervals. The Moving Ensemble Average Program (MEAP) [7] was used to extract cardiovascular features of interest including pre-ejection period, cardiac output, and total peripheral resistance. This program performs a moving ensemble average of the data from each of interval and works similarly to even-related potential analysis used in electroencephalogram (EEG) research. The pre-ejection period (PEP) is the time from the onset of the heart muscle depolarization to the opening of the aortic valve. When PEP decreases, ventricular contractility (VC) increases. VC has been used extensively as a measure of engagement or attentiveness to a task [8,9,10,11]. Cardiac output is the amount of blood pumped in liters per minute. Total peripheral resistance (TPR) reflects vasodilation (increased blood flow) and vasoconstriction (decreased blow flow) which are related to parasympathetic and sympathetic activity. It has been shown that TPR unambiguously increases in a threat state and decreases in a challenge state, whereas CO either remains unchanged or decreases in a threat state and increases in a challenge state [12]. These three cardiovascular features were transformed to percent change from the average of last sixty seconds of baseline data.

3 Results

3.1 Task Engagement

VC was examined to assess changes in task engagement over the duration of the experiment. A repeated measures ANOVA was run with time as a within-subjects factor reflecting each of the last thirty seconds from all games (Practice 1 – Game 4) and experiment group (IB or CT) as the between-subjects factor. Figure 2B illustrates the change in VC as a function of time for the IB and CT conditions. IB participants showed an increase in VC during the later portion of the experiment whereas CT participants show the opposite effect, $F(5,170) = 2.6$, $p < .05$, partial $\eta 2 = .071$.

Engagement state was also assessed pre- and post-task using the DSSQ factor of task engagement. A one-way ANOVA was run in order to assess the effects of the task on changes in pre- and post-task subjective engagement. Time was a within-subjects factor contrasting pre and post-task state. The between-subjects factor was experiment group (IB or CT). Figure 2B illustrates the change from pre- to post-task state for task engagement as a function of the IB and CT conditions. On average, task engagement decreased. Although not significant, there is a trend towards the CT condition exhibiting lower levels of post-task engagement. Task Engagement was significantly lower after the experiment than before, $F(1,38) = 43.82$, $p < .001$, partial $\eta 2 = .54$. Participants in the IB condition exhibited a trend for higher levels of engagement both before and after the task.

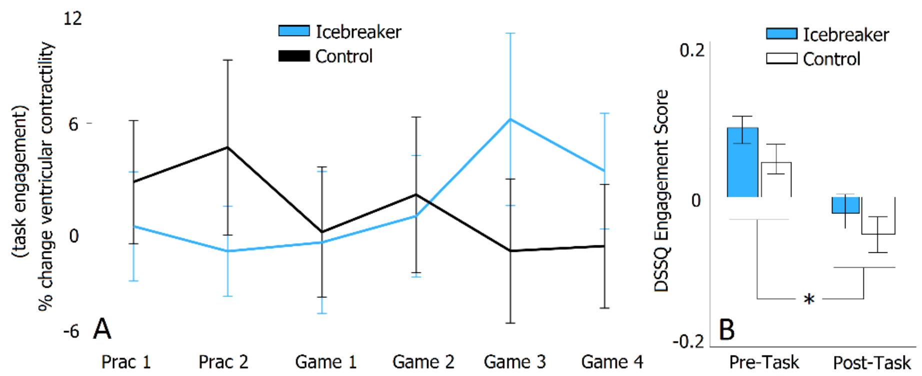

Fig. 2. A. Change in ventricular contractility between the IB and CT groups as a function of experiment time. B. Pre to post-task changes in subjective engagement state. Error bars show one standard error of the mean for both A and B.

3.2 Challenge and Threat State

TPR and CO were converted to a unitary index of challenge and threat according to [13]. This transformation was applied because changes in these measures reflect the same underlying sympathetic and parasympathetic nervous system activation. The unitary index was created by summing participants' percent change in TPR and CO values after assigning CO a weight of 1 and TPR a weight of -1, so that greater positive values corresponded to greater challenge. Paired samples *t*-tests were than computed between experiment role (manual reader and bomb diffuser) and experiment group (IB or CT). Although no significant differences were found, it is clear that the data trend to show that IB participants exhibit greater challenge across all experiment intervals than CT participants (see Figure 3).

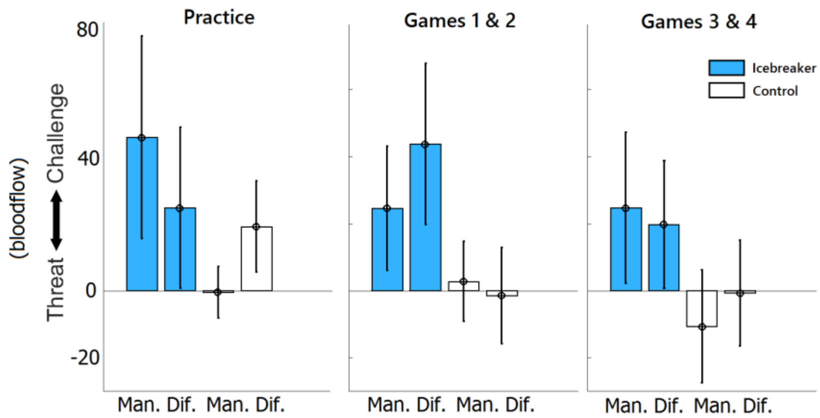

Fig. 3. Change in blood flow and state between the IB and CT groups as a function of game role and experiment time. Error bars show one standard error of the mean.

3.3 Machine Learning

We attempted to classify subjects as belonging to the CT or IB group using only their cardiovascular features from the final thirty seconds of data from the first real game after practice. A leave-one-out cross validation method was using where each model was trained on the data from seventeen subjects and tested on the unseen remaining subject. This process was repeated for all subjects and the final model performance was considered as the average performance across iterations. A principle components analysis (PCA) was also performed on the cardiovascular features and the training and testing procedure was repeated to create a new set of models for comparison.

Five different machine learning algorithms were tested: bagged tree ensemble, decision trees, support vector machine (SVM), Naïve Bayes, and a k-nearest neighbor model (KNN). The cardiovascular features submitted to each model included those previously discussed (PEP, CO, TPR) as well as heart rate, stroke volume, and mean arterial pressure. The classification performance of each model is shown in Table 1. For all models, the inclusion of PCA resulted in a performance loss of at least 10%.

The greatest classification performance was with the KNN model. For this model, we set $k = 2$ and used an exhaustive search with Euclidian distance. Prior probabilities for each class (IB or CT) were set to 0.5 each.

Table 1. Classification accuracy for each machine learning method.

Machine Learning Method	**Average Classification Accuracy** ($n = 18$)
200 Bagged Tree Ensemble	44%
Decison Tree	61%
SVM	72%
Naïve Bayes	72%
K – Nearest Neighbors	88%

4 Discussion

The purpose of this investigation was to examine changes in team behavior within a two-person dyadic scenario. Specifically, this study measured the effects of teammate familiarity on level of engagement, challenge/threat state, and physiological response. Overall, the data suggest that intra-team intimacy leads to increased task engagement and the exhibition of a challenge state, whereas the lack of team intimacy shows decreased task engagement and decreases in measures related to a challenge state.

4.1 RQ1 – State depends on team intimacy

Our first research question was focused on assessing whether familiarity with a teammate would result in cardiovascular changes that are related to being in a challenge or threat state. Personal knowledge of the teammate was manipulated by allowing half of the teams to participate in an informal icebreaker conversation prior to the task (i.e., the IB condition). We recorded EKG, impedance cardiography and pulse to assess changes in physiology in relation to the task, which was a simulated bomb defusal scenario. First, we found that ventricular contractility increased significantly over time for IB subjects and decreased for CT subjects. This suggests that individuals who knew each other became more engaged with the task the longer they participated in the experiment.

The results of the DSSQ show that participants reported feeling significantly less engaged after the experiment compared to before the experiment. This effect has been found in a number of studies [14,15,16] At face value this appears to conflict directly with the physiological results, but two things must be considered. The first is that the questionnaire was taken after the end of experiment when participants had just experienced failing the game six times in a row. Therefore, it may be the case that participants were simply tired and ready to be done with the experiment. Second, the physiological measure, namely percent change in VC, was computed during the final thirty seconds when the game was the most intense and stressful. Although the DSSQ is measuring engagement, it is a subjective measure whereas the physiological data require no conscious information from the participant and reflect changes on a much finer time scale. Thus researchers must consider what aspect of engagement they are trying to measure when considering the use of either a questionnaire or physiological measure.

We also report a trend that IB participants exhibited a greater challenge response than CT participants. The Biopsychosocial Model of Challenge and Threat [17] describes a threat state as occurring when perceived resources are less than task demands, and a challenge state as occurring when perceived resources are greater than or equal to task demands. In this experiment, task demands were fixed for all participants regardless of experiment group (i.e., IB or CT) and since no participant had played the game before, resources upon entering the experiment room were also the same. The primary manipulation of team intimacy, namely by the inclusion of an icebreaker session for half of the teams, may have resulted in the perception of increased resources from the stronger social interaction in comparison to control groups. In other words, when teammates know each other, they might feel like they can rely on the other to help them in a high-stress scenario, and this can be captured through cardiovascular measures.

4.2 RQ2 – Group classification through machine learning

Finally, we show that participants' data are well-classified through machine learning. The KNN model was able to classify unseen data as belonging to the IB or CT groups with an average accuracy of 88%. This is interesting because KNN is a relatively simple algorithm and suggests that the combination of features used are quite distinct between the experiment groups. However, as this model was only trained and tested on one of the games, further work is necessary to determine how robust the model is to classifying data from other game sessions.

5 References

1. Neubauer, C, Woolley, J.; Khooshabeh, P., Scherer, S. Getting to Know You: A Multimodal Investigation of Team Behavior and Resilience to Stress. In press. International Conference on Multimodal Computing. *Tokyo, Japan*
2. Blascovich, J., Seery, M. D., Mugridge, C. A., Norris, R. K., & Weisbuch, M. (2004). Predicting athletic performance from cardiovascular indexes of challenge and threat. *Journal of Experimental Social Psychology*, *40*(5), 683-688.
3. Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. *Advances in experimental social psychology, 28*, 1-51.
4. Blascovich J. (2008). Challenge, threat, and health. *Handbook of motivation science*. New York: Guildford, 481–493.
5. Kane, B., Fetter, B., & Pestaluky, A. (2015). Keep Talking and Nobody Explodes. [Steel Crate Games]. Ottawa: Canada.
6. Matthews, G., Campbell, S.E., Falconer, S., Joyner, L., Huggins, J., Gilliland, K., Grier, R., & Warm, J.S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress and worry. Emotion, 2, 315- 340.
7. BIOPAC Systems, Inc., Goleta, CA, USA.
8. Newlin, D. B., & Levenson, R. W. (1979, nov). Pre-ejection Period: Measuring Beta- adrenergic Influences Upon the Heart. Psychophysiology, 16(6), 546–552.
9. Richter, M., & Gendolla, G. H. E. (2009, may). The heart contracts to reward: Monetary incentives and preejection period. Psychophysiology, 46(3), 451–457.
10. Spangler, D. P., & Friedman, B. H. (2015). Effortful control and resiliency exhibit different patterns of cardiac autonomic control. International Journal of Psychophysiology, 96(2), 95–103.
11. Seery, M. D. (2011, jul). Challenge or threat? Cardiovascular indexes of resilience and vulnerability to potential stress in humans. Neuroscience and biobehavioral reviews, 35(7), 1603–10.
12. Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. Journal of Personality and Social Psychology, 83(1), 63–72.
13. Seery, M. D., Weisbuch, M., & Blascovich, J. (2009). Something to gain, something to lose: The cardiovascular consequences of outcome framing. International Journal of Psychophysiology, 73(3), 308-312.
14. Neubauer, C., Matthews, G., Langheim, L.K., & Saxby, D.J. (2012a). Fatigue and voluntary utilization of automation in simulated driving. Human Factors, 54, 734-746.
15. Neubauer, C., Matthews, G., & Saxby, D.J. (2012b). The effects of cell phone use and automation on driver performance and subjective state in simulated driving. Proceedings of the Human Factors and Ergonomics Society, 56, 1987-1991.
16. Neubauer, C., Matthews, G., Saxby, D.J., & Langheim, L.K. (2010). Simulator methodologies for investigating fatigue and stress in the automated vehicle. Advances in Transportation Studies: An International Journal, Special Issue, 7-18.
17. Seery, M. D. (2013, sep). The Biopsychosocial Model of Challenge and Threat: Using the Heart to Measure the Mind. Social and Personality Psychology Compass, 7 (9), 637-653.