Tallinn University

School of Digital Technologies

Human-Computer Interaction

DESIGNING AN EXPERIMENT TO EFFICIENTLY STUDY THE RELATIONSHIP BETWEEN TRUST AND EMOTIONS

Master thesis

Author: Kulichenko Olena

Supervisors: Sonia Sousa, PhD

Ilkka Kosunen, PhD

Author: _____ "_____" 2018

Supervisor: _____ "_____" 2018

Supervisor: _____ "_____" 2018

Head of the institute: _____ "_____" 2018

Tallinn 2016

Author's Declaration

I declare that apart from work whose authors are clearly acknowledged, this manuscript is a product of the author's original work, and it has not been published before, neither is not currently being considered for publication elsewhere for any other comparable academic degree.

This master thesis document has been supervised by PhD Sonia Sousa (Tallinn University, Estonia) and Ilkka Johannes Kosunen(Tallinn University, Estonia)

Author: Kulichenko Olena

Date: 1.05.2018

……...

(signature)

……..

(data)

Abstract

Studies show that trust plays a crucial role in user's decision to adopt a technology or in decision-making process while doing an online-shopping. But, despite the increased interest on the topic of trust, still, there is a gap in the literature in human-computer interaction research, especially in what concerns studying the effects of users trust in technology. Specifically studying how trust interplays with emotions, as they also showed to be a factor that affects decisions. Most of the existing studies that explore how trust and emotions interplays have focused on sociological, psychological and political aspects and put little evidence on the effects of technology or on the effects of users' interactions with technology. This work attempts to describe and demonstrate how to design a study that explores how trust and emotions interplays from a Human-Computer Interaction perspective. We will report it opportunities strengths, weakness and constraints when designing such studies. We will end by discussing future approaches on how to foster the design of such studies. Our focus is to investigate the common practices of building an experiment to study the interplay between two complex and multidisciplinary topics like emotion and Trust. First, a literature review was done to better understand those topics multidisciplinary and complexity. This was followed by reporting possible ways to measure them, finally ending by eliciting the importance of such approach on tackle nowadays technology shift, where we find ubiquitous, and non-deterministic characteristic most of the time supported by artificial intelligence features. The second part of this thesis focuses on reviewing the various researches approaches and on understanding how trust and emotions can be studied from an experimental approach. The main methodology used was based on the contextual inquiry and uses participatory design approaches where the author worked closely together with the researchers. Multiple meeting was done to efficiently discuss, design and plan an experiment, then two pilot studies were performed, and highlight results are reported together with the needed planning changes. The result shows that is efficient to use the experimental design approach to investigate the interplay between two complicated topics, and not only gives valid answers to the questions posed by researchers but also provokes the creation of new knowledge and new topics to research. However, it has some constraints that researchers should be aware of to keep the study going in the initial direction.

Kokkuvõte

Tänapäeval muutuvad inimesed olulisemaks teadlikuks, eriti kui käsitletakse kasutajate usaldust tehnoloogiasse. Teine teema, mis huvitab HCI-d, on emotsioon ja selle mõju kasutajakogemusele ja kuidas see mõjutab kasutaja otsuseid tehnoloogia kasutuselevõtu kohta. Olles teadlik olemasolevatest teemadest, on selle väite eesmärgid uurida, kuidas võiks uurida kahe teema koostoimet. Oleme teadlikud, et seda teemat on uuritud teistes uurimisvaldkondades nagu sotsioloogia, psühholoogia ja poliitika. Kuid me mõistame liiga vähe mõjust, mida kasutajad suhtlevad tehnoloogiaga. Käesolev väitekiri uurib, kuidas kavandada eksperimentaaluuringuid, et uurida tehnoloogia ja emotsioonide vastastiku mõju. Selles töös püütakse kirjeldada ja illustreerida keerukaid teemasid käsitlevaid tugevaid külgi ja piiranguid inimese arvutitevahelise suhtluse seisukohast. See annab ülevaate protsessist, kuidas doktorant kujundab eksperimentaalse uuringu, et uurida emotsioonide mõju kasutajate usaldamisotsuses kasutada Google'i abistajat. Samuti annab aru pilootuuringu esialgsetest tulemustest ning arutleb selliste keerukate uuringute kujundamisel võimalike tulevaste lähenemisviiside üle. Kokkuvõttes keskendume uuringule, kuidas korraldada eksperimentaaluuringuid, milles kasutatakse psühhomeetrilisi ja füsioloogilisi mõõtmisi. Kõigepealt alustame kontekstuaalse ülevaatega põhiteemast. Sellele järgnes aruanne katsete kavandamisel võetud sammude kohta. Lõpuks selle üldise tugevuse ja piirangute ilmnemisega. Kasutatud peamine metoodika põhines kontekstuaalsel uurimisel, kus autor jälgis tähelepanelikult iga sammul uuringusse kaasatud teadlasi. Peamisteks uurimismeetoditeks olid vaatlus ja intervjuu. Sellele lisati ka kaks pilootuuringut, et protsessi paremini mõista ja esile tuua tulemusi. Üldiseks järelduseks on see, eksperimentaalne lähenemine näib olevat kõige tõhusam lähenemisviis. Seda protsessi raskendas kaks peamist tegurit: (1) tehnoloogia usaldusväärsuse uurimisel on tehtud vähe uuringuid; (2) Eriti kui kasutatakse laboris saadaolevaid füsioloogilisi mõõtmisi või instrumente. Teadlased vajavad täiendavat uuringuid, et oleks võimalik sügavalt mõista füsioloogiliste mõõtmiste kasutamise eeliseid sellistes uuringutes.

Acknowledgements

I express my deep sense of gratitude to my supervisor Sonia Sousa of the School of Digital Technologies, University of Tallinn for her kind support and supervision. I am truly grateful for guidance and support on every step of my research intentions through two years of my master studies and possibility to participate in this research. I would also thank Ilkka Johannes Kosunen of the School of Digital Technologies for valuable feedback and supervision.

I would like to express my sincere appreciation to Ighoyota Ben for his constant guidance and knowledge, that he shared with me. I would also like to thank all experts who participated in this study and gave their valuable feedback.

I am very much thankful to my parents and to my spouse for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis.

Table of Contents

INTRODUCTION

Trust is a key factor in influencing user's decision to adopt a technology or make an online purchase. Nowadays, with the growing use of artificial intelligence to support user modeling trust plays more and more dominant role in supporting users to adopt a technology. Such technologies tend to reflect more non-deterministic behaviors, mimic more human characteristics, leading people to assume it as having more human-friendly characteristics and willing to help. On the other hand, latest researchers had shown that technology can assume deceiving behavior as well. For instance, in one study by Cowan et al. (2016) of Siri author point that users did not trust the system to do complex tasks like writing emails or calling someone down to an apprehension that the system would not get the task done correctly, on the other hand, people see Siri willingness to help and embed it with human-like characteristics.

Trust plays a dominant role in e-commerce, some researchers state it to be the key to success in Internet business as the establishment of trusted transaction processes where e-sellers create an environment in which a prospective consumer can be relaxed and confident about any prospective transactions. The study by Söllner & Leimeister (2013) showed that consumer trust positively affects consumer's decision to purchase.

At the same time, emotions also showed to make an effect on a decision-making process (Lerner & Keltner, 2001), prior research has identified a significant relationship between moods and emotions and normatively unrelated judgments (Forgas & George, 2001). Studies in sociology and politology showed that affect complex decisions, such as the decision whether to trust a stranger, a politician or a potential competitor (Dunn & Schweitzer, 2005). Some studies suggest that people in positive moods will be more trusting. People in positive moods will have an easier time accessing memories where trusting behavior led to a positive outcome, they also are more altruistic and make optimistic predictions, especially concerning the future actions of others. The study by Forgas (1998) shows that positive mood makes participants more cooperative in negotiating situations. In line with above ideas, the following sections present the research problem and underline the research goal, research question, research objectives and the description of the research methodology.

1.1 RESEARCH PROBLEM AND SIGNIFICANCE

Trust, is defined by Rousseau et al. (1998) an individual's willingness to depend on another party because of the characteristics of the other party. It is a multifaceted concept and a subject studied in many fields (e.g. economy, political science, sociology, psychology). In a technologically enhanced context, trust is generally described as an important key factor to facilitate the mediation between users and computers. Current studies that focus on studying the trust relation with technology tend to either provides us a more operational approach (Benantar, 2001, Manchala, 2000) to the subject or focus on a more social interaction approach (Hubert et al. ,2018; Vance, Elie-Dit-Cosaque & Straub, 2008; Mcknight et al., 2011; Kim et al., 2008). For instance, some studies narrow their approach to studying more technical aspects of trust like reputation, privacy, security, and other (Kim et al., 2008). Others face the challenge of today's fast technological developments and study its impacts on user's behaviors and interactions. For instance, in studying how non-deterministic technologies affects users trust perceptions. Technologies like an online-recommendation system, or that include artificial intelligence algorithms that adapt to user's profiles. Or technologies that assume more human-like [1]characteristic.

Above paragraph illustrates how complex studying trust can be, described the need as well to study it, as technologies are shifting from presenting a more static, predictable and behaviors towards presenting more anthropomorphic and non-deterministic behaviors.

Studies, also showed that trust plays an essential role in user's decision to adopt a technology (e.g. make an online-purchase, delegate a task to Intelligent Personal Assistant (IPA)). In other words, trust has an impact on users decision-making process (Cowan et al., 2016; Kim et al., 2008).

At the same time, we know that emotions are also proved to influence humans' decisions (Dunn & Schweitzer, 2005; Lerner & Keltner, 2001). We also are aware how emotions are complex subject to study.

Our main aim is to investigate and describe how those topics can be interplayed and studied in Human-Computer Interaction. We are aware of several researchers conducted in

[1] measuring technology trust using the human-like trust constructs of integrity, ability/competence, and benevolence that researchers have traditionally used to measure interpersonal trust

other fields that already study the interplay between Trust in technology (Personal intelligent assistant) and Emotions, such as psychology or politology, and this study showed a relation between this two notions (Dunn & Schweitzer, 2005; Lerner & Keltner, 2001). Most have used an experimental approach to use it (Lankton & McKnight, 2011; Vance, Elie-Dit-Cosaque & Straub, 2008). There is a gap in the literature and a lack of clear evidence how emotions affect trust in a field of technology, especially when addressing it from Human-Computer Interaction perspective. Therefore, we propose to address this problem by further research how we can investigate this subject. In sum, this work intends to explore novel research approaches into further explore the interplay between trust in technology and emotion in Human-computer Interaction. It aims to further report on the needed steps to design an experiment that studies this interplay from two measurements a physiological computing instruments and psychometric measurements instruments.

1.2 RESEARCH GOAL AND MOTIVATION

The main goal of this research is to clarify the process of designing a research experiment to efficiently study the relationship between trust in technology and emotions. In other words, it aims to investigate what are the strengths and constraints of design such complex experimental research studies.

This work contributes to this understanding in two main ways:

1. From one hand on providing a **systematic** description of how to design an experiment that explores the interplay between emotions and trust in technology (Personal intelligent assistant) using two distinct measurement approaches physiological and psychometric.
2. From another, contribute to illustrate and explore a more in-depth reflection on the constraints of designing such research approaches in Human-Computer Interaction.

As addressed before, despite similar studies already existing, few studies report and or explore that interplay using two distinct measurement approaches physiological and psychometric.

1.3 RESEARCH QUESTION

The main research questions of this study are:

RQ1. How can we use physiological computing measurements to study the interplay between trust in technology and emotions?

RQ2. How can we use psychometric computing measurements to study the interplay between trust in technology and emotions?

RQ3. What are the constraints of using experimental approaches to design such complex research study?

RQ4. How effective is this research design approach?

1.4 RESEARCH PROCEDURE

Steps of the research	**Research aim and method**	**Method**
Understand the context	Understand, ● How emotions and trust have been studied in Human-Computer interaction? ● How it affects users trust in technology? ● What are the research approaches?	Literature review
Experimental-design first phase	To study how researchers, conceptualize an experimental design approach that explores the interplay between trust and emotion	Observations techniques and interviews, meeting with experts.
Experimental-design second phase	To study how researchers prototype/implement their initial concept, evaluation	pilot study, observation, interviews .

Experimental-design third phase	How researchers re-design it and perform the final implementation steps.	observation, interviews
Conclusion and overall discussion	Analysis of what are the strengths and constraints of design such complex research studies	

Table 1: Overall Research Procedure

2 CONTEXTUAL BACKGROUND

The contextual background section aims to contextualize the reader to the main topics addressed in this work. It starts by illustrating the multidisciplinary nature of the researched terms emotions and trust. Further describes its nature, how to do emotion elicitation from a Human-Computer interaction perspective, and compare the two techniques, such as emotion elicitation using pictures and video clips. This is followed by a brief description of existing studies that explore emotions and present the best practices to study and measure emotions.

The second part of this section focus on trust and its nature, describe its role of trust in Human-Computer Interaction. Present the work of some researchers performed when studying Trust in Human-Computer Interaction, and how trust can be measured. The final part is followed by authors final overview of similar studies designed to study the interplay between emotions and trust in technology.

2.1 EMOTION IN TECHNOLOGY

The term Emotion is difficult to define because it complements a multi-complex term like feelings, behavior, physiological change and cognition. An emotional state also occurs within a context. Overall emotions comprise three main components (Scherer, K. R., 1993):

1. A subjective experience (e.g., feeling angry),
2. An expressive component (e.g., severe frown), and

A physiological component. To that others add to it the motivational state or action tendency or cognitive processing.

To William James that started a debate in 1884, and Scherer & Johnstone (2001), emotions are an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism.

Lang (1995), on the other hand, provides a more general biological definition of emotions, he states that emotions are biologically based action dispositions that have an important role in the determination of behavior.

However, this two definitions barely shade light to the understanding of emotions, especially if we try to understand how it affects technology usage. Much less if we want to examine its influence on user's trust in technology.

But above definitions demonstrate how emotions can be a multi-component term. Argue for the need to distinguish between emotions such as anger, fear, sadness, happiness, disgust, and surprise. Suggesting as well that they are present from birth, and have distinct adaptive value, and differ in important aspects such as appraisal, antecedent events, behavioral response, and physiology (Ekman, P., 1992).

What differentiates above terms are the dimensions we focus on. Proponents of this dimensional view have suggested that all emotions can be in a two-dimensional space as coordinates of valence and arousal (Lang, P. J, 1995). The valence dimension refers to the hedonic quality or pleasantness of affective experience and ranges from unpleasant to pleasant (Scherer, K. R., 2005).

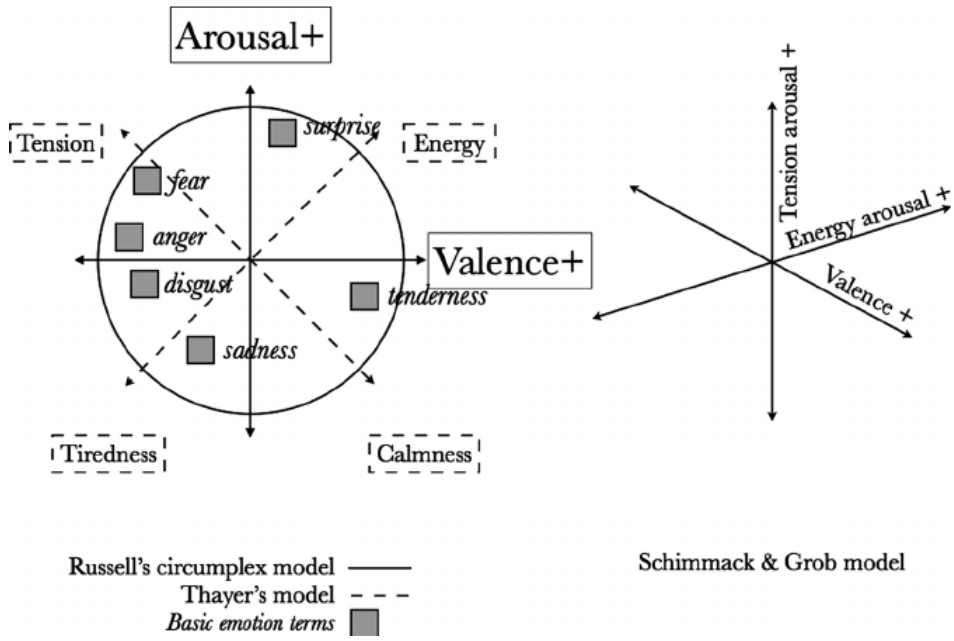

Figure 1: Russell's circumplex model. source Russell's (1980)

2.1.1 UNDERSTANDING EMOTIONS

According to Scherer (2005) emotions can be measured as:

> (1) a **continuous** change in appraisal processes at all levels of central nervous system processing (i.e. the results of all the appraisal checks, including their neural substrata),
>
> (2) a **response** patterns generated in the neuroendocrine, autonomic, and somatic nervous systems,
>
> (3) a **motivational** change produced by the appraisal results, action tendencies (including the neural signatures in the respective motor command circuits),
>
> (4) a **facial and vocal expression** patterns as well as body movements, and
>
> (5) the nature of the subjectively experienced **feeling** state that reflects these component changes.

In other words, emotions can be expressed psychologically and physiologically and they represent and adaptation to changing environmental.

- Psychologically, emotions alter attention, shift behaviours.
- Physiologically, emotions organize the responses of disparate biological systems (e.g., facial, muscular tone, voice tone).

An example of a Psychologically state of emotion is the participant reaction to external or internal stimuli. It reaction provide a feedback response reflected psychologically. It reaction can also be manifested verbally in a form of feelings. But, above classification help us to better understand emotions but does not explain how to measure it. Illustrates although how complex it is to do it as the emotional state of a subject varies and depends on other conditions like:

- Environmental conditions (quiet, comfortable seat, light intensity, room temperature), or
- Subject mental set, and so on.

For instance, delays, even short ones, between the activation of emotion and the assessment of emotion by an experimenter can introduce measurement error (Coan & Allen, 2007). Further compounding this error is the prospect that the time course of an emotional response varies by emotion response system (e.g., a facial expressive behavior may have a faster offset than emotion experience).

In literature, we can find several methods that help us to do this measurement. Those include approaches like self-report data or body physiological variations. The most common practice of assessing it is using self-report to stimulus. In this case, the subject emotional state can be labeled by the subject by self-reporting his or her cognitional states (e.g. the Geneva Emotion Wheel (Coan & Allen, 2007)).

Another approach is to manipulate subject feelings in diverse directions by using the emotion elicitation techniques. Although the reliability of these methods depending on environmental conditions. The costs of delay are well illustrated by the common practice of assessing self-report responses to films retrospectively using questionnaires. As time elapses between the film's (to stimuli emotion states) and the questionnaires completion, we can face the problem that the elicited affect is likely to fade and be distorted by errors or systematic biases in recall (Levenson, 1988).

To avoid problems associated with delayed retrospective reports, and to obtain continuous measures of experience that parallel continuous measures of other response domains (e.g., behavior and autonomic psychophysiology), there has been a growing interest in rating dial methodologies, which afford continuous measures of emotion experience, in either online or cued-review rating formats (Levenson & Gottman, 1985).

To help better understand concepts in the next paragraphs we will describe most commonly used methodologies in the field of Human-Computer Interaction (HCI).

2.1.2 RESEARCH IN EMOTIONS

In HCI, physiological computing has gained popularity as a method to study user experience and users' reaction to interaction, and also as a technique that can be used particularly in interactive systems, as a field in which human physiological data act as system input in real time (Levenson, 1988).

In other words, emotion measured by peripheral variables (autonomic system):

1. Heart Rate (HR),
2. Electrodermal (EDR, GSR),
3. Electromyography (EMG),
4. Temperature,

5. Electroencephalography (EEG).

Electroencephalography (EEG) can indicate Theta variations (4-7 c/s) can be an indicator of emotion variations. Changes in the Heart Rate (HR), as well as variation in skin Temperature, can also be an indicator of emotional stimuli.

Although, those measurements are the only indicator and need to be examined together with the environmental conditions.

For example, facial EMG has been found to be a successful method primarily in discriminating positive emotions from negative ones (Ravaja, 2004). In research, facial EMG activity is usually recorded over three facial muscle areas: the zygomaticus major (the cheek muscle area that activates during smiling), corrugator supercilii (the brow muscle area that activates during frowning), and orbicularis oculi (the periocular muscle area that activates during the so-called 'enjoyment smile' (Cacioppo, Tassinary & Berntson, 2000). A large body of evidence shows that the processing of pleasant emotions is associated with increased activity within the zygomaticus major muscle area and that processing of unpleasant emotions evokes higher activity in the corrugator supercilii muscle area during affective imagery (Ravaja, 2004).

On the one hand EMG has proved itself to be a good method to detect emotions, especially in contexts wherein little social interaction is involved, such as viewing of emotion-eliciting material (e.g., emotional images and news items) on a computer screen, on the other it is important to distinct weather it is real emotion, or it is a social signal without any connection to emotional experience.

In addition to EMG, signals captured from the brain in the central nervous system have been proved to provide informative characteristics in responses to the emotional states. Power spectra of the EEG were often assessed in several distinct frequency bands, such as delta (δ: 1–3 Hz), theta (θ: 4–7 Hz), alpha (α: 8–13 Hz), beta (β: 14–30 Hz), and gamma (γ: 31–50 Hz), to examine their relationship with the emotional states (Lin et al., 2010). EEG has a considerable advantage as researchers can continuously gather information about the users' affective states. As was mentioned, it's also useful in cases where people don't perform any emotional movement or facial expressions. A significant amount of work has been conducted by Picard et al. at the Massachusetts Institute of Technology (MIT) Laboratory, showing that certain affective states may be recognized by using physiological

data, including heart rate, skin conductivity (SC), temperature, muscle activity, and RSP velocity. They used personalized imagery to elicit target emotions from a single subject who had two years of experience in acting and they achieved an overall recognition accuracy of 81 percent for eight emotions by using hybrid linear discriminant classification. Thatcher et al. (2011) used movie clips based on the study of Vance, Elie-Dit-Cosaque & Straub (2008) for eliciting target emotions from 29 subjects and achieved an emotion classification accuracy of 83 percent using the Marquardt Backpropagation algorithm (MBP). In the study by Gefen, Karahanna & Straub (2003), the IAPS photoset (Kim & Benbasat, 2003) is used to elicit target emotions with positive and negative valence and variable arousal level from a single subject. The arousal and valence dimensions of the emotions were classified separately using a neural network classifier, and recognition accuracy rates of 96.6 percent and 89.9 percent, respectively, were achieved.

Another way to assess emotions and level of arousal is questionnaires. And one of the most popular among the existing self-reporting tools is the Self-Assessment Manikin (SAM) proposed by Kassim, Jailani, Hairuddin & Zamzuri (2012), this method is widely used in psychological studies, as well as in marketing and advertising. SAM is a scale that measures the dimensions of pleasure, arousal, and dominance (also called “PAD”) using

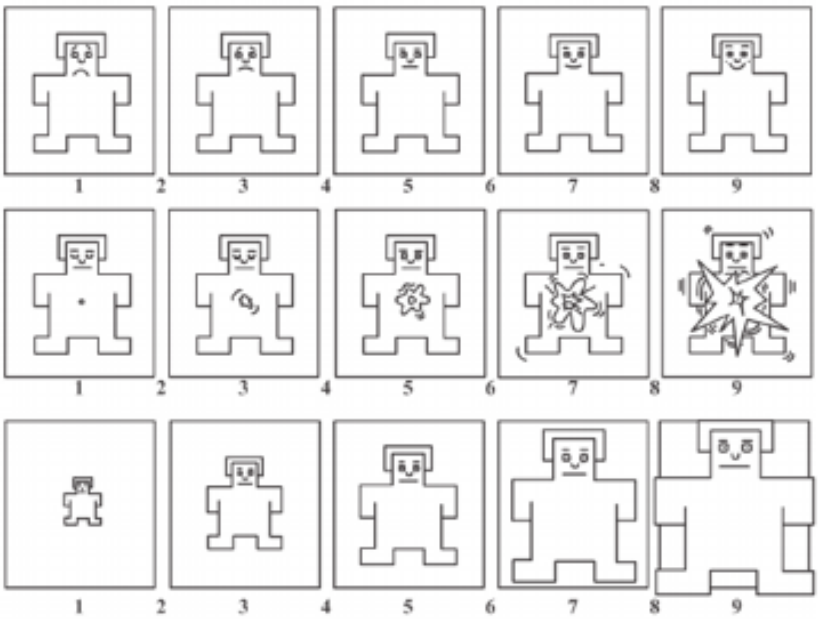

Figure 2: Self-Assessment Manikin is a scale. Source Bradley & Lang (1994)

A series of graphic abstract characters horizontally arranged according to a 9-point scale. This system was widely used in psychology and emotion research and have a lot of citation in Google Scholar. Alberto Betella et al. created a new generation of SAM called Affective Slider. The AS is composed of two sliders that measure basic emotions regarding pleasure and arousal on a continuous scale that author systematically calibrated to the SAM. This system has a lot of advantages, one of them is the possibility of collecting ratings on continuous scales that allow for more accurate high-resolution measurements, as opposed to the SAM which records data upon a relatively condensed Likert scale.

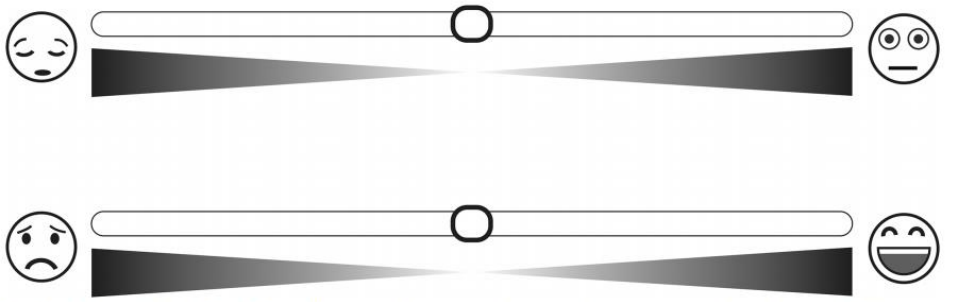

Fig 2. The "Affective Slider" (AS) is a digital self-reporting tool composed of two sliders that measure arousal (top) and pleasure (bottom) on a continuous scale. The AS does not require written instructions and it is intentionally displayed using a neutral chromatic palette to avoid bias in ratings due to the emotional connotation of colors. See text for more details.

Figure 3: Affective slider. Source Betella & Verschure (2016)

Also, the widely-known method to measure emotion is the emotional wheel that places all different emotional constructs in a multidimensional space. This space has two-dimensional arrows - one that indicates user's valence emotional state (negative, positive) and another that indicates user's arousal states (low to high).

That dimensional approach was first proposed by Wilhelm Wundt in 1905, in his attempted to develop a structural description of subjective feeling as it is accessible through introspection. He suggested that these feelings could be described by their position in a three-dimensional space formed by the dimensions of valence (positive-negative), arousal (calm–excited), and tension (tense–relaxed).

This method is commonly used to ask for the self-report instrument of emotion by the respondent. This method had proved to be simple and straightforward and quite reliable (Betella & Verschure 2016). Nonetheless, this method has also the drawbacks, one of them is the difficulty of knowing whether the valence dimension describes the intrinsic quality of an eliciting object or the quality of the feeling. For instance, extremely intensive anger is likely to be characterized by high arousal whereas intense sadness may be accompanied by very low arousal.

This was the basic principle of the new generation wheel - the Geneva emotion wheel, in which to separates emotions in four quadrants: negative/low control, negative/high control, positive/low control, and positive/high control. Is applied as a self-report mechanism where users in a dimensional response format, respondents may indicate their feelings based on these four abstract dimensions.

Those abstract dimensions that are thought to constitute the basic structure of emotions (valence, arousal, dominance) an example of such instrument is the Self-Assessment Manikin Test developed by Bradley & Lang (1994). Their respondents are requested to indicate their feelings about a limited number of discrete emotions (e.g., anger, happiness) on a scale corresponding to the intensity of the feelings.

This format has proven to be easy to use because discrete emotion terms correspond to the natural way of talking about emotions. Also, the results can readily be interpreted (Sacharin, Schlegel & Scherer, 2012)

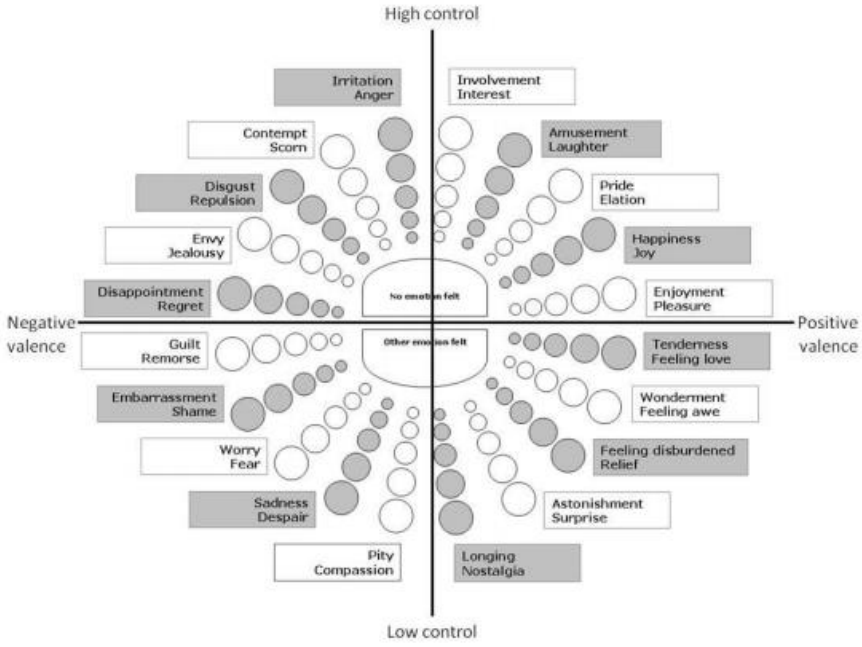

Figure 4: Geneva emotion wheel. source Sacharin, Schlegel & Scherer (2012)

As it is easy to implement and had proved to be simple and straightforward and quite reliable (Colomo-Palacios et al., 2011) method, it also had been commonly adopted in Human-Computer Interaction contexts. Nonetheless as all methods that measure complex constructs such as emotions it also includes some drawbacks, for instance, the difficulty of knowing whether the valence dimension describes the intrinsic quality of an eliciting object or the quality of the feeling. For instance, extremely intensive anger is likely to be characterized by high arousal whereas intense sadness may be accompanied by very low arousal.

Evidence exists that these methods are used in the field of Human-computer interactions to elicit and measure emotions.

2.1.3 EMOTION ELICITATION

Several techniques have been used to elicit the emotions in the laboratory, like.

- Including hypnosis (i.e., Bower, 1983);
- imagery (e.g., Lang, 1979),
- music (e.g., Sutherland et al., 1982),
- facial muscle movements (Ekman et al., 1983),

- interaction with trained confederates (e.g., Ax, 1953),
- the Velten/self-statement technique (repeating phrases with emotional content, Velten, 1968) and even
- drugs and sleep deprivation (cf. Martin, 1990; Hagemann et al., 1999).

Although as some of these methods can be associated with ethical problems such as drug intake, the most widely adopted methodology to elicit emotions is the International Affective Picture System (Coan & Allen ,2007).

PICTURES FOR EMOTION ELICITATION

The International Affective Picture System is a set of static images based on a dimensional model of emotion. It has been shown that the startle reflex can be inhibited by the viewing of positive pictures and accentuated by the viewing of arousing negative pictures, an observation that reveals the differential role of positive versus negative affect in the modulation of attention and orienting (Lang, Bradley & Cuthbert, 1997). This method has proved itself as one that can be used in a cross-cultural studies as anger, disgust, fear, sadness, and enjoyment are the basic universal emotions that can be decoded from the facial cues (Mikels et al., 2005).

FILM CLIPS FOR EMOTION ELICITATION (IAPS)

In addition to the IAPS, film clips have also been widely used as well and are accepted as stimuli in the field of emotion research. Films have been used for many decades with the first use of film clips for emotion elicitation described in 1930, in a study of the effects of anger, fear and sexual arousal on blood pressure, for instance in one study by Scott in 1930. Since then, many studies have used films as emotion elicitors to study phenomena such as sad mood induced smoking behaviour (Fucito & Juliano, 2009), emotional modulation of the acoustic startle reflex, the readiness for affective reactions related with frontal brain asymmetry and the effect of emotion on eating behaviour (Coan & Allen, 2007). Some researchers state that films could be even preferable that pictures in elicitation of emotions, as in contrast to pictures films are dynamic and thus thought to be more like real life. For instance, Gross and Levenson (1995) state: “Films also have a relatively high degree of ecological validity, in so far as emotions are often evoked by dynamic visual and auditory stimuli that are external to the individual”.

However, in the study by Coan & Allen (2007) authors revealed, that film clips were less effective than pictures in producing negative emotions. This corresponds to the finding that negative film clips led to lower emotional arousal than at least the negative 3 pictures (three consecutive images of the same valence and with similar content presented) variant.

Should notice, that the important determinant of participants' responses to film stimuli is the physical setting in which films are presented. Emotional reactivity to films has been associated with mundane aspects of the experimental situation such as room lighting (Knez, 1995), larger display size (Detenber & Reeves, 1996), warmer room temperature (Anderson, Deuser, & DeNeve, 1995), and color (as opposed to black and white: Detenber, Simons, & Reiss, 2000). Even within the context of single-subject paradigms, subtle changes in the physical arrangements may influence reactivity via the implied social presence of others (Fridlund, 1992). For example, the presence of video recording equipment in a participant room may increase self-consciousness that dampens or enhances behavioral responses (Coan & Allen, 2007).

2.2 TRUST IN COMPUTING

Like emotions, trust itself is a complex term to define. Literature gives a lot of knowledge to what trust is, in psychology and sociology and last years there was a huge interest in studying the effects of trust in technology. To understand the importance of trust in human-computer interaction and technology we first need to define it. McAllister (1995) in his work gives an overview of several studies that define trust as a cognition-based concept in that "we choose whom we will trust in which respects and under what circumstances, and we base the choice on what we take to be 'good reasons,' constituting evidence of trustworthiness". Simmel states that the amount of knowledge necessary for trust is somewhere between total knowledge and total ignorance. Given total knowledge, there is no need to trust, and given total ignorance, there is no basis upon which to rationally trust. People make emotional investments in trust relationships, genuine express care and concern for the welfare of partners, believe in the intrinsic virtue of such relationships, and believe that these sentiments are reciprocated (McAllister, D. J. 1995)

As said before, trust has been studied across several disciplines, including economics, sociology, and psychology. And across these disciplines, trust was defined differently, and

in this work, we take definition given by Rousseau, Sitkin, Burt, and Camerer (1998) that integrates a multidisciplinary approach to trust: Trust is the willingness to accept vulnerability based upon positive expectations about another's behavior.

2.2.1 UNDERSTANDING TRUST

The previous study was mostly focused on trust in technological perspective, addressing such issues as privacy, security, infrastructure, etc. But with the development of the technology and widespread of online e-commerce, online interaction systems and especially artificial intelligence, in other words, human-like technologies, researchers began to also address the behavioral aspects of trust.

The first view is that of Friedman et al. (2000) who opine “people trust people, not technology”. They argue that unlike humans, technology does not possess moral agency and the ability to do right or wrong and hence it should be viewed as being a participant in a trust-distrust relationship between a user and the person who programmed the technology. This view is also held by Olson and Olson (2000), who argue that when people interact through technology, it is not the technology that needs to be trustworthy. Instead, the trust-distrust relationship is between two humans independent of whatever technology they would use.

Essentially, these views can be distilled to the following two points, which are:

- People cannot enter a relationship with technology, and
- The question of people trusting the technology does not arise as people cannot develop a trusting relationship with technology because it lacks volition and moral agency (Gulati, Sousa & Lamas, 2017).

Firstly, research has demonstrated that computers can act as social actors and people can enter into relationships with and respond to them in a way comparable to responding to other people (Benbasat & Wang, 2005). Studies have shown that people assign personalities (Nass et al., 1995), gender (Nass, Moon & Green, 1997) and readily form team relationships with computers and consider them as teammates (Nass, Fogg & Moon, 1996). Similarly, users can enter trusting relationships with a technological artifact. A study by McKnight et al. (2011) showed that people could and do develop a trusting

relationship with an IT artifact such as Microsoft access or excel.

Concerning the role of trust within HCI, it has shown to be a key factor in reducing risk and uncertainty associated with a technological interaction, creating positive and meaningful experiences with technology and is crucial in helping a user adopt and maintain a gradual and steady relationship with the system (McKnight et al., 2011; Lankton, McKnight & Tripp, 2015; Söllner et al., 2012).

In fact, researchers have shown trust in technology to influence acceptance of various technologies such as online recommendation agents (Benbasat & Wang, 2005), business information systems, e-commerce portals, and knowledge management systems (Thatcher et al., 2011). Some researchers have measured technology trust using the human-like trust constructs of integrity, ability/competence, and benevolence that researchers have traditionally used to measure interpersonal trust (Vance, Elie-Dit-Cosaque & Straub, 2008). For instance, research on trust in software agents has employed interpersonal trust beliefs (i.e., competence, integrity, and benevolence) to represent trust in technology because software agents have some human-like characteristics, such as giving advice and interacting with the user on-screen (Benbasat & Wang, 2005). In contrast, other researchers have measured technology trust using system-like trust constructs such as reliability, functionality, and helpfulness (Mcknigh et al., 2011) to use. For example, people interface with other people on Facebook, but they neither obtain advice directly from Facebook itself nor interact with Facebook as a person or quasi-person. While trust in social networking websites research has generally examined interpersonal trust attributes, people may trust Facebook in other ways. For example, McKnight (2005) explains that people may trust a technology because it provides specific functionality, operates reliably, and is helpful to its users. Thus, people may be willing to depend on Facebook (or any technology) because it has these technology-related attributes that make it trustworthy (Lankton & McKnight, 2011). In the same study by Lankton & McKnight (2011), researchers claim that technologies are not helpful in terms of volition or moral agency and demonstrate its helpfulness through help functions that help to achieve a goal and individuals who perceive that a technology can provide the help needed will perceive fewer risks and uncertainties associated with technology use.

Overall, trust has shown to be a key factor in technology adoption and user satisfaction with technology (Kassim, Jailani, Hairuddin & Zamzuri, 2012) and its absence could lead

to a user proceeding more cautiously when using the technology and taking unnecessary time to think through their actions, which will ultimately lead to dissatisfaction, failure to continue using the system and the user not being able to fully realize the potential which the technology has to offer (Constantine, 2006).

2.2.2 RESEARCH IN TRUST

In information systems domains trust is usually examined and studied in terms of the trust to people, where technology is used as a mediator, without the regard to trusting the technology itself. Information systems trust research primarily examines how trust in people affects IT acceptance. For example, trust in specific Internet vendors (Hubert et al., 2018; McKnight, Choudhury & Kacmar, 2002; Kim & Benbasat, 2003) has been found to influence Web consumers' beliefs and behavior. In general, Internet research provides evidence that trust in another actor (i.e., a Web vendor or person) and/or trust in an agent of another actor (i.e., a recommendation agent) influences individual decisions to use technology. Comparatively, little research directly examines trust in a technology, that is, in an IT artifact. Some researchers have conceptualized and measured trust in technology as if the technology were a human. That is, they have measured technology trust using the human-like trust constructs of integrity, ability/competence, and benevolence that researchers have traditionally used to measure interpersonal trust (Lankton, McKnight & Tripp, 2015). From the other side, researchers have measured technology trust using system-like trust constructs such as reliability, functionality, and helpfulness (Mcknight, Carter, Thatcher & Clay, 2011). It is reasonable for users to associate human-like trusting beliefs with an online recommendation agent that has voice and animation. In his recent studies McKnight (McKnight, Choudhury & Kacmar, 2002; McKnight, Choudhury & Kacmar, 2002; Mcknight, Carter, Thatcher & Clay, 2011) split trust to two subcategories: system-like trust vs. human-like trust, and state that researchers should correctly differentiate each of them during the research as it could make dramatic impact on a result.

Table 1. Major Trust in Technology Constructs Used	
Human-like trusting beliefs	**Corresponding system-like trusting beliefs**
Definition	**Definition**
Integrity: the belief that a trustee adheres to a set of principles that the trustor finds acceptable (Mayer, Davis, & Schoorman, 1995).	**Reliability:** the belief that the specific technology will consistently operate properly (McKnight et al., 2011).
Ability: the belief that the trustee has the group of skills, competencies, and characteristics that enable them to have influence within some specific domain (Mayer et al., 1995). **Competence:** the belief that the trustee has the ability to do what the trustor needs to have done (McKnight et al., 2002).	**Functionality:** the belief that the specific technology has the capability, functions, or features to do for one what one needs to be done (McKnight et al., 2011).
Benevolence: the belief that the trustee will want to do good to the trustor, aside from an egocentric profit motive (Mayer et al., 1995).	**Helpfulness:** the belief that the specific technology provides adequate and responsive help for users (McKnight et al., 2011).
Exemplar Studies	**Exemplar Studies**
Vance et al. (2008): m-commerce portal Wang & Benbasat (2005): online recommendation agent	Lippert & Swiercz (2005): human resource information system McKnight et al. (2011): spreadsheet Muir & Moray (1996): simulated pump mechanism Thatcher et al. (2011): knowledge management system

Table 2: Human-like characteristics vs System-like. Source McKnight et al. (2015)

Usually, during the research that measure trust the survey methodology and physiological measurement are used. In a study by McKnight (2015) researchers used a survey methodology to test the hypotheses about technology humanness and trusting beliefs. Researchers state that by surveying rather than controlling the social context in an experiment, they could capture differences in the two technologies' humanness and detect relationships among constructs in naturally occurring situations. It can be important to detect what features of humanness result in human-like and system-like trust having influence on the outcomes. They analyzed whether, in a naturally occurring environment, perceived humanness differs between two technologies. They also tested whether these differences in humanness result in certain trusting beliefs having influence on the outcome variables.

In another study by Lankton & McKnight (2011), researchers also used a survey to test their hypotheses. The survey used social networking websites as the target technology. The study participants were junior and senior business college students in a required introductory information systems course at a Midwestern U.S. public university. College students are an appropriate sample for investigating Facebook trusting beliefs because a sizeable percentage of Facebook users are college-aged. 40% of unique Facebook users

were age 18-24 in 2006, and 29% were age 18-24 in 2007.

In a study by Hubert et al. (2018) physiological measurements were used. The aim of this study was to investigate how the consumer personality trait impulsiveness influences trustworthiness evaluations of online-offers with different trust-assuring and trust reducing elements Kim & Benbasat (2003) by measuring the brain activity of consumers. Based on an analysis of a fMRI data set, researchers identified the impact of consumer impulsiveness differentiated into hedonic and prudent shoppers on trustworthiness evaluations within an online environment; and revealed differences about the neural correlates of interaction between consumer's impulsiveness and trustworthiness evaluations.

2.2.3 MEASURING USERS TRUST TOWARDS TECHNOLOGY

One of the factors underpinning the research problem is the challenge in defining and measuring trust due to its multidimensional nature. Researchers present different models to measure trust and like emotions trust can be measure using different dimensions. Similar to emotions perceptual trust is measured mainly using self-report and psychometric questionnaires. One of the earliest works towards a psychometric instrument to measure trust in HCI is that of Madsen and Gregor (2000) who developed a human-computer trust scale but there is no full validation which has been reported on this, and the empirical validity of the scale is questionable because of its low sample size. There have been other empirical attempts to come up with statistical models and scales to measure trust (Gulati, Sousa & Lamas, 2017). However, these are proposed within a context such as e-commerce or social networking, and the results are difficult to generalize.

Users trusting beliefs in technology can be measured using three human-like trusting beliefs: Integrity, competence, and benevolence (Gulati, Sousa & Lamas, 2017). Ability/competence is the belief that a person has the skills, competencies, and characteristics that enable them to influence some specific domain. Benevolence is the belief that a person will want to do good to the trustor aside from an egocentric profit motive. Integrity is the belief that a person adheres to an acceptable set of principles. Researchers have used these human-like trusting beliefs to study trust in technology because people tend to anthropomorphize technologies and ascribe to them human motivation or human attributes (Lankton & McKnight, 2011).

For example, Wang & Benbasat (2005) studied trust in online recommendation agents (RAs) and found that these human-like trusting beliefs significantly influenced individuals' perceived usefulness and intention to use RAs. Vance et al. (2008) use these beliefs to study m-commerce portals (Söllner & Leimeister, 2013).

Sousa et al. (2014) proposed a model for the human-computer trust that depicts trust as a construct informed by seven individual qualities, such as motivation, willingness, reciprocity, predictability, honesty, benevolence, and competence, and determines the extent to which one relates with one's social and technical environment.

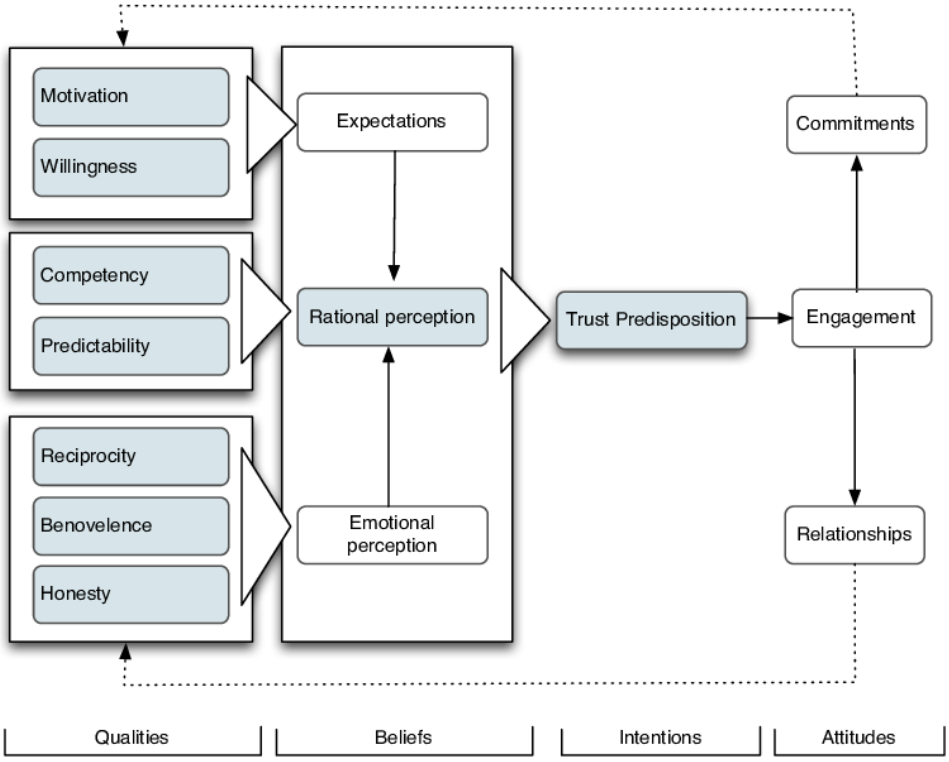

Figure 5: The Human-Computer Trust model, source Sousa et al. (2014)

Besides, scholars started to investigate the topic from a biological perspective during the past decade. These biological researches can be classified into three categories: genetics, endocrinology, and brain functionality. One of the reasons for such research is that such methods as fMRI (functional magnetic resonance imaging) became available. A further reason is an insight that all human behaviour that varies among individuals is associated, at least partly, with biological factors, in particular, those related to the nervous system. So far studies that use fMRI revealed several brain regions associated with trust. Moreover, endocrinological studies show that some hormones affect trust. Finally, recent gene-based research has demonstrated that at least a moderate degree of human trust behaviour is

genetically predetermined (Riedl & Javor, 2012). In the same study researchers unified the most widely used methods to measure trust:

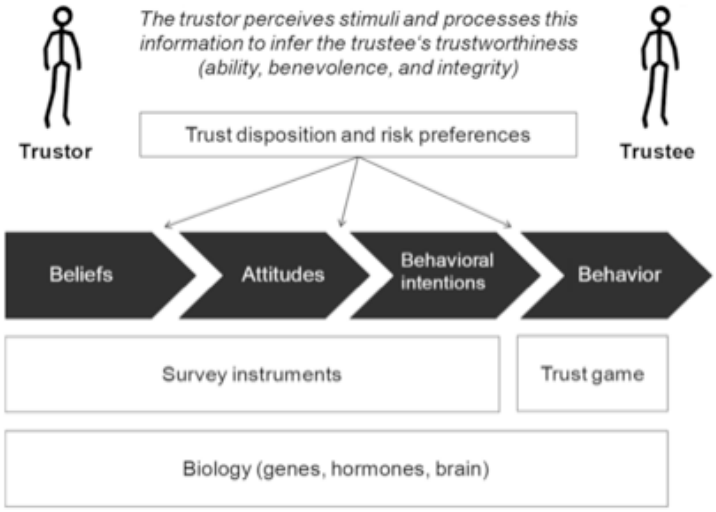

Figure 6: Structure of trust situation, source Riedl, R., & Javor, A. (2012)

The *trust game* was developed to measure both trust and trustworthiness as actual behaviour of players in an economic exchange game (Riedl & Javor, 2012). Same study reports and unifies some studies, where fMRI and trust game is used to study a neural correlate of trust. The table below illustrates studies of human trust behaviour using trust games and fMRI:

Table 1
FMRI Studies On Human Trust Behavior and Associated Brain Regions

	Baumgartner et al. (2008)	Delgado et al. (2005)	King-Casas et al. (2005)	Krueger et al. (2007)	Winston et al. (2002)	Dimoka (2010)	Riedl et al. (2010)
Sample size (female/male)	49 (0/49)	12 (5/7)	96 (n.a.)	44 (22/22)	14 (6/8)	15 (6/9)	20 (10/10)
Experimental paradigm (stimulus)	Trust game	Trust game	Trust game	Trust game	Faces	eBay websites	eBay websites
Mental processes/brain regions							
Reward							
Striatum	●	●	●			●	●
Thalamus	●		●		●		●
Uncertainty, risk, ambiguity, fear/Memory							
Amygdala*	●				●	●	
Insular cortex	●	●			●	●	●
Hippocampus, parahippocampus gyrus*		●			●		●
Cognitive conflict							
Cingulate cortex	●	●	●	●		●	●
Mentalizing and deliberate thinking							
Frontal cortex				●		●	●

Note. The table lists brain regions that are mentioned in at least three of the seven studies. The brain regions are categorized along classes of mental processes that are important in trust situations: (a) reward processing, (b) uncertainty, risk, ambiguity, and fear processing, (c) processing of cognitive conflict, and (d) mentalizing and deliberate thinking. Note that we do not claim that a particular brain region is exclusively associated with one mental process. Rather, the relation between brain regions and mental processes is best described by a "many-to-many-concept." That is, activation in a specific brain region might be associated with several mental processes, and a specific mental process might be associated with activation in several brain regions (Price and Friston 2005). The asterisk (*) indicates that the amygdala, together with the hippocampus and parahippocampus gyrus, is also a crucial brain structure for emotion-based memory and learning, which are of particular importance for trust decisions independent of the experimental paradigm. Finally, note that although five out of the seven studies used a mixed-gender sample (with the exception of the study by Baumgartner et al., and with information unavailable in King-Casas et al.), only the Riedl et al. study has a focus on gender differences in brain activation in trust decisions.

Figure 7: FMRI studies on Human Trust Behaviour. source Riedl, R., & Javor, A. (2012)

2.2.4 THE INTERPLAY BETWEEN TRUST AND EMOTIONS

Trust plays a big role in a decision-making process, for instance, it is a key role in making online purchases (Kim, Ferrin & Rao, 2008), at the same time emotions also have been shown to affect a variety of decision-making processes, and complex decisions, such as the decision whether to trust a stranger, a politician, or a potential competitor, are particularly likely to be influenced by a person's emotional state (Dunn & Schweitzer, 2005). In the last few decades, some findings and theoretical models have suggested that emotion could have a powerful influence over cognition and decision-making, with different emotions having different effects (Lerner & Keltner, 2001).

Some studies suggest that people in positive moods will be more trusting. People in positive moods will have an easier time accessing memories where trusting behavior led to a positive outcome, they also are more altruistic and make optimistic predictions, especially concerning the future actions of others. They may take their positive mood as information about the trustworthiness of others, and about the possible consequences of engaging in trusting behavior, another study by Forgas (1998) shows that positive mood makes participants more cooperative in negotiating situations. However, recent studies by Capra (2004) finds no effect of mood on a play in the trust game. In a study that was already mentioned (Lerner & Keltner, 2001) research has shown that distinct emotions with the same valance can produce wildly divergent effects on cognition and decision-making. Capra (2004) induces subjects to feel either positive or negative moods using the AEMT and then has them play a series of economic games including the trust game. She finds no effect of the mood induction on a play in the trust game. While this shows that positive or negative valance by themselves have no measurable effect on trust, it does not differential between specific negative and positive emotions. These leaves open the possibility that other dimensions of emotional experience can have an impact on trusting behavior. At the same time, some studies have shown that emotions with the same valence, but different control appraisals, have different effects on judgment, for example, researchers (Lerner & Keltner, 2001) found that fear and anger, two emotions with negative valence, had significantly different effects on risk assessments. They found that people feeling angry had more optimistic risk assessments than did people feeling fear and those different appraisals of control mediated this relationship.

2.3 RESEARCH APPROACH AND METHOD

In this section, we briefly address the main research principles and illustrate how the experimental research approach fits on the case study described in this thesis.

Research design principles

Research usually are the design from two main principles, Inductive and or deductive nature. In the inductive principle, the researcher selects the instrument for data collection, analysis the results and see if any patterns emerge. The inductions principle leads to a Hypothesis/theory to be tested.

On the other hand, when the deductive principle assumes the main principle of research, researchers usually start with a Working theory or Hypothesis and then move towards hypothesis testing principle. He or she tries to verify if the principle is confirmed, refuted or modified (Gray, 2014). This assumes experimental design research in nature.

In other words, the deductive principle, therefore, elaborates a set of principles or ideas that are then tested through empirical observation or experimentation.

According to Gray (2014), the main stages in the deduction process includes

- Organizational mission Theory
- Hypothesis
- Operationalize
- Testing by corroboration or attempted falsification
- Examine outcomes
- Modify theory (if necessary)

But, before such experimentation can take place, underlying concepts must be operationalized (made measurable) as so they can be used to confirm refuted or modified.

Research through design

Research through Design (RtD) is an approach to conducting scholarly research that employs the methods, practices, and processes of design practice with the intention of generating new knowledge. For the first glance, this approach to research can look like design practice, but it is more systematic and reflective process of review the detailed documentation that was done during the design process. Key to this work of capturing and translating primary design knowledge into broader academic knowledge is how design processes are documented; design documentation is a key raw material out of which such

knowledge is constructed (Riedl & Javor, 2012). One challenge was noted in 1989 by Caroll & Kellogg: the thing proceeds theory instead of theory driving the creation of

new things. They said that mouse was invented first, and then a lot of studies were conducted to prove that this is a good decision for direct manipulations of the GUI. In response to this challenge, we see RTD as a way for many new things to enter into HCI that can spawn new theory. At the same time, these new things can be informed by current theory, creating an ongoing dialogue between what is and what might be (Zimmerman, Forlizzi & Evenson, 2007).

Experimental and quasi-experimental research

Emphasizes on:

- Reproducing the techniques of the laboratory experiment with highly structured methods.
- The generation of initial hypotheses.
- The control of variables.
- Accurate (quantitative) measurement of outcomes.
- Generalization from samples to similar populations.

2.3.1 THE EXPERIMENTAL APPROACH

We as researchers have a variety set o options to choose from, but most derive from two main principles. More quantitative or qualitative research design in nature. In this case, the use of experimental methodology assumes a more qualitative nature and follow a more deductive principle as described by David in his book (2014). Although the research design described as followed, in spite of being experimental do not assume that there is a control group and a test group. Instead, it aims to explore the relationship between two or more variables through a correlational analysis. It intends to determine if and to what degree the variables (trust and emotions) are related. And does not necessarily assumes that one causes the other. The goal of a research project is to understand a specific situation and develop localized solutions, generalizability is not necessary or desirable. Instead, other models of scientific merit are emphasized, including the notion of cross-contextual transfer of solutions and dependability of the research results (Hayes, 2011).

Utilizing cycles of inquiry that include planning, action, reflection, was action being undertaken is continually designed and evaluated with research results emerging throughout these cycles. Design research can incorporate multiple methods and welcomes the use of both qualitative and quantitative methods. The methodology is open-ended and aims to find the best possible answer for a given hypothesis.

3 THE RESEARCH DESIGN CASE STUDY

The purpose of this case study was to describe the process of designing an experimental research design approach. As said before it main goal is to measure efficiently the relationship between two variables, trust, and emotions. The experiment was designed at Tallinn University as part of a Ph.D. study. Due to the complexity of the topics, the design time frame lasted 6 months. Both research team, specialists and the author started working in it at the end of November and the first two pilot tested were run in March. This case study aimed to,

- from one hand clarify and report the process of designing a research experiment to efficiently study the relationship between trust and emotions. on another
- to investigate what are the strengths and constraints of design such complex research studies.

As referred above this deductive approach, to find the interplay between trust and emotions was elaborated in a set of principles or ideas that needed to be tested through empirical observation or experimentation.

As Gray (2014) sated the main stages of this deduction process includes

- Organizational mission Theory
- Hypothesis
- Operationalize
- Testing by corroboration or attempted falsification
- Examine outcomes
- Modify theory (if necessary)

This case study reports the first 3 above items. Describes the process of selecting the Theory, building the Hypothesis and understand how to operationalize it. Most of the efforts on designing such experimentation were on operationalization and trying to understand the underlying concepts that made measurable the variables emotions and trust from two main data collection instrument psychometric measurement and physiological measurements. As so they can be used to confirm refuted or modified. Which include the

final 3 above mention items, testing by corroboration or attempted falsification, examine outcomes, modify the theory.

As addressed before, this work proposes to contribute by systematizing such experiment design process. Illustrating the readers on how to approach when examining the interplay between emotions and trust in technology using quantitative approaches and collecting data through two distinct operationalization measurement physiological and psychometric.

Mainly aims to seek answers to the research question like:

- RQ1. How can we use physiological computing measurements to study the interplay between trust and emotions?
- RQ2. How can we use psychometric computing measurements to study the interplay between trust and emotions?
- RQ3. What are the constraints of using experimental approaches to design such complex research study?
- RQ4. How effective is this research design approach?

This part will end by providing final considerations and a description of the main outcomes and a reflection on further steps.

3.1 RESEARCH METHODS

The author illustrates and describes in this chapter the research process of how the researchers designed the experiment. This process involved three main stages. Those are described in the following lines.

- **First step:** included an initial exploratory approach to understand the theory and main concepts associated with studying the effects of Trust and emotions in technology. Its findings reflect, as well how researchers initially conceptualize this experimental design and formulate the hypothesis.
- **Second step:** represents the operationalization.
- **Third step:** describes the examine outcomes and the proposed redesign for providing more accurate measurements.

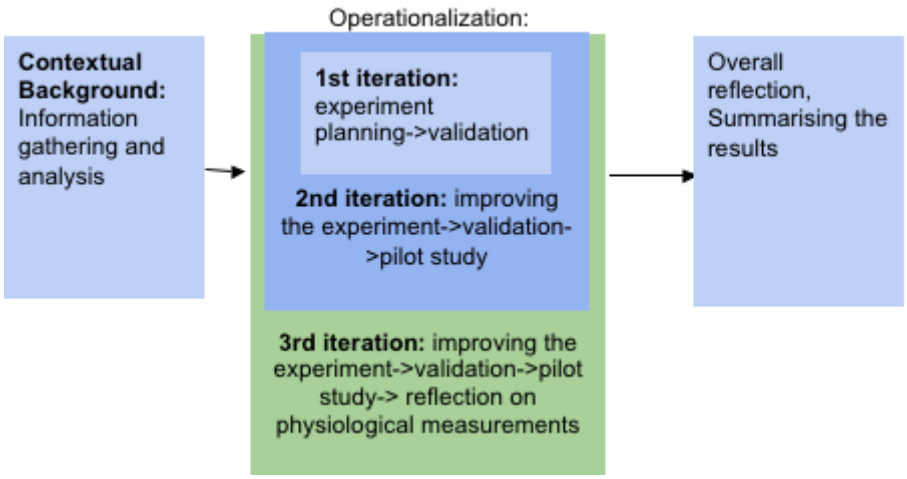

Figure 8: The experimental design process

The main research questions are:

RQ1. How can we use physiological computing measurements to study the interplay between trust in technology and emotions?

RQ2. How can we use psychometric computing measurements to study the interplay between trust in technology and emotions?

RQ3. What are the constraints of using experimental approaches to design such complex research study?

RQ4. How effective is this research design approach?

A contextual inquiry approach was used to observe and understand above questions. The author was involved on most key planning decisions. Punctually (in every design process step), outside experts were interviewed to brainstorm about the process and give specific contributions. Two pilot studies were as well performed to better understand what are the constraints of using experimental approaches. As well as how effective is this research design approach.

This helped the author to have a deeper overview of the overall strengths and weakness of the process. For instance, it helped to better understand:

- the overall implications of designing and experimental study,
- to understand how we can validate the design process; and
- to understand how emotion and trust in technology can be measured using physiological computing measurements and psychometric measurements.

- to understand as well how can we extract the data and how to formulate the hypothesis.

In the next lines, we illustrate each design step and present our findings and final considerations.

3.2 EXPERIMENTAL DESIGN: STEP 1

As addressed before the experimental design first step aimed to explore the main theoretical frameworks associated with the topics of trust and emotions. This included a search for similar studies performed in Human-computer interactions and understand how were they designed.

This initial information gathering process was important, as it revealed existing approaches in the research literature that studied the topic (i.e. the interplay between emotions and trust in technology). Also, helped to better understand what were the best measurement instruments and data analysis approaches.

The major questions that arise in this phase were:

- How to elicit and measure emotions?
- What technological artefact use as a trust stimuli?
- How to measure Trust towards the technological artefact?

For instance, from existing similar studies researcher found four main ways to elicit emotions by pronouncing emotional phrases, or writing down emotional memories, using video clips or images.

The researchers, in the end, opted to use videos instead of images, or phrases as the video database provided by Coan & Allen (2007) was clear to understand and provide more scientific evidence on how emotions were elicited. Literature review, also showed that to measure emotions and the level of arousal is better to use self-assessment and questionnaire measurements using a Likert scale of 1 to 5 points. Same regards the measurement of trust, they decided to use a self-report psychometric measurement.

Same regards to assessing user trust predisposition towards technology. The technological artifact chosen was google assistant as this artifact is classified as human-like technology.

This means that besides including anthropomorphic characteristics that make it close to human characteristics (e.g. voice) also included artificial intelligence that added non-deterministic characteristics. Both key features of trust in technology. This was based on a study done by Cowan et al. (2016), where researchers studied intelligent personal assistants and what influence the frequency of usage.

Additionally, as researchers were willing to use physiological measurements as well. They explore in literatures the different possibilities, but that become a hard decision to make. As besides the existing laboratory provided few alternatives, the existing ones could not measure trust directly. The option was to use the electroencephalography (EEG)[2] for measure the frontal cortex and Emphatical E4 wristband to measure variations of heart rate.

Main contributions from this phase included the hypothesis formulation, those were:

H1: Does positive and low arousal emotion influence trust in technology?

H2: Does positive and high arousal emotion influence trust in technology?

H3: Does negative and low arousal emotion influence trust in technology?

H4: Does negative and high arousal emotion influence trust in technology?

As well as the selection of the three main emotional states are to elicit.

- anger - negative and low arousal
- sadness - negative and high arousal emotion
- amusement - positive and low arousal emotion
- tenderness - positive and high arousal emotion

This first cycle ended by assembling a possible approach to operationalize the overall experimental process. This included 4 main parts.

- an initial part to measure participants initial trust levels (known as a pre-test measurement).
- a set of tasks to introduce participants to the artefact.
- a set of videos to induce 4 emotional states in the participants (emotion induction stimuli - based on an emotion elicitation principle).

[2] EEG Nautilus wireless research grade high quality brain wave sensor, this device tracks EEG data using 8 sensors. Data came as delta (δ: 1–3 Hz), theta (θ: 4–7 Hz), alpha (α: 8–13 Hz), beta (β: 14–30 Hz), and gamma (γ: 31–50 Hz) waves.

The data collection instrument included a set of psychometric measurements, and physiological measurements and where applied throughout the process.

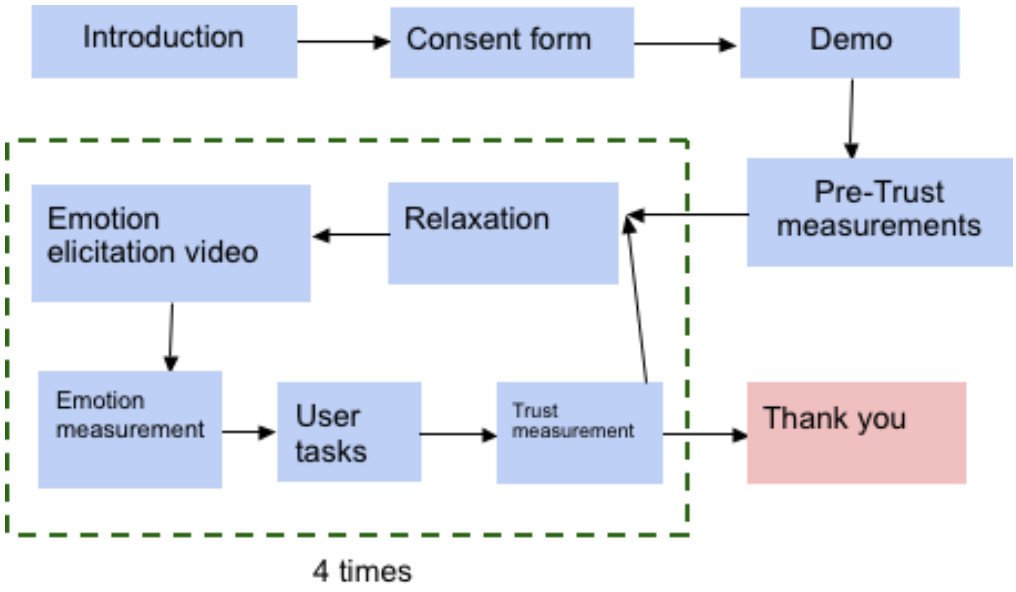

Figure 9: overall experimental design procedure, phase 1

As you can see from the above schema, the experiment design operationalization procedure started with:

- an introduction, where for the participant was presented a brief info about study, the procedure of the experiment and some privacy info,
- then participant signed a consent form.
- This was followed by a demo so participants get acquainted with the intelligent personal assistant in case they didn't use it before.
- Followed by pre-test measurement questionnaire to measure participant initial trust levels before inducing any emotional state. This trust measurement will be used during data analysis as a constant.

Next step was to present a random video to the participant (it worked as an emotional stimuli)

- Then the participants were asked to answer a set of questions with the help of google assistant.
- Finally, they needed to answer to an emotional and trust psychometric questionnaire.

This process of induction of emotional - the user tasks - and measurement was repeated 4 times. As participants needed to go through 4 emotional states: sadness, anger, enjoyment, tenderness. The reasoning for selecting the 4 emotional states are based on the emotional wheel (Sacharin, Schlegel & Scherer, 2012) principles. For further information please read above contextualization section. In sum, they represent the 4 main dimensions of emotion, valence (negative to positive) and control (low to high).

The emotional elicitation was achieved through a video. The questionnaires aimed to measure participant emotional state, and trust predisposition after seeing the video and playing the game. The game principal will be described in detail in the second cycle phase, below.

3.2.1 MAIN RESULTS AND FINDINGS

Researcher and experts meet 3 times during this phase to discuss and validate the design process. These were the main topic discussed during those meetings,

First: Researchers questioned experts if eliciting all 4 emotional states to the same participant was advisable. Experts stated that it is possible but to avoid any contamination of data we could provide a relaxation session in between. This will avoid any possible emotional biased.

Second: Experts were concerned with nature of the stimuli used to elicit trust in technology. The tasks used by Cowan et al. (2016) could not be used as stimuli, as the main aim was to measure trust in using a personal intelligent assistant and the task provided did not create a risk control situation. This cognitive-based approach could not be enough to foster strong predispositions to trust in the technological artifact. The solution provided was to introduce a risk, control situation needed to measure trust (using the game theory principle of control and risk as stimuli). In sum, the solution needed to include the game theory principles and needed to illustrate as much as possible the principle of control and risk situation. Where participant felt, they lost something for not trusting a personal intelligent assistant for instance.

Third: Again, this complex experimental design was also questioned if the game could influence participant's emotional states. A solution was to measure participant emotion after playing the game to validate if his or her emotions were influenced by the game.

Fourth: Experts pointed that result of the experiment could be biased by emotions participant experienced before coming to the experiment. One of the possible solutions was to add relaxation session before the experiment.

3.2.2 FINAL CONSIDERATIONS

The game trust stimuli designed needed to be built specifically for this experiment. And provided both risk and control situations and was based on decision making process principle. Where participants needed to decide between trust in themselves (by choosing the 50/50 chance), an audience or google assistant to help them to answer the game questions. In case of choosing the wrong answer, they lose points and eventually a chance for a cinema ticket.

As there is a gap in a literature, on how to use EEG physiological measurement or sensor to measure trust, researchers decided to use that data as a complementary data collection process. Mainly to see if there are major variations in users' biometrics when submitted to emotional stimuli and when taking the decision to choose google assistant or not when answering the game questions.

The main measurement source of studying the interplay between emotion and trust in technology was done using psychometric measurements. The Geneva Wheel (Sacharin, Schlegel & Scherer, 2012) and the Human-computer Trust measure (Gulati, Sousa & Lamas, 2017). From a physiological perspective in this first cycle was not clear yet how to operationalize the measurement variables emotion and trust using EEG.

3.3 EXPERIMENTAL-DESIGN: STEP 2

After the game "who wants to be a millionaire" was designed and the main emotional and trust measurement instruments were defined the aim was to focus on operationalizing the experiment process and run the first pilot study.

The purpose of pilot study aimed to tackled the following goals:

- to validate the overall procedure including relaxation session.
- to validate the game principle (i.e. see if included a risk control variables).
- to verify if the psychometric measurement measures of trust and emotion
- to verify possible software errors
- to further explore how operationalize the physiological measurements using EEG.

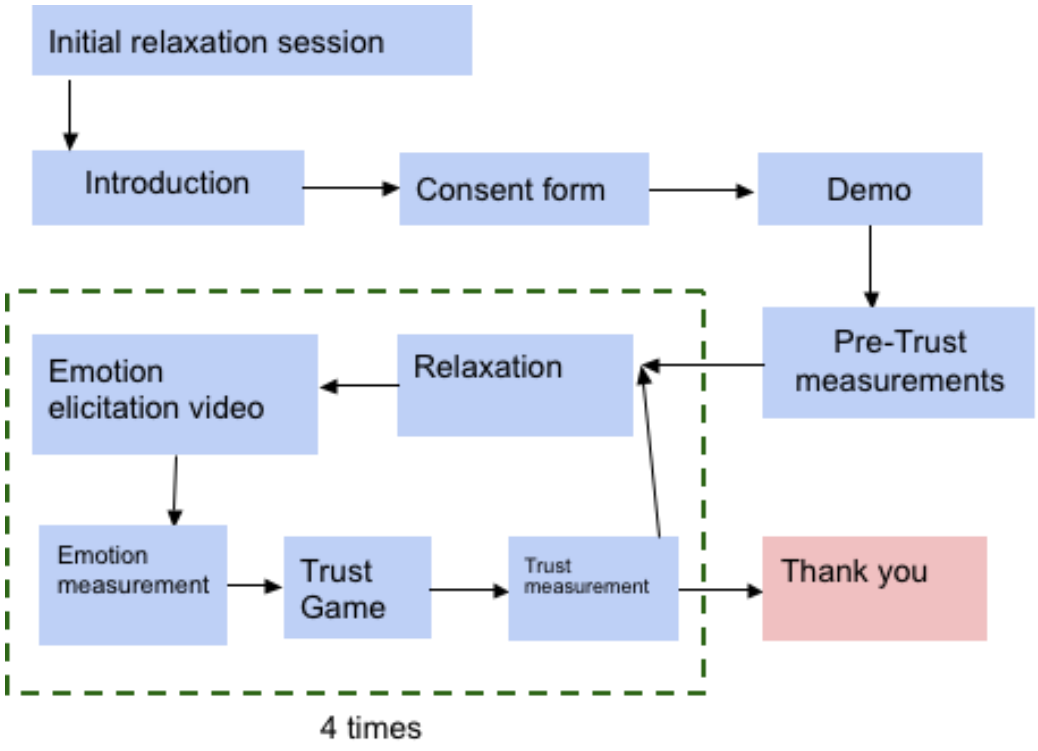

Figure 10: Overall experimental design procedure, stage 2

In sum to operationalize and implement all steps as shown in the figure above and identify the weakness of the design procedure.

How the process was operationalized

The emotion elicitation and measurements was operationalized using *Psychopy* software. *Psychopy* is an open source psychology platform independent tool developed with python programming language. It is used for running psychological experiments including neuroscience experiments. It enables data synchronization when using different data collection instruments (Peirce, 2007).

Apparatus

The experiment was designed using a high-end mini mac-i5 pc equipped with headsets, keyboard, mouse, and 15inch LED display. The headsets delivered the emotion eliciting video clip sound to subjects, the keyboard and mouse are input devices for participant to navigate through the experiment, and the LED display makes the video visible to participants. All throughout the experiment, subjects wear 16 electrode EEG sensor, and Emphatical E4 wristband sitting in front of the experiment computer (one running psychopy another for filling trust questionnaire) in the experiment room mounted with cameras all for measuring subjects physiological state during each emotional episode.

3.3.1 OPERATIONALIZATION

As addressed above, to create a trust control/risk situation and submit participants to a trust-based decision-making problem, the game "Who wants to be a millionaire" was created.

The idea of this game is to gain money by responding a set of questions. With every correct answer participant gain some points, when the answer is wrong he loses everything and game are finished. For answering the question, each participant has 30 seconds and 4 variants of respond, only one is correct. Questions are sorted from easy ones to very hard. Also, there are three ways to ask for a help: call to a friend, half of the possible responses is removed and help from the audience.

Figure 11: screenshot of the game

For this study game was slightly modified. Participants were rewarded with movie tickets, one movie ticket for a correct answer, in case of incorrect answer the score was settled to zero. Also, possibility to call a friend was changed to a possibility to ask the personal assistant. Questions were not sorted from easy to hard ones and were from the beginning very specific and hard, so chance that participant knows the correct answer is very low.

In the original game, the participant has a possibility to take gain and leave, but obviously, researchers decided to skip this possibility.

That fact that participant can make a choice how to answer a question is crucial. As we know, that his decision was not guided by the necessity, and he also has other options to gain a movie ticket, so we can assume that participant is trusting to the personal assistant. Also, there is a risk to lose tickets in case of the wrong answer.

After the game “who wants to be a millionaire” was designed and the main emotional and trust measurement instruments were designed and in addition, as referred before a relaxation session was added to avoid emotional bias before participant arrived at the experiment.

The main aim of this initial pilot study was to validate the overall process and test the apparatus. The conducted pilot study revealed many additional problems needed to be addressed, mostly related to the software and implementation. For instance, the glitch in

the use of EEG as a physiological measurement, time synchronization errors, software errors. In the following section, we describe the pilot study results in more detail.

3.3.2 First pilot Study

A pilot study was carried out with one participant. The participant was female, age 23. All instruments were in place except for the Emphatical E4 wristband that was with another researcher.

The study protocol was set in place and all the instruments were verified as the figure below demonstrate. Before experiment started researchers synchronized time on every computer to sync the milliseconds. The experiment was done in a small room where the participant was left alone to avoid research interference.

Figure 12: The room where the experiment was operationalized

The room included the following technical setup:

1. One device with Google assistant
2. One computer that runs the game tasks
3. One computer that runs PsychoPy
4. One computer that runs software for EEG Nautilus

5. One big screen to display emotion questionnaire and video clips
6. One screen to display a game
7. Between sessions user was offered to fill in a trust questionnaire on the paper

3.3.3 MAIN FINDINGS

In the beginning, the researchers had some difficulties in understanding how to set the EEG Nautilus but solved it. After implementing the pilot study researchers identified several major problems, that needed to be solved:

- The PsychoPy crashed due to video format incompatibility.
- Researcher found out the needed to have more control over the experiment procedure so he could map participants progress in real time.
- To make Nautilus EEG device sensors sit properly participants had to have big amount of gel in their hair. What was uncomfortable for the participants, as it leaves wet and sticky marks on a hair, in addition, this is hard to remove.
- EEG Nautilus is quite tight as well, so participants with sensitive necks also experience discomfort.
- Additionally, researcher still need to understand how to extract and analyse the EEG data.

3.3.4 FINAL CONSIDERATIONS

In this phase, we could get a deeper grasp of how can we use physiological computing measurements to study the interplay between trust in technology and emotions (RQ1); or how psychometric computing measurements can be used to study the interplay between trust in technology and emotions (RQ2).

The results of this initial pilot study clearly show that still are many things to consider before starting running the experiment.

Also after interviewing the Ph.D. researcher, we concluded that from the pilot study results several considerations were formed to improve the re-design of this experiment. Those included:

Technical issues like:

- The psychometric questionnaires need to be time-stamped to avoid mixing of data. A possible solution is to use online survey or added to the PsychoPy software, as we already use this software. tool and avoid the paper.
- the game design, the time sorted in the game database need to be synchronized with the EEG database. So it is possible to grab the exact moment in time that participants make the decision to trust Google assistant.

Data analysis procedures

- Research still needed to clearly understand how to interpret the results and correlate them with the pre-and post-test questionnaire.

In what regards the use of EEG, many issues still need to be solved, those range to:

- learn how to properly collect data to avoid being intrusive to the participant. For instance, EEG Nautilus was uncomfortable to use.
- Learn how to extract and analyze data after extraction. An example was in addition to useful data captured, the Nautilus also produced a lot of garbage data. Was important to know how to clean it and how to correlate those results with participants' initial trust predisposition and monte participants made the decision to trust Google assistant.

Or **implementation** issues like:

- participants need to make a lot of movements during the process switching from different instruments.
- The cognition effort to fill all questionnaires, watch the video, perform tasks was stressful.

3.4 EXPERIMENTAL-DESIGN: STEP 3

In this phase, all technical problems reported in the previous section were resolved. For instance, synchronization of data was solved in the following way:

1) Nautilus data logged in an Excel-like format with timestamps
2) Game data logged in Excel-like format. In this document, several variables defined: helpline, question answered, time (timestamp) it was answered.

3) The psychometric questionnaires time stamped were collected through the PsychoPy software.

3.4.1 SECOND PILOT STUDY RESULTS

With this pilot study, the researchers aimed to verify if the technical implementations made in the experiment were solved.

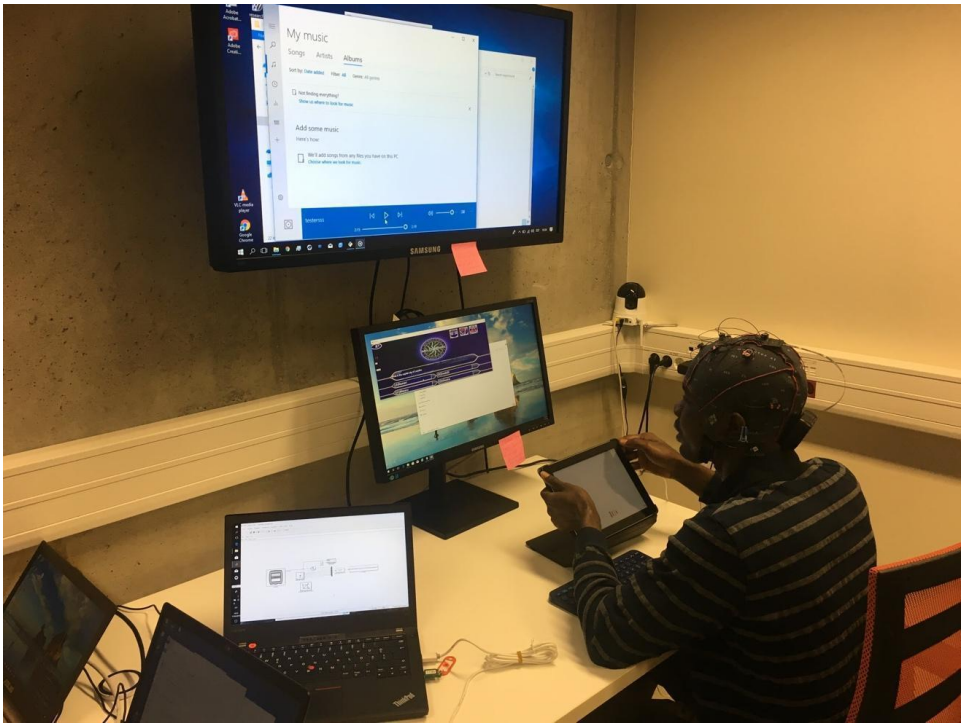

Figure 13: Subject playing a trust game

The implementation protocol was the same as previous. The main difference was that researcher could see the room through a webcam when they were away.

The participant was male, and successfully done the 4 full circles of emotion elicitation - trust game – measurements. It took about one hour without the introduction part.

The participant was using the Nautilus EEG device, as figure below illustrates. But not Emphatical E4 wristband as this instrument was still not available for researchers to use.

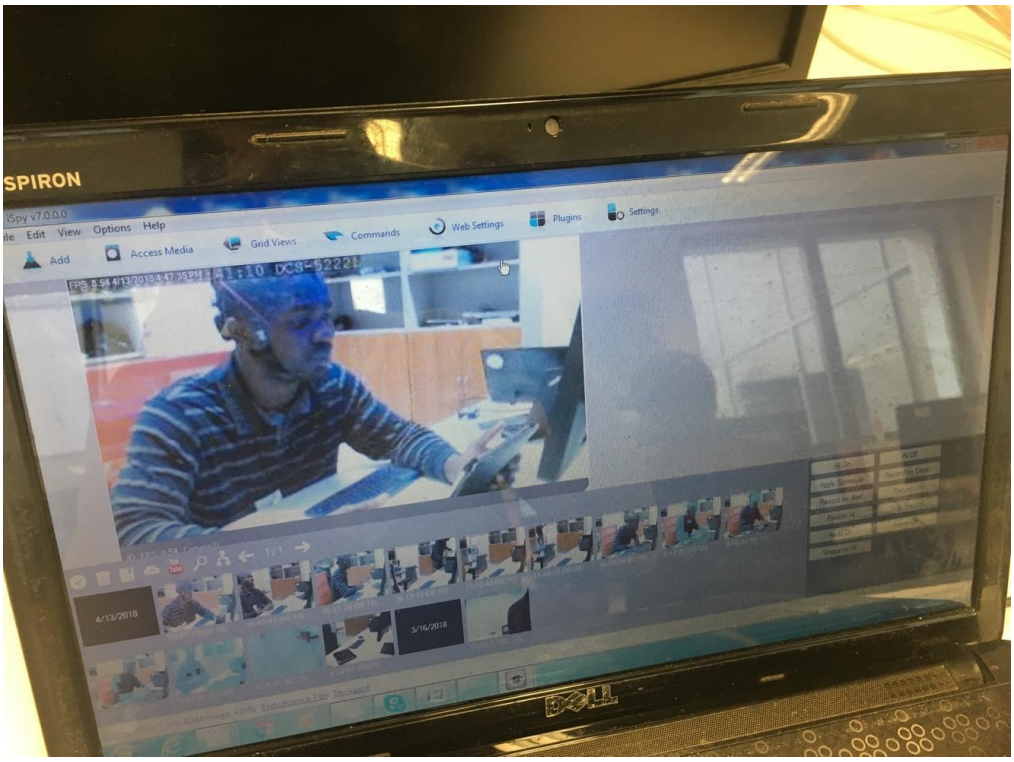

Figure 14: Software to follow the experiment process

RESULTS

This study enables to identify additional issues that were not noticed the first pilot study:

- The experiment duration is still too long and participant got tired in the end, this caused a lack of attention and as a result, the participant chose only such variants as 50/50 or help from the audience.
- the study revealed that sometimes participant can forget to choose the helpline, but still use the intelligent personal assistant to answer the question, so in the end, results could be biased
- In the middle of the experiment, the software crashed
- on video camera situated in front of the participant, it is not visible if something went wrong with the setup or software. Potentially it could be improved by introducing to the participant a special gesture that can be used in case he needs a help.
- the participant reported Nautilus to be extremely uncomfortable, both concerning the big amounts of gel and tightness in the throat.

3.4.2 MAIN FINDINGS

In what regarding the data analysis procedure those were the proposed solutions:

- **On trust, psychometric measurements:** to clearly understand how to interpret the results and correlate them with the pre-and post-test questionnaire. Researchers will use the following formula to calculate users trust level before and after submitted to an emotional state.

$$Trust\ Level\ = \frac{Points\ Gained\ *\ 100\%}{180}$$

The questionnaire has a total of 35 5 Likert scale questions, from which we can gather a maximum 180 score, assuming the highest rate will be 5.

The results will be correlated and significant correlation between the following hypothesis will be highlighted:

H1: Does positive and low arousal emotion influence trust in technology?

H2: Does positive and high arousal emotion influence trust in technology?

H3: Does negative and low arousal emotion influence trust in technology?

H4: Does negative and high arousal emotion influence trust in technology?

This potentially will give more insight to the interplay of trust and emotion and will facilitate data analysis from a psychometric perspective.

- **On the use of physiological measurements:** this regards most of EEG data extraction and analysis.

In what concern extracting the data, researcher where able to extract it ins spite of finding some software versions incompatibility. But we're not able to clean the data for further analyses as they expected as it seems (on a first glance) that data was not being recorded properly. So, further research needs to better understand how the EEG works, before doing another attempt to record records data from participants.

In what regards how to do the data analysis after extracting the EEG data to the MathLab software, there is still a not clear answer on how to do proceed with examining the possible correlation between trust and emotions. The researcher is now studying different statistical

ways on how to describe and explore data. As well as reviewing similar studies possible approaches to the problem.

A possible solution provided by a researcher could be to measure trust indirectly using EEG frontal cortex data. And maybe correlate it with the moment when participants decide to trust google assistant to help them to answer the game questions. But as the EEG data is so unreliable and at the same time collects data that is not needed. That more clarification needs to be done to understand how to better clean the data provided from the one it is not necessary.

It will need a deeper research to better understand how what physiological measurements could, in fact, be helpful to the study.

Another useful result from this phase was provided on the last feedback session. In this meeting, the external experts were present as well.

The main results included the following issues to be addressed:

- One expert pointed out that it is not efficient to use video clips for emotion elicitation, we should use images instead. This will shorten the study time, and pictures proved to be more efficient in emotion elicitation. This method also proved to be less biased.

- The same expert proposed to reduce the number of emotions elicit from 4 to 3 and make a shorter break between sessions (up to 1 minute). This to minimize the time needed to run the experiment and therefore minimize the cognitive effort from the participants.

- In addition, they proposed to help in further investigation how the interplay between trust and emotions from biological perspective could be measured. Suggested that a possible solution could somehow link it to the decision making or cognitive overload processes. The main problem addressed here is that as mentioned above, in the contextual background section. There is evidence of being possible to measure trust from a biological nature but this was mainly using fMRI and hormones (Riedl & Javor, 2012). Not using EEG measurements or Emphatical

E4 wristband. But using fMRI was not possible as besides being expensive is not available in the University.

So, a possible solution is to do an exploratory research looking for evidence on potentially indirect effects.

3.4.3 FINAL CONSIDERATIONS

In sum, the researcher had a clearer idea on how to analyze the data gathered from the psychometric questionnaire. But no clear answer was found so far, in what concerns "RQ1. How can we use physiological computing measurements to study the interplay between trust in technology and emotions?"

On one hand, the overall study procedure still needs to be better operationalized. The second pilot study still revealed major problems that needed to be solved. For instance, in the game, we could minimize the cognitive workload of the participants by adding a new feature that facilitates participant to answer the questions.

On the other hand, it is too risky to use physiological measurements as EEG data recording is captured with a lot of garbage data that need to be cleared. In addition to the clear evidence that researcher still doesn't know how it could be analyzed to find a possible interplay between trust and emotions.

4 CONCLUSIONS

This chapter analyses the major findings and results of the study reviews the chosen methods used in the thesis, explains the limitations of the study and gives input for the possible further research.

The main goal of this research was to clarify the process of designing a research experiment to efficiently study the relationship between trust and emotions. In other words, it aimed to investigate what are the strengths and constraints of design such complex research studies. This thesis reached its goal by providing a systematic description of how the experiment and study were initially designed and validated. Further literature illustrates how complex these topics can be and revealed that few studies explore this especially in

the field of human-computer interaction. Also, revealed that trust in addition to behavioral and psychological nature has also a biological nature and can be measured using fMRI. In addition, providing insights on how we can design such complex process. How to explore new measurements instruments, how to elicit emotions and create scenarios for exploring users trust predispositions in a specific technology.

The iterative process described in this study revealed that involving researchers and experts from fields can be useful in connecting two broad notions. It also showed some drawbacks, however, they could be solved by thorough control from the lead researchers in the study. Data gathered during this study also helped to create some insights that will facilitate researchers to build similar studies in future.

4.1 FURTHER INSIGHTS

The findings of this work aimed to clarify the process of designing a research experiment to efficiently study the relationship between trust in technology and emotions. It presents the strengths and constraints of using experimental approaches to design such complex research study, and discuss how effective is this research design approach.

It addresses the constraints as well of using EEG computing measurements to study the interplay between trust in technology and emotions.

Overall, as these two notions studied are complex by nature, the design process involved many people with different expertise, but none had a clear idea of what was the best approach.

This document describes the three main design stages: an initial exploratory approach, followed by the hypothesis formulation and operationalization and a final validation process where researchers examine outcomes of a pilot study and proposed the redesign for providing more accurate measurements.

The three design phases described above revealed major constraints and strengths of using this experimental research design approach. For instance, a quantitative data approach can provide more reliable results but when implemented in a complex scenario it can be very difficult to design and implement.

Moreover, such approach also showed its strength to bring a new approach to address such complex topics. For instance, the feedback provided by experts and the pilot studies done helped to:

- to fine-tune the overall design and hypothesis formulation;
- reveal that studying the biological nature of trust combined with self-psychometric measurement could potentially bring valuable results. An argument supported by the external experts' feedback and from literature.
- Create an initial landmark for this study, avoiding possible biases in the experiment design and data collection and analysis.
- provide insight into the further design the trust stimuli, starting from uncertain user tasks and finished with a such a powerful tool as game theory to create a risk/control situation.
- To identify useful technical constraints and issues with the pilot studies. That otherwise might be missed.

In sum in such novel approaches, using this interactive approach combined with external feedback helped the team visualize potential ways for further research. And using innovative approaches to address the problem of understanding how to measure trust from a biological nature without using expensive and not reachable instruments like fMGI.

Also, this study showed that there is a point to involve the limited number of people at every stage and presence of several skilled and confident researchers from the start to the end of the study. As such experiments are very complicated and require involving many people with different opinions, moderation of such process is an important feature. It is also reasonable to have researchers who are very keen on topics, in the current study, there were experts both in trust and emotions, from the very beginning of the study.

In the end, as overall remarks could include the following advice.

The study also revealed constraints using physiological measurements and how those complicate the process. As so when designing such study it is important to bring specific competencies to the team. And, people with those they competencies should be involved since the beginning.

- Experts from both fields (Emotion and Trust) during the whole process, who support the study and validate the feedback.
- A team leader should be assigned to moderate the process.
- Limit the number of external experts. Look for those who can provide specific insight into the study.
- Break big tasks into smaller ones and validate after every iteration.

We identified technical and methodological limitations in what concern using physiological data with resources we have available in a laboratory. Or that, the Emphatical E4 wristband was not available for the researcher to explore. And few insights were provided on how to properly use the EEG Nautilus.

Methodological, for instance, the researcher still doesn't know how these physiological instruments can help to grasp the interplay between trust and emotions. More exploration needs to be done on what concern the use of physiological measurements to better grasp its usefulness. Also, there is no clear evidence how this two devices could be replaced, as fMGI is very expensive and unavailable at this stage. Secondary, data analysis remains unclear, as Nautilus produces a lot of data that is not needed.

5 REFERENCES

McAllister, D. J. (1995). Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. Academy of management journal, 38(1), 24-59.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. Academy of management review, 23(3), 393-404.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of management review, 20(3), 709-734.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. Information systems research, 13(3), 334-359.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: a trust building model. The journal of strategic information systems, 11(3-4), 297-323.

Kassim, E. S., Jailani, S. F. A. K., Hairuddin, H., & Zamzuri, N. H. (2012). Information system acceptance and user satisfaction: The mediating role of trust. Procedia-Social and Behavioral Sciences, 57, 412-418.

Constantine, L. L. (2006, June). Trusted interaction: User control and system responsibilities in interaction design for information systems. In International Conference on Advanced Information Systems Engineering (pp. 20-30). Springer, Berlin, Heidelberg.

Friedman, B., Khan Jr, P. H., & Howe, D. C. (2000). Trust online. Communications of the ACM, 43(12), 34-40.

Olson, J. S., & Olson, G. M. (2000). i2i trust in e-commerce. Communications of the ACM, 43(12), 41-44.

Gulati, S., Sousa, S., & Lamas, D. (2017, September). Modelling Trust: An Empirical Assessment. In IFIP Conference on Human-Computer Interaction (pp. 40-61). Springer, Cham.

Benbasat, I., & Wang, W. (2005). Trust in and adoption of online recommendation agents. Journal of the association for information systems, 6(3), 4.

Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, D. C. (1995). Can computer personalities be human personalities?. International Journal of Human-Computer Studies, 43(2), 223-239.

Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. Journal of applied social psychology, 27(10), 864-876.

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates?. International Journal of Human-Computer Studies, 45(6), 669-678.

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. ACM Transactions on Management Information Systems (TMIS), 2(2), 12.

Lankton, N. K., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. Journal of the Association for Information Systems, 16(10), 880.

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. ACM Transactions on Management Information Systems (TMIS), 2(2), 12.

Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., & Leimeister, J. M. (2012). Understanding the formation of trust in IT artifacts.

Benbasat, I., & Wang, W. (2005). Trust in and adoption of online recommendation agents. Journal of the association for information systems, 6(3), 4.

Thatcher, J. B., McKnight, D. H., Baker, E. W., Arsal, R. E., & Roberts, N. H. (2011). The role of trust in postadoption IT exploration: An empirical examination of knowledge management systems. IEEE Transactions on Engineering Management, 58(1), 56-70.

Vance, A., Elie-Dit-Cosaque, C., & Straub, D. W. (2008). Examining trust in information technology artifacts: the effects of system quality and culture. Journal of management information systems, 24(4), 73-100.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. MIS quarterly, 27(1), 51-90.

Kim, D., & Benbasat, I. (2003). Trust-related arguments in internet stores: A framework for evaluation. J. Electron. Commerce Res., 4(2), 49-64.

Lankton, N. K., & McKnight, D. H. (2011). What does it mean to trust Facebook?: examining technology and interpersonal trust beliefs. ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 42(2), 32-54.

Hubert, M., Hubert, M., Linzmajer, M., Riedl, R., & Kenning, P. (2018). Trust me if you can–neurophysiological insights on the influence of consumer impulsiveness on trustworthiness evaluations in online settings. European Journal of Marketing, 52(1/2), 118-146.

Madsen, M., & Gregor, S. (2000, December). Measuring human-computer trust. In 11th australasian conference on information systems (Vol. 53, pp. 6-8).

Sousa, S., Lamas, D., & Dias, P. (2014, June). A model for Human-computer Trust. In International Conference on Learning and Collaboration Technologies (pp. 128-137). Springer, Cham.

Söllner, M., & Leimeister, J. (2013). What we really know about antecedents of trust: A critical review of the empirical information systems literature on trust.

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. Decision support systems, 44(2), 544-564.

Dunn, J. R., & Schweitzer, M. E. (2005). Feeling and believing: the influence of emotion on trust. Journal of personality and social psychology, 88(5), 736.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. Journal of personality and social psychology, 81(1), 146.

Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., ... & Bandeira, N. (2017, September). What can i help you with?: infrequent users' experiences of intelligent personal assistants. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (p. 43). ACM.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. Academy of management review, 23(3), 393-404.

Scherer, K. R. (2005). What are emotions? And how can they be measured?. Social science information, 44(4), 695-729.

Coan, J. A., & Allen, J. J. (Eds.). (2007). Handbook of emotion elicitation and assessment. Oxford university press.

Betella, A., & Verschure, P. F. (2016). The affective slider: A digital self-assessment scale for the measurement of human emotions. PloS one, 11(2), e0148037.

Sacharin, V., Schlegel, K., & Scherer, K. R. (2012). Geneva emotion wheel rating study.

Coan, J. A., & Allen, J. J. (Eds.). (2007). Handbook of emotion elicitation and assessment. Oxford university press.

Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., & Reuter-Lorenz, P. A. (2005). Emotional category data on images from the International Affective Picture System. Behavior research methods, 37(4), 626-630.

Benantar, M. (2001). The Internet public key infrastructure. IBM Systems Journal, 40(3), 648-665.

Manchala, D. W. (2000). E-commerce trust metrics and models. IEEE internet computing, 4(2), 36-44.

Riedl, R., & Javor, A. (2012). The biology of trust: Integrating evidence from genetics, endocrinology, and functional brain imaging. Journal of Neuroscience, Psychology, and Economics, 5(2), 63.

Hayes, G. R. (2011). The relationship of action research to human-computer interaction. ACM Transactions on Computer-Human Interaction (TOCHI), 18(3), 15.

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007, April). Research through design as a method for interaction design research in HCI. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 493-502). ACM.

Bardzell, J., Bardzell, S., Dalsgaard, P., Gross, S., & Halskov, K. (2016, June). Documenting the research through design process. In Proceedings of the 2016 ACM Conference on Designing Interactive Systems (pp. 96-107). ACM.

Colomo-Palacios, R., Casado-Lumbreras, C., Soto-Acosta, P., & García-Crespo, Á. (2011). Using the affect grid to measure emotions in software requirements engineering.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. Journal of neuroscience methods, 162(1-2), 8-13.

Forgas, J. P., & George, J. M. (2001). Affective influences on judgments and behavior in organizations: An information processing perspective. Organizational behavior and human decision processes, 86(1), 3-34.

Forgas, J. P. (1998). On feeling good and getting your way: Mood effects on negotiator cognition and bargaining strategies. Journal of personality and social psychology, 74(3), 565.

Scherer, K. R. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. Cognition & Emotion, 7(3-4), 325-355.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). Appraisal processes in emotion: Theory, methods, research. Oxford University Press.

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. American psychologist, 50(5), 372.

Ekman, P. (1992). An argument for basic emotions. Cognition & emotion, 6(3-4), 169-200.

Levenson, R. W. (1988). Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity. Social psychophysiology: Theory and clinical applications.

Levenson, R. W., & Gottman, J. M. (1985). Physiological and affective predictors of change in relationship satisfaction. Journal of personality and social psychology, 49(1), 85.

Allanson, J., & Fairclough, S. H. (2004). A research agenda for physiological computing. Interacting with computers, 16(5), 857-878.

Ravaja, N. (2004). Contributions of psychophysiology to media research: Review and recommendations. Media Psychology, 6(2), 193-235.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2000). Psychophysiological science. Handbook of psychophysiology, 2, 3-23.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), 49-59.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. NIMH Center for the Study of Emotion and Attention, 39-58.

Fucito, L. M., & Juliano, L. M. (2009). Depression moderates smoking behavior in response to a sad mood induction. Psychology of Addictive Behaviors, 23(3), 546.

Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K., Duann, J. R., & Chen, J. H. (2010). EEG-based emotion recognition in music listening. IEEE Transactions on Biomedical Engineering, 57(7), 1798-1806.

Bower, G. H. (1983). Affect and cognition. Phil. Trans. R. Soc. Lond. B, 302(1110), 387-402.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. Science, 221(4616), 1208-1210.

Ax, A. F. (1953). The physiological differentiation between fear and anger in humans. Psychosomatic medicine, 15(5), 433-442.

Velten Jr, E. (1968). A laboratory task for induction of mood states. Behaviour research and therapy, 6(4), 473-482.

Hagemann, D., Naumann, E., Maier, S., Becker, G., Lürken, A., & Bartussek, D. (1999). The assessment of affective reactivity using films: Validity, reliability and sex differences. Personality and Individual differences, 26(4), 627-639.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. Cognition & emotion, 9(1), 87-108.

Knez, I. (2001). Effects of colour of light on nonvisual psychological processes. Journal of environmental psychology, 21(2), 201-208.

Detenber, B. H., & Reeves, B. (1996). A bio-informational theory of emotion: Motion and image size effects on viewers. Journal of Communication, 46(3), 66-84.

Anderson, C. A., Deuser, W. E., & DeNeve, K. M. (1995). Hot temperatures, hostile affect, hostile cognition, and arousal: Tests of a general model of affective aggression. Personality and Social Psychology Bulletin, 21(5), 434-448.

Detenber, B. H., Simons, R. F., & Reiss, J. E. (2000). The emotional significance of color in television presentations. Media psychology, 2(4), 331-355.

Fridlund, A. J. (1992). The behavioral ecology and sociality of human faces.

Williamson, O. E. (1993). Calculativeness, trust, and economic organization. The Journal of Law and Economics, 36(1, Part 2), 453-486.

Capra, L. (2004, May). Towards a human trust model for mobile ad-hoc networks.

Russell, J. A. (1980). A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161.

Appendix A

A Sub appendix section

Grounded theory

Grounded theory is a method of qualitative research that aims to produce new theories during the research. It was originally identified in sociology field as the product of close inspection and analysis of qualitative data(Glaser and Strauss, 1967). Three key principles that distinguish grounded theory methodology are the principle of emergence, constant comparative analysis and theoretical sampling. Later, Strauss and Corbin(1990) used the term to refer to a data collection and analysis technique. The methodology combines systematic levels of abstraction into a framework of interpretation of a phenomenon, which is iteratively verified and expanded throughout the study.(Strauss and Corbin, 1990). The theory emerges from the data after rigorous application of the grounded theory method. This theory works by providing the researcher "with relevant predictions, explanations, interpretations and applications".[Glaser, B.G. and Strauss, A. 1967. The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine, Chicago, Illinois.] The main feature of grounded theory that it does not require a prior hypothesis for focusing the research(Strauss et al. 1964).

But today the grounded theory was often implicated with some problems in HCI research. In [Adolph, S., Hall, W., and Kruchten, P. 2008. A methodological leg to stand on: lessons learned using grounded theory to study software development.], authors say they and many other researchers that claimed to follow the grounded theory only followed several practices from this method. This is similar for nursing, as it was first developed in the medical field and often used in nursing research.

[Glaserian and Straussian grounded theory: similar or completely different?]

Action Research

Action research is an approach to research that involves engaging with a community to address some problem or challenge and through this problem solving to develop scholarly knowledge. AR is explicitly democratic, collaborative, and

interdisciplinary. The focus when conducting AR is to create research efforts “with” people experiencing real problems in their everyday lives not “for,” “about,” or “focused on” them.

[Knowing by Doing: Action Research as an Approach to HCI.Gillian R. Hayes]. In a nutshell action research it is something that unites the practical concerns of people and academic goals of science, “by joint collaboration within a mutually acceptable ethical framework” (Rapoport, 1970).

Doing good action research means do it empirically and cyclically, which means that actions are undertaken are responsive to emerging evidence. In addition, research should be cyclical, in which actions follows the planning, and the followed by reflections.

Research Through Design

Research through Design (RtD) is an approach to conducting scholarly research that employs the methods, practices, and processes of design practice with the intention of generating new knowledge. For the first glance this approach to research can look like design practice, but it is generally more systematic and reflective process of review the detailed documentation that was done during the design process. Key to this work of capturing and translating primary design knowledge into broader academic knowledge is how design processes are documented; design documentation is a key raw material out of which such knowledge is constructed.[Documenting the Research Through Design Process, Jeffrey Bardzell et al.]. One challenge was noted in 1989 by Caroll and Kellogg, the thing proceeds theory instead of theory driving the creation of

new things. They said, that mouse was invented first, and then a lot of studies were conducted to prove that this is a good decision for direct manipulations of the GUI. In response to this challenge, we see RtD as a way for many new things to enter into HCI that can spawn new theory. At the same time, these new things can be informed by current theory, creating an ongoing dialog between what is and what might be.[Research Through Design in HCI,John Zimmerman and Jodi Forlizzi].

Experimental Research

To describe what is experimental design we need to think about variables, the researcher usually manipulates by one or more independent variables and measuring the effect of this variables on a dependant variables. Also a researcher may have a hypothesis about relationship between all this variables. And finally, the researchers may involve random participants to experiment. However, experimental research is useful not only for generating hypothesis-driven knowledge and theoretical advancement but also for informing practical and applied goals.

This experimental methodology have a number of advantages over other HCI research methods, one of the major is internal validity, or the extent to which the experimental approach allows the researcher to minimize biases or systematic error and demonstrate a strong causal

connection.

[Experimental Research in HCI. Darren Gergle and Desney S. Tan]