Analysis of Cognitive Load in Human Robot Interactions

Simon Bluett, Alfie Sargent, Natalia Gonzalez Cadiente, Xuan Wang, Jinze Ding
University of West England

***Abstract*—The aim of the project was ... Include background, methods, results and a conclusion.**

***Index Terms*—cognitive load, ambiguity, human-robot interaction**

I. Introduction

Background, motivation, aims and solutions

In this study, we aim to measure the change in cognitive load which occurs as a result of introducing a physically embodied robot when interacting with a human to complete a complex task.

II. Related Work

The following section details pre-existing work in the two main areas that are being investigated by this user study: Human Robot Interaction vs Video Displayed Agents, and Cognitive Load Theory and its physiological effects in Human Robot Interaction.

A. Cognitive Load Theory

Remove this: *For the study being described to be meaningful, prevailing cognitive load theory must be examined.* Based on prevailing cognitive load theory [1, 2, 3] there are 3 types of cognitive processing when learning a new skill: extraneous processing, intrinsic processing & germane processing.

Extraneous processing is created dependent upon the way the information is presented to the individual such as having a badly laid out book that requires the reader to go back and forth between pages. Intrinsic processing is the cognitive load needed to understand the material that is being learned; the inherent load required to learn something. Germane processing is a deeper cognitive learning process that involves the individual mentally organising what has been learned for cognitive access later; this is often described in papers as being organised into a schema [4]. **NEED MORE HERE**

B. Cognitive Load Measurements

This study investigates the effects that human robot interaction has on cognitive load for simple tasks & as a result it is important to understand how to accurately measure this. Cognitive Load & by extension the varieties of cognitive load are difficult to measure because a variety of conditions effect it such as the task being done, how its being done, amount of physical activity required etc. To generate the most reliable results for this study, the environment must be as controlled as possible.

The paper studied [1] investigated the cognitive load of 155 college students with limited engineering knowledge when attending a lesson on electric motors. The lesson itself involved a secondary task where the participants were required to respond to a distraction task throughout the experiment as well as self-report their cognitive load throughout the experiment. The cognitive load was measured in a few ways: a preliminary questionnaire, response time to the secondary task, 8 self-report scales throughout experimentation & a final questionnaire. The results created are consistent with the prevailing theory at the time [5]; that different features of cognitive load can be accessed by different measurements of cognitive load. However, there was not much variance in the type of cognitive load & this could potentially be due to one method of measurement; the self-reporting of cognitive load. By reporting cognitive load during the experiment, itself, the results are no longer reliable because the participants cognition is being manipulated during the task by making them do this report. The measuring techniques for cognitive load in this study will emulate some parts of the researched experiment but it will not require participants to report their own cognition this many times during the study to avoid tampering with the other measurements. Furthermore, the experiment discussed does not use any physiological or biometric measurements of cognition & therefore the conclusions drawn are not as impactful; the study detailed in this paper will use biometric data to further support any statements made.

A variety of papers have examined the different physiological effects that cognitive load can have on individuals such as eye position tracking [6], relative blood flow [7] & pupil size [8]. However, the most prominent measurement which suggests a proportionality with cognitive load is the galvanic skin response (GSR) of an individual which measures skin conductivity [9][10]. The paper that has been reviewed [7] consisted of a study with 13 male US air force volunteers aged between 18-21 with each participant being studied individually for 1 hour. The experiment had participants select fractions that appeared on the left of a computer screen & moved to the right. The participant was tasked with selecting fractions greater than a 1/3 before it hit the right side of the screen. The participant would gain points for choosing correctly, lots points for not selecting in time & also lose points for selecting the wrong fraction. The task also had 2 ”difficulties”, 1 with only 16 fractions that were simple and one with 56 fractions that were more complex.

The GSR results show that the device was very capable of detecting a decrease in sweat cell arousal as the complexity of the task increased. However, the study itself used a bespoke physiological measuring device with very little detail on its development. The results gathered from this device have not been proven to be reliable; the study described throughout in this paper will use a professional grade sensor to ensure reliability in results. Furthermore, the sensor itself was always applied to the participants left hand, with no indication of which hand was preferred by the user, potentially changing the results especially considering the experiment required participants to actively use their hand; this will be asked in the study described in this paper as well as ensuring the participants hand is stationary to ensure reliable results.

C. HRI vs Video Displayed Agents

Investigating how humans interact with robots in comparison to a computer is one of the main targets for this study & thus current research must be investigated prior to the study being run.

One study in particular investigated how a robots physical presence can affect a humans judgement when handling physical objects [11]. In the experiment, 65 university staff, graduates & undergraduates were tasked with relocating books around a room; 22 had the robot physically with them, 22 had a live feed of the robot & 21 had an augmented video of the robot. The results of the experiment show that the participants had a greater ”respect” for the robots sense of space & would actively go out of their way to avoid the robots space. Furthermore, the study showed an increase in ”trust” if the instructions for the task were being given by a physical robot.

Although this experiment was focused on social interaction & not cognitive load of participants, the results gathered are still relevant to the user study being created. The mention of personal space & trust being greater with the presence of a robot may be indicative of cognitive load as well; this will be studied further in the experiment being designed.

III. Methods

A. Hypotheses

The main research question of the study is: Does the introduction of a physically embodied robot increase cognitive load when interacting with humans to complete a complex task? Our hypothesis is that introducing a robot into the learning process of a complex task will increase cognitive load and therefore hinder the ability to complete the task.

B. User Study Design

The study was designed to take into consideration the three categories of cognitive load. The task consisted of operating a oscilloscope and waveform generator, whose complex array of buttons and dials should provide an intrinsic cognitive load to people who are not familiar with the equipment. The extraneous cognitive load comes from the demands by the teacher or instructions; how the information is presented. This type of cognitive load inhibits the humans ability to learn and successfully complete tasks and is therefore the type of load a robot may influence. We deliver some instructions via a robot, and others via a computer to analyse the changes in cognitive load in human-robot interaction. A within user study design is chosen due to the limited number of users available to participate within the time frame. This is to allow us to gather sufficient results for analysis. The study works by alternating which test condition the user is exposed to first; robot or computer. 9 participants were used in this study. Participants had a variety of backgrounds and experiences with both the oscilloscope equipment and robot interaction.

C. User Study Procedure

The participant is guided through the task of creating a waveform using the equipment provided in 8 steps. Four steps are presented by the robot, and four by the computer. Throughout this, the participant must complete an awareness task.

1) Awareness Task: We set a cognitive load baseline by using an awareness task. The users are asked to pay attention to a computer whose screen changes colour periodically. When the screen changes to black, the participant must press the spacebar. The users heart rate and skin temperature during this time, and the number of correct and incorrect key presses by the user is measured. This task is used before the test to create a baseline, and during both computer and robot stages of the study.

2) Questionnaires: Three questionnaires are used throughout the study:

- Pre-study Questionnaire: Questions regarding age, demographic, equipment experience, and robot experience
- Mid- and Post-study Questionnaire: The same questionnaire presented separately, after the robot condition and after the computer condition. Questions regarding frustration levels, understanding of the task, how well they have learned the task. (NASA TLX style)
- Mid- and Post-study Interview: Pre-written questions regarding opinions of the robot/computer, suggested improvements, and thoughts on the task itself.

D. Dependent Measures

- NASA TLX style questionnaire [12]
- Galvanic Skin Response (GSR) and Heart Rate
- Awareness Test
- Time taken to complete the tasks
- Number of times the instructions had to be repeated
- The successful completion of each of the tasks

IV. Datasets and Results

The experiment was conducted at the Bristol Robotics Laboratory and it was attended by randomly selected 5 men and 4 women. All participants ages are between 21-25 and 66.7% of them are English native speaker. The pre-survey questionnaire showed that three of them had no relevant human-robot interaction experience, neither heard of the studied robot NAO. But only one of them has not contacted with home service robots before. Due to the participants academic background also influence their cognitive ability in this task, the familiarity to the mission devices have also been taken into consideration, more than half of them are unfamiliar with the task devices like oscilloscope and so on. The duration of each trial was about 15 minutes. During the course of the experiment, the basic physiological characteristics of the participants, such as heart rate, skin conductance, and skin resistance, were recorded and saved using a shimmer sensor. Referring to these data and the post questionnaire the psychological changes in this user study can be assessed.

A. *r1*

B. *Post Experiment Interview*

Five participants claim that there is no significant difference between the instructions from robot and computer, due to the complexity of the mission, they even have not pay attention to the robot. In contrast, the questionnaire also show that the computer instructions need less attention, one person prefer to do the interaction with computer. The other two state that they have better understanding with the help of NAOs friendly gesture.

V. Discussion

VI. Conclusion

References

[1] Krista E DeLeeuw and Richard E Mayer. A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of educational psychology*, 100(1):223, 2008.

[2] Ton De Jong. Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional science*, 38(2):105–134, 2010.

[3] John Sweller. Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4): 295–312, 1994.

[4] Fred Paas, Alexander Renkl, and John Sweller. Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, 38(1):1–4, 2003.

[5] Paul Ayres. Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and instruction*, 16(5):389–400, 2006.

[6] Oskar Palinko, Andrew L Kun, Alexander Shyrokov, and Peter Heeman. Estimating cognitive load using remote eye tracking in a driving simulator. In *Proceedings of the 2010 symposium on eye-tracking research & applications*, pages 141–144. ACM, 2010.

[7] Curtis S Ikehara and Martha E Crosby. Assessing cognitive load with physiological sensors. In *Proceedings of the 38th annual hawaii international conference on system sciences*, pages 295a–295a. IEEE, 2005.

[8] Timo Partala and Veikko Surakka. Pupil size variation as an indication of affective processing. *International journal of human-computer studies*, 59(1-2):185–198, 2003.

[9] Yu Shi, Natalie Ruiz, Ronnie Taib, Eric Choi, and Fang Chen. Galvanic skin response (gsr) as an index of cognitive load. In *CHI'07 extended abstracts on Human factors in computing systems*, pages 2651–2656. ACM, 2007.

[10] Nargess Nourbakhsh, Yang Wang, Fang Chen, and Rafael A Calvo. Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. In *Proceedings of the 24th Australian Computer-Human Interaction Conference*, pages 420–423. ACM, 2012.

[11] Wilma A Bainbridge, Justin W Hart, Elizabeth S Kim, and Brian Scassellati. The benefits of interactions with physically present robots over video-displayed agents. *International Journal of Social Robotics*, 3(1):41–52, 2011.

[12] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.