Trust and Cognitive Load in the Text-Chat Environment

Author:
Khawaji, Ahmad

Publication Date:
2016

DOI:
https://doi.org/10.26190/unsworks/3044

Trust and Cognitive Load in the Text-Chat Environment

Ahmad Khawaji

A thesis in fulfilment of the requirements for
the degree of doctor of philosophy in
computer science and engineering

Supervisors: Dr. Fang Chen, Dr. Nadine Marcus and
Dr. Jianlong Zhou

Oct 2016

ORIGINALITY STATEMENT

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed:

Date: 24/Oct/2016

To my wife and my son for their support and their patience during the PhD program.

To my parents for their upbringing (My Lord! Bestow on them Your Mercy as they did bring me up when I was young).

To the loving memory of my grandmother.

Acknowledgement

I would like to wholeheartedly thank my supervisors Dr Fang Chen, Dr Nadine Marcus and Dr Jialong Zhou for their wisdom, support, guidance, patience and encouragement during the PhD program.

Also, special thanks for Dr Yang Wang, Mr Ronnie Taib, Dr Asif Khawaja and Dr Kun Yu for the time they spent with me when I consulted with them on the design of my experiments.

Thanks also to Ms Michele Money for proofreading my thesis and correcting the linguistic expressions that I used in my thesis.

Publications by the Author

During the course of this research, many publications have been made:

1- **Ahmad Khawaji**, Fang Chen, Nadine Marcus, and Jianlong Zhou. Trust and Cooperation in Text-Based Computer-Mediated Communication (CMC). In Proceedings of OzCHI Conference, 2013. *This is a main part of my thesis.*

2- **Ahmad Khawaji**, Fang Chen, Jianlong Zhou, and Nadine Marcus. Trust and Cognitive Load in the Text-Chat Environment: The Role of Mouse Movement. In Proceedings of OzCHI Conference, 2014. *This is a main part of my thesis.*

3- **Ahmad Khawaji**, Jianlong Zhou, Fang Chen, and Nadine Marcus. Using Galvanic Skin Response (GSR) to Measure Trust and Cognitive Load in the Text-Chat Environment. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2015). *This is a main part of my thesis.*

4- Fang Chen, Jianlong Zhou, Yang Wang, Kun Yu, Syed Arshad, **Ahmad Khawaji,** and Dan Conway. Robust Multimodal Cognitive Load Measurement. (Springer (Book) - Human Computer Interaction Series 2016). I participated in chapter 9 *"Mouse Based Measure"* and chapter 13 *"Trust and Cognitive Load"*. The link: http://www.springer.com/us/book/9783319316987. *This is a main part of my thesis*.

5. Jianlong Zhou, Jinjun Sun, Fang Chen, Yang Wang, Ronnie Taib, **Ahmad Khawaji**, and Zhidong Li. Measurable Decision Making with GSR and Pupillary Analysis for Intelligent User Interface. ACM Transactions on Computer-Human Interaction (ToCHI, 2015). *This provided me the inspiration to use "GSR" in the thesis.*

6. Jianlong Zhou, Constant Bridon, Fang Chen, **Ahmad Khawaji**, and Yang Wang. Be Informed and Be Involved: Effects of Uncertainty and Correlation on User's Confidence in Decision Making. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2015). *This isn't a part of my thesis.*

ABSTRACT

Over the last decade, text messages have become one of most popular forms of communication. A diverse array of software to enable the exchange of text messages has been developed and deployed on computers and smart phones which is one of reasons for the increased use of text messages for communication purposes. However, there are several disadvantages for communicators when using text messages. For example, in the text-chat environment, there is a lack of trust between communicators (Bos et al., 2002) and communicators may face a high cognitive load (Thirunarayanan et al., 2002). Therefore, we conducted several studies to improve communication between people in the text-chat environment where important elements for communication, such as facial expressions, do not exist. We examined various data collected from people with different levels of interpersonal trust and cognitive load, including chat contents (e.g., the number of assent words used), mouse movements (e.g., distance travelled by the mouse) and physiological signals (skin response). These data proved to be useful indicators for measuring the level of trust of the communicators and the cognitive load to which the communicators were exposed. These findings have implications for enhancing the communication and relationships between communicators in a text-chat environment and will assist system developers to design applications to measure the conditions the communicators are under when they use the text-

chat environment to complete tasks in a business or government context and provide them with suitable assistance.

In addition to using the previously mentioned data to measure trust and cognitive load, we further examined the text-chat environment by allowing communicators to see the actions of their partners in real time when they were trying to solve tasks. Our results show that when communicators are able to see the behaviors of their partners when they chat in the text-chat environment while completing tasks, it improves the level of interpersonal trust between them. Finally, in this thesis, we examined the effects of cooperation and cognitive load on interpersonal trust. The findings show that trust is significantly improved between communicators in the text-chat environment when the communicators act in a cooperative way and when they are under a low cognitive load.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Research Motivations

The emergence of the Internet has dramatically changed the way people communicate, and there is now a strong reliance on text communication. The social network service Facebook, one of the most popular online services, provides a text communication environment which can be used for both social and business purposes. As an example of the influence of Facebook, the most recent statistics show that every day in June 2015, there were 968 million active Facebook users and 844 million using Facebook from their mobile phones (Facebook Reports Second Quarter 2015 Results, 2015). Another similar text communication service is offered by Twitter, with the difference being that Twitter does not allow users to write messages longer than 140 characters. Twitter's website shows that there are 316 million active users per month and that each day, 500 million tweets are sent (Twitter Usage, 2015). Another popular application which can be deployed on smartphones is WhatsApp. Recent statistics show that there are 600 million users of WhatsApp every month (Koum, 2014). In addition, Skype, an application which can be deployed on computers, has 4.9 million daily users (Skype Statistics and Facts, 2015).

People today regularly use computer-mediated communication (CMC), such as video conferencing, email and text messaging more so than face-to-face communication (Quan-Haase et al., 2005). People often use this media because it is more convenient and less expensive than face-to-face communication which needs a mutually convenient location and time for a meeting to take place. The text-chat environment is one form of the CMC which enables people to exchange messages in a synchronized and asynchronous manner. A study showed that, in an organization, workers used the text-chat environment to communicate significantly more than telephone communication (Quan-Haase et al., 2005). However, the research showed that the level of trust between people engaged in CMC is low due to the absence of face-to-face communication, as people usually find it difficult to trust someone they cannot see (Riegelsberger et al., 2003). Furthermore, people who communicate via CMC are sometimes from different contexts or cultures which may also affect trust (Riegelsberger et al., 2003).

It has been found that trust between people is weaker in the text-chat environment compared with face-to-face communication and also compared with other forms of CMC, such as video and audio (Bos et al., 2002). Also, it has been found that in the text-chat environment, people may be exposed to a high cognitive load (Thirunarayanan et al., 2002). One of the most important questions is how can communication between people in the text-chat environment be improved by reducing the negative effects of weak trust and a high cognitive load, which adversely impacts the performance and productivity of team workers in organizations which use this media for communication. Therefore, we chose the text-chat environment for further investigation.

1.2 Research Methodology

In this thesis, we analyse various data (e.g., chat contents) from 134 participants in different conditions of interpersonal trust and cognitive load across experiments. Specifically, the data which is analysed relates to linguistic features (e.g., positive emotion and assent words), time responses (e.g., pauses), keyboard buttons (e.g., *Backspace* button), mouse movement (e.g., distance travelled) and galvanic skin response (e.g., the average of GSR). We also use questionnaires to measure cognitive load and interpersonal trust. In addition, the social dilemma game is used in the experiments to measure interpersonal trust. This social dilemma game is an investment game which depends on trust in order for the participants to earn acceptable profits, therefore it is used to measure interpersonal trust (Scissors et al., 2008). To collect the galvanic skin response data, we use a device called the ProComp Infiniti System from Thought Technology Ltd.

1.3 The Significance of the Thesis

The growth in the number of text-chat environments has recently attracted a large amount of research attention to better understand the factors which affect people who communicate in this way and to examine ways to enhance text-chat communication.

As previously discussed, two issues have been found to have an adverse impact on communication in the text-chat environment, a low level of interpersonal trust between people and being exposed to a high cognitive load (Bos et al., 2002; Thirunarayanan et al., 2002). Hence, the primary goal of this research is to consider how these media may be enhanced so that they are as effective as face-to-face communication, which will improve the performance of organisations and individuals who mainly use the text-chat environment.

In this thesis, the findings have theoretical and applied implications that could be used to improve communication between people in the text-chat environment. On the theoretical side, we explore the factors that affect interpersonal trust, an important motivator of behaviour. For instance, the results found in this thesis show that trust is established in the text-chat environment when the participants are cooperative. On the applied side, we propose that interpersonal trust and cognitive load can be measured by analysing the linguistic features and mouse movements in the data collected from the communicators in the text-chat environment. These measures can be used to develop applications and systems to monitor the trust and cognitive load of communicators who use text-chat media in an organization to accomplish their tasks and provide them with the appropriate support to work effectively. Also, chat systems can be developed with new features to allow communicators to see the actions of their partners in real time when they are trying to complete a particular task online because this enhances the trust between the communicators. Figure 1-1 summarizes the thesis, showing the negative factors that may arise between people, that is low trust and high cognitive load, in the text chat environment and it also shows the procedures and steps followed to address these and mitigate the effects on the communicators.

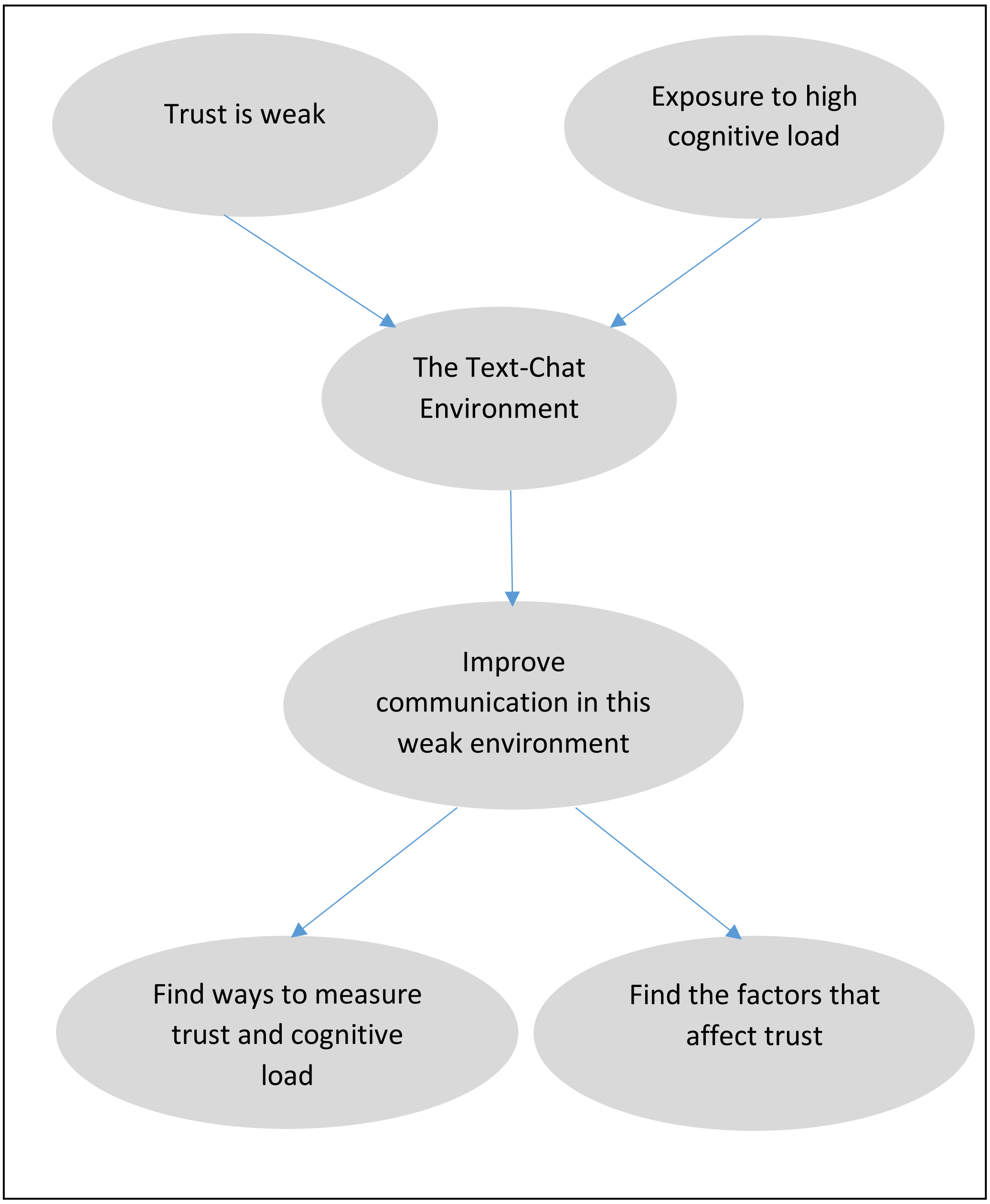


Figure 1-1: The problems which face people in the text-chat environment and the steps followed in the thesis to enhance communication.

1.4 Thesis Structure

This section provides a brief description of each chapter in this thesis:

- Chapter 1 details the motivation for this research which is due to the rapid increase in the use of text communication and also because of the negative factors which have been found (e.g., low trust (Bos et al., 2002)). It also overviews the research methodology and explains how the data collected from the participants (e.g., chat content) is used to measure interpersonal trust and cognitive load. Finally, the significance of the thesis and its role in enhancing the text-chat environment is detailed, for example, developing systems based on the data collected from the communicators and providing them with assistance.
- Chapter 2 reviews the background and related work, beginning with definitions of trust and cognitive load and their effects. This chapter also overviews the related research and the tools used in this research, such as the role of linguistics in measuring interpersonal trust and cognitive load, the trust test methodology (social dilemma game) and the Linguistic Inquiry Word Count tool (LIWC).
- Chapter 3 presents an experiment where different levels of cognitive load are manipulated, which can affect interpersonal trust. The results show that the participants trust their partners more when they are exposed to a low cognitive load. Also, the experiment examines the hand movements of participants by specifically looking at how the participants move the mouse (e.g., distance travelled and

movement count). The results show that the participants use the mouse more actively under a low cognitive load. In addition, the chat contents (e.g., word count and message count) are examined in this environment and the results also show that the participants chat more with their partners when they are exposed to a low cognitive load.

- Chapter 4 describes how different behaviours can affect interpersonal trust. We design two chat systems which reply automatically to the participants, so the participants chat with automatic systems rather than a real partner. The first chat system, called a cooperative partner, replies with stored cooperative messages and displays cooperative behaviour and the other chat system, called a competitive partner, replies with stored competitive messages and displays competitive behaviour. The results show that participants had higher trust when chatting with the system which replies with cooperative messages. In addition, we show how certain linguistic features (positive emotion and assent words) are positively associated with interpersonal trust.
- Chapter 5 presents an experiment to investigate the effects of overlapping conditions between interpersonal trust and cognitive load on the galvanic skin response (GSR), a physiological measurement that indicates the degree of nervousness of people and

the amount of sweat on their skin (Peuscher, 2012; Westerink et al., 2008) where a high GSR value indicates an increase in nervousness and the amount of sweat. Also, we examine the effects of an overlap between trust and cognitive load on the hesitation features in writing messages (using the *Backspace* and *Delete* buttons on the keyboard). The experiment results indicate that there is only an interaction effect between interpersonal trust and cognitive load on the GSR values when the GSR values of the participants are at the lowest level when the participants trust their partners and are exposed to a low cognitive load at the same time.

- Chapter 6 examines how adding the feature of shared visual information (which allows communicators to see the actions of their colleagues online when they complete tasks) in the design of the text-chat environment enhances interpersonal trust. We allowed communicators to solve several tasks with and without shared visual information and the findings show that the trust of the communicators was enhanced significantly with shared visual information compared to those who worked with a partner without shared visual information.
- Chapter 7 examines several algorithms to predict the level of cognitive load with linguistic features (e.g., word count and message count) and mouse features (e.g., distance travelled and movement

count) and we also combine these linguistic and mouse features to determine the ability of fusing/combing for cognitive load classification but none of these attempts shows a high accuracy for load classification. However, we examine the hesitation features (*Backspace* and *Delete* buttons) for the classification of interpersonal trust and the results show that measuring the use of the Backspace button can be useful in the prediction accuracy when classifying interpersonal trust. This chapter shows that when participants trusted their partners, they were less cautious, as we noted a reduced use of the *Backspace* button.

- Chapter 8 summarizes the findings of our experiments and also details how these findings can be used. This chapter also shows how this work can be extended by applying other strategies to improve communication in the text-chat environment.

Chapter 2 Background and Related Work

In this chapter, we explore the background of trust and cognitive load and we also examine studies which have been conducted on CMC which are related to our research in order to identify the gaps in the existing literature and find new ideas to help us reach our goal of improving communication in the text-chat environment (such as improving communication via measuring interpersonal trust between communicators).

2.1 Trust and Cognitive Load

Mayer et al. (1995) define trust as *"a willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party"*. Trust is simply a feeling of one individual towards another individual which is built on several factors, such as cooperation and transparency. The researchers examined the effects of different behaviors (cooperation and competition) on trust and they found that a person's trust in a competitive partner was less than their trust in a cooperative partner (de Melo et al., 2013). In addition, Norman et al. (2010) studied trust with transparency and found that transparency had a positive impact on trust.

The existence of trust is important for individuals and organizations. For instance, trust enhances perceptions and performance in an organizational environment (Dirks et al., 2001). Also, in relation to trust in the workplace and employee communication, Kelly (2013) stated that "*Establishing trust can increase profitability, boost market value, add competitive advantage, lower* *costs, provide efficiencies, improve morale, and result in lower turnover, improved productivity, and increased job satisfaction*".

Several studies have examined trust in CMC. Bos et al. (2002) studied the trust of people in different communication conditions, namely text chat, audio, video and face-to-face. Specifically, they compared the existence of trust in these four conditions and found that less trust was built in the text chat condition. In addition, it was found that the trust that was built in the audio and video conditions was similar and as good as face-to-face communication. Zheng et al. (2002) examined the existence of trust between people in different scenarios in the text-chat environment where before the chatting, they were allowed to see a photo of their partners, read the personal information of their partners and meet face-to-face for several minutes. They found that trust in the text chat environment was enhanced when people met face-to-face before chatting. Pai and Gasson (2008) examined different cultures (collectivist cultures and individualistic cultures) in relation to trust in CMC. The researchers found that people from collectivist cultures may more reconcile with their partners after experiencing untrustworthy behavior when their partners gave them an explanation than people from individualistic cultures.

The interpersonal trust between people can be measured by asking them several questions. Butler (1991) provides a list of questions for measuring trust. These questions are classified under different categories such as integrity (e.g. "*He always tells me the truth?*") and loyalty (e.g. "*He is likely to take advantage* *of me?*”). Participants were asked to respond to these questions on a five-point Likert scale (from “*Strongly Agree*” to “*Strongly Disagree*”).

Cognitive load is the amount of mental load imposed on a human’s working memory when attempting to accomplish a task (Chandler and Sweller, 1991). People usually have limited memories which differ from one to another, and when these memories are exposed to a heavy load when attempting to complete a task, it becomes more difficult. Cognitive load has an effect on individuals and organizations. For instance, a negative effect of a high cognitive load is the decreased performance of people engaged in tasks (Paas et al., 2004) and also decreased performance in terms of reading and understanding text (Clevinger, 2014). Also, cognitive load was defined as “*the state of mentally attending to one or more tasks peripheral to the task at hand*” and this may have a negative effect on the performance when completing a target task (Psychwiki Website, 2010). Thirunarayanan et al. (2002) studied cognitive load in the text chat environment and found that sometimes, cognitive load increased when people communicated via the chat environment.

To measure cognitive load levels, various data such as speech (Khawaja et al., 2007) and writing (Kun et al., 2011) can be collected and used for cognitive load measurement. These various data can be classified under categories (e.g., behavioural measurements and physiological measurements), which is similar to the categories and classification used in Khawaji’s thesis (2010). Cognitive load levels can also be measured using ratings (surveys). We reviewed the previous studies to identify the classifications that have been used (for example, speech data is classified as a behavioral measurement), which are discussed in detail in the following paragraph.

For behavioural data Khawaja et al. (2014) examined the effects of cognitive load on speech and found that there was a correlation between longer sentences and increased negative words with increased cognitive load. For physiological data, the relationship between cognitive load and sweat from the skin (measured using galvanic skin response) was examined, the results showing a correlation, as demonstrated in the study by Nourbakhsh et al. (2013) which showed that GSR can be used to measure cognitive load. Another study (Engstrom et al., 2005) investigated the link between cognitive load and gaze focus while driving, finding that gaze focus increased towards the center of the road with increased cognitive load.

Another way to measure cognitive load is via questionnaires. After completing a task, participants were asked to indicate the level of cognitive load required by the task. Nasa (1986) provides a list of questions that can be used to measure cognitive load, such as "*How hard did you have to work (mentally and physically) to accomplish your level of performance?*" using a Likert scale from "*low*" to "*high*"; and questions such as "*How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)*?" and "*How satisfied were you with your performance in accomplishing these goals?*" using a Likert scale from "*good*" to "*poor*".

2.2 Measuring Trust in the Text-Chat Environment

Previous studies have examined the measurement of trust in the text chat environment using linguistic features. The first study, conducted by Scissors, Gill and Gergle (2008) evaluated trust and its relationship to linguistic mimicry in text-based (CMC). The authors reviewed studies on the relationship between trust and mimicry, finding that mimicry between people is increased if they trust each other. They wanted to test this relationship between the participants in the text chat environment, expecting that high-trusting pairs will mimic each other's words more than low-trusting pairs.

In this study, the authors used the investment social dilemma game to divide the participants into high-trusting pairs and low-trusting pairs (high-trusting pairs who cooperated more in the investment social dilemma game and low-trusting pairs who defected more in the investment social dilemma game). They found that: 1) in within-chat sessions, high-trusting pairs repeated each other's words more than low-trusting pairs; and 2) in across-chat sessions, low-trusting pairs repeated each other's words more than high-trusting pairs but these words were standard, such as *okay* and *yeah*.

The second study, conducted by Scissors et al. (2009), where they examined trust and linguistic similarity (the same data collected in the previous study).

Linguistic similarity refers to the use of different words to refer to the same meaning, such as *soccer* and *football* and *apartment* and *flat* (Scissors et al., 2009). In this study, the authors wanted to see how trust is associated with words in different language categories between high-trusting pairs and low-trusting pairs. The categories of linguistic similarity are content (positive emotions: joy, negative emotion: angry), structural (interjections: hello, present tense verb: play) and stylistic (emoticons: >:], >: D) (Scissors et al., 2009). In addition, the authors wanted to compare the word count used by each participant and the words per sentence used by each participant.

They found that: 1) there was no difference in the word count between the two groups; 2) there was no difference in the number of words per sentence between the two groups; 3) high-trusting pairs used more future and past tense verbs than low-trusting pairs; 4) high-trusting pairs used more positive emotion words than low-trusting pairs; 5) low-trusting pairs used more negative emotion words than high-trusting pairs; and 6) high-trusting pairs used more emoticons than low-trusting pairs.

The research in this section is directly related to our research topic because it discusses trust in the written communication (text) environment and provides encouraging evidence about how linguistic features can be used to determine the extent of trust between people, whether high or low in text-based CMC.

2.3 Measuring Cognitive Load

Many studies have measured levels of cognitive load by collecting different modalities of data, such as behavioural data (e.g., talking) and physiological data (e.g., eye blinks) (e.g, Khawaja et al., 2013, Syed et al., 2013, Nourbakhsh et al., 2013 & Chen et al., 2016). Specifically, using behavioural data, Khawaja et al. (2013) analyzed the language used by trainee bushfire operators in Australia and found that they used a higher number of different words under a low cognitive load level. Also, Khawaja et al. (2007) examined the effects of cognitive load on the speech on twenty-four subjects and found the length of pauses increased with a high cognitive load. Syed et al. (2013) also examined different levels of cognitive load using pauses in the mouse movements of 88 subjects, finding that the number of pauses increased with a high cognitive load. Kun et al. (2011) examined the different levels of cognitive load using the participants' writing, finding that certain features, such as the speed of writing and the pressure of writing are indicators of cognitive load. In relation to the physiological data, Nourbakhsh et al. (2012) conducted a study on the galvanic skin response (GSR) of twenty-five subjects and found evidence that GSR can be used to measure cognitive load. In addition, Nourbakhsh et al. (2013) combined the results of the GSR with the eye blinks of thirteen subjects to identify cognitive load, producing acceptable results in predicting the levels of cognitive load. Also, Shi et al. (2007) studied the association between cognitive load and galvanic skin response (GSR), finding that the value of GSR increased with a high cognitive load.

2.4 Emotions and Trust

2.4.1The Effect of Emotions on Trust

Dunn and Schweitzer (2005) conducted five experiments on emotions and trust. The first experiment investigated the influence of anger, sadness and happiness on trust on 120 people at a train station. The second experiment investigated the effects of emotion on trust, conducted on 64 people at a train station. The third experiment investigated the relationship between emotions and trust, conducted on 161 people at a train station. The fourth experiment investigated the link between emotions and trust judgments, conducted on 112 undergraduate students. The fifth experiment investigated the relationship between emotions and trust, conducted on 181 undergraduate students.

In these experiments, the authors asked the participants to complete a questionnaire or describe things related to particular emotions, such as things that make you the most happy, sad or angry. The authors analysed the data from the five experiments and found that positive emotions, such as happiness and gratitude, increase trust between people, negative emotions, such as anger, decrease trust between people but emotions do not affect the trust between people when they know each other very well or when they know the source of the emotions.

This study shows how emotions impact trust, therefore we use these emotions with control participants who express different emotions to the other participants in the text chat environment to see how this affects trust between other participants.

2.4.2 Embodied Agent Emotions

Previous studies have examined the effects of embodied agent emotions on people. The first study, conducted by Melo, Zheng and Gratch (2009), examined the attitudes of people toward agents. They developed two agents: a cooperative agent who expresses particular emotions to encourage people to cooperate; and a neutral agent who expresses a neutral emotion to allow people to choose to cooperate or defect without attempting to influence their decision. These agents were used to determine the effect of facial expressions on a participant's decision making. The former agent expresses anger, gratitude, remorse and distress emotions with the avatar faces shown whereas the latter agent only shows a neutral emotion.

The participants play the prisoner dilemma game with the agents where the police do not have enough evidence to convict the participant or the agent. The rules of the dilemma game in this study are as follows: if the participant and agent remain silent, both are placed in prison for three months; if both testify against each other, they will be placed in prison for one year; if the participant remains silent and the agent testifies against the participant, the participant will be placed in prison for three years and the agent will be freed; and if the agent remains silent and the participant testifies against the participant, the participant will be freed and the agent will be placed in prison for three years.

The authors found that the participants cooperated more with the cooperative agent than the neutral agent and they thought the cooperative agent was more human-like than the neutral agent.

The second study, conducted by Melo, Carnevale and Grath (2011a), examined how the use of different emotions affects cooperation (using cooperative and individualistic agents).

The authors developed two computer software agents to play the investment social dilemma game and tested them with real participants: the first software agent was the cooperative agent who shows emotions (facial expressions) to encourage participants to cooperate with them. For example, if a participant cooperates and the cooperative agent doesn't cooperate in any round in the game, then the agent will receive more points than the participant, so the cooperative agent will show a sad avatar face to the participant as regret for what happened. This sad face usually encourages participants to cooperate with the agent in the next rounds. The second software agent was the individualistic agent who shows emotions (facial expressions) to encourage participants to play individually with them (such as a joyful avatar when the agent receives more points than the participant).

The authors found that that participants who played with the cooperative agent were more cooperative than the participants who played with the individualistic agent.

The third study, also conducted by Melo, Carnevale and Grath (2011b), examined the effects of happiness and anger emotions on people when using computer software agents. The authors wanted to investigate if people usually make more concessions in negotiations when they negotiate with people who express angry emotions compared with those who express happiness and neutral emotions.

The authors developed computer software agents to negotiate with participants where the agent represents the role of a buyer and the participant represents the role of a seller. These negotiations were conducted on three offers: the price of a mobile phone, the duration of a service contract and the duration of a warranty. The agent expresses emotions to the participants in two ways: verbally and in text form, for example, if the agent is angry, the software will show "*I am starting to get really angry*"; and in the non-verbal method the agents use avatar faces to express their emotions.

The authors found that the participants made more concessions when they negotiated with agents with emotions of anger compared with happiness and neutral emotions.

The striking thing in these three studies is that the emotions of agents affected cooperation, despite the participants being aware of the fact that they weren't dealing with real humans. These three studies are related to our research because, as we know, there is an associated correlation between cooperation and trust and one reflects the other (Nahapiet & Ghoshal, 1998). We can also use the idea of the embodied agent's emotions in the automated chat systems in the form of emotional statements rather than using avatar faces as used in these three studies to try and improve trust between communicators using these emotions.

2.5 Low Trust and High Cognitive Load in the Text Chat Environment

Bos et al. (2002) conducted a study to evaluate the trust between participants when they communicate with each other in different conditions. The authors compared trust between participants in face-to-face communication and three forms of CMC, video, audio and text chat.

The authors used the investment social dilemma game to measure trust in the different conditions. In addition, they used a questionnaire survey.

The authors found that the trust between participants in the audio and video communication environment was almost as high as the level of trust in face-to-face communication, unlike the level of trust in the text chat environment which was low. In addition, trust was high from the beginning of the game in face-to-face communication, unlike trust in the audio and video mediums where building trust between participants was slow.

This study provides evidence to show that trust in communication between people via text chat is low compared with other mediums and there is need to further investigate ways to enhance trust in this medium.

Thirunarayanan et al. (2002) conducted a study on cognitive load with thirty-four students. The authors asked the students to discuss a topic from a book they had read at school and the students answered questions on this topic, all via a text-chat environment. After this, the students completed a survey about cognitive load which included questions about their level of confusion and the degree of difficulty they had in focusing on their text-chat, the results showing that cognitive load sometimes increases in the text-chat environment.

2.6 The Effects of Design and Shared Visual Information in the Text Chat Environment

Gergle et al. (2004) conducted a study on how different chat system designs can affect the way tasks are completed.

The participants were divided into pairs, each pair consisting of two participants, one being the helper and the other being the worker to solve the puzzle task. The helper is given the correct form of the puzzle and describes the pieces and their positions to the worker to complete the puzzle. This study was conducted under different conditions: small dialogue history, large dialogue history, with shared visual information (allows the helper to see the work area of the worker) and without shared visual information.

The authors found that a chat system which has a large dialogue history has more impact on the speed of completing the puzzle game than a chat system which has a small dialogue history. Furthermore, a chat system which has shared visual information has more impact on the speed of completing the puzzle game than a chat system which has none. In addition, shared visual information has more impact on the speed of completing the puzzle game than dialogue history. Finally, shared visual information and a large dialogue history were more effective when the puzzle game tasks were more complex.

2.7 Improving Trust using a preliminary meeting in the Text Chat Environment

Zheng et al. (2002) examined the trust in the text chat environment. The authors divided the participants into five groups. In the first group, there was no communication between participants. In the other four groups, the authors asked the participants to do four pre-task activities: the first group communicated face-to-face for 10 minutes and discussed a topic; the second group communicated via text chat for 15 minutes and discussed a topic; in the third group, each participant only saw their partner's photo; and in the fourth group, each participant only read his partner's personal information.

The authors found the following results: trust between the participants was highest when the participants engaged in a face-to-face pre-task activity for several minutes before the task, and trust was good when the participants engaged in a text chat pre-task activity (text chat pre-task activity: the participants talked for several minutes using a text chat before the task). A surprising result was found in the photo pre-task activity where participants showed a similar level of trust as those who had engaged in the chat pre-task activity. Using the personal information pre-task activity was not effective in building trust between the participants.

2.8 Improving Trust using linguistic politeness in the Text Environment

Linguistic politeness is a well-accepted method of analysing the social interaction between people as it provides rich information about interlocutors such as social distance, solidarity and deference (Scollon 1983).

Brown and Levinson (1987) divided linguistic politeness into the following four categories, with examples from the text chat environment where the participants are solving mathematical problems (where these examples were taken from Park (2008), the definitions were from Brown and Levinson (1987) and Park (2008)):

1) direct speech: this kind of speech is used between people who are close each other.

For example:

*a group member [K–12 student] Ya**, help us out :-)***

*Moderator [adult] No comment, sorry**:-)***

2) positive politeness: this kind of politeness reflects the solidarity between people. Positive politeness strategies include seeking common ground, the use of an informal speech style, seeking agreement and the use of small talk.

For example:

*AEL **I like this way of sharing ideas***

NIP ***ya I'm all for it***

For example, use of informal speech style:

ORB I don't like having to write up my solution

MEA ***m 2***

3) negative politeness: negative politeness strategies include being hesitant, giving deference, being apologetic and the use of a formal speech style.

For example:

NIN btw. . .did any of you get this week's POW?

CMP the parallelogram?

NIN ya

CMP Yes

Moderator ***I am sorry, could you please*** *answer my question?*

4) indirect speech: one of the strategies used in indirect speech is the inclusion of hints.

For example:

ARE try posting ur picture on the URL

CPM ***I'm too new at that.***

However, a study conducted by Lam (2011) evaluated the impact of linguistic politeness in emails on trust between supervisors and subordinates in the workplace to determine whether trust increases between people who are more polite in their conversations.

In this study, the author wrote 20 emails containing requests from supervisors to subordinates using different forms of linguistic politeness, optional moves (supportive move: "*I know you are really busy, but could you send me the file as soon as you can?*") and levels of directness (direct: "*Send me the file as soon as you can*" and indirect: "*Could you send me the file as soon as you can*?") Lam (2011).

These emails appear to be authentic, have a date and time and are printed on separate sheets of papers. These papers were given to the participants to read. After this, the authors gave the participants a questionnaire about trust to determine whether the participants would trust the supervisors who wrote these emails, if they were in the position of the subordinates.

This study was conducted on 115 students at the Midwestern University and the results were as follows: the participants trusted the supervisors who used downgrader and supportive moves more than those who used aggravating moves; there was no difference in the trust of the participants in the supervisors when they read direct or indirect requests; and finally, downgrader moves with a direct request had a significant impact on building trust.

This section shows clearly how we can increase trust in the text environment by using linguistic politeness.

2.9 Linguistic Categories Tool

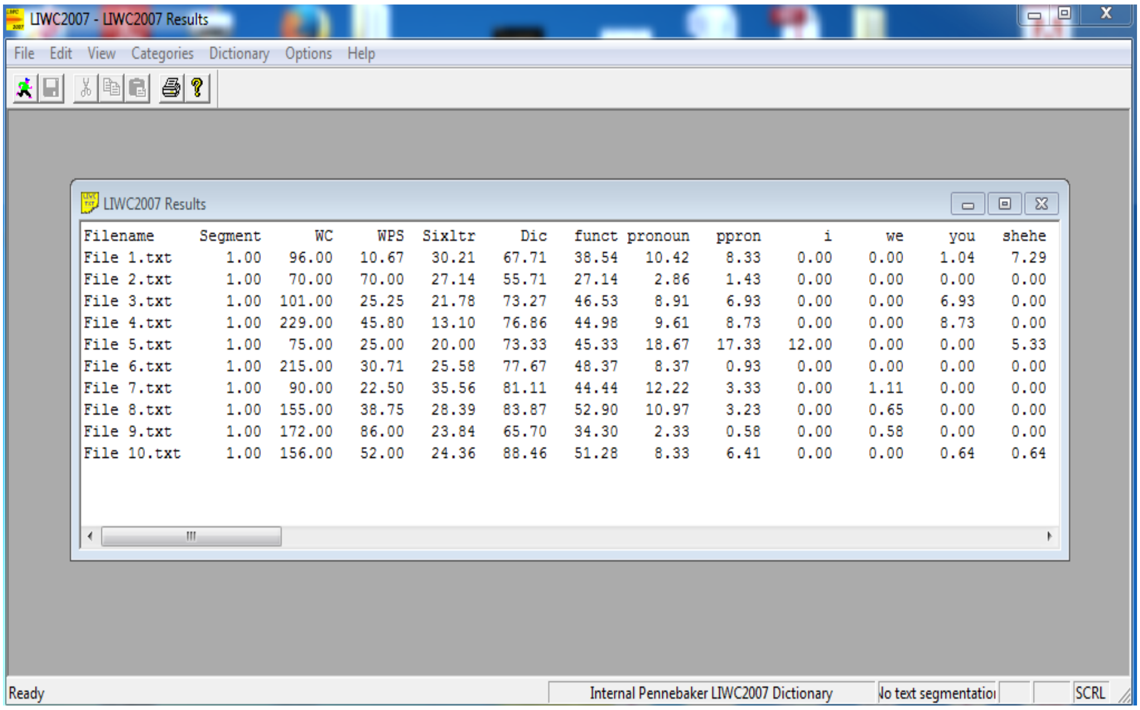
There are several websites and various software that analyse texts and split them into specific categories and provide a list of linguistic categories with the number of words used for each category. A well-known software package, Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al., 2007), has been used in previous work such as Ramirez-Esparza et al. (2008), Gill et al. (2008), Khawaja et al. (2009), Khawaja et al. (2012), Khawaja et al. (2014) and Nagarajan et al. (2009). LIWC divides text into various linguistic categories, such as social processes, affective processes and cognitive processes, with the majority of categories having subcategories, for example, a sub-category of social processes is family and friends, as shown in the screenshot in Figure 2-1 (LIWC Website). LIWC calculates the percentage of the occurrence of words from the total number of words for each category, as shown in Figure 2-2, when a text file from this research was tested using LIWC software.

2.10 Trust Test Methodology

Riegelsberger, Angela & McCarthy (2003) discussed and critiqued using the social dilemma game and also they presented previous studies which use the social dilemma game to measure trust. As shown in Riegelsberger, Angela & McCarthy (2003), the Prisoner's Dilemma game requires participants to play a number of rounds where participants have two options in every round, to cooperate or to defect and the game has three rules: 1) if both participants cooperate, they will each receive the same number of points (or the same amount of the payoff); 2) if one cooperates and one defects, the defecting participant will receive more points than the cooperating participant; and 3) if both participants defect, they will each receive the same number of points but this will be less than the points they would have received had they cooperated with each other, see Table 2-1, adapted from Riegelsberger, Angela & McCarthy (2003). This dilemma game can measure trust by evaluating the cooperation and defection actions, where more cooperation indicates higher trust and less cooperation indicates lower trust.

Swear words	swear	Damn, piss, fuck	53		.65/.48
Psychological Processes					
Social processes	social	Mate, talk, they, child	455		.97/.59
Family	family	Daughter, husband, aunt	64	.87	.81/.65
Friends	friend	Buddy, friend, neighbor	37	.70	.53/.12
Humans	human	Adult, baby, boy	61		.86/.26
Affective processes	affect	Happy, cried, abandon	915		.97/.36
Positive emotion	posemo	Love, nice, sweet	406	.41	.97/.40
Negative emotion	negemo	Hurt, ugly, nasty	499	.31	.97/.61
Anxiety	anx	Worried, fearful, nervous	91	.38	.89/.33
Anger	anger	Hate, kill, annoyed	184	.22	.92/.55
Sadness	sad	Crying, grief, sad	101	.07	.91/.45
Cognitive processes	cogmech	cause, know, ought	730		.97/.37
Insight	insight	think, know, consider	195		.94/.51
Causation	cause	because, effect, hence	108	.44	.88/.26
Discrepancy	discrep	should, would, could	76	.21	.80/.28
Tentative	tentat	maybe, perhaps, guess	155		.87/.13
Certainty	certain	always, never	83		.85/.29
Inhibition	inhib	block, constrain, stop	111		.91/.20

Figure 2-1: Screenshot taken from the LIWC website for several linguistics categories and subcategories with examples (from LIWC Website).



Filename	Segment	WC	WPS	Sixltr	Dic	funct	pronoun	ppron	i	we	you	shehe
File 1.txt	1.00	96.00	10.67	30.21	67.71	38.54	10.42	8.33	0.00	0.00	1.04	7.29
File 2.txt	1.00	70.00	70.00	27.14	55.71	27.14	2.86	1.43	0.00	0.00	0.00	0.00
File 3.txt	1.00	101.00	25.25	21.78	73.27	46.53	8.91	6.93	0.00	0.00	6.93	0.00
File 4.txt	1.00	229.00	45.80	13.10	76.86	44.98	9.61	8.73	0.00	0.00	8.73	0.00
File 5.txt	1.00	75.00	25.00	20.00	73.33	45.33	18.67	17.33	12.00	0.00	0.00	5.33
File 6.txt	1.00	215.00	30.71	25.58	77.67	48.37	8.37	0.93	0.00	0.00	0.00	0.00
File 7.txt	1.00	90.00	22.50	35.56	81.11	44.44	12.22	3.33	0.00	1.11	0.00	0.00
File 8.txt	1.00	155.00	38.75	28.39	83.87	52.90	10.97	3.23	0.00	0.65	0.00	0.00
File 9.txt	1.00	172.00	86.00	23.84	65.70	34.30	2.33	0.58	0.00	0.58	0.00	0.00
File 10.txt	1.00	156.00	52.00	24.36	88.46	51.28	8.33	6.41	0.00	0.00	0.64	0.64

Figure 2-2: Screenshot of an analysis of files by LIWC software showing the number of words used for each linguistic category.

		Participant 2	
Participant 1		Cooperation	Defection
	Cooperation	Participant 1: 2 points Participant 2: 2 points	Participant 1: 0 points Participant 2: 3 points
	Defection	Participant 1: 3 points Participant 2: 0 points	Participant 1: 1 point Participant 2: 1 point

Table 2-1: The points in the dilemma game (adapted from Riegelsberger, Angela & McCarthy (2003)).

However, drawing from the literature, this thesis explores how communication in the text chat environment can be improved via various experiments. These experiments examine the factors which may affect interpersonal trust. Specifically, we examine the effects of different behaviours (cooperative and competitive behaviours) on the interpersonal trust of communicators. We examine the effects of distractions and a reduction of attention and focus (using high cognitive loads) on the interpersonal trust of communicators and we also examine the effects of shared visual information on the interpersonal trust of communicators. In addition, this thesis examines how various data from the communicators, such as simple and short messages and the length of hesitation in responses, can be used as indicators of different levels of interpersonal trust and cognitive load. Moreover, this thesis examines physiological data (that is, data on the amount of sweat produced by the skin as measured by the GSR) and compares it with different levels of trust and cognitive loads and shows the relationship between them. Finally, this thesis discusses several attempts to use various machine learning algorithms (such as Random Forest) to predict interpersonal trust and cognitive load levels with the collected data.

Chapter 3 The Effects of Cognitive Load on Trust: The Role of Mouse Movements and Chat Contents

3.1 Chapter Contributions

In this chapter, we examine the effects of different levels of cognitive load on the extent of interpersonal trust. Mouse movements (such as distance travelled) under different levels of cognitive load were examined when the participants were chatting. In addition, as a different method by which to measure cognitive load, the chat content (such as the number of messages) was examined.

3.2 Chapter Organisation

Section 3.3 introduces the chapter, including definitions of trust and cognitive load and explains the experiment that will be conducted on trust and cognitive load. Section 3.4 presents the previous work on trust and cognitive load and overviews the hypotheses. The data collection procedure and the results are presented in section 3.5 and 3.6, respectively. The discussion which includes the justification for the findings and the conclusion to the chapter are presented in section 3.7 and 3.8, respectively.

3.3 Introduction

Trust refers to a situation when someone can predict how others will behave and what will occur from their behaviors (Starker, 2008). It is also defined as "*a willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to* *the trustor, irrespective of the ability to monitor or control that other party*" (Mayer et al., 1995). However, researchers have found that a lack of trust exists between interlocutors in the text-chat environment (Bos et al., 2002), but despite a lack of trust, this chat medium is commonly used. For instance, it has been found that within an organization, the chat medium is used between workers significantly more than telephone calls and face-to-face communication (Quan-Haase et al., 2005).

Cognitive load refers to the amount of mental load imposed on a human's working memory when a person attempts to accomplish a task (Chandler and Sweller, 1991). In the text chat environment, researchers have found that people were exposed to different cognitive load levels (Thirunarayanan et al., 2002).

An increased amount of new information has a significant impact on the way people behave, for example, people engaged in a low cognitive load task use a greater variety of words when speaking compared with people engaged in a high cognitive load task (Khawaja et al., 2010). The language they use changes as cognitive load increases, with people using more negative words and longer sentences (Khawaja et al., 2014). However, previous research showed that if people were given extra time (15 minutes) to chat in the text-chat environment, it built a higher level of trust between them when they chatted again later via the same medium compared with people who didn't chat for additional time (Zheng et al., 2002). This finding raises a question about what happens to the trust between people in this medium when their attention is distracted from the communication. To explore this question, we examine the effects of cognitive load on trust under two different conditions: low load and high load tasks, to find out whether the building of interpersonal trust can be affected. We also examine a novel approach, namely mouse movement measures, which are a set of indicators to track the mouse cursor, to measure the cognitive load level in this chat medium. Also, we examine the chat contents to explore different features to measure cognitive load.

The findings of this study could have implications for improving communication in the text-chat environment. We are interested in whether a high cognitive load can have a negative effect on building trust, and whether mouse movement and chat data can be used to monitor cognitive load levels between team members.

3.4 Background Literature

Previous research has demonstrated that the trust between people in the text-chat medium can be affected by giving them additional time to communicate (Zheng et al., 2002). To the best of our knowledge, there are no existing studies which investigate the effects of cognitive load on trust in CMC. However, Biros et al. (2004) examined the automated systems and trust and found that people tend to count on the automated systems when they experience a high load.

The behaviours of people vary significantly under different levels of cognitive load. For example, it was found that the length of pauses in the speech of people who were under a high cognitive load was longer (Khawaja et al., 2008), and people speak more, use more disagreement terminology and more plural pronouns under a high cognitive load (Khawaja et al., 2009, Khawaja et al., 2012). Khawaja et al. (2009) also found that the length of words, the words count and the words per sentence were increased with the high load. In relation to using mouse movements as indicators to measure cognitive load, an existing study has investigated the relationship between a user's cognitive load and their mouse activities and proposes one indicator for measuring cognitive load (pauses) as there was a strong correlation between an increased numbers of pauses in mouse activity and a high cognitive load (Arshad et al., 2013). However, in this study, we expect to find that distracting the attention of people who are communicating will hinder the building of trust between them in the text-chat medium. In addition, we expect to find that requiring people to undertake complex tasks makes them concentrate more on solving the task, resulting in less mouse movements and keyboard use. In order to determine whether there are significant differences in the mouse movements and the amount of chat in which people engage when they are communicating under high and low cognitive loads, novel measures which we have developed, such as mouse movements (distance travelled), will be used. These measures are described in detail in the method section. If these measures show significant differences, they can be used to distinguish the level of cognitive load. The hypotheses of this study are:

(H1) The establishment of trust will increase with a lower level of mental load.

(H2) An increase in mouse movements (such as distance travelled) is associated with a lower level of mental load.

(H3) The chat between people (as measured using linguistic features such as the number of words and the number of words per message) will increase with a lower level of mental load (Khawaja et al., 2009).

3.5 Method

3.5.1 Participants

Twenty participants were recruited for this study (13 males and 7 females, aged between 22 and 40). All the participants were university students and none of them had met each other prior to the task. The participants were randomly assigned to chat with their partner.

3.5.2 Procedure

We collected the data using the DayTrader task (Bos et al., 2002; Scissors et al., 2009) which requires players to communicate with each other to play an investment game. This investment game follows the rules of the Prisoner's Dilemma game. To obtain high and satisfying rewards, players must trust each other. The data collected from this game can be used to measure the extent of interpersonal trust between people, therefore, for this reason, it was chosen.

Each participant chatted and played with one other participant only. Therefore, there were ten pairs of partners in the study. The total chat time was thirty minutes in duration, divided into six sessions. The participants played the investment game and in each session, the participant and their partner chatted for five minutes about how much they would invest. At the end of each session, the participants commenced investing in the market and they were not able to chat again until they had finished making their investment. The participants had to invest five times with their partners in each of the six sessions, so the total number of rounds for investment was thirty. In each round, the participants were given $60 to invest and they could invest an amount between $0 and $60. After each round, the participants received a payoff as follows: the money invested in the market was multiplied by three and was split equally between both participants, while the money which was not invested by each participant was only multiplied by two and was calculated separately for each participant. However, after each investment round, a random amount of money of between -$3 and +$3 was given to the participants for their payoff to increase the defections and cheating between participants (Scissors et al., 2009; Zheng et al., 2002), as, by changing the amount of the payoff, the aim was to make the participants think that their partner had not entered the agreed amount of money. Also, after each investment round, the participants were not able to see their partner's payoff until the end of the game (Scissors et al., 2009).

Each participant was exposed to two cognitive load conditions, low load and high load, but only during their chats with each other. We asked the participants to sum random numbers in their heads, without using pen and paper or a calculator, and enter the total of the numbers at the end of each session. In the low cognitive load condition, the participant summed small random numbers, either 1 or 2, but in the high cognitive load condition, the participants summed large random numbers between 100 and 300. During each five-minute chat session, different numbers were shown eight times in pop-up boxes in the chat window. Each pop-up box was displayed for 15 seconds and then closed automatically, unless the participant closed it. As the participants chatted for six sessions, five pairs of partners were firstly given a low cognitive load for three chat sessions followed by a high cognitive load for three chat sessions; while the other five pairs of partners were firstly given a high cognitive load for three sessions followed by a low cognitive load.

However, the participants were told before the game that they would earn between $10 and $22 based on their performance to motivate them to take the investment game more seriously and sum the numbers correctly.

3.5.3Mouse Motion

In this study, the movements of the mouse cursor in the graphical user interface were recorded only when the participants were chatting (that is, the mouse movements were not recorded when they invested). During chatting, the participants move the mouse and perform the following tasks: 1) read all messages exchanged using the horizontal scrollbar; 2) check all investment payoffs from the sessions which have been completed using the horizontal scrollbar; 3) put the mouse cursor in the text field to write a new message; and 4) close the pop-up boxes which display the random numbers to be summed.

3.5.4Cognitive Load and Trust Measures

We used a post-questionnaire to check our approach in relation to cognitive load to make sure there is a clear difference in the level of mental load imposed on the participants. Another post-questionnaire was also used in conjunction with the investment game to measure the participant's level of trust in their partners. Each questionnaire, either the cognitive load questionnaire or the trust questionnaire, was given to each participant twice, once after the low cognitive load sessions and the other after the high cognitive load sessions. The cognitive load questionnaire comprised one question adapted from Nasa (1986): "Please rank the mental effort you had to expend while summing these numbers". The trust questionnaire comprised several questions: e.g., "I feel my partner didn't do anything to cause me to have less money than them." These questions were adapted from Butler (1991), which provides a long list of questions to measure trust.

3.5.5Mouse Movement Measures

During the chat sessions, when the mouse cursor moved, the time stamps and coordinates *(X, Y)* of the mouse cursor were recorded. For each two sequential pairs of coordinates *(X, Y)* and *(X, Y)* which constitute a line (in other words, a movement), we called these two *A (AX, AY)* and *B (BX, BY)* to carry out the calculation. We calculated a set of measures for the mouse movements. These measures are:

- Distance: The total distance travelled which are between each two sequential pairs of coordinates.

$$Distance = \sqrt{dx^2 + dy^2} = \sqrt{(AX - BX)^2 + (AY - BY)^2}$$

$$(Math\ Open\ Reference, 2015)$$

- Slope (both positive and negative slopes): The total steepness of the straight lines which are between each two sequential pairs of coordinates.

$$Slope = \frac{BY - AY}{BX - AX} \quad (Math\ Open\ Reference, 2015)$$

- Line (both horizontal and vertical lines): The total number of horizontal and vertical lines which are between each two sequential pairs of coordinates. Where these features are similar to *Slope* but BY and AY have the same value for the horizontal and lines:

$$Slope = \frac{0}{BX - AX} \quad (Math\ Open\ Reference, 2015)$$

  and for the vertical lines, both BX and AX have the same value:

$$Slope = \frac{BY - AY}{0} \quad (Math\ Open\ Reference, 2015)$$

- Movement Count: The total number of lines which are between each two sequential pairs of coordinates.
- Duration: The total length of time when the mouse cursor isn't moving.

3.5.6Chat Content Measures

Previous studies found that the total number of words, the number of words in each sentence and the length of words increased with a high cognitive load (Khawaja et al., 2009). However, we used these features but we expect to find the opposite results (as shown in our hypothesis above) because we believe that when people face a high cognitive load, they will have less time to complete tasks or engage in any other activities and consequently, they will talk less. We separated the messages of each participant from his/her partner. For each participant, we calculated the following features (some of them also came from Avrahami and Hudson (2006) where they were used for different purposes to predict the interpersonal relationships between people from the communication characteristics of chat while in this study, they were used to measure cognitive load in the text-chat environment):

- ***WN****: The total number of words written by the participant.*
- ***CN****: The total number of characters used by the participant.*
- ***MN****: The total number of messages written by each participant.*
- ***CPW****: The average number of characters in the words written by the participant.*
- ***CPM****: The average number of characters in the messages written by the participant.*
- ***WPM****: The average number of words in the messages written by the participant.*

- ***TT***: *The total number of turns of each participant.*
- ***MPT***: *The average number of messages in each of the turns of each participant.*

3.6 Results

We analysed and compared the data for each participant independently from their partner in the low load condition and the high load condition, using a dependent-sample two-tailed t-test with α=0.05.

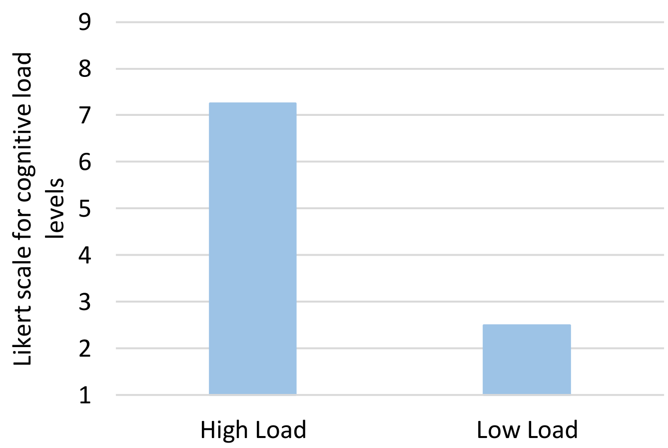


Figure 3-1: The average of ranking the cognitive load level to check the manipulation of the cognitive load levels (using a Likert scale from 1 to 9 where 1 indicates a low load and 9 a high load).

3.6.1Manipulation Results of Cognitive Load

The participants showed differences in the evaluation of the summing numbers task. The results reveal that mental load increased significantly ($t(19)=9.99$, $p<0.000$) from a mean value of 2.5 (SD=1.64) under a low cognitive load condition to a mean value of 7.25 (SD=1.41) under a high cognitive load condition (Figure 3-1).

3.6.2 Trust Results

The questionnaire results show that the level of trust increased significantly ($t(19)=2.18$, $p=0.039$) from a mean value of 19.9 (SD=6.42) under a high cognitive load condition to a mean value of 25.1 (SD=8) under a low cognitive load condition. However, in some cases, the payoff from the investment game didn't illustrate the trust between participants as shown in previous studies (e.g., Scissors et al., 2008; Scissors et al., 2009) as they didn't use payoff to measure trust in their analysis. We found this to be the case in our study as the results were not significant. The reason for this is because the rules of the game rely on high payoff and whenever the payoff is high, the trust will be high, but in fact, even those who have high trust may reap a low payoff. For example, if two participants agree to invest $40 each but one invests $40 and other invests $38, and another two participants agree to invest $20 each and they both invest $20, the first group which invested $40 and $38 will receive a higher payoff than the second group despite the existence of cheating, unlike the second group which kept their promises because of the existence of trust, this is similar to the justification in Scissors et al. (2009).

3.6.3 Mouse Movement Results

Figure 3-2 shows an example of the mouse data which was collected and stored in text file for evaluation and Figure 3-3 shows an example of the java code which was used to calculate mouse movements.

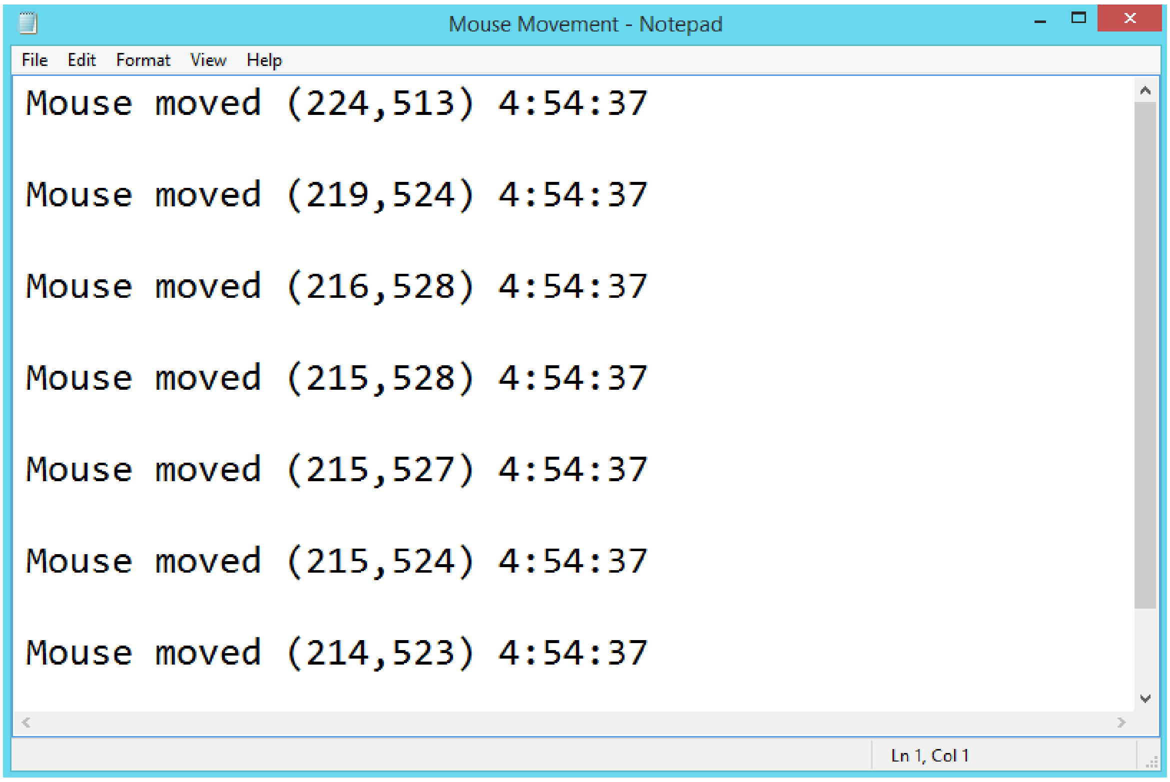


Figure 3-2: A participant sample for time stamps and (*X,Y*) coordinates recorded during chat sessions.

```
X_coordinates =AX - BX;
Y_coordinates =AY - BY;
this_distance = (int) Math.sqrt(((X_coordinates *
X_coordinates) + (Y_coordinates * Y_coordinates)));
total_distance= total_distance + this_distance;
```

Figure 3-3: An example of the Java code to calculate the total distance travelled by the mouse (Calculate the distance between two points in Java, 2013).

The distance travelled by the participants' mouse was significantly higher (p<0.0025) when the participants' mental load was low (M=25026 pixels) compared with when the participants' mental load was high (M=16954 pixels). Similarly, the total steepness of lines for positive and negative slopes increased significantly ($p<0.0025$, $p<0.0025$) when the participants' mental load was high (M=+478, M=-1058) compared to when the participants' mental load was low (M=+861, M=-1490). Also, the total number of horizontal lines and vertical lines (M=267 horizontal lines, M=266 vertical lines) under a low mental load were higher significantly than the total number of horizontal lines and vertical lines under a high mental load (M=161 horizontal lines, M=184 vertical lines) with a p value <0.01 and <0.05, respectively. In addition, the total number of mouse movements significantly increased ($p<0.025$) with low mental load (M=809 movements) more than with high mental load (M=556 movements). Finally, the total length of time that the mouse cursor stopped failed to show significant results ($p>0.05$). These results with standard deviation values (SD) and statistical values (t) are summarized in Table 3-1.

Measure	**High Load** ***Mean(SD)***	**Low Load** ***Mean(SD)***	***t***	***p***
Distance (pixels)*	16954(6549)	25026(10055)	3.93	<0.0025
Positive Slope (+)*	478(299)	861(497)	3.74	<0.0025
Negative Slope (-)*	1058(440)	1490(451)	4.32	<0.0025
Horizontal Lines (No.)*	161(95)	267(172)	3.20	<0.01
Vertical Lines (No.)*	184(82)	266(144)	2.38	<0.05
Movement Count(movements)*	556(236)	809(378)	2.89	<0.025
Duration (seconds)	655(231)	710(172)	0.90	>0.05

Table 3-1: Summary of mouse movement measures at high and low cognitive load ("*" indicates the significant features).

3.6.4Chat Content Results

Figures 3-4 and 3-5 show a sample of the chat between the same pair of participants under a low and high cognitive load during one session (5 minutes).

It can be seen that under a low cognitive load, the participants usually used more words in their messages.

Participant (B): hi
Participant (A): what is your stratagy 50%
Participant (B): yes its better. Isnt it?
Participant (A): yup lets go for it
Participant (B): awesome
Participant (A): if we go like this it will give high equile balance i like that
Participant (B): agreed
Participant (B): fairness for everyone
Participant (A): how was your day we have 5 mint to start
Participant (B): i think we have 5 mins. may be 2 mins
Participant (A): Im planning to go to the Hoyte cenima
Participant (B): cool
Participant (A): know any good movies
Participant (B): I dont know.. May be The gravity
Participant (A): thanks will go for it rating is high
Participant (A): hope to get good seats

Figure 3-4: Chat content between two participants during one session (five minutes) under a low cognitive load.

The results of the linguistic features showed that participants behave differently under both low and high cognitive load conditions. The results of the total number of words (WN), the total number of characters (CN), the total number of messages (MN) and the total number of turns (TT) were significantly higher under a low load (where the mean was as follows: WN (M=114.2), CN (M=432.35), MN (M=26.9) and TT (M=18.95)) compared with a high load (where the mean with the p value was as follows: WN (M=87.95, $p<0.05$), CN (M=330.4, $p<0.05$), MN (M=23.55, $p<0.0025$) and TT (M=16.6, $p<0.01$)). On the other hand, the results of the average number of characters in the words

(CPW), the average number of characters in the messages (CPM), the average number of words in the messages (WPM) and the average number of messages in the turns (MPT) did not show significant results (p>0.05). The results of the linguistic features are summarized in Table 3-2.

Participant (B): hi
Participant (A): Hi
Participant (B): any sugesstions?
Participant (A): what is your stratagy
Participant (B): 30$s
Participant (A): but $60
Participant (B): why?
Participant (A): maximize profit
Participant (B): ok then 60
Participant (A): okay
Participant (A): hard to count numbers
Participant (B): omg yeah
Participant (A): lost track
Participant (B): :)
Participant (A): hahahhahh
Participant (A): :)
Participant (B): we need a cal
Participant (B): :)
Participant (A): im jumping randomly
Participant (B): ha ha that is classic
Participant (A): ahhhahah
Participant (A): I feel stopped
Participant (A): :(
Participant (B): :)
Participant (B): dont give up
Participant (A): yah
Participant (B): awesome
Participant (A): we need to do thi three times
Participant (B): hell yeah..

Figure 3-5: Chat content between two participants during one session (five minutes) under a high cognitive load.

Measure	High Load *Mean(SD)*	Low Load *Mean(SD)*	*t*	*p*
WN *	87.95(51.44)	114.2(64.24)	2.34	<0.05
CN *	330.4(200.13)	432.35(240.78)	2.17	<0.05
MN *	23.55(10.28)	26.9(10.53)	3.67	<0.0025
CPW	3.7(0.29)	3.79(0.29)	0.96	>0.05
CPM	13.89(5.26)	15.49(5.25)	0.98	>0.05
WPM	3.70(1.26)	4.06(1.22)	0.97	>0.05
TT *	16.6(7.44)	18.95(7.97)	3.17	<0.01
MPT	1.47(0.32)	1.42(0.2)	0.65	>0.05

Table 3-2: Summary of linguistic measures at high and low cognitive load ("*" indicates the significant features).

3.7 Discussion

It was noted that when the participants summed large random numbers, they faced an extreme load on their working memories which was reflected directly in their attitudes and feelings toward their partners and their way of moving the mouse. The trust results reveal support for hypothesis H1 that the level of cognitive load affects the building of trust when people communicate in the chat medium. These results are consistent with another study which demonstrated that extra time spent in communication builds trust between people (Zheng et al., 2002), which is similar to what happened indirectly in the low cognitive load sessions, where the participants were in more communication with each other which led to building higher trust. However, this may mean that cognitive load may affect the length of time of the communication and the length of time of the communication may affect the building of trust. Therefore, this study shows that another independent factor (length of time of communication) may impact trust.

In relation to mouse movements, the results also provided support for hypothesis H2. The relationship between mouse movements and cognitive load was observed as there was less mouse movement under the high mental load condition compared with the low mental load condition, indicating the versatility of hand movements at the low level of mental load. The reason for this was because the participants were preoccupied by summing large numbers which required more thinking and focus, thus hindering their mouse movements. The findings in relation to mouse movements are also compatible with other studies on the effects of cognitive load on people's movements, for example, the mean stride length and velocity of people while walking was less with a high cognitive load task compared to a low cognitive load task (Martin & Bajcsy, 2011).

The measures of distance, slope and movement count varied substantially between high and low cognitive load sessions and indicate the role of the mouse in distinguishing mental load level. In the case of duration, this was not a significant factor by which to measure cognitive load, as the results showed that in both high and low cognitive load sessions, the participants moved the mouse cursor an equal amount of time, however, there was a significant difference in the speed of this movement, where the mouse cursor was moved more quickly in the low cognitive load sessions, resulting in a greater distance and a higher slope and movement count than in the high cognitive load sessions.

Examining mouse movements in relation to cognitive load is a hot research topic which requires further study. In addition to our study, another study conducted by Arshad et al. (2013) showed that there is a relationship between mouse pauses and cognitive load. In light of two studies which show there is a correlation between cognitive load levels and mouse movements, it can be said that there is reasonable evidence indicating the effectiveness of using the mouse movements to measure cognitive load.

Similar to the mouse movements findings, we found the participants chatted more when they faced a low cognitive load, as indicated by the results of the total number of words (WN), the total number of character (CN), the total number of messages (MN) and the total number of turns (TT). These findings confirmed our hypothesis (H3) that the chat between people will increase with a lower level of cognitive load. The reason for this is because with a low cognitive load (summing small numbers), the participants were less preoccupied, and consequently, they used the keyboard to write more. However, we noted that three linguistic features, the average number of characters in the words (CPW), the average number of characters in the messages (CPM) and the average number of words in the messages (WPM) was also higher with a low load but this difference was not significant, therefore, these features can be used to support the significant features in measuring levels of cognitive load.

The number of words in our study was low under a high cognitive load whereas the number of words in Khawaja et al.'s (2009) study was high under a high cognitive load. A possible reason for this may be because in our study, the participants were writing to their partners using a computer keyboard and consequently, under a high cognitive load, their body movements may be reduced (including using their fingers for writing). However, in Khawaja et al.'s (2009) study, the type of data they used (speech) was different and was recorded using microphones, and consequently maybe people talked more under stress (high cognitive load).

The results of this study have implications for the possibility of improving communication in the text-chat environment. This work demonstrates an optimal way to build trust between individuals in the chat medium by avoiding high cognitive load which has a negative effect on the process of building trust. In addition, the mouse data and chat contents can be used to develop interfaces and applications to monitor the different levels of cognitive load between team members and to support them.

3.8 Chapter Conclusion

This chapter presented encouraging evidence for how to establish interpersonal trust between people in the chat medium. As trust has already been found to be weak in this medium (Bos et al., 2002), it is possible that this study will show that a higher cognitive load will worsen the situation in relation to trust building. Moreover, based on the present findings, mouse movements and the size of the messages proved to be reliable indicators of the level of cognitive load.

Chapter 4 Trust and Cooperation

4.1 Chapter Contributions

In this chapter, we examine how different behaviours can affect trust in the text-chat environment to determine who can improve interpersonal trust via the behaviours of people. We designed two automated chat systems: one behaves cooperatively and the other behaves competitively. Thirty participants took part in this study and the results revealed that the trust between the participants who chatted with a cooperative partner was significantly higher than the trust between the participants who chatted with a competitive partner. Also, this chapter examines the associations between trust and the chat content which result from the different behaviours, the results showing that when the participants trusted their partner, they used more assent and positive emotion words. This finding emanating from the chat content can be used as an indicator of the level of interpersonal trust in the text-chat environment.

4.2 Chapter Organisation

The introduction of this chapter is presented in section 4.3. In section 4.4, we present the relevant previous work which includes the effects of behaviour on trust and the associations between trust and the chat contents. The hypotheses for this study are detailed in section 4.5. The procedure by which the data was collected and how the automated chat systems were developed to influence the participants are outlined in section 4.6. The analysis of the results showing the differences between the participants is given in section 4.7. In section 4.8, the findings are discussed with reasons for these findings. Finally, section 4.9 concludes this chapter.

4.3 Introduction

There are different types of behaviours which people can display when they interact with others to complete a particular task, for example, they can cooperate or they can compete. These diverse behaviours have different effects on the motivation, performance, attitudes and reactions of others who are also involved in these particular tasks. For example, cooperative people show kindness to their partner while competitive people show hostility to their partner (Deutsch, 2006). The behaviours of cooperative and competitive individuals have a significant impact on the degree of trust which exists with their partner. Cooperative behaviours (exchanging information and caring about one's partner's interests) build trust (Butler, 1995) but the reverse is true for competitive behaviours (when there is an incentive to encourage someone to win against their partner) which decreases trust (Harbring, 2010).

Researchers have found that the level of trust between people in the text-chat environment is weaker than the trust displayed between those who engage in face-to-face communication and is also weaker than the trust displayed between those who communicate via other CMCs, such as video or audio communication (Bos et al., 2002). Also researchers studying the use of text-chat for collaboration purposes in an organization found that workers used text-chat to exchange messages more than engaging in face-to-face and telephone communication (Quan-Haase et al., 2005). This raises the question as to whether it is possible to build interpersonal trust between individuals who use text-chats, which is a medium of communication where trust is weaker, and if so, which cues used in their communication serve as evidence of the extent of trust. To answer these questions, we examine the degree of trust between participants under two different conditions: cooperation and competition, due to their significant impact on trust. In addition, this study examines the chat content to identify differences in the attitudes and reactions of the participants in the text-chat environment when dealing with a cooperative and a competitive partner and how this relates to the establishment of trust.

4.4 Background Literature

4.4.1 Trust in Text-Based CMC

In less rich communication environments that do not transmit important cues such as facial expressions and speech tone, the establishment of trust is low compared to other environments. Previous research has demonstrated that the text-chat environment is a weaker medium of communication in relation to the establishment of trust compared with face-to-face, video and audio communications (Bos et al., 2002). It has been found that when a person sees their partner's photo, this helps them to establish trust (Zheng et al., 2002). The ability to measure trust in the text-chat environment has been demonstrated in previous research. Individuals who have a high level of trust in each other repeated each other's words whereas individuals who have a low level of trust in each other mostly repeated standard words such as okay (Scissors et al., 2008). In addition, Scissors et al. (2009) found that people who trusted each other used more leisure words and also optimism words (positive emotion words) while people who had less trust in their partners used more negative emotion words. Finally, Kalman et al. (2010) found an association between less trust and longer pauses.

4.4.2 Cooperation, Competition and Trust

The attitudes and reactions of individuals vary, depending on whether their behaviour towards each other is cooperative or competitive (Deutsch, 2006). For example, cooperative individuals are more friendly with each other, agree more with each other and there is respect between them, while competitive individuals mislead their partner and have negative and hostile attitudes to each other. In addition, Khawaja, et al. (2012) found co-operating towards a shared complex task leads to the use of more agreement than disagreement words. In relation to trust, a recent study which compared the level of trust of people under two different conditions (cooperation and competition) showed that people trusted cooperative partners more than competitive partners (de Melo et al., 2013), they examined the effects of emotions (the emotions were presented for the people by the computer program) on the trust where we found that we used similar emotions and a similar way but for the text-chat environment and we examined the chat contents. However, to the best of our knowledge, there are no existing studies which investigate the different effects of cooperative and competitive behaviour on trust in a less rich medium such as the text-chat environment. This chapter aims to investigate the possibility of enhancing trust between people who communicate in the text-chat environment using different behaviours and to identify the differences in the chat content which might be associated with building and reducing trust, resulting from the effects of different behaviours.

4.5 Hypotheses

As people cooperated more when they encountered a cooperative agent compared with an individualistic (competitive) agent (de Melo et al., 2012), we expect that this will be reflected in the participants and that they will have greater trust in a cooperative partner rather than a competitive partner. We also expect that there will be linguistic cues which can be associated with building or reducing trust, resulting from the different behaviours.

We used the four linguistic categories: positive emotion, negative emotion, assent and dissent (negations) from LIWC (Pennebaker et al., 2007), which is software to analyze text files, to find the percentage of occurrence of words under specific categories. We put the text messages of each participant in a separate text file and entered these text files in the LIWC software to calculate the percentage of the occurrence of different types of words. Examples of the categories and associated words are as follows:

- *Assent*: ***agree, ok, yes, etc.;***
- *Dissent*: ***no, not, never, etc.;***

- *Positive emotion*: ***awesome, cool, great, etc.;***
- *Negative emotion*: ***bad, hate, unhappy, etc.;***

We expect that the participants who have a high level of trust in their partners will use more assent and positive emotion words as a result of their satisfaction in the behaviour of their partners and vice versa, that is, the participants who have a low level of trust in their partners will use more dissent and negative emotion words as a result of their dissatisfaction in the behaviour of their partners. Also, we expect the time taken to reply can be associated with the level of trust where the participants who have a high level of trust in their partners will not be hesitant, therefore they will take less time to reply compared with those who have a lower level trust (Kalman et al. (2010) anticipated the relationship between longer pauses and less trust in their study while we expect to find a relationship between longer pauses and less trust resulting from the effects of different behaviours, cooperation and competition).

4.6 Method

4.6.1Participants

Thirty participants (16 males and 14 females) were recruited for this study. Fifteen participants (8 males and 7 females) were assigned to play with a cooperative partner and the same number and gender were assigned to play with a competitive partner.

4.6.2 Procedure

We developed software to allow the participants to engage with the computer (referred to as 'the partner') to play an investment game and chat (similar to an automated chat) under two conditions, cooperation and competition. The details are described below.

The participants played an investment game with their partners. This game was based on the same rules as the prisoner's dilemma game, following the approach of Kiesler et al. (1996) and de Melo et al. (2012). In the investment game, the participants are offered two different investment choices, each with a different payoff and this is repeated a number of times. In our study, we used two investment choices, properties and shares, where properties represent trust (and cooperation) and a desire to share money and shares represents selfishness and a desire to reap money (Table 4-1). The participants played 25 rounds of the investment game, choosing either properties or shares in each round. For both conditions, cooperation and competition, the partner made the same choice as the participant in the previous round (tit-for-tat), e.g. if the participant chose shares in round 7 then in round 8, the partner chose shares. However, except for the first five rounds where the partner chose properties, properties, shares, shares and then properties, the participants would find it difficult in guessing the strategy of their partners, which forces them to try to understand their partners' attitude from their text messages (de Melo et al., 2012).

Investment Profits		Your Partner	
		Properties	Shares
You	Properties	You: $7 Your partner: $7	You: $4 Your partner: $9
	Shares	You: $9 Your partner: $4	You: $5 Your partner: $5

Table 4-1: Payoff matrix for investment game.

After the participants and the partners chose an investment type, either properties or shares in each round, the participants chatted with their partners once, where the partners first sent a message and the participants replied. The participants chatted with either a cooperative or competitive partner where the partners displayed a different emotion for each action in the investment game (Table 4-2). These emotions were displayed in the form of written statements in our study, following the approach by de Melo et al. (2012) which used these emotions but in the form of facial displays for each cooperative and competitive agent where they used this approach with the prisoner's dilemma game as we did in our study. These emotions in the previous study were shown to have an impact on the participants as the participants cooperated more with the cooperative agents compared with the competitive agents. This was calculated using the prisoner's dilemma game in the previous study but this game can also be used to measure trust (Riegelsberger et al. (2003) presented previous works that use this type of the games in different ways for evaluating trust but most works did not use the binary decisions in the games and also Riegelsberger et al. (2003) provided a critique) as we have now done in this chapter. Therefore, we used the approach of de Melo et al. (2012) but for trust measurement. Also, de Melo et al. (2013) used similar emotions to affect people's trust by using competitive and cooperative emotions where trust was improved with cooperation, we found that we used a similar method but for the text-chat environment with collecting chat contents to analyse them. However, to summarize the scenario of the investment game and the chatting: 1) the participant and their partner in the investment game choose an investment type, either properties or shares; 2) the partner shows an emotional statement to the participant; 3) the participant reads the statement expressing the emotion of their partner; and 4) the participant replies to his partner in one statement to express what's on his mind.

Cooperative Condition			
	Partner		
Participant	Action	Properties	Shares
	properties	happy	embarrassed
	shares	angry	sad
Competitive Condition			
	Partner		
Participant	Action	Properties	Shares
	properties	neutral	happy
	shares	sad	sad and neutral [1]

Table 4-2: The emotional states in the cooperative vs. competitive conditions (four states for each condition) (de Melo et al., 2012).

The written emotional statements differ, according to whether they are in relation to the cooperation or competition condition. For example, in a case where a participant plays with a cooperative partner and they both cooperate (that is, they both choose properties), the cooperative partner will show a happy

[1] Sad and neutral: in each "emotion" state, we used ten different statements but in a case where both the competitive partner and the participant chose shares, we used 5 sad statements and 5 neutral statements, following a similar approach to de Melo et al. (2012).

emotional statement such as “Wow, wonderful choice for us” because both of them receive an equal amount of money; on the contrary, a competitive partner will show a happy emotional statement “Great! I earned $9!” when the competitive partner chooses shares and the participant chooses properties because the competitive partner receives more money. In the written emotional statements, we used the words “we” and “us” with the cooperative partner which aim to interest the other party and the words “I” and “me” with the competitive partner which aim to interest the interlocutor itself (see Figure 4-1 for an example from the automated chat system we developed).

The participants were not aware that their opponent was not a human to ensure they dealt honestly with their opponent. To ensure the participants believed their partners were real people, ten different emotional statements were stored for all four states in both conditions, therefore the total number of emotional statements was 80. Also, after the participant and the partner had chosen an option, the system showed the message “Please wait until your partner has finished writing”, to ensure the partner did not reveal his reply too soon. In addition, information regarding the time taken for the participant to reply and the participant’s name was displayed, similar to a real chat system (the name of the partner used in this experiment was Mike). Finally, prior to playing the game, the participants were informed that they were allowed to reply to their partner in any way they deemed fit, but they were not allowed to ask a question. To check the design of our experiment, after the game had ended, we asked the participants if they thought they had played with a cooperative person or a competitive person.

4.6.3Trust Measures

After the participants finished the investment game and chatting, we asked them to complete a questionnaire on trust to measure the participants' level of trust in their partners. These questions were adapted from Butler (1991), and has a wide range of items in the trust inventory. This questionnaire comprised eight questions (e.g. I feel my partner was usually honest; I feel that I can trust my partner) to which the participants were required to respond on a 7-point Likert scale where 1 indicates "Strongly Disagree" and 7 indicates "Strongly Agree". In addition, the dilemma game, that was used for cooperation measurement, was also used for trust measurement as we mentioned above.

4.7 Results

We measured the participants' trust with their partners by calculating how many times the participants cooperated by choosing properties. We also measured the participants' level of trust in their partners using a post-questionnaire. To analyze whether the establishment of trust is associated with linguistic cues, we used the LIWC tool to investigate several linguistic categories. We also examined how time might be associated with trust by counting the number of seconds the participant took to reply.

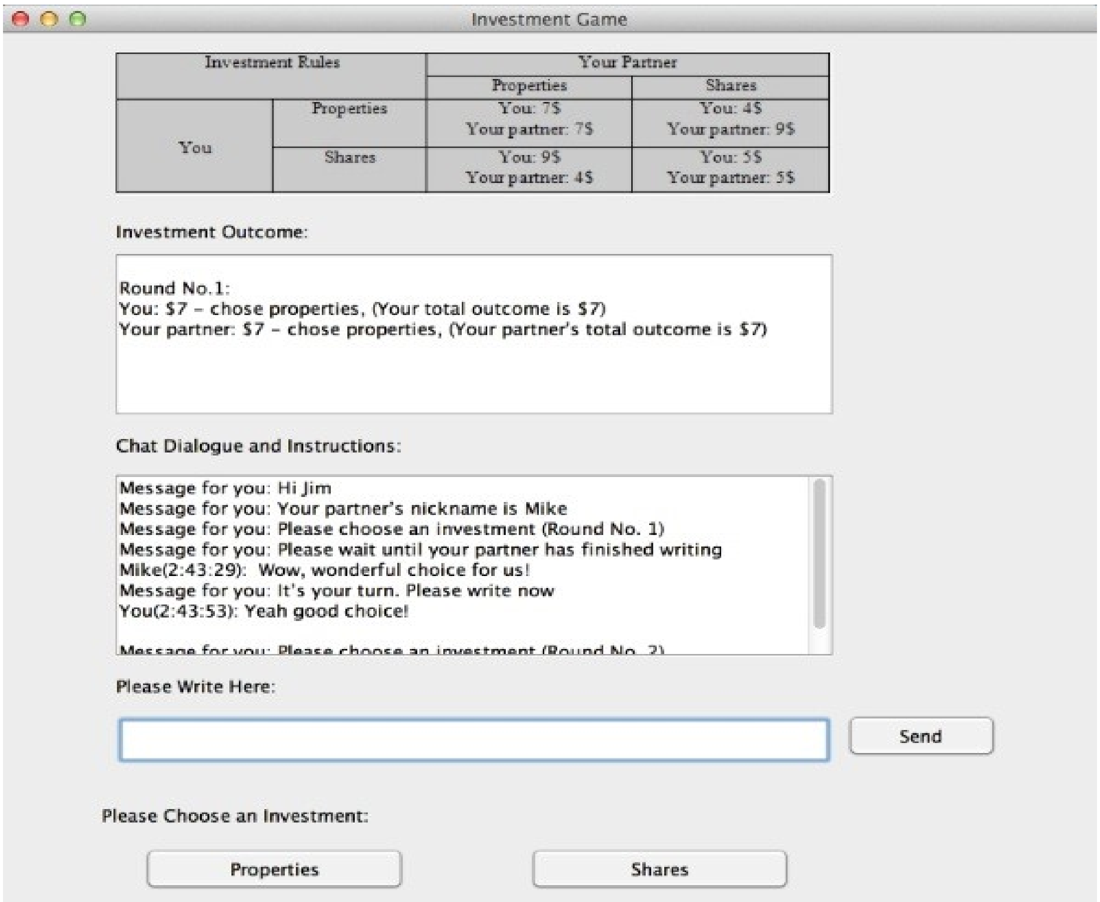


Figure 4-1: Automated chat system

Thirty participants played twenty-five rounds of the investment game and in each round, the participants reply only once, therefore 750 messages were collected from the participants. We used emotional statements to indicate the reaction of the partners. The participants used a significant number of words from the four linguistic categories to indicate whether they approved or disapproved of their partner's behaviour. For example:

- *Assent: 1)* ***Yeah****..both of us got 7million!, 2)* ***Yes*** *indeed.*
- *Dissent: 1)* ***Not*** *so good this time, 2)* ***No****, I just got 4.*
- *Positive emotion: 1)* ***Fantastic*** *keep going, 2)Yes* ***great****.*
- *Negative emotion:1) I* ***hate*** *your choice, 2) Very* ***bad****.*

At the conclusion of the game, we asked the participants if they thought they had faced a cooperative or a competitive partner, the results showing a significant difference between the two conditions ($p=0.028$, using a two-tailed t-test ($\alpha=0.05$)). This validates the design of our cooperative and competitive partners and also indicates that the participants were able to predict the type of partner they had, this being either cooperative or competitive.

We analyzed also the differences between the participants using a two-sample two-tailed t-test, $\alpha=0.05$ (see Table 4-3). The results revealed that the level of trust (from the trust game, based on choosing the properties option) increased significantly ($p=0.038$) from the competitive condition ($M=45.3\%$) to the cooperative condition ($M=60.5\%$). Similarly, the level of trust as measured by the analysis of the questionnaire responses, significantly increased from a mean value of 46.6% to 79.6% between the competitive and cooperative conditions ($p=0.000$), respectively. The occurrence of assent and positive emotion words increased significantly ($p=0.047$, $p=0.036$) from the low trust group ($M=5.9\%$, $M= 15.9\%$) to the high trust group ($M=11.8\%$, $M=22.5\%$) but the occurrence of dissent and negative emotion words did not significantly increase, at $p=0.206$ and $p=0.154$, respectively. Finally, the time the participants took to reply increased significantly from a mean of 448 seconds in the low trust group to 621 seconds in the high trust group ($p=0.010$).

Measure	Competitive Condition Mean(SD)	Cooperative Condition Mean(SD)	*p*-value
Trust Rate (from the game)	45.3%(17.2%)	60.5%(20.9%)	0.038
Trust Rate (from the questionnaire)	46.6%(18.2%)	79.6%(18.7%)	0.000
Assent	5.9%(3.9%)	11.8%(10.1%)	0.047
Dissent	3.9%(1.8%)	2.9%(2.4%)	0.206
Positive Emotion	15.9%(6.3%)	22.5%(9.7%)	0.036
Negative Emotion	7%(4.9%)	10.7%(8.3%)	0.154
Time for Replying	448 sec(163 sec)	621 sec(183 sec)	0.010

Table 4-3: Summary of measures under competitive vs. cooperative conditions.

We then divided each condition (cooperation and competition) into two levels of trust (low and high) to examine the extent of the association between the different levels of trust and the significant features (assent and positive emotion words) in each condition.

Based on their responses to the eight questions in the questionnaire, the participants were divided into two groups, either the cooperation or competition condition as follows: 1) participants who responded to the eight questions with an average of 5 or higher on the 7-point Likert scale were placed in the high trust group; and 2) participants who responded to the eight questions with an average of less than 5 on the 7-point Likert scale were placed in the low trust group.

In the cooperative condition (see Table 4-4), the results of the two-sample two-tailed t-test (α=0.05) show that the number of assent words (M=15%) and positive emotion words (M=25.7%) under a high trust level increased significantly compared with the number of assent words (M=2.9%, p=0.034) and positive emotion words (M=13.8%, p=0.031) under a low trust level. In the competitive condition, the results of the two-sample two-tailed t-test (α=0.05) show that there was no significant difference in the number of assent and positive emotion words between the two trust levels (Table 4-5).

Measure	**High Trust Mean(SD)**	**Low Trust Mean(SD)**	***p*-value**
Assent	15%(9.9%)	2.9%(2.6%)	0.034
Positive Emotion	25.7%(8.9%)	13.8%(6.5%)	0.031

Table 4-4: Summary of measures between two levels of trust (high and low) under cooperative condition.

Measure	**High Trust Mean(SD)**	**Low Trust Mean(SD)**	***p*-value**
Assent	8.7%(3.5%)	5.4%(3.9%)	0.290
Positive Emotion	16.9%(4.9%)	15.7%(6.6%)	0.819

Table 4-5: Summary of measures between two levels of trust (high and low) under competitive condition.

4.8 Discussion

As expected, the results of the questionnaire and the game showed that the trust of the participants was affected greatly by the different behaviours of their partners. The participants showed the highest trust in the cooperative partner compared with the competitive partner. This finding is encouraging evidence that a sense of cooperation optimizes the establishment of trust between people in the text-chat environment which is regarded as a weak medium in which to establish trust.

In the case of assent, Deutsch (2006) showed that one of the characteristics of cooperation is that people agree with each other's ideas and values which is compatible with our finding that the participants agreed significantly with the views of the partners in whom they had greater trust, which was a common reaction when both the participants and their partners chose the investment type of properties. In relation to dissent, the participants who had a low level of trust in their partner used slightly more rejection and disapproval words in response to the views of their partners compared to the participants who had a higher level of trust in their partner. After the chat contents were analysed, we found that both the groups of participants which had the lowest and the greatest level of trust used the dissenting word "no" in their responses significantly but the difference was that many of the responses of the participants who had greater trust were expressing consensus (e.g. It's good that no one has lost money) unlike the responses of the low trust participants which were expressing disagreement (e.g. No, bad choice). In the case of positive emotions, the participants who had a higher level of trust in their partner used signficantly more positive emotion words compared to the participants whose trust levels were low. The reason for this was because, in the cooperative condition, the properties option was chosen more frequently than the shares option, causing the partner to show happy feelings frequently, which influenced the participants to also show happiness frequently. In relation to negative emotions, we found the results were not as expected in that there were no significant differences and also we found that the participants who had a higher level of trust expressed slightly more negative emotions. The possible reason for this is that the cooperative participants were more credible and interactive when they expressed negative feelings, unlike the competitive participants where, as shown by Deutsch (2006), their communication was weak. Interestingly, we expected that the participants whose trust levels were low would be hesitant and would need more time to reply to their partner compared to the participants whose trust levels were high. However, we found the opposite to be significant where the participants who had higher trust needed more time to reply, a finding which was contrary to another study which found a correlation between longer pauses and lower levels of trust (Kalman et al., 2010). In this study, there was no chance for open discussion between participants and their partners (e.g. to become acquainted) to reduce social distance as the participants didn't chat with their real partners and their replies were restricted to the topic of the task, which is a possible explanation as to why the participants who had a higher level of trust needed more time. However, it was shown that with cooperative behaviours, both interpersonal trust and linguistic features (assent and positive words) increased.

It was observed that the cooperative condition, where the participants trusted their partners, is a stimulating environment in which to use assent and positive emotion words unlike the competitive condition, where the emotions of the participants may be conflicted toward their partners, and consequently, this may weaken the number of assent and positive words they use in their writing.

The results of this study have important implications. As this study shows that cooperative behaviours impact trust, this informs chat system designers of the benefits of designing chat systems which are able to encourage interlocutors to behave cooperatively which will have a positively effect on trust (e.g. chat systems which show a large dialogue history in the chat have a more positive impact on collaboration compared with chat systems which only show a small dialogue history (Gergle et al., 2004)). In addition, as some linguistic cues are associated with the establishment of trust as shown in this study, it is possible to develop applications to support relationship establishment between interlocutors located in different places (Scissors et al., 2008).

4.9 Chapter Conclusion

This study has provided encouraging evidence to show that trust can be established in the less rich medium of the text-chat environment, where cooperative behaviours significantly enhance trust. Moreover, it has been found that there is a correlation among cooperation, trust and particular words. Linguistic features (assent and positive emotion words) are shown to be indicators of the extent of trust between people that results from the different behaviours.

Chapter 5 The Interaction between Cognitive Load and Trust with GSR and Hesitation

5.1 Chapter Contributions

In this chapter, we examine the Galvanic Skin Response (GSR: physiological signals) and hesitation features (by monitoring the *Backspace* and *Delete* buttons on the keyboard) at overlapping conditions between interpersonal trust and cognitive load. We try to understand how these two factors together (trust and cognitive load) affect the physiological signals and hesitation and how these two factors can be measured.

5.2 Chapter Organisation

In section 5.3, we provide a general introduction to the physiological signals (GSR) and hesitation. In section 5.4, we discuss the related work and propose hypotheses which were derived from the previous work. The data collection procedure is described in Chapter 5.5. The results of our experiment and a discussion of our results are presented in Chapter 5.6. Finally, we present, in chapter 5.7, a summary of our attempts to examine the two factors together.

5.3 Introduction

Galvanic Skin Response (GSR) is a physiological signal captured easily and cost effectively via the skin (Zhou et al., 2014). These signals reflect changes in the skin's ability to conduct electricity and are used to indicate the extent of nerve response (Peuscher, 2012). 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 (Westerink et al., 2008), thus, GSR is considered as a credible physiological measure for the level of sweat in the skin.

Zhang et al. (2002) define hesitation as when "*the users answer questions in hesitant or uncertain mode*" and provide examples of hesitation such as "*incomplete answer (e.g. I saw the ?yellow part?)"*. Also, *"no hesitation means that you are not allowed to stop and say "um – er" when you cannot think what to say next"* (Peter, 2009).

The text-chat environment is a form of CMC which is low cost compared with face-to-face communication. As discussed previously, it has been found that there is a lack of interpersonal trust between the communicators in a text-chat environment compared to other CMC forms such as video (Bos et al., 2002) and also communicators may face different levels of cognitive load (Thirunarayanan et al., 2002). These findings raise questions about identifying ways to measure the impact of a combination of factors (such as interpersonal trust with cognitive load) which negatively affect communicators for the purpose of providing support to them. Our research focuses on analysing the physiological and hesitation data of communicators in the text-chat environment to see how this data can be affected by certain levels of trust and cognitive load. Specifically, in this study, we analyze GSR signals and hesitation features (by examining the use of the *Backspace* and *Delete* buttons on the keyboard) in different and overlapping levels of trust and cognitive load to examine how the signals emitted from the skin of the subjects' fingers and their degree of hesitation in writing messages can be used as an indicator of trust and cognitive load for each situation to which the subjects were exposed.

5.4 Background Literature

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 (Leichtenstern et al., 2011). Also, the ability to measure cognitive load via GSR values has been explored. Previous research demonstrates that people's GSR values increase when they are exposed to a high cognitive load which is related to stress (Shi et al., 2007). However, stress isn't only associated with cognitive load but also with trust, as demonstrated by a study which shows that people whose trust level was high had low stress, while on the contrary, people whose trust level was low had high stress (Costa et al., 2001). 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 chapter examines GSR signals with four gradients and overlapping conditions of interpersonal trust and cognitive load to determine how two factors, trust and cognitive load, influence GSR signals and also to 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).

Previous research has demonstrated that the extent of trust between communicators in the text-chat environment can be measured. For instance, when people trusted each other, they repeated chat abbreviations in their messages (Scissors et al., 2008). Cognitive load can also be measured via language. A study conducted by Khawaja et al. (2010) and found that people used a wider variety of words when they were exposed to a low cognitive load. Another study investigated the connection between trust and the speed of response. Kalman et al. (2010) found that, in the text-chat environment, an increase in the number of pauses in communication occurred when the communicator's trust in their partners was low. Arshad et al. (2013) found that there is also a connection between cognitive load and the number of pauses in mouse movements as when people faced a high cognitive load, the number of pauses increased. Also, Khawaja et al. (2008) found the pauses people's speech increased when they were exposed to a high cognitive load. In summary, these two findings show two associations: one between interpersonal trust and the speed of response and the other between cognitive load and the speed of response. Therefore, as previously stated, we examine four overlapping conditions of interpersonal trust and cognitive load and anticipate that:

Hesitation (that is the number of times the Backspace and Delete buttons are pressed) will be at its highest level when participants' trust is low and when they also experience a high cognitive load (H3). Conversely, hesitation (the number of times the Backspace and Delete buttons are pressed) will be at its lowest level when participants' trust is high and when they also experience a low cognitive load (H4).

5.5 Method

5.5.1 Participants

The GSR and hesitation data were collected from twenty-eight students from the University of NSW and from National ICT Australia (NICTA) who were recruited for this study, their ages ranging from 18 to 40 years (18 males and 10 females).

5.5.2Procedure

We examined four conditions in this chapter and in each condition, there are two independent factors together (trust and cognitive load). Specifically, two independent factors are examined in our experiment: cognitive load (low and high) and trust (low and high), hence, these two independent factors are examined together in relation to each of the four conditions.

Each participant was assigned randomly to one partner during the experiment, hence 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)
- 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 (Zheng et al., 2002). These participants were given three brainstorming tasks from (Wang et al., 2011):

- if people have an extra thumb on each hand
- if people have a third eye on the back of their head
- if people have two wings on their back

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, similar to the same idea used in (Zand, 1972):

> *"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"*

In contrast, to decrease trust, we didn't allow the participants to meet or see each other (Zheng et al., 2002) and also a different paragraph to encourage distrust between the partners was included as a preface to the instructions to the game (Zand, 1972):

> *"Caution is a normal behavior in people and it is evident when someone deals with strangers to avoid problems. Caution takes several forms, such as distrust in others. Logic always says that you cannot trust someone if you do not know them or have never dealt with them before. As you can see in the procedure of this game, you will deal with strangers"*

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.

Figure 5-1: Setup of the experiment (GSR device: ProComp Infiniti System from Thought Technology Ltd).

The GSR and the hesitation data (Figure 5-1) were collected during the DayTrader task (Bos et al., 2002, Scissors et al., 2009), where the partners were allowed to chat with each other using instant messaging while playing the investment game in the same window. The rules of this investment game were adapted from the Prisoner's Dilemma task, where the payoff resulting from the investment game was used to measure trust (trust was measured by the payoff of each participant where an increase in the payoff indicates an increase in trust and vice versa because when each participant trusts their partner, they will invest more money and this will increase their payoff.). 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?) to determine how the participants perceived each cognitive load level (NASA, 1986).

The participants were paid between \$10 and \$18 based on their performance of earning money from the investment game and summing the random numbers. However, the hesitation features which were collected were the number of times the *Backspace* and *Delete* buttons were pressed when the participants were writing messages. Also, when the participants were writing messages, thousands of GSR values were collected from each participant. The number of GSR values collected from the participants varied because some participants finished chatting earlier than others. Therefore, using this data, we calculated the average of the GSR values for each participant and the average of the peaks of the GSR values for each participant and then we examined these features statistically using t-test and ANOVA. We also examined the minimum of the GSR statically (using t-test and ANOVA), where for the minimum of the GSR, we selected only one value for each participant (the lowest value) during chat time.

5.6 Results

5.6.1Manipulation Results

The results of our approach in relation to manipulating trust and cognitive load levels were as expected. As our study includes four overlapping conditions, we used a two-way ANOVA to examine the two levels of trust and also the two levels of cognitive load. To measure trust, we calculated the payoff for each participant and then we compared them. The subjects' trust was the 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. Also, the cognitive load was the highest when the subjects summed large random numbers. In relation to the trust results, the two-way ANOVA showed that the trust rate from the payoff of each participant in the high trust condition (M=$858.55, SD=$299.91) was significantly higher than the trust rate in the low trust condition (M=$566.95, SD=$399.21, $p<0.01$). The two-way ANOVA did not show any significant interaction on the trust rate ($p>0.05$) and this means cognitive load did not significantly affect the trust rate which indicates this is a valid manipulation for trust. The result of interaction on the payoff was not significant ($p>0.05$, ANOVA) which indicates cognitive load does not affect trust. The reason for this may be due to the design of experiment where both trust and cognitive load together were manipulated, and consequently this may mitigate the impact of cognitive load on the level of trust. In relation to the results of cognitive load, we used a questionnaire adapted from Nasa (1986) to measure cognitive load (e.g., how accurately do you think you summed the numbers? with using a nine-point Likert scale ranging from accurately and not accurately). The two-way ANOVA showed that the participants were exposed to a high cognitive load significantly under the high cognitive load condition (M=30.4, SD=7.2) compared to the participants under the low cognitive load condition (M=21.1, SD=8.04, $p<0.000$). Also, the two-way ANOVA did not show significant interaction on the extent of cognitive load ($p>0.05$) which indicates that trust does not affect the extent of cognitive load and our manipulation for the levels of cognitive load was valid. The results above confirm that we manipulated trust and cognitive load as expected. From this result, it is clear that different levels of cognitive load affect the results of the survey (the survey was used to measure cognitive load) whereas the different levels of trust affect the payoff (the payoff of the investment game was used to measure trust).

5.6.2 GSR Results

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

Figure 5-2 shows that the average of the GSR values decreased in the HTLCL condition more than the other conditions, the averages of the four conditions being, from largest to smallest, LTLCL (M=2.5E-5, SD=9.8E-6), LTHCL (M=2.2E-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, the results showing 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 a significance difference. The results showed that the HTLCL averages were significantly lower than the averages of the other three conditions: the comparison between HTLCL and HTHCL was t=2.6 and p=0.02, between HTLCL and LTHCL was t=3.5 and p=0.001, and between HTLCL and LTLCL was t=4.1 and p<0.00. The results of the other comparisons between the conditions LTHCL, HTHCL and LTLCL were not significant (p>0.05).

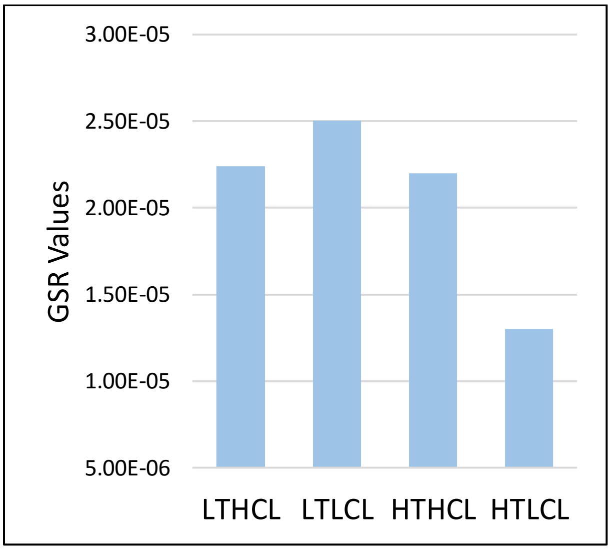


Figure 5-2: The average of the GSR values.

The minimum of the GSR values was also calculated for each participant, the results are as follows (Figure 5-3): HTHCL (M=1.6E-5, SD=9.6E-5), HTLCL (M=9.4E-6, SD=4.9E-6), LTHCL (M=1.8E-5, SD=7.8E-6) and LTLCL (M=2E-5, SD=8.9E-6). To examine the interaction effect between cognitive load and trust on the minimum of the GSR values, we used a two-tailed two-way ANOVA and the results were not significant (F=3.04, p=0.087).

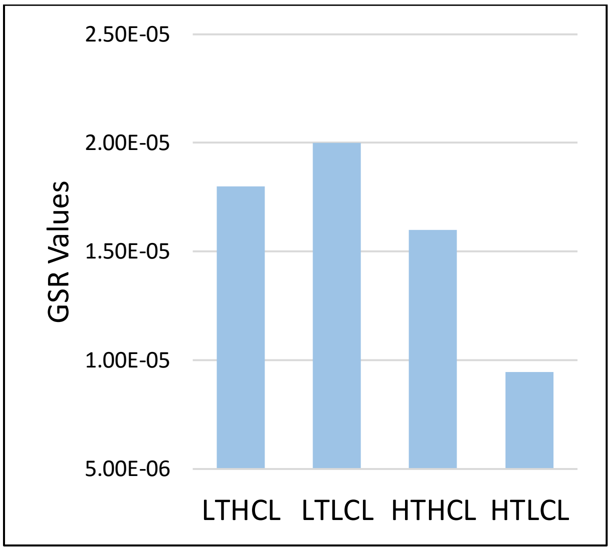


Figure 5-3: The average of the minimum of the GSR values.

The average of the peaks of the GSR was also examined and the results were the same as the average of the GSR (where the HTLCL condition was less than other three conditions). Figure 5-4 illustrates the average of the peaks for each participant where the results were as follows: LTLCL (M=2.5E-5, SD=9.8E-6), LTHCL (M=2.2E-5, SD=8.5E-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 average of the peaks, the results showing significant interaction between them (F=5.4, p=0.02). A post-hoc two-tailed t-test was also used to compare the average of the peaks. Six comparisons were made and, similar to the results for the average of GSR, three showed significant differences. The average of the HTLCL peaks were significantly lower than the average of the peaks of the other three conditions: the comparison between HTLCL and HTHCL was t=2.6 and p=0.02, between HTLCL and LTHCL was t=3.5 and p=0.002, and between

HTLCL and LTLCL was t=4.1 and p<0.00. There is no a significant difference when we compared between these conditions LTHCL, HTHCL and LTLCL (p>0.05).

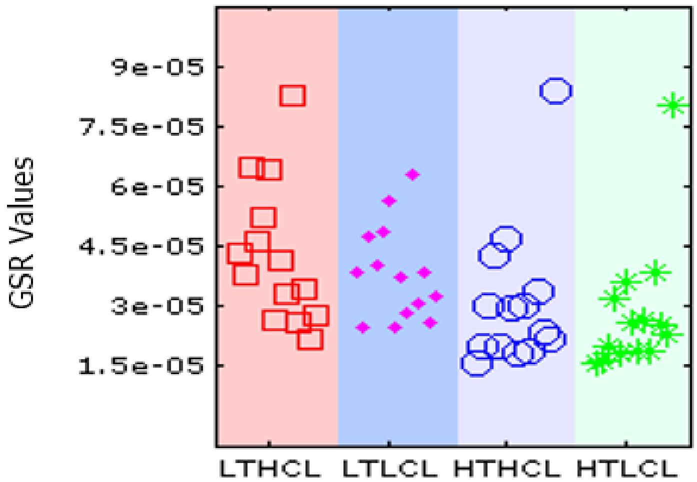


Figure 5-4: The average of the peaks for each participant (each symbol represents one participant).

5.6.3Hesitation Results

The number of times the *Backspace* and *Delete* buttons were pressed was examined. Figure 5-5 shows the average number of times that the *Delete* button was pressed in the four conditions are as follows: HTHCL (M=0.29, SD=0.83), HTLCL (M=0.78, SD= 1.85), LTHCL (M=0.93, SD=2.02) and LTLCL (M=0.71, SD=1.89). We used a two-tailed two-way ANOVA to examine the interaction effect between trust and cognitive load on the number of times the *Delete* button was used and the result was not significant (F=0.061, p>0.05).

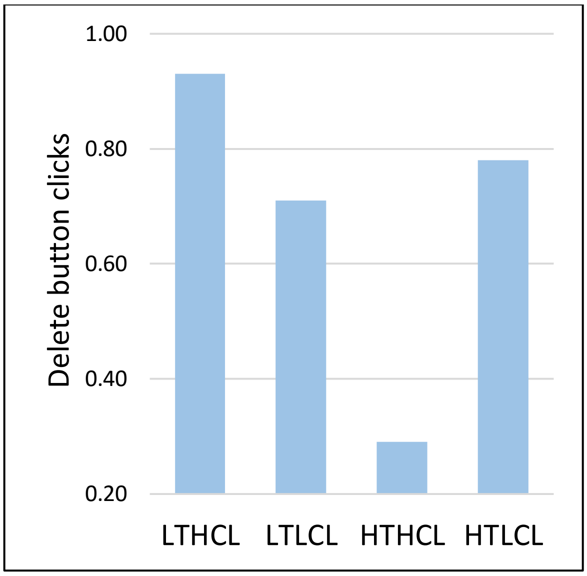


Figure 5-5: The average of the number of times the Delete button was used (as shown, none of these conditions reach average one time to click Delete button)

In relation to the *Backspace* feature (Figure 5-6), the results of the number of times the *Backspace* button was used is as follows: HTHCL (M=61, SD=34.91), HTLCL (M=63.35, SD= 31.03), LTHCL (M=72.2, SD=18.56) and LTLCL (M=75.79, SD=14.54). A two-tailed two-way ANOVA was used and did not show any interaction effect between trust and cognitive load on the number of times the *Backspace* button was used (F=0.008, p=0.93). We noted with cognitive load, the average is not stable in the four conditions (increased and also decreased) unlike the stable results for trust where each low trust condition (LTHCL or LTLCL) had a higher average than each high trust condition (HTHCL or HTLCL), as shown in the results above. A comparison was made of the number of times the *Backspace* button was used in the low trust condition (LTHCL (+) LTLCL) with the number of times the *Backspace* button was used in the high trust condition (HTHCL (+) HTLCL). The results were as follows: the average of low trust was 74 (SD=16.51) and the average of high trust was 62.18 (SD=32.43) with a p value of 0.09, which suggests a near significant result.

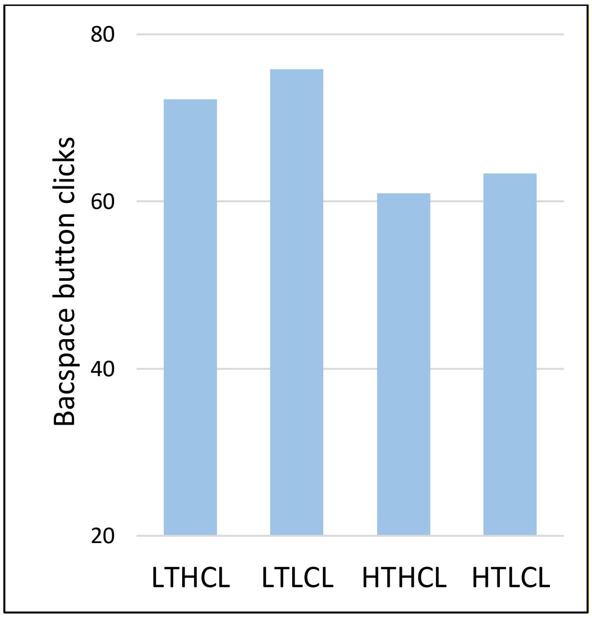


Figure 5-6: The average of the number of times the Backspace button was used.

5.7 Discussion

In this study, we examined the following GSR features: the average of the GSR values, the average of the peaks of the GSR values and the minimum of the GSR values. The results of the average of the GSR values and the average of the peaks of the GSR values were significantly lower when trust was high and cognitive load was low. This study has provided promising findings in relation to the indicators to determine the level of trust and cognitive load. Specifically, the results of the 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 the average of the GSR values and the average of the peaks of the GSR values in 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. However, the design of the experiment in Chapter 3 is different from the design of the experiment in this chapter. Chapter 3 examined the effects of cognitive load on trust whereas this chapter examines the effect of cognitive load and trust together on the GSR data. The GSR data was affected by both trust and cognitive load together when cognitive load was low and trust was high (HTLCL condition) causing the GSR data to be significantly low whereas the GSR data did not reach this low level in the other conditions (LTHCL, LTLCL and HTHCL).

In the case of hesitation features, the data do not support the hypothesis (H3) which states the number of times the *Backspace* and *Delete* buttons are used will be at its highest level when participants' trust is low with a high cognitive load. In relation to the *Delete* feature, after examining the data, we found the majority of participants (75%) did not use this button and the number of times the *Delete* button was used was low, being between one and seven times. This is a possible reason for the weakness of the *Delete* feature in relation to distinguishing between the overlapping levels of trust and cognitive load. In relation to the *Backspace* feature, after examining the data, we found a conflict in the results for the number of times the *Backspace* button was used in the different levels of cognitive load. It was found that when the cognitive load was high, there was a high use of the *Backspace* button in some cases, and also when the cognitive load was high in some other cases, there was a low use of the *Backspace* button. Specifically, the number of times the *Backspace* button was used in the LTHCL condition (high load) was higher compared to its use in the HTLCL condition (low load), whereas the opposite occurred for HTHCL condition (high load) and LTLCL condition (low load) as the number of times the *Backspace* was used was lower in the HTHCL condition compared to the LTLCL condition.

These findings show that with a high cognitive load, the number of times the *Backspace* button was used increased and with a low cognitive load, it decreased. A possible reason for this is that with a high cognitive load, participants may make mistakes when writing messages due to the difficulty of the task and consequently, they need to use the *Backspace* button more frequently to correct their messages. Likewise, participants with a low cognitive load are more comfortable when writing messages and have more time to interact with their partners because of the small numbers they are required to sum and consequently, this may result in participants using more keyboard buttons (including the *Backspace* button). This might be the reason why we found a weakness in the results for the *Backspace* feature to distinguish the overlapping levels of trust and cognitive load. However, similarly, we found the results do not support H4 (that the number of times the *Backspace* and *Delete* buttons are clicked will be at its lowest level when the trust of participants is high with a low cognitive load).

However, the results of the GSR signals have implications which can be used. For instance, through the development of a 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 are exposed in real time and provide them with suitable assistance.

5.8 Chapter Conclusion

This chapter examined the effects of overlapping conditions between interpersonal trust and cognitive load on GSR signals and hesitation features in the text-chat environment. The findings show that GSR signals were affected by interpersonal trust and cognitive load which is encouraging evidence to use physiological data as indicators for the negative factors which may arise between communicators in the text-chat environment, such as a high cognitive load.

Chapter 6 Shared Visual Information and Trust

6.1 Chapter Contributions

In this chapter, we examine the interpersonal trust of communicators in the text-chat environment when shared visual information is provided. We compare the level of trust of the participants when they chatted with and without shared visual information. The findings show that shared visual information plays a positive role in improving the level of interpersonal trust between communicators in the text-chat environment.

6.2 Chapter Organisation

Sections 6.3 and 6.4 provide a summary of the use of shared visual information from previous studies and why this feature can be useful in building trust in the text-chat environment. Sections 6.5 and 6.6 describe the method used to collect the data from the participants and the hypotheses of this study, respectively. Section 6.7 presents the results of this study and Section 6.8 discusses the results, including a justification of these results. Section 6.9 presents a summary of our findings which includes a discussion of the role of shared visual information.

6.3 Introduction

When performing tasks in the text-chat environment, it is possible to allow the communicators to see the behaviors and actions of the other person in real time. Kraut et al. (2002) called this technique shared visual information and they described it as follows: "*shared visual information is important for maintaining an awareness of the current state of the collaborative task in relation to an end goal. This awareness helps a pair plan how to proceed towards the goal, what instructions need be given, and how to repair incorrect actions. Shared visual information provides the ability to monitor specific actions*". Shared visual information was successfully incorporated into the text-chat environment at NORAD, so that employees could check aircraft distribution in airspace and communicate via text chats with airport staff about this (Gergle, et al., 2004).

In our study, shared visual information is examined in the text-chat environment to determine its impact on people's trust when they are able to watch the behaviors of their partners. Trust will be measured in two contexts in this chapter: the level of the participants' trust in their partners when performing tasks when they can see their partners' behaviors and actions in real time; and the level of the participants' trust in their partners when performing tasks when they can't seeing their partners' behaviors and actions in real time however, their partners have access to this feature. In the text-chat medium, which is considered weak in the area of building trust (Bos et al., 2002), we believe that providing shared visual information can enhance trust between people and also that trust can be built even if only one of the communicators has access to shared visual information.

6.4 Background Literature

Robert (2016) examined trust based on monitoring virtual teams and the results were as follows: when internal monitoring was high and performance was high, trust was affective, and when external monitoring was high and performance was low, trust was affective. In addition, the degree of trust that people had in the information from the agents was examined and it was found that people trusted information from agents when the body of the agent has a physical form (Takeuchi et al., 2011). Also, Wang (2010) examined different levels of automation using a human-telerobot system in a firefighter experiment and found that when the level of automation was high with less visual information, the level of trust decreases.

Previous research (Kraut & Gergle 2003; Gergle, Kraut & Fussel 2004; Gergle et al 2004; Gergle, Kraut & Fussel 2006; Gergle, Rose & Kraut 2007) shows the importance of shared visual information which can be used in the text-chat medium to enable tasks to be completed more easily and to facilitate more efficient communication. This existing research has also shown that shared visual information has many benefits, such as reducing the amount of time needed to complete tasks, reducing miscommunication and reducing misunderstanding. For instance, Gergle et al. (2004) designed an experiment which involved the Puzzle game and helper and worker roles. The helper (who describes the puzzle) , the worker (who solves the puzzle). Gergle et al. (2004) found that the participants completed the puzzle game faster when the helper was able to see the worker's board and the worker's actions immediately and when the helper chatted with the worker compared to when the helper could not see the worker's board or actions.

It is possible to affect trust in different ways, as shown in previous studies. For example, it has been found that negative emotions negatively affect the establishment of trust (Myers et al., 2011). In addition, other factors, such as transparency, can affect trust. For example, it was found that the followers' trust in the leaders improves when communication is transparent between leaders and followers (Norman et al., 2010).

We believe that using shared visual information (which can inform the communicators of the current situation in real time) can increase transparency when information is being transferred from one person to another and communication is taking place. This transparency is associated with the building of trust, as described above and consequently, we believe that shared visual information also can increase the degree of trust between people when they communicate online. Therefore, shared visual information in the text-chat medium needs further investigation to examine its correlation with interpersonal trust.

However, in this study, we used another factor (the level of familiarity between the partners) in addition to the feature of shared visual information as it has been found that when familiarity is developed between people (for example, talking face-to-face before chatting in the text-chat environment), this enhances the establishment of trust in the text-chat environment (Zheng et al., 2002). The reason for including the level of familiarity in this study is to understand the extent of the impact of shared visual information under different conditions. The participants in this study were exposed to four overlapping condition: 1) with shared visual information and familiarity between the participants; 2) with shared visual information and no familiarity between the participants; 3) without shared visual information and familiarity between the participants; and 4) without shared visual information and no familiarity between the participants. We examined the level of trust in these four conditions to determine the role that shared visual information plays in affecting the relationship between people in the text-chat environment.

6.5 Method

6.5.1Participants

Fifty-six participants were recruited for this study from a commercial market for mobile phones. These participants were retailers and technicians in the mobile phone industry and customers (students and employees in government and private institutions). Their ages ranged between 18 and 36 years.

6.5.2Procedure

In this study, the task which the participants are required to complete includes two parts, text chatting and solving a puzzle, which has been used previously (Gergle, et al., 2004). We followed the procedures used in the previous study, which are described as follows: two participants (the *Helper* and the *Worker*) were paired to complete the task. The *Helper* is able to see the completed puzzle and also can view the *Worker's* staging area in real time but the *Worker* is not able to see any real time information on the Helper. The *Helper* must communicate with the *Worker* via the text-chat environment to guide them to complete the target puzzle by correctly arranging the pieces of the puzzle. The previous study compared the completion of the tasks between the *Helper* and the *Worker* when the *Helper* was able to see the *Worker's* staging area in real time (using shared visual information feature) with the completion of the tasks between the *Helper* and the *Worker* when the *Helper* was unable to see the *Worker's* staging area (in this case, the *Helper* attempts to assist the *Worker* to complete the puzzle without being able to see the *Worker's* progress, in other words, without shared visual information). Whereas the previous study was conducted to determine the effect of shared visual information on completing tasks, we used this strategy to examine the interpersonal trust between people when using shared visual information.

The fifty-six participants were divided into two groups to solve the puzzle in our study: one group (twenty-eight participants) was given the role of *Sender* and the other group (twenty-eight participants) was given the role of *Recipient* (a total of twenty-eight pairs). Similar to the *Helper* in the previous study, the *Sender* in this study is able to see the completed puzzle and also can view the *Recipient's* actions in real time, and similar to the *Worker* in the previous study, the *Recipient* receives the information from the *Sender* and uses this to correctly arrange the puzzle pieces. The *Sender* must convey information to the *Recipient* using the text-chat environment so that the task can be completed, and the *Recipient* must follow the instructions of the *Sender* and is allowed to ask questions in the text-chat environment to clarify the instructions.

Each pair must assemble two jigsaw puzzles in a time period of ten minutes per puzzle. One jigsaw puzzle contains a photo of a young boy wearing traditional cultural clothing (Figure 6-1) and the other jigsaw puzzle contains three photos of a pink galah (Figure 6-2). Fourteen pairs commenced the task using the jigsaw of the boy and the other fourteen pairs started with the jigsaw of the cockatoo. Each jigsaw comprised thirty-five pieces and was made using the jigsaw puzzle creator software, Astra Gift Maker (Astra Gift Maker, 2010).

Fourteen pairs had access to shared visual information, where the *Sender* was able to see the *Recipient* move their pieces in real time but the *Recipient* was only aware of this and their task was to listen to the *Sender's* instructions in order to complete the puzzle (so, the *Sender* benefits directly from the feature of shared visual information while the *Recipient* benefits indirectly from the feature of shared visual information). For the other fourteen pairs, the participant had no access to shared visual information.

(a)

(b)

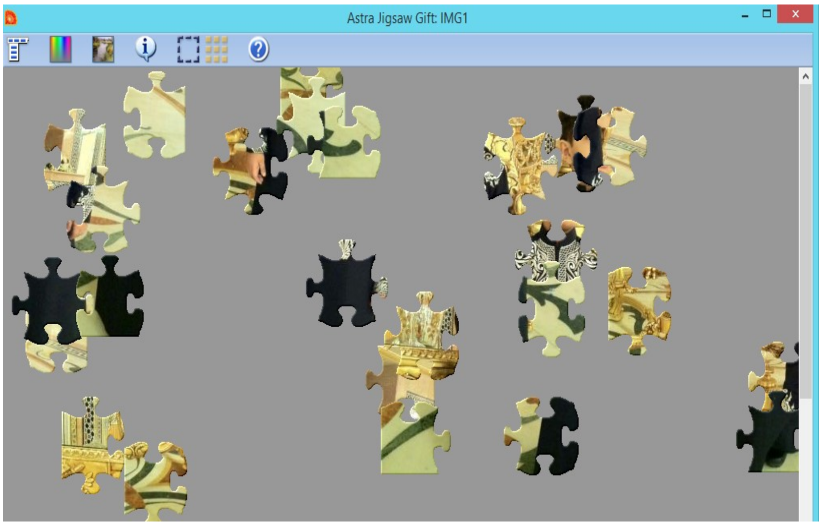


Figure 6-1: (a) Photo of a boy wearing traditional cultural clothing which was turned into a 35-piece jigsaw puzzle using Astra Gift Maker; and (b) The pieces of the jigsaw puzzle which the Recipient must arrange to complete the task.

6.5.2.1 The Familiarity between the Participants

For all twenty-eight pairs, we added the factor of familiarity where a previous study shows that trust between communicators in the text-chat environment is improved if they meet face-to-face before they start chatting (Zheng et al., 2002). Hence, we asked each pair of participants to meet face-to-face for ten

(a)

(b)

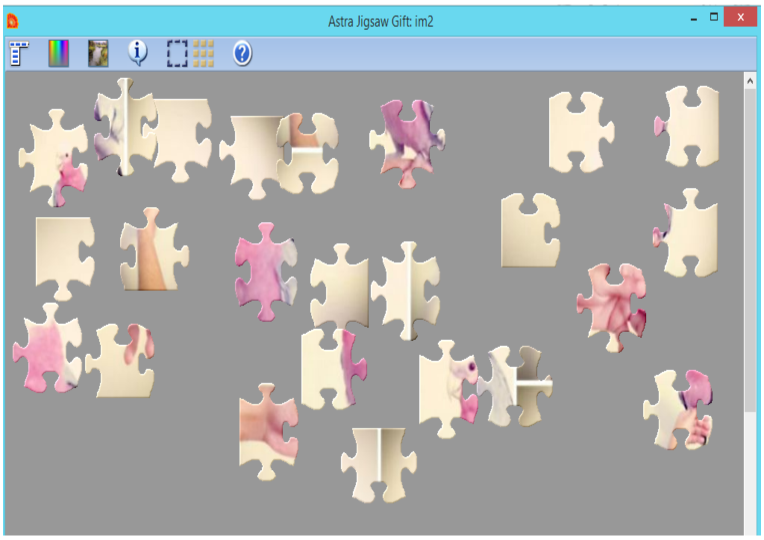


Figure 6-2: (a) Jigsaw of three photos of a pink galah which was turned into a 35-piece jigsaw puzzle using Astra Gift Maker; and (b) The pieces of the jigsaw puzzle which the Recipient must arrange to complete the task.

minutes before they start chatting online (Zheng et al., 2002) to orally brainstorm the advantages and disadvantages of the following (Wang et al., 2011):

- if people have an extra thumb on each hand

- if people have a third eye on the back of their head
- if people have two wings on their back

Each of the twenty-eight pairs was tasked with completing two jigsaw puzzles, one task conducted with familiarity between the participants and the other without familiarity between the participants as follows: before beginning the task (chatting and solving the puzzle) each pair of participants met face-to-face and were told that they would work with each other to complete the first puzzle but they would complete the second puzzle with a different partner, however, this was not true and in actual fact, the pair of participants completed both tasks with each other. To make sure the participants didn't discover this, they had to adhere to the following rules: 1) they were not allowed to use their names when chatting in the text-chat environment; 2) they were not allowed to chat outside the scope of the task; 3) they were not allowed to see previous chat messages; and 4) they were not allowed to chat about the first jigsaw puzzle task during the completion of the second jigsaw puzzle task. Half of the pairs were told that they would chat with the partner they had met face-to-face before the first puzzle task while the other half of the pairs were told they would chat with the partner they had met face-to-face before the second puzzle task.

Table 6-1 summarizes the four conditions in our study (that is, the overlap between shared visual information and familiarity between participants):

Conditions	Familiarity	Shared Visual Information
F_S	√	√
WF_S	Χ	√
F_WS	√	Χ
WF_WS	Χ	Χ

Table 6-1: shows the four conditions to which the participants were exposed. F indicates with familiarity, S indicates with shared visual information, WF indicates without familiarity and WS indicates without shared visual information.

6.5.2.2 The medium used in this study

For this study, Gmail accounts are used, as they allow people to chat via text messages and also, this service is provided by the well-known and reputable company, Google. In our study, when the participants were chatting in the text-chat environment, they were in different rooms and were alone. We used the Share Screen feature offered by Gmail (Figure 6-3) to provide the participants with the feature of shared visual information.

6.5.2.3 Measuring Trust

To measure trust, we used a questionnaire with a 9-point Likert scale where 1 indicates "*Strongly Disagree*" and 9 indicates "*Strongly Agree*". The questions in this questionnaire were adapted from Butler (1991) and measured the extent of the trust of the participants in their partners (e.g, "*I feel that my partner dealt honestly with me?*" and "*If I play again, I would prefer to play with the same partner?*"). The same questionnaire was given to *Senders* and

Recipients twice, after the first puzzle had been completed and again after the second puzzle had been completed.

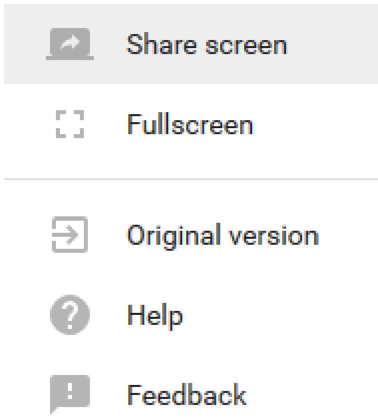


Figure 6-3: a list of the features in Gmail accounts including the Share Screen feature which allows a desktop to be seen by other people (taken from Gmail accounts).

6.6 Hypotheses

The hypotheses of this experiment are:

(H1) The participants who have access to shared visual information will have a higher level of trust than the participants who do not have access to shared visual information.

(H2) Participants will have the highest level of trust when they are familiar with each other and have access to shared visual information.

(H3) The Senders will have a higher level of trust than the Recipients under the shared visual information condition because the Sender is able to see the Recipient's actions in real time whereas the Recipient only knows the Sender is able to watch his actions (that is, the Sender benefits directly from shared visual information while the Recipient benefits indirectly).

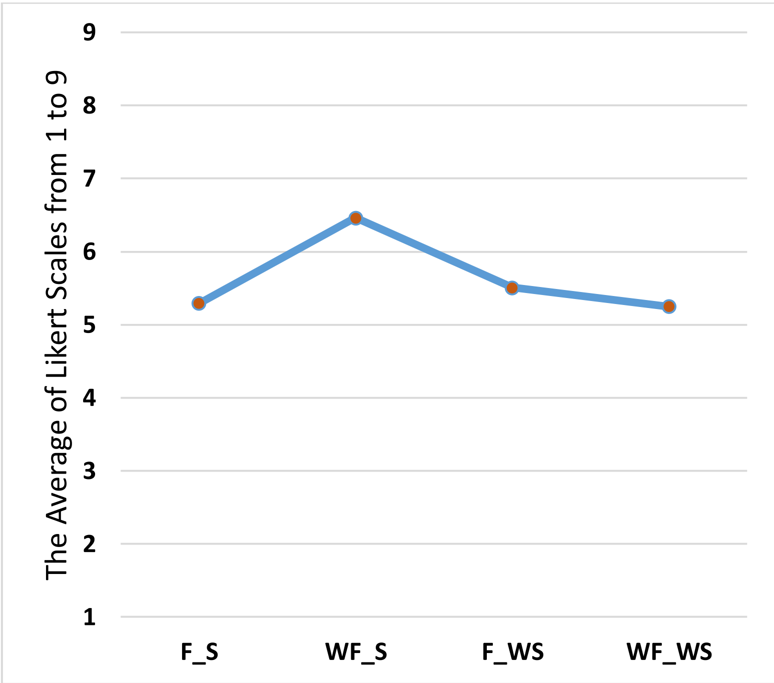


Figure 6-4: The average of the responses to the Likert scales for the Recipients under the four conditions: 1) *F_S* (familiarity + shared Visual Information), 2) *WF_S* (without familiarity + shared Visual Information), 3) *F_WS* (familiarity + without shared Visual Information) and 4) *WF_WS* (without familiarity + without shared Visual Information).

6.7 Results

6.7.1Recipient Results

We analyzed the results of the questionnaire on the trust of the *Recipients* and used a two-tailed two-way ANOVA to examine the interaction effect between the four overlapping conditions and to also compare the two samples (with and without the shared visual information, regardless of the familiarity factor). The results of the questionnaire for the interaction effect between the four conditions was not significant ($p>0.05$). The mean and standard deviation values were as follows: F_S ($M=5.29$, $SD=1.84$), WF_S ($M=6.46$, $SD=1.88$), F_WS ($M=5.51$, $SD=1.85$) and WF_WS ($M=5.25$, $SD=1.19$), see Figure 6-4.

In addition, the results of the comparison of the results of the two samples (between with the shared visual information condition and without the shared visual information condition) was not significant ($p>0.05$). The mean and standard deviation values were as follows: with the shared visual information condition ($M=5.88$, $SD=1.92$) and without the shared visual information condition ($M=5.38$, $SD=1.53$), see Figure 6-5.

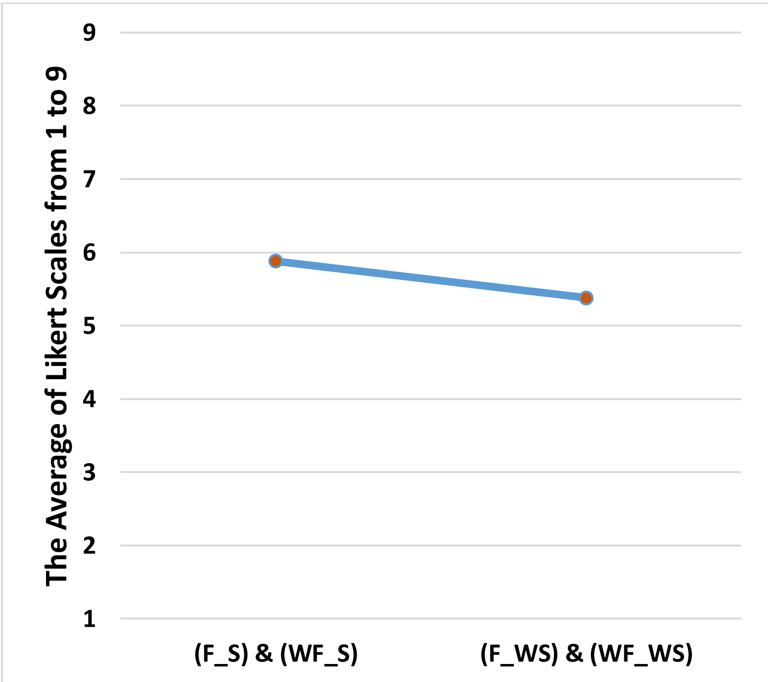


Figure 6-5: The average of the responses to the Likert scales for the Recipients under the two conditions (1) *(F_S)&(WF_S)*: with shared visual information and 2) *(F_WS)&(WF_WS)*: without shared visual information, regardless of the familiarity factor).

6.7.2 Sender Results

The two-tailed two-way ANOVA was used to examine the interaction effect between the four overlapping conditions on the trust of the *Senders* and also to compare the two samples (with and without the shared visual information, regardless of the familiarity factor). The results for the interaction effect were significant between the four conditions ($p<0.05$). The mean and standard deviation values were as follows: F_S ($M=7.88$, $SD=0.98$), WF_S ($M=6.84$, $SD=1.28$), F_WS ($M=4.26$, $SD=1.98$) and WF_WS ($M=5.07$, $SD=1.12$), see Figure 6-6. After finding significant results between the four conditions, a post-hoc two-tailed t-test was performed between the four conditions, finding that trust in the F_S condition was significantly higher than the other three conditions WF_S, F_WS and WF_W, $p<0.05$.

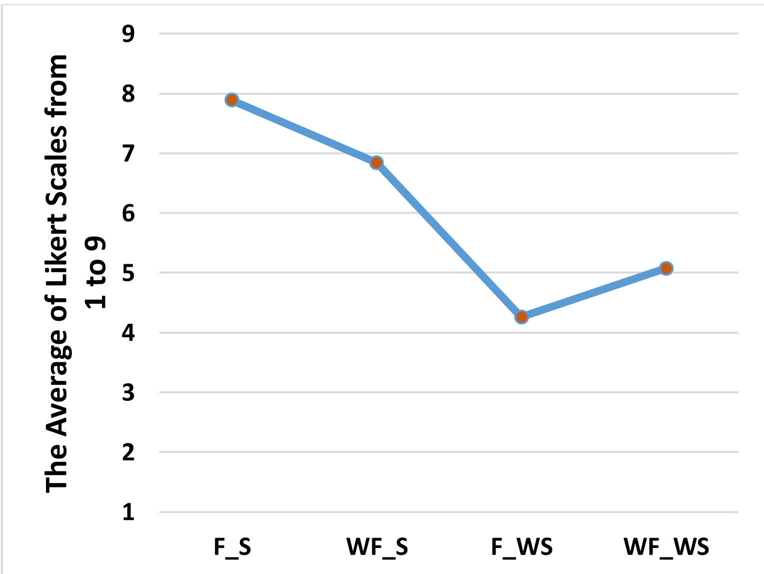


Figure 6-6: The average of the responses to the Likert scales for the Senders under the four conditions, showing the interaction effect between the four conditions was significant ($p<0.05$): 1) *F_S* (familiarity + shared Visual Information), 2) *WF_S* (without familiarity + shared Visual Information), 3) *F_WS* (familiarity + without shared Visual Information) and 4) *WF_WS* (without familiarity + without shared Visual Information).

However, a comparison of the results of the questionnaire on the trust between the two samples (with and without the shared visual information) showed that the trust of the sample with shared visual information ($M=7.36$, SD=1.24) was significantly higher than the sample without shared visual information (M=4.67, SD=1.63, $p<0.05$), see Figure 6-7.

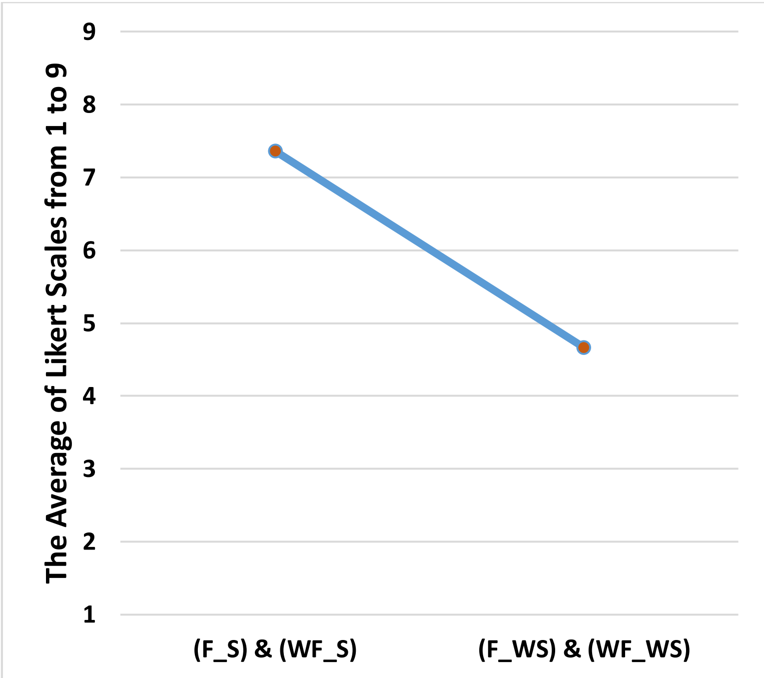


Figure 6-7: The average of the responses to the Likert scales for Senders under the two conditions (1) *(F_S)&(WF_S)*: with shared visual information and 2) *(F_WS)&(WF_WS)*: without shared visual information, regardless of the familiarity factor), showing a significant difference (p<0.05).

6.7.3 Within Shared Visual Information Tasks

We compared the level of trust of both the *Senders* and *Recipients* under the shared visual information condition (that is, where the *Sender* can see the *Recipient's* actions in real time when completing the puzzle, whereas the *Recipient* only knows that the Sender can see their actions, so, the *Recipient* benefits indirectly from the shared visual information) to determine whether the *Senders* or the *Recipients* had the higher level of trust. The results of the shared visual information tasks show that the level of trust of the *Senders* (M=7.36, SD=1.24) was significantly higher than the level of trust of the *Recipients* (M=5.88, SD=1.92, p<0.05), see Figure 6-8.

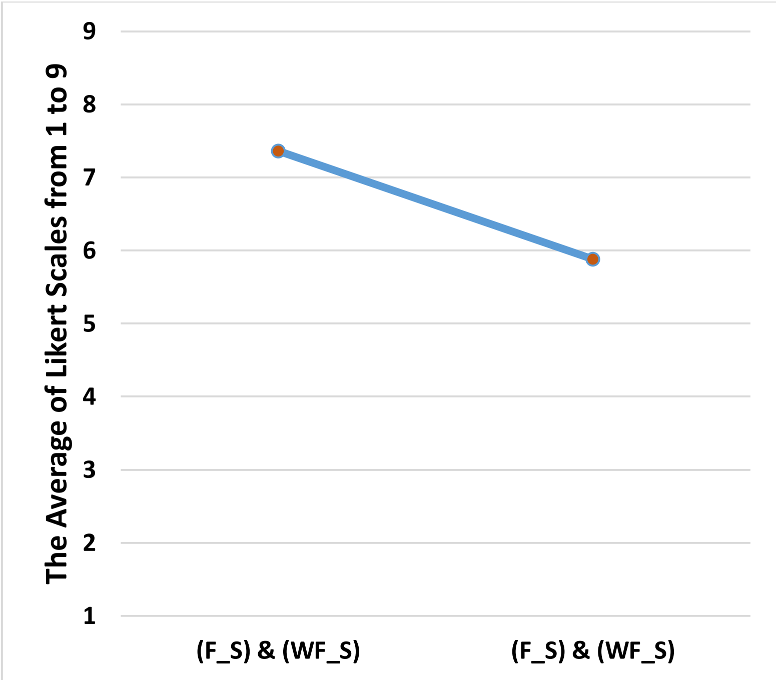


Figure 6-8: The average of the responses to the Likert scales for the Senders (shown on the left) and the Recipients (shown on the right) under the shared visual information condition only, showing a significant difference (p<0.05).

6.8 Discussion

In this study, we examined the importance of shared visual information in relation to building trust. For the *Recipients*, the shared visual information with or without the familiarity factor did not assist the establishment of trust in the text-chat environment. This finding does not support the first hypothesis (H1, shared visual information will increase trust) nor the second hypothesis (H2, participants will have the highest level of trust when they are familiar with each other and have access to shared visual information). One possible reason for this is due to the fact that the *Recipient*, under the shared visual information condition, benefits indirectly from the feature of shared visual information unlike the *Sender*, resulting in the *Recipients* having a low level of trust and consequently, there was no difference between their levels of trust in the different conditions (with and without shared visual information).

However, the results for the *Senders* support the first hypothesis (H1, trust will improve with shared visual information). We think the reason for this is that increasing transparency by allowing the participants to have access to shared visual information increases trust, as it has been found that transparency has a positive effect on enhancing trust (Norman et al., 2010). In addition, the findings support the second hypothesis (H2, participants will have the highest level of trust when they are familiar with each other and have access to shared visual information) as the level of trust of the *Senders* was highest in this condition. A possible reason for the higher trust in this condition may be because there were two significant factors together: the familiarity factor which helped to build trust before they chatted in the text-chat environment (where the meeting face-to-face improved the trust in the text-chat environment later as shown in a previous study (Zheng et al., 2002)) as well as the shared visual information factor which increased transparency between communicators and consequently increased trust. Therefore, these factors may be complementary to each other in relation to trust building.

A comparison of the results of the *Senders* and *Recipients* in the shared visual information condition shows that the *Senders* had a higher level of trust than the *Recipients*, which supports the third hypothesis (H3, *Senders* will have a higher level of trust than *Recipients*). This finding shows that those who benefit indirectly from shared visual information (such as the *Recipient*) have a lower level of trust compared to those who benefit directly (such as *Sender*).

This study has implications that can be used in the text-chat environment. The developers of text-chat systems can design systems which include shared visual information for specific work domains, for example, a similar feature to shared visual information was used in NORAD so that employees could watch the distribution of aircraft in airspace and use a text chat to discuss this with airport staff (Gergle, et al., 2004). Implementing the technique of shared visual information can enhance the relationship between communicators in the text-chat environment.

6.9 Chapter Conclusion

In this experiment, we examined the effects of shared visual information on enhancing trust between communicators in the text-chat environment. The results suggested that shared visual information plays a positive role in the establishment of trust. In addition, this study found that trust was high when the communicators had developed a degree of familiarity with each other and shared visual information was provided. However, the results showed that when the communicators received an indirect benefit from shared visual information in the text-chat environment (as with the *Recipients* in this study), it has no effect on building trust.

Chapter 7 Cognitive Load and Trust Classification

7.1 Chapter Contributions

In this chapter, we discuss attempts to use the data from the text-chat environment to predict the level of interpersonal trust and cognitive load that people feel and face using Machine Learning (ML) algorithms (such as Naïve Bayesian). Our attempts to use (ML) algorithms provide encouraging evidence that text-chat data can be used in real applications to benefit the community and business. For example, applications can be developed for workers in organizations who predominantly rely on the text-chat environment to accomplish their tasks that can alert communicators in real time via pop-up boxes containing written messages to the current situation and warn the communicators to be careful of what they write. This research has also shown that when algorithms are used with text-chat data, this can be integrated with other systems for use in critical situations. For example, it can be incorporated into the Beidou Navigation System (BDNS) (Other Regional Systems, 2014) which is widely used on ships to exchange text messages between the end users on the ships and the control station on the land to allow the receiver in the control station to judge whether the emergency situations reported via the text messages are serious or not in order to better judge the need for further action.

7.2 Chapter Organisation

In section 7.3, we present the introduction and background and show how algorithms have a beneficial use in the field of human computer interaction (such as cognitive load classification via eye blinks). Section 7.4 overviews our attempts to predict the level of cognitive load that people can face in the text-chat environment using several linguistic and mouse features. Section 7.5 discusses the level of accuracy of the prediction of interpersonal trust between people in the text-chat environment using hesitation features. Section 7.6 presents the summary of this chapter.

7.3 Introduction and Background

With the recent acceleration in the development of communication technology, textual communication has become one of the most commonly used methods of communicating, as evidenced by the software deployed on smart phones these days. One of the most prominent ways to connect with others is via a text chat, which enables people to exchange messages in both a synchronous and asynchronous manner. One important example of the use of text messages is the exchange of information between people on ships and people on the land via the navigation system, BDNS (Other Regional Systems, 2014). The question is, how can these types of systems be improved in order to determine important and sensitive information, such as the level of trust? To do this, it is important to examine ML algorithms, such as Naïve Bayesian algorithms with the message data and see if these algorithms show a higher prediction rate. If the algorithms increase predictions rate, then they can be incorporated into systems which are used widely.

A previous study showed that the words' number, the words per sentence and the words' length were increased under a high cognitive load (Khawaja et al., 2009) and also that people used a lower variety of words under a high cognitive load (Khawaja et al., 2010). It has also been found that pauses increased with low trust (Kalman et al., 2010).

Several studies have classified people's behaviours and attitudes in the field of human-computer interaction. Nourbakhsh et al. (2013) classified cognitive load using a combination of GSR data, which reflects the level of electricity (nerve responses) emitted from the skin (Peuscher, 2012), and eye blinks, resulting in some reasonable results. Yu et al. (2011) conducted a study on the digital pen, using pressure and the orientation of the pen to classify cognitive load levels. Castillo et al. (2011) classified and predicted the credibility of tweets automatically. They used a set of features extracted from Twitter, such as whether the tweets have question marks "?" or if they contain smile emoticons. Some of the features tested showed significant differences in the measurement of tweets and the automatic classification showed good performance in classifying tweets as credible or not. Al-Eidan et al. (2010) examined the credibility of information on Twitter and found there was a high correlation between accurately determining the credibility of tweets and similarity with the news from news sources.

This chapter details several attempts to use ML algorithms to predict the level of interpersonal trust and cognitive load. Weka software[2] (Hall et al., 2009), which has various classifiers, such as Naïve Bayesian and k-Nearest Neighbors, was used as it is able to extract several features from the text-chat data, such as the total number of words and messages. The classification results for cognitive load were weak and were only moderately accurate in terms of predicting the level of cognitive load, even though some of the data showed significant differences between the levels of cognitive load after being examined by a statistical analysis method (t-test). On the other hand, our attempts to classify interpersonal trust levels had a high prediction accuracy, reaching 82.14%.

We used Weka software (Hall et al., 2009) to classify the levels of cognitive load and trust in this chapter. The data that was used in this chapter were the data on the linguistic features and the mouse movements as discussed in Chapter 3 and the data on hesitation (using the backspace and delete buttons) which is discussed in Chapter 5. The data on linguistic features and mouse movements were used in this chapter to classify cognitive load while the hesitation data were used in this chapter to classify interpersonal trust.

In this chapter, we calculated two measures for each algorithm: accuracy and F-measure. These measures were calculated automatically using Weka

[2] The default 10-fold cross-validation was used in the settings of Weka software for all algorithms. For the attribute evaluator "CfsSubsetEval", we also used 10-fold cross-validation and we selected any measure which reach 10% or higher (the percentage of folds) for classification.

software. The accuracy measure shows the percentage of accuracy of the prediction. For example, if we have ten instances and the accuracy is 70%, this means that the algorithm accurately predicted seven instances from the ten instances. However, the F-measure was defined as "*the harmonic mean of precision and recall*" (Intro to AI Website).

We selected these algorithms, Random Tree, Random Forest, Naïve Bayesian and k-Nearest Neighbors in our study because these algorithms are considered advanced algorithms which are able to work with complex data better than simple algorithms. The difference between these algorithms is that these advanced algorithms belong to different classifiers as follows: the Random Tree and Random Forest algorithms belong to the tree classifier, the Naïve Bayesian algorithm belongs to the Bayes classifier and the k-Nearest Neighbors algorithm belongs to the Lazy classifier. These algorithms work differently, for example, the Random Tree algorithm builds a model like a tree then it compares the data with the model to predict the type of class.

7.4 Cognitive Load Classification

7.4.1Method

The data used for cognitive load classification was from the same experiment as detailed in Chapter 3. In summary, twenty participants chatted for thirty minutes, using instant messaging, and as they chatted, they faced a low cognitive load (LCL) which involved summing small numbers and a high cognitive load (HCL) which involved summing large numbers (cognitive load was classified into two conditions: low cognitive load and high cognitive load). The participants were requested to answer a question adapted from (Nasa, 1986), "*Please rank the mental effort you had to expend while summing these numbers*", to make sure the participants faced a high cognitive load with the task involving large numbers and a low cognitive load with the task involving small numbers. The results of the t-test were compatible with the expected results for the large and small numbers (p<0.000, two-tailed t-test), further details are in Chapter 3.

Linguistic features (such as the total number of messages) and mouse activity (such as distance travelled) were used to classify cognitive load. The linguistic features and mouse activity were examined statistically, the results showing significant differences between low cognitive load and high cognitive load, as detailed in Chapter 3.

In general, we used the default 10-fold cross-validation in this chapter to classify the two levels of cognitive load, low load and high load.

We also classified cognitive load levels via only best attributes. Therefore, we used the evaluator of attribute "*CfsSubsetEval*" (with the default 10-fold cross-validation) in the Weka software to select the best attributes for classification. We selected each attribute which reached 10% or higher for classification.

7.4.2Classification Results

The cognitive load classification was conducted using mouse activity and linguistic features independently, and then both were combined to better understand the extent of the effect on cognitive load classification of each one separately and in combination.

7.4.2.1 Using Mouse Activity

Three attempts were made to use mouse activity in cognitive load classification: 1) all mouse measures; 2) the mouse measures which showed only statistically significant differences (between low and high load conditions); and 3) the attribute evaluator, "*CfsSubsetEval*" in the Weka software, which evaluates the importance of each attribute to select the best measures for cognitive load classification.

For the first attempt, all mouse measures, namely Distance, Vertical Lines, Horizontal Lines, Positive Slope, Negative Slope, Movement Count and Duration, were used. Table 7-1 summarizes the performance of using all mouse measures, showing that most results were weak except for the Naïve Bayesian algorithm which showed a moderate accuracy (70%) for the classification of the two levels of cognitive load.

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.47	47.5%
Random Forest	0.63	62.5%
Naïve Bayesian	0.70	70%
k-Nearest Neighbors	0.53	52.5%

Table 7-1: The classification performance of all mouse measures.

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.47	47.5%
Random Forest	0.65	65%
Naïve Bayesian	0.70	70%
k-Nearest Neighbors	0.48	47.5%

Table 7-2: The classification performance of mouse measures which have statistically significant differences.

Table 7-2 shows the performance of those measures which only had significant differences ($p<0.05$, two-tailed t-test), these being Distance, Vertical Lines, Horizontal Lines, Positive Slope, Negative Slope and Movement Count. The results were similar to the results of using all measures, ranging from 47.5% to 70%.

Table 7-3 shows the results of the measures which were selected by the attribute evaluator "*CfsSubsetEval*", these being Distance, Positive Slope, Negative Slope and Movement Count. The results were not better than the previous attempts, showing a small improvement with the Naïve Bayesian algorithm (2.5%) where accuracy reached 72.5%.

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.47	47.5%
Random Forest	0.63	62.5%
Naïve Bayesian	0.72	72.5%
k-Nearest Neighbors	0.45	45%

Table 7-3: The classification performance of best mouse measures which were selected by the attribute evaluator.

7.4.2.2 Using Linguistic Features

The same procedure which was used for the mouse measures was also used for the linguistic features in relation to cognitive load classification. Table 7-4 shows the performance of using all linguistic features, which are total number of messages, total number of words, total number of characters, average number of words in messages, average number of characters in messages, average number of characters in words, total turns and average number of messages in turns (as adapted from Avrahami and Hudson (2006)). The prediction accuracy was low, as shown in Table 7-4, when these linguistic features were used, ranging from 35% to 52.5%.

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.53	52.5%
Random Forest	0.35	35%
Naïve Bayesian	0.49	50%
k-Nearest Neighbors	0.38	37.5%

Table 7-4: The classification performance of all linguistic measures.

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.40	40%
Random Forest	0.45	45%
Naïve Bayesian	0.55	55%
k-Nearest Neighbors	0.52	52.5%

Table 7-5: The classification performance of linguistic measures which have statistically significant differences.

The results for some of the linguistic features which only showed statistically significant differences ($p<0.05$, two-tailed t-test), was summarized in Table 7-5. These linguistic features were total number of messages, total number of words, total number of characters and total number of turns. The results for accuracy were similar to the results obtained when using used all linguistic features. Also, the performance when using the linguistic features that were selected by the attribute evaluator "*CfsSubsetEval*" was low, as shown in Table 7-6 (only one feature was selected by the attribute evaluator, total messages).

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.38	37.5%
Random Forest	0.40	40%
Naïve Bayesian	0.45	45%
k-Nearest Neighbors	0.40	40%

Table 7-6: The classification performance of the best linguistic measures which were selected by the attribute evaluator.

7.4.2.3 Using a Combination of Mouse Measures and Linguistic Features

An existing study used a combination of features to predict the level of cognitive load (Nourbakhsh et al., 2012), combining the number of eye blinks and GSR data, producing good results in terms of the classification of cognitive load. Similarly, we examined two different measures, mouse activity and linguistic features. Table 7-7 summarizes the performance when using a combination of all linguistic features and mouse activity. The linguistic features were total number of messages, total number of words, total number of characters, average number of words in messages, average number of characters in messages, average number of characters in words, total number of turns and average number of messages in turns, while the mouse activity measures were Distance, Vertical Lines, Horizontal Lines, Positive Slope, Negative Slope, Movement Count and Duration. The prediction accuracy was low for most algorithms Random Tree, Random Forest and k-Nearest Neighbors, ranging between 40% and 57.5%, while the Naïve Bayesian algorithm had moderate accuracy (70%).

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.57	57.5%
Random Forest	0.55	55%
Naïve Bayesian	0.70	70%
k-Nearest Neighbors	0.40	40%

Table 7-7: The classification performance of combining all mouse and linguistic measures.

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.58	57.5%
Random Forest	0.60	60%
Naïve Bayesian	0.70	70%
k-Nearest Neighbors	0.25	25%

Table 7-8: The classification performance of combining mouse and linguistic measures which have statistically significant differences.

In relation to statistically significant differences, Table 7-8 summarizes the performance of combining the significant measures of the linguistic features ($p<0.05$, two-tailed t-test) with the significant measures of the mouse activity features ($p<0.05$, two-tailed t-test). For the linguistic features, the significant measures were total number of messages, total number of words, total number of characters and total number of turns, while for the mouse activity, the significant measures were Distance, Vertical Lines, Horizontal Lines, Positive Slope, Negative Slope and Movement Count. However, the results were not better than using a combination of all measures.

The evaluator "*CfsSubsetEval*" did not select any attributes from the linguistic features, whereas from the mouse activity, the evaluator selected four attributes, Distance, Positive Slope, Negative Slope and Movement Count. Therefore, we used only these four measures. Similar to the previous results, there was no improvement in the prediction accuracy for the classification of cognitive load when using this combination of measures (see Table 7-9).

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.47	47.5%
Random Forest	0.63	62.5%
Naïve Bayesian	0.72	72.5%
k-Nearest Neighbors	0.45	45%

Table 7-9: The classification performance of combining the best mouse and linguistic measures as selected by the attribute evaluator.

7.5 Interpersonal Trust Classification

7.5.1 Method

In order to measure the prediction accuracy of interpersonal trust classification, we used the hesitation data (frequency of use of the *Backspace* and *Delete* buttons) which was examined statistically in Chapter 5 where the data collection procedure was detailed. In summary, the twenty-eight participants were randomly assigned a partner, making fourteen pairs of partners. Half of these pairs had their trust manipulated in a positive way so that they had a high level of trust in their partners and half of the pairs had their trust manipulated in a negative way so that they had a low level of trust in their partners. Trust was manipulated in two ways. To create the high trusting groups, we allowed the partners to meet each other face-to-face to talk for ten minutes (adapted from Zheng et al., 2002) before chatting in the text-chat environment and also, a paragraph was included in the instructions of the experiment to encourage trust (adapted from Zand, 1972). To create the low trusting groups, the partners were not allowed to meet each other face-to-face (adapted Zheng et al., 2002) and a paragraph was included in the instructions of the experiment to encourage distrust (adapted from Zand, 1972). The results of the measurement of interpersonal trust (between the two conditions: high trust and low trust) were as expected, these being that the participants in the former group trusted their partners significantly more compared with the participants in the latter group ($p<0.01$, two-way ANOVA), further details are in Chapter 5.

In this chapter, we use the hesitation features obtained from the participants as they wrote their text messages and the algorithms to classify the two levels of interpersonal trust, high trust and low trust (two classes).

Similar to the steps we followed in the classification of cognitive load, we employed the same steps for the classification of trust. We used the default 10-fold cross-validation in all our attempts to classify the two levels of trust, high trust and low trust. Also, in one of our attempts for classification, the attribute evaluator "*CfsSubsetEval*" was chosen with 10-fold cross-validation to find the best attributes and use them for the classification of trust. In relation to selecting the best attributes via the "*CfsSubsetEval*", any attribute reaching 10% or higher was chosen for the classification of trust.

7.5.2 Classification Results

The hesitation measures used by the participants when writing their text-chat messages, that is, the frequency of using the *Backspace* and *Delete* buttons, were used to evaluate their performance in accurately classifying the two level of interpersonal trust (low and high). The results show that these two measures had a high degree of prediction accuracy, at 80% and higher, as follows: Random Tree (82.14%), Random Forest (82.14%) and k-Nearest Neighbors (80.36%), as shown in Table 7-10. We used the t-test to select the best feature to examine it with the algorithm but neither the *Backspace* feature nor the *Delete* feature showed significant results.

The evaluator "*CfsSubsetEval*" showed that only the performance of the *Backspace* feature is impressive in terms of trust classification, therefore, we removed the *Delete* feature from the data set and only used *Backspace* feature in the next attempt to classify trust. Table 7-11 shows that this resulted in a slight improvement in prediction accuracy, as the accuracy of Naïve Bayesian and k-Nearest Neighbors increased by about 5% and 2%, respectively, compared to when the two features were combined. The most important finding is that three of the four algorithms (k-Nearest Neighbors, Random Tree and Random Forest) reached a higher prediction accuracy of 82.14%, 82.14% and 82.14%, respectively.

Classification Algorithm	**F-Measure**	**Accuracy**
Random Tree	0.82	82.14%
Random Forest	0.82	82.14%
Naïve Bayesian	0.64	64.29%
k-Nearest Neighbors	0.80	80.36%

Table 7-10: The classification performance of hesitation measures which measured the frequency of using the Backspace and Delete buttons.

Classification Algorithm	F-Measure	Accuracy
Random Tree	0.82	82.14%
Random Forest	0.82	82.14%
Naïve Bayesian	0.69	69.64%
k-Nearest Neighbors	0.82	82.14%

Table 7-11: The classification performance of best hesitation measure which was selected by the attribute evaluator (the total number of Backspace clicks).

7.6 Discussion

In this chapter, we examined the ability of ML algorithms to classify the level of cognitive load and interpersonal trust. Several different attempts were made to accurately predict the level of cognitive load. Although there were statistical differences ($p<0.05$) for some measures, as shown above, the classification results were not high, the reason for this being due to the large variation in the data of participants which consequently causes difficulties for cognitive load classification. These results encourage further investigation to try to improve prediction accuracy, such as examining each person's data individually rather than using all the data on all the people together by developing an algorithms that is able to split each person's data into parts and examine these parts to classify cognitive load.

In relation to the classification of interpersonal trust, the hesitation measures from the participants' writing (specifically looking at their use of the *Backspace* button) showed a high ability to predict the level of trust between participants. The reason for this is that at a high trust level (where the number of times the *Backspace* button was used was low), the majority of participants were close to each other and also we found the same thing at a low trust level (where the number of times the *Backspace* button was used was high). A previous study (Kalman et al., 2010), which found that delays in responses were associated with low trust between people when chatting, supports our finding where our study found the increase use of the *Backspace* button was associated with a low level of trust.

All the advanced algorithms showed low accuracy for cognitive load classification when we used the linguistic and mouse data. However, most of the advanced algorithms showed a high accuracy to predict the trust level when we used the data on hesitation. In addition, no one algorithm was significantly better than all algorithms for the classification of trust which means that the most advanced algorithms may work effectively in similar situations.

7.7 Chapter Conclusion

The results of the measurement of hesitation when the participants were writing messages are encouraging evidence that these ML algorithms are able to predict the level of interpersonal trust. Trust classification via hesitation features can be applied in systems which mainly use the text-chat environment, such as the Beidou Navigation System (Other Regional Systems, 2014), which is used on ships to exchange messages with people on the land, to help in deciding the credibility of emergency information which may be sent via messages from the ship to the control station on the land. However, it was noted that one hesitation feature (*Backspace*) was prevalent when participants had a low level of trust as they showed hesitation, hence this is the reason for the high accuracy of the trust classification in this chapter.

Chapter 8 Conclusions and Future Directions

8.1 Conclusions

One of most popular ways for people to communicate today is through the use of text messages. There is a diverse array of software deployed on computers and smart phones which facilitates this kind of communication. When using text messages, communicators do not see or hear each other, which can affect the efficiency of this form of communication. In this thesis, we examined the factors which affect the communication between people in the text-chat environment and also we examined the data of the communicators collected during their chatting to better understand what problems the communicators face and how they behave to help them use this environment more effectively. This thesis examined the effects of different levels of cognitive load on the interpersonal trust of the communicators. Our results found that increasing cognitive load has a negative impact on the building of trust. The reason for this may be due to the fact that communicators do not usually communicate when under a high cognitive load which adversely impacts the level of trust between the communicators. In order to determine the different levels of cognitive load, we analysed two indicators, the amount of chatting that took place and the mouse movements of the communicators, the results revealing that both these indicators were higher with a low cognitive load.

This thesis also examined the effects of different behaviours (behaving cooperatively and competitively) on the interpersonal trust of the communicators in the chat medium, finding that these behaviours have a significant impact on the trust. The results showed that behaving cooperatively in order to accomplish tasks and establishing the feeling that the communicators are one team enhances the establishment of trust, whereas behaving competitively and trying the defeat the other party decreases the level of trust of the communicators. It was also found that there is an association between interpersonal trust and certain types of words, for example, high trust is associated with assent and positive emotion words which encourages harmony and respect between the communicators when they trust each other.

Our findings in this thesis show that physiological signals, which are indicators of stress (for example, the amount of sweat captured from the fingers) are affected significantly by the overlap between trust and cognitive load, showing that when there was trust between the communicators who were exposed to a low cognitive load, their physiological signals were low.

This thesis also examined the role of shared visual information in building the trust relationship, the results showing that having access to shared visual information enhances trust between people in the text-chat environment. Finally, ML algorithms were proven to be effective in predicting the issues the communicators may face in the chat medium. Specifically, it was found that the algorithms (with the hesitation features from the chat sessions) were able to predict the level of interpersonal trust between the communicators.

8.1.1 Implications

The results presented in this thesis have multiple implications and these implications can be divided into two types: theoretical and applied. In relation to the theoretical implications, this thesis showed that the difficulties which people face when they write text messages influence their interpersonal trust and also that different behaviours impact interpersonal trust. This tells us about appropriate ways that people can behave in order to build a better relationship in the text-chat medium.

In relation to the applied implications, there are many useful implications which can be extracted from the findings in this thesis. Making the text-chat environment more usable for people may enhance the building of relationships between communicators. For instance, previous work showed that collaboration between people in the chat medium was enhanced when they were shown a large dialogue history (Gergle et al., 2004). Chat mediums can be developed that can encourage people to cooperate as cooperation has a significant influence on the building of interpersonal trust, as shown in our results.

We believe that the text-chat environment is a convenient medium of communication between colleagues who work as it allows messages to be exchanged in both a synchronous or asynchronous manner. It is also cost and time effective as it spares the need for people to travel to meet their colleagues. However, the question is, how can the text-chat environment be improved, particularly in the government and business context? Analysing various features which can be collected from this rich medium, such as the amount of chatting, the exchanged emotions and the way of moving the mouse which were shown to be indicators of interpersonal trust and cognitive load, adaptive systems can be developed to determine the issues which employees face and provide them with suitable assistance, such as warning them during a chat session about an inappropriate situation which may be arising and reminding them to be careful in their communications with their colleagues. Also, these adaptive systems can pass on information about employees to supervisors so that they can better understand the issues facing employees, take the necessary action to support them and enhance communication between colleagues in the workplace. Finally, in relation to implications, the feature of shared visual information can be added to the design of text-chat systems to allow communicators to see the actions of their partners in real time when they solve or complete tasks, as this will enhance communication, as shown in this thesis, where the feature of shared visual information has a positive relation with trust building between communicators. A real example of using shared visual information can be seen at NORAD where the workers watch the distribution of aircraft in airspace and use the text-chat environment to communicate with airport staff about these aircraft (Gergle, et al., 2004).

8.2 Future Directions

There are many other ideas than can be explored to enhance the use of the text-chat environment. As shown in this thesis, linguistic features are associated with interpersonal trust and cognitive load, which was encouraging evidence to use them as indicators of the issues people face in the text-chat environment. However, linguistic politeness was not used to measure interpersonal trust in the text-chat medium. We believe that there may be an association between trust and linguistic politeness, this is another direction for future research. Brown and Levinson (1987) divided speech into direct and indirect speech. Park (2008) showed some examples for these types of speech: 1) *direct speech:* someone says something in a direct way ***"Ya, help us out :-)"*** *;*2) *indirect speech:* if someone needs something "*try posting ur picture on the URL*" but his partner replies in an indirect way ***"I'm too new at that".*** We anticipate that if a study was conducted to investigate the relationship between interpersonal trust and linguistic politeness in the text-chat environment, it will be found that participants who trust their partners will being closer and have more solidarity with each other during the chatting which will be reflected in their use of direct speech, and specifically, we think they will avoid the use of hints and will use informal speech.

Another research direction is to examine the extent of the relationship between eye gazes and the level of interpersonal trust in the text-chat environment. Specifically, we think that people who have low trust in their partners will check their text messages more than people who have high trust in their partners (the checking of text messages can be captured by using eye gazes). However, it would be useful to conduct further studies on the text-chat environment to find other ways which are complementary to our studies in enhancing the communication between people in this environment.

REFERENCES

Arshad, S., Wang, Y. & Chen, F. Analysing Mouse activity for Cognitive Load detection. In Proc. OzCHI 2013, 115-118

Avrahami, D. and Hudson, S. Communication Characteristics of Instant Messaging: Effects and Predictions of Interpersonal Relationships. In Proc. CSCW 2006, 505-514.

Al-Eidan, R., Al-Khalifa, S., Al-Salman, A.: Measuring the Credibility of Arabic Text Content in Twitter. In: 5th International Conference on Digital Information Management, pp. 285-291. IEEE Press, New York (2010).

Astra Gift Maker (Free jigsaw puzzles creator). http://www.lenagames.com/astra-gift.htm. Release date 1-Nov, 2010, last accessed: Dec-2015.

Bos, N., Olson, J., Gergle, D., Olson, G. & Wright, Z. Effects of four computer-mediated communications channels on trust development. In Proc. CHI 2002, 135-140.

Brown, P. & Levinson, S. 1987, Politeness: Some universals in language usage, Cambridge, United Kingdom: Cambridge University Press.

Biros, D., Daly, M. & Gunsch, G. The Influence of Task Load and Automation Trust on Deception Detection. Group Decision and Negotiation, 13, 2(2004), 173-189.

Butler, J. Behaviors, trust, and goal achievement in a win-win negotiating role play. Group & Organization Management, 20, 4 (1995), 486-501.

Butler, J. Toward understanding and measuring conditions of trust: Evolution of conditions of trust inventory. Management, 17, (1991), 643-663.

Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S.Z., Khawaji, A. and Conway, D. Robust Multimodal Cognitive Load Measurement. 2016.

Clevinger, A. Test Performance: the Influence of Cognitive Load on Reading Comprehension. Psychology Theses, 2014.

Chandler, P. & Sweller, J. Cognitive load theory and the format of instruction. Cognition and Instruction, 8, 4(1991), 293-332.

Castillo, C., Mendoza, M., Pobletey, B.: Information Credibility on Twitter. In: 20th Interna-tional Conference on World Wide Web, pp. 675-684. ACM Press, New York (2011).

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.

Calculate the distance between two points in Java. http://stackoverflow.com/questions/14431032/i-want-to-calculate-the-distance-between-two-points-in-java. 2013, last accessed: May-2016.

Dirks, K. & Ferrin, D. The Role of Trust in Organizational Settings. Organization Science, vol. 12, no. 4 , pp. 450-467, 2001.

Deutsch, M. Cooperation and competition. The Handbook of Conflict Resolution: Theory and practice, (2006), 23-42.

de Melo, C., Carnevale, P., and Gratch, J. People's biased decisions to trust and cooperate with agents that express emotions. In Proc. AAMAS 2013.

de Melo, C., Carnevale, P., and Gratch, J. The impact of emotion displays in embodied agents on emergence of cooperation with people. Teleoperators & Virtual Environments, 20, 5(2012), 449-465.

de Melo, C., Carnevale, P. & Grath, J. 2011a, 'The impact of emotion displays in embodied agents on emergence of cooperation with people', Journal of Presence: Teleoperators and Virtual Environments, volume 20, issue 5, pages 449 – 465.

de Melo, C., Carnevale, P. & Grath, J. 2011b, 'The effect of expression of anger and happiness in computer agents on negotiations with humans', in Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems, pages 937-944, 2011.

de Melo, C., Zheng, L. & Gratch, J. 2009, ' Expression of moral emotions in cooperating agents', in Proceedings of the 9th International Conference on Intelligent Virtual Agents (IVA'09), pages 301 - 307, 2009.

Engstrom, J., Johansson, E. and Ostlund, J. Effects of visual and cognitive load in real and simulated motorway driving, 2005.

Facebook Reports Second Quarter 2015 Results. http://investor.fb.com/releasedetail.cfm?ReleaseID=924562. (Retrieved Aug 10, 2015).

Gergle, D., Millen, D., Kraut, R. & Fussell, S. 2004, 'Persistence matters: making the most of chat in tightly-coupled work', in Proceedings of the 22th SIGCHI Conference on Human Factors in Computing Systems (CHI), pages 431 - 438, Vienna, Austria, 2004.

Gill, A., French, R., Gergle, D. and Oberlander, J. The language of emotion in short blog texts. In Proceedings of the 2008 ACM conference on Computer supported cooperative work (CSCW), 2008. ACM, New York, NY, USA, 299-302.

Gerven, P., Paas, F., Merrienboer, J., Hendriks, M. & Schmidt, H. The efficiency of multimedia learning into old age. British Journal of Educational Psychology, 73, 4(2003), 489-505.

Harbring, C. On the Effect of Incentive Schemes on Trust and Trustworthiness. Institutional and Theoretical Economics, 166, 4 (2010), 690-714.

Intro to AI Website, http://cse3521.artifice.cc/classification-evaluation.html. Last accessed: 5/10/2016.

Koum, J. (2014). Twitter Page. https://twitter.com/jankoum?lang=en. (Retrieved Aug 13, 2015).

Kalman, Y., Scissors, L., and Gergle D. Chronemic aspects of chat, and their relationship to trust in a virtual team. In Proc. MCIS 2010.

Kelly, N. Building Trust in Communications, 2013.

Kraut, R., Gergle, D., Fussell, S. The Use of Visual Information in Shared Visual Spaces: Informing the Development of Virtual Co-Presence. In Proc. CSCW 2002.

Kiesler, S., Waters, K., Sproull, L.: A prisoner's dilemma experiment on cooperation with human-like computers. Soc Psy 70, (1996), 47-65.

Khawaja, A., Ruiz, N. & and Chen, F. Potential speech features for cognitive load measurement. In Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces (OZCHI), 2007, 57-60.

Khawaja, A., Chen, F. & Marcus, N. Using Language Complexity to Measure Cognitive Load for Adaptive Interaction Design. In Proc. IUI 2010, 333-336.

Khawaja, A., Chen, F. & Marcus, N. Measuring Cognitive Load using Linguistic Features: Implications for Usability Evaluation and Adaptive Interaction Design. The Journal of Human Computer Interaction, 30, 5(2014), 343-368.

Khawaja, M. A., Ruiz, N. & Chen, F. Think before you talk: An empirical study of Relationship between speech pauses and cognitive load. In Proc. OzCHI 2008, 335-338.

Khawaja, A., Chen, F. & Marcus, N. Analysis of Collaborative Communication for Linguistic Cues of Cognitive Load. The Journal of the Human Factors and Ergonomics Society, 54, 4(2012), 518-529.

Khawaja, A., Chen, F. Owen, C. and Hickey, G. Cognitive Load Measurement from User's Linguistic Speech Features for Adaptive Interaction Design. In Proc. INTERACT 2009, 485-489.

Khawaja, A. Cognitive Load Measurement using Speech and Linguistic Features, thesis, 2010.

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.

Lam, C. 2011, 'Linguistic Politeness in Student-Team Emails: Its Impact on Trust Between Leaders and Members', Journal of IEEE Transactions on Professional Communication, volume 54, number 4, pages 360 -375.

LIWC Website. LIWC2007 Output Variable Information. http://liwc.net/descriptiontable1.php. (Retrieved Sep 14, 2015).

Myers, D., and Tingley, D. The Influence of Emotion on Trust, 2011.

Mayer, R. & Davis, J. An Integrative Mode of Organizational Trust. Academy of Management Review, 20, 4(1995), 709-734.

Math Open Reference. http://www.mathopenref.com/tocs/coordpointstoc.html. Last Accessed: 15-Sep-2015.

Martin, E. & Bajcsy, R. Analysis of the Effect of Cognitive Load on Gait with off-the-shelf Accelerometers. In Proc. COGNITIVE 2011, 1-6.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.

Nourbakhsh, N., Wang, Y. & and Chen. F. GSR and Blink Features for Cognitive Load Classification. In Proceedings of INTERACT 2013, 159-166.

Nourbakhsh, N., Wang, Y., Chen, F. and Calvo R. Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. In Proceedings of the 24th Australian Computer-Human Interaction Conference (OzCHI), 2012, 420-423.

Nagarajan, M. and Hearst, M. An Examination of Language Use in Online Dating Profiles. Proceedings of the 3rd International Conference on Weblogs and Social Media, AAAI Press (2009).

Nahapiet, J. & Ghoshal, S. 1998, 'Social Capital, Intellectual Capital and the Organizational Behaviour', The Academy of Management Review, volume23, number 2, pages 242-266.

Norman, S., Avolio, B., and Luthans, F. The impact of positivity and transparency on trust in leaders and their perceived effectiveness. The Leadership Quarterly, volume 21, issue 3, June 2010, pages 350-364.

NASA. Nasa Task Load Index, Ver. 1.0, 1986.

Other Regional Systems. http://www.navipedia.net/index.php/Other_Regional_Systems. (Retrieved Oct 13, 2014).

Peuscher, J. Galvanic Skin Response, Ver. 3, 2012.

Pai, S., Gasson, S. Effects of Cultural Differences on Trust Reparation in a Computer Mediated Communication Setting, 2008.

Psychwiki Website, http://www.psychwiki.com/wiki/Cognitive_Load. 2010, Last Accessed: 4/10/2016

Pennebaker, J., Chung, C., Ireland, M., Gonzales, A., and Booth, R. The development and psychometric properties of LIWC2007, 2007.

Peter (2009), without hesitation, repetition or deviation. http://www.listen-to-english.com/?id=461. Last Accessed: 1/12/2015.

Park, J. 2008a, ‘Linguistic politeness and face-work in computer-mediated communication’, Part 1: A theoretical framework. Journal of the American Society for Information Science and Technology, volume 59, issue 13, pages 2051-2059.

Park, J. 2008b, ‘Linguistic politeness and face-work in computer-mediated communication’, Part 2: An application of the theoretical framework. Journal of the American Society for Information Science and Technology, volume 59, issue 14, pages 2199-2209.

Paas, F., Renkl, A. & Sweller, J. Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. Instructional Science, vol. 32, no. 1-2, pp. 1-8, 2004.

Quan-Haase, A., Cothrel, J. & Wellman, B. Instant messaging for collaboration: A case study of a high-tech firm. Computer-Mediated Communication Journal, 2005.

Robert, L. Monitoring and Trust in Virtual Teams. In Proc. CSCW 2016.

Riegelsberger, J., Angela, M. & McCarthy, J. 2003. The researcher's dilemma: evaluating trust in computer-mediated communication. International journal of Human-Computer Studies: Trust and Technology, volume 58, issue 6, pages 759 – 781.

Ramirez-Esparza, N., Chung, C., Kacewicz, E. and Pennebaker, J. The Psychology of Word Use in Depression Forums in English and in Spanish: Testing Two Text Analytic Approaches (2008).

Skype Statistics and Facts. http://expandedramblings.com/index.php/skype-statistics/. (Retrieved Aug 13, 2015).

Scollon, R. & Scollon, S. 1981, Narrative, Literacy and Face in Interethnic Communication.

Scissors, L., Gill, A., and Gergle, D. Linguistic mimicry and trust in text-based CMC. In Proc. CSCW 2008, 277-280.

Scissors, L., Gill, A., Geraghty, K., and Gergle, D. In CMC we trust: the role of similarity. In Proc. CHI 2009, 527-536.

Starker, D. (2008). What is Trust?. http://www.changingminds.org/explanations/trust/what_is_trust.htm. (Retrieved Dec 23, 2013).

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.

Takeuchi, Y. and Naito, H. How Do Real or Virtual Agent's Body and Instructions Contribute to Task Achievement?. Human Interface, Part I, HCII, 2011, 142-151.

Thirunarayanan, M, Ryan, C. and Perez-Prado, A. An Exploratory Study of Cognitive Load in Instructional Chat Rooms. In Proc. ICCE 2002.

Twitter Usage. https://about.twitter.com/company. (Retrieved Aug 12, 2015).

Thirunarayanan, M, Ryan, C. and Perez-Prado, A. An Exploratory Study of Cognitive Load in Instructional Chat Rooms. In Proc. ICCE 2002.

Thirunarayanan, M, Ryan, C. and Perez-Prado, A. An Exploratory Study of Cognitive Load in Instructional Chat Rooms. In Proc. ICCE 2002.

Van Gog, T., Paas, F., Marcus, N., Ayres, P. & Sweller, J. The Mirror Neuron System and Observational Learning: Implications for the Effectiveness of Dynamic Visualizations. Educational Psychological Review, 21,1(2009), 21-30.

Wang, M. Trust, situation awareness and automation use: exploring the effect of visual information degradations on human perception and performance in human-telerobot. Graduate Theses and Dissertations (2010).

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.

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.

Yu, K., Epps, J., and Chen, F. Cognitive Load Measurement with Pen Orientation and Pressure. 2011. Proceedings of MMCogEmS

Yu, K., Epps, J., Chen, F. Cognitive load evaluation of handwriting using stroke-level features. In Proceedings of the 16th international conference on Intelligent user interfaces In Proc. IUI 2011, 423-426.

Zheng, J., Veinott, E., Olson, J. & Olson, G. Trust without touch jumpstarting long-distance trust with initial social activities. In Proc. CHI 2002, 141-146.

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.

Zhang, T., Hasegawa-Johnson, M., and Levinson, S. A Hybrid Model for Spontaneous Speech Understanding. American Association for Artificial Intelligence, 2002.

Zand, D. Trust and Managerial Problem Solving. Administrative Science Quarterly 17, 2 (1972).

UNSW Approval No ()

A Cognitive Load Questionnaire

Please rank the mental effort you had to expend while summing these numbers (using a number from 1 to 9 where 1= “Almost no effort”, 5= “Neutral” and 9= “Extreme effort”):

Almost no effort				Neutral				Extreme Effort
1	2	3	4	5	6	7	8	9

UNSW Approval No ()

B Trust Questionnaire

Please rate the following items using a number from 1 to 9 (1= "Strongly Disagree", 5= "Neutral" and 9= "Strongly Agree").

1) I feel my partner did not take advantage of me to maximize his/her money:

Strongly Disagree				Neutral				Strongly Agree
1	2	3	4	5	6	7	8	9

2) I feel my partner didn't do anything to cause me to have less money than him/her:

Strongly Disagree				Neutral				Strongly Agree
1	2	3	4	5	6	7	8	9

3) I feel my partner dealt fairly with me:

Strongly Disagree				Neutral				Strongly Agree
1	2	3	4	5	6	7	8	9

4) I usually know how much he/she will invest in each round:

Strongly Disagree				Neutral				Strongly Agree
1	2	3	4	5	6	7	8	9

5) I feel my partner always tells me the truth:

Strongly Disagree				Neutral				Strongly Agree
1	2	3	4	5	6	7	8	9