Universidade do Minho
Escola de Engenharia
Departamento de Informática

André Francisco Soares Carvalho Alves Teixeira

The Influence of Emotions
on Human Computer Interaction

Master's Dissertation
Integrated Master's in Informatics Engineering

November 2018

Universidade do Minho

Escola de Engenharia

Departamento de Informática

André Francisco Soares Carvalho Alves Teixeira

The Influence of Emotions on Human Computer Interaction

Master's Dissertation
Integrated Master's in Informatics Engineering

Master dissertation
Master Degree in Computer Science

Dissertation supervised by
Paulo Novais
Davide Carneiro

November 2018

ACKNOWLEDGEMENTS

First of all I would like to thank my beloved parents, Francisco and Cristina, who have supported me and encouraged me through the great endeavour of writing a master's dissertation in engineering, something that they know was a big challenge for me. Without them I simply would not have made it. This extends to my dear sister Beatriz, my inspirational and ever optimistic uncle Jorge, my grandparents and the rest of my family. I would also like to thank my supervisors Paulo Novais and Davide Carneiro, for without their guidance and help this project would have never seen its goals completed, nor this dissertation would have ever been written. Without their aid I would not be finishing this chapter in my life, and for that I thank them deeply.

Some thanks to essential friends are also mandatory. To José Nuno Macedo, for all the times we sat down with some coffee and worked though my numerous troubles with the project. To Francisco Ribeiro, Alexandre Pinto and Bruno Saraiva, for being comrades always ready to encourage me and help me vent my frustrations. A special hug to Andreia Faria and Joana Silva. To all my friends in CAUM, who helped me deal with stress by singing. To the people of my ADAUM for believing in me, to the colleagues in CeSIUM for helping me feel part of my chosen course all these years. To all the others who greeted me and wished good luck. Finally I would like to thank all the people who helped me in my testing phase or in important little things, by name and individually. Thank you Clara, Marta, Alexandra, Daniel, Gustavo, Sérgio, Malhadas, Sofia, Andreia, Bruno, Maria, Olicarpo, Cientista, Afonso, Nuno, Ana Cláudia, Diogo, Rui, Bárbara, António, Esquível, Catarina, Carolina, Humberto, Joana, Hélio, Quesado, Inês, Sara, Xico and Xana. Without you all I would have no data, no conclusions, no dissertation. I thank you for your helpfulness, your support, your friendship.

There are a lot more people I should thank that are not there, and I should thank all the people listed here a lot more, but I would never move on if I did. As such I will end by acknowledging that all big projects, like dissertations, are never made alone. People need people in order to create and work and investigate and inspire and succeed. All my failures are my own, but all my successes are built on the shoulders of others. And for that I give you all my thanks.

ABSTRACT

Emotion is an essential part of what means to be human, but it is still disregarded by most technical fields as something not to be taken into account in scientific or engineering projects. However, the understanding of emotion as an aspect of decision making processes and of modelling of human behaviour is essential in order to create a better connection between humans and their tools and machines. This dissertation focuses on the measurement of emotion of users through the use of non-intrusive methods, like measuring inputs and reactions to stimuli, along with the creation of a tool that measures the emotional changes caused by visual output created by the tool itself. Usage of the tool in a test environment and the subsequent analysis of the data obtained will allow for conclusions about the effectiveness of the method, and if it is possible to apply it to future studies on human emotions by investigators in the fields of psychology and computation.

RESUMO

Emoção é uma parte essencial do que nos faz humanos, mas mesmo assim há muitos nos campos mais técnicos e científicos que a desconsideram como algo que não deve ser tido em conta em projetos científicos ou de engenharia. A compreensão da emoção é no entanto um aspeto chave para o estudo dos processos de decisão humanos, assim como para a modelação de comportamentos, algo essencial para a criação de uma melhor ligação entre o homem e as suas máquinas e ferramentas. Esta dissertação foca-se na medição da emoção de utilizadores através de métodos não-intrusivos, como medições de inputs em computadores e reações a estímulos, assim como na criação de uma ferramenta que meça as mudanças emocionais causadas nos utilizadores por certos estímulos visuais provenientes da própria ferramenta. Utilização da ferramenta num ambiente de teste e a subsequente análise dos dados obtidos permitirão a tomada de conclusões sobre a eficiência deste método, assim como sobre se é viável a aplicação do mesmo a estudos futuros da emoção humana realizados por investigadores nas áreas de psicologia ou computação.

CONTENTS

LIST OF FIGURES

LIST OF LISTINGS

ACRONYMS

D

DI Departamento de Informática.

H

HCI Human Computer Interaction.

I

IAPS International Affective Picture System.

M

MIEI Mestrado Integrado em Engenharia Informática.

U

UM Universidade do Minho.

INTRODUCTION

This dissertation describes the Master's work developed in the context of the Mestrado Integrado em Engenharia Informática (MIEI) held at Departamento de Informática (DI), Universidade do Minho (UM).

1.1 MOTIVATION

Emotion is a key aspect of the Human psyche. Although humans can be considered rational and logical creatures there are many options and attitudes people choose to follow that cannot be understood if the emotion of the individual in question is not taken into account. More than a detail or peripheral aspect of our mentality, emotion is a vital part of why we do what we do, why we are what we are. Emotion binds us together and breaks us apart, emotion can ravage the greatest of logical minds and break the greatest of strategists, and it can give rise to the best and the worst of any given individual. As such an essential aspect of the causes and reasons behind human behaviour this phenomenon must be better understood. The understanding of emotion and of both its expression and unconscious experiencing has been the goal of a great many investigators, in areas like philosophy, psychology and most recently neuroscience. Initially relegated to the work of natural thinkers due to its immaterial nature and difficulty to measure, the emotive part of humans is now being researched by more precise and exact sciences, since there is now a possibility to use modern technology to try and measure this phenomenon, be it through behaviour analysis be it through biological changes. Although there exist many systems that attempt to use emotion as its main component, none of these are technological. Politicians try to use emotion to capture voters, judicial systems of the past and present use the emotion of juries and judges to manipulate their decisions, casinos and marketplaces try to influence their customers to spend as much money as possible through promises of emotional peace and respite through consumerism and thrill. There are numerous examples of people using emotion as a basis with which to create systems that they consider beneficial either to society or more commonly, to themselves. It is therefore strange that scientific and technical usages of this component of what it means to be alive and human have eluded us

for as long as they have. Much of it has to do with the difficulty inherent to the specification and modulation of what exactly is emotion, and how can we measure or deconstruct it.

This phenomenon is defined by the American Psychological Association as a complex pattern of changes, including physiological arousal, feelings, cognitive processes, and behavioural reactions, made in response to a situation perceived to be personally significant. Although difficult to grasp, this definition allows us to come to some conclusions. The first is that all emotion, including its type and intensity, is entirely dependent on stimuli to the individual, be it external like a smell or sound or internal like a memory. The second is that said stimuli must have personal meaning for the subject in order to not simply be ignored with the passive bureaucracy our minds resort to in order to go through the massive information surges they have to deal with. As such, emotion is something that can be studied in a non-intrusive way through the application of stimuli to individuals, and subsequent analysis of response and reaction to that stimuli. It is therefore essential to better understand emotion as a real and influential component of the human psyche, something that many in the scientific field have tried to ignore as something very unscientific and unworthy of scientific study and scrutiny. Patterns of emotion can be built this way, if both the system used to create the pattern and the behaviour of the subject in question are fully understood.

The value of such systems in informatics is obviously enormous, since human machine interaction has always depended on a logical cause-effect relation, discarding the emotional component or simply delegating it to a secondary place where it is treated simply as another variable to be divined by the established algorithms. This has created a very stable and control machine environment, but also a stale one, where machines are not truly able to adapt to the users state of mind, and therefore unable to provide emotional bridges so that the user can empathize and feel understood by the machine. By including the measurement of the user's emotion a system would be able to adapt to any single user in a way never seen until now, since most systems are completely blind in this regard. This would benefit the user due to their needs being better met, and the system due to being able to better synchronize with the user, therefore increasing efficiency. Although there is still no effective and trustworthy way to measure emotion without invading the user's biology and therefore their personal and bodily privacy, there is a great potential to create systems that can learn the emotions of a human through the way they interact with the system itself, since most communication of emotion is non-verbal in nature. Better understanding of emotions from the part of machines can also be beneficial in the attempt to create a virtual entity that can think for itself, the singularity all artificial intelligence lovers always have yearned for. There are theories that true conscience can only be achieved through a mixture of self-preservation, rational logic and emotional responses. Thesis that state that emotional

pain is essential in the early stages of human development in order to create a functional mind corroborate the idea that without emotion there can never be a truly able-minded being. All things point to the fact that emotion is essential for human interaction, and is therefore essential for the interaction of humans with all other things. It is therefore necessary that more efforts are put in place in order to try to better understand emotion and to better input that understanding into our intelligent systems. Technologic integration of this human phenomenon is essential moving forward.

1.2 SCOPE OF THE DISSERTATION

Past work done at ISLab by teams led by both advisors of this thesis suggests that auditory emotion recognition is both possible and an efficient way to measure the phenomenon. This was confirmed in thesis such as this one and in investigative projects pursued by senior investigators in the field, in conjunction with investigators in other areas of study relevant to the study of thought, emotion and stress. There is however a lack of work in the learning and creation of emotional patterns simply by natural interaction, like the act of writing a text, or simply using a mouse to perform mundane actions and tasks.

The idea is then to create a system based on existing psychology models of emotion that can induce emotive responses on the subject through visual stimuli and create a model of their interaction with the machine that can be used to measure the subject's emotion in different situations. The creation of such models would allow for a better understanding of emotional changes in individuals, therefore being an asset in both the field of psychology and informatics. If one could detect emotional changes in users through non-intrusive means, that is, without capturing live or past data on the user's being, interests or activities, based solely on the users interaction with the machine itself in the current session of usage, then there could exist high potential to profile users and their actions and therefore respond to them in real time, seamlessly. This would be a huge step since there would be no information on the user itself being stored, protecting the individual from the invasion of privacy people unfortunately seem to be getting used to. This would also allow for a highly adaptive system, not based on individuals and their characteristics but on their impact on the world that surrounds them through their actions and their actions alone.

The first step is therefore to be able to perform such a detection without obtaining intrusive data. As such, and since the idea of using sound has been successful in similar projects, the idea of using visual stimuli to provoke emotion came forth. Most models of prediction in computation rely on models of learning, where the machine models an expected behaviour based on past interactions. Since the main intention of this thesis of non-intrusion and adap-

tation of systems while respecting privacy is independence on data from past events it is necessary to obtain such a model while the user is performing its tasks. Since it is however impossible to predict behaviour without past cases the usage of images to provoke emotion and variation in the performance of the user in mundane tasks came forth. If subjected to an image that is expected to create a specific emotion the user will react in a certain way, changing its behaviour in response to that emotion. If the system knows the emotion the image is associated with then it can detect the behaviour detected when such emotion is present, therefore detecting the emotion itself. While this is a highly flawed method, since it directly involves showing users images so that their reaction can be saved and analysed for future usage, it is still very respectful of the person's individuality and privacy, since no data on the person will be saved, only data on their reactions to the predicted visual stimuli.

The project will then focus on the study of emotional models, on human computer interaction and on the construction of a tool that can explore and implement this model of learning through the study of the reaction to the visual stimuli. Although not directly applicable in commercial or marketplace products, this tool would allow for the confirmation of the thesis that non-intrusive methods can be used to detect emotions and react to them, opening the door to future applications or research projects that could afterwards proceed to a more practice implementation of the thesis for commonplace usage.

1.3 OBJECTIVES AND GOALS

Intentions and motivations cannot be transformed into result and put into action without the verbalization of what it is one wishes to achieve. Effort without direction is more than often wasted sweat. I have therefore tried to plan and put into concise words the main objectives of the project, in order to better focus on the milestones that are essential for its success. As such this project and subsequent dissertation have as its main objectives the following:

- The better understanding of emotion and of its importance for intelligent systems and specifically in human-computer interaction;

- The creation of a different method of measuring emotion through non-intrusive means;

- The development of a tool that has the ability to study emotional changes in an individual using it;

- The exploration and treatment of data obtained through such tool, in order to judge the effectiveness of the method;

These four main objectives are the guidelines for the thesis and represent the goals that need to be met for it to be considered successful. This project involves both a technical and an investigative approach to the issue at hand, considering the proposed development of a tool that can be used by other investigators of different areas of study to detect and evaluate emotional impacts of images and pictures, and the uncertainty of the results that can be obtained with the tool itself, since there is no assurance that the results obtained after the usage of the tool in a test environment will be favourable to the proposed theory that emotions and changes in them can be modelled and subsequently detected through nothing more than the normal usage of the computer, through non-intrusive methods. This thesis will also be pursued under the assumption that the library of pictures to be used can and will provoke a strong and detectable emotional change in the test subject, since without emotional response there is nothing to be detected, even if the detection method itself is successful, exact and accurate.

As such, the goals proposed in this section will be the guidelines used to judge the successfulness of the project, since positive results that support the proposed method of emotion detection can simply be non-existent. As with all investigative projects the value is to be obtained through the journey of research more than with the treasure found at the end of the road. That being said the research will be results oriented and shall proceed having assumed that the proposed hypothesis is right, since there would be no chance of obtaining true results and proving the hypothesis true otherwise.

1.4 METHODOLOGY OF WORK

In order to accomplish the objectives of this work, enumerated in the section above, the Action-Research methodology was and will be followed. This methodology starts by identifying the problem so that a hypothesis can be formulated on which the development will be based. The data is then recompiled, organized and analysed, continuously building a proposition for solving the identified issue. Conclusions can finally be made based on the results obtained during the projects research. For this research model to be followed, six complementary stages have been defined to achieve the planned goals, according to the pre-established description of this established work mentality. The stages themselves are described in firm and concise sentences below:

- Specification of the problem and its characteristics, necessary to evaluate the issue at hand and plan for a solution.

- Constant and incremental update and review of the state of the art, since only by using established knowledge can innovation be attained.

- Idealization and interactive development of the proposed model, in order to allow the creation of a non-virtual solution.

- Experimentation and implementation of the solution through the development of a prototype, necessary for the confirmation of the proposed hypothesis.

- Result analysis and formulation of conclusions, based on that which could be ascertained by the usage of the prototype.

- Constant diffusion of knowledge, results obtained and experiences with the scientific community, in order to benefit future research and investigators in the field.

2

EMOTION

2.1 WHAT IS EMOTION

Emotion is a difficult thing to describe or define. All across human history it has been present as a vital component of our psyche, a part of what makes us human. We all recognize its existence and its relevance in both personal and social contexts, but that's about where the consensus ends. "The only thing certain in the emotion field is that no one agrees on how to define emotion," says Alan Fridlund, associate professor of psychological and brain sciences at the University of California, Santa Barbara. In psychology, emotion is often defined as a complex state of feeling that results in physical and psychological changes that influence thought and behaviour. Emotion is associated with a range of psychological phenomena, including temperament, personality, mood, and motivation. According to author David G. Meyers, human emotion involves "...physiological arousal, expressive behaviours, and conscious experience."(Myers, 2004)

We can recognize its existence in other species of beings, so it makes sense to believe that emotions are not something created by rationale or through social conditioning, but something present in our beings since birth. It is not something that supports human exceptionalism but rather something that supports our position as instinct oriented beings. Emotion can be hidden and it can change, it can evolve and it manifests in different ways for different people, but it cannot be avoided and cannot be disabled. It stands as the basis of expressionism, of artistic creation and vision, even of scientific advancement, as it often is traits and desires of certain individuals that fuel discovery, not the benefits of the research itself. Many try to use it to manipulate others, and many succeed in this endeavour, be it through invoking specific reactions to pieces of art or through using fears and desires to draw support for political movements.

Emotion as a psychic phenomenon of reaction to internal or external stimuli doesn't seem to change with time, at least not during recorded human history, but our ways of dealing with it and expressing said emotion are highly variable and almost completely social-oriented.

There is very little done to integrate emotion in the more technical fields of study and research, which is understandable due to the difficulty of defining it as a palpable and usable variable. Emotion still seems to draw attention from various researchers and fields of study, even though most of them are connected to natural science, psychology or philosophy. Only recently has emotion begun be studied in fields like neuroscience and artificial intelligence, but it tends to be used more as something one wishes to evoke than something one wishes to work with or adapt to. There are still many models of study of emotion that are extremely relevant to anyone trying to better comprehend the subject at hand, even if most of these have no connection to any hard science.

2.2 EMOTIONAL MODELS AND THEORIES

There are many ways of exploring emotion, but most can be divided in one of three general categories (Beck), namely physiological, neurological and cognitive. **Physiological** theories suggest that responses within the body are responsible for emotions. **Neurological** theories propose that activity within the brain leads to emotional responses. Finally, **cognitive** theories argue that thoughts and other mental activity play an essential role in forming emotions. Again, these are models of explaining emotion and that seek to demystify the phenomenon, not models of usage or interaction, but one cannot really hope to make use of something one does not understand. As such, we will be reviewing some of the major theories and models related to emotions.

Naturalist Charles Darwin was the first to effectively explore and propose that emotions appeared and evolved because they were adaptive and necessary for the survival of humans and animals. Feelings of love and affection lead beings to seek others and reproduce, and bonds created through these emotions make these beings more probable to defend their mates and their territory. Fear, on the other hand, causes beings to better deal with danger, since they flee or fight the source of said danger. These emotions are useful in many ways to ensure the maintenance and continuing of species. According to the **Evolutionary Theory of Emotion**, our emotions exist because they serve an adaptive role. Emotions motivate people to respond quickly to stimuli in the environment, which helps improve the chances of success and survival. This theory even states that our quest of better understanding other people's emotions our even our own is nothing more than a mechanism to ensure that we are better prepared to deal with events caused by said emotions, in a way that knowledge about a specific behaviour allows for preparation to said behaviour. It is possible to deduce that a foe will attack if the emotional state of the foe is distressed or angry, and therefore it is easier to prepare for it. As all purely evolutional theories that try to explain human behaviour this theory does no contemplate the complexities of one's psyche and the varia-

tions that emotion can experience from person to person, especially having in account that said emotion can sometimes stop reproduction and harm the chances of survival, and as such seems to be incomplete, even though its biological basis for the necessity of emotion seems likely correct.

The **James-Lange Theory of Emotion** is one of the best-known examples of a physiological theory of emotion. Independently proposed by psychologist William James and physiologist Carl Lange, the James-Lange theory of emotion suggests that emotions occur as a result of physiological reactions to events. According to this theory, witnessing an external stimulus leads to a physiological response. Your emotional reaction depends on upon how you interpret those physical reactions. This connects emotion to bodily function in a structural level. Following this theory, we do no experience changes in our bodies in response to emotions, we experience emotions due to changes in our bodies. Therefore, you are not trembling because you are frightened, you feel frightened because you are trembling. This inversion of what we take for granted has interesting implications, but again fails to take into account self-inflected emotion, caused not by outside stimuli but by our own thoughts. If emotion can come from within then it may not be linked to physiological causes. Thoughts are of course caused by electrical connections in the brain, and have biological backgrounds, but these are not considered classic physiological events. While modern researchers largely discount the James-Lange theory, there are some instances where physiological responses do lead to experiencing emotions. Developing a panic disorder and specific phobias are two examples. One major criticism of the theory was that neither James nor Lange based their ideas upon anything that remotely resembled controlled experiments. Instead, the theory was largely the result of introspection and correlational research, and as such it is still being debated to this day. While it seems as if this theory should be nothing more than something you might study due to its historical significance, it maintains its relevance today because researchers continue to find evidence that supports at least some parts of its original ideas

Still in the realm of physiological theories we inevitably have to address the **Cannon-Bard Theory of Emotion**. This theory can be easily describe as contradicting the previous one. In fact, this theory was formulated as a reaction to the James-Lange theory of emotion. Where James-Lange theory represented a physiological explanation for emotions, the Cannon-Bard theory represents and neurobiological approach. Walter Cannon suggested that people can experience physiological reactions linked to emotions without actually feeling those emotions. Cannon also stated that emotional responses occur much too quickly for them to be simply products of physical states. Cannon first proposed his theory in the 1920s and his work was later expanded on by physiologist Philip Bard during the 1930s. According to this

theory emotion are not caused by physical reactions, but are experienced in conjunction to them, at the same time. As such, one does not cause the other, or vice-versa, being simply two sides of the same coin and more specifically it states that emotions occur when the thalamus sends a message to the brain in response to a stimulus, resulting in a physiological reaction. At the same time, the brain also receives signals triggering the emotional experience. This theory suggested that emotions could be experienced even when the body does not reveal a physiological reaction, and that physiological reactions to different emotions can be extremely similar, and therefore cannot be linked to specific emotions as the starter in a cause-effect relation. Although more balanced than the previous theory this one still fails to explain the origins of the emotion itself, the causes for the impulse that starts the whole experience, since they still base this theory in the reaction to outside stimuli, disregarding emotion that starts in one's self, even though this could simply be described as a late reaction to a physical reaction of the past more than a different experience altogether. This seems to be a recurring assumption, one that seems difficult to contest without direct scientific knowledge of the biological processes occurring during the emotion itself.

Now reaching for a cognitive theory, we have the **Schachter-Singer Theory of Emotion**. This theory suggests that the physiological arousal occurs first, and then the individual must identify the reason for this arousal to experience and label it as an emotion. A stimulus leads to a physiological response that is then cognitively interpreted and labelled which results in an emotion. There is therefore a three part journey to an emotion, against James's two part explanation. This might seem like only a slight deviation, but it is actually extremely different, since it is the first theory explored until know that accepts the possibility of emotion being also a cause of actual personal choice and experience, albeit a subconscious one. Schachter and Singer's theory draws on both the James-Lange theory and the Cannon-Bard theory of emotion. It accepts that emotion is derived from a physiological event or arousal, but it states that the essential part of the emotional equation is the cognitive interpretation of the event, which then opens the path for the emotion itself. This would allow for different people to experience different stimuli differently, since it depends on cognitive functions and subjectivity is a very important factor of cognition. This goes along the hypothesis defended by the Schachter-Singer theory, which also suggests that similar physiological responses can produce varying emotions.

Now, according to the **Cognitive Appraisal Theory of Emotion**, thinking must occur first before experiencing emotion. Richard Lazarus was a pioneer in this area of emotion, and this theory is often referred to as the Lazarus theory of emotion. This theory goes a step further than the previous one, not only linking cognitive interpretation of stimuli to the emotion but conscious thought itself. According to this theory, the sequence of events first

involves a stimulus, followed by thought which then leads to the simultaneous experience of a physiological response and the emotion. This opens up incredible explanations for human interactions, such as an explanation for the phenomenon we call love. Even though many physical reactions indicate attraction or sexual desire, humans do not experience love for all others that provoke said reactions. Instead of looking for a specific set of characteristics that elicit love this theory defends that love can come of cognitive deliberation and interpretation of the stimuli it is presented, followed even by conscious thought about the matter, more than the nature of the stimuli itself. Emotion can then be something one's mind decides to experience in reaction to something, more than something it is obligatorily coded to do in certain situations. This allows for an appraisal of the value of emotion that's much higher than if we consider them to be mere biological responses. This theory can explain emotions not caused by physical events, since thought and subconscious deliberation about the matter is ultimately what causes the emotion. Seemingly built on top of all other theories, only with the extra step of "thinking" added on top, this theory seems to be the more rounded up of the group, even if it does not approach the subject on a manner scientific enough to justify enough consideration from the more technical part of the scientific community.

There are many other theories and models of explaining and assessing emotion (Cherry), like the facial-feedback theory of emotions, for example, that suggests that facial expressions are connected to experiencing emotions in a natural level, meaning a very smiling person tends to be happier than a frowning one. Since this chapter intends to allow for an overview of emotion as a phenomenon and not a complete list of theories we will keep from referencing them, as they would not add anything specific to the projects goals. It is however important to comment the apparent lack of ways to measure emotion, since most theories come from a psychology standpoint and simply do not intend to measure emotion on a numeral or countable scale. The most defended methods of measuring emotion tend to be performed by the individual itself, in an autonomic examination of their feelings in a specific moment. There are systems that intend to help this examination, like the Self-Assessment Manikin (SAM), a non-verbal pictorial assessment technique that directly measures the pleasure, arousal, and dominance associated with a person's emotional reaction to a wide variety of stimuli. Still, the amount of actual measurements in the field of emotion is thin, a barrier that had to be overcome in order to lead this project into fruitful conclusion.

2.3 IAPS AND CHOSEN MODEL

There are obviously many different models of studying, describing and measuring emotion, but none of these are directly related to computation or the proposed method of connecting images to specific emotional output in order to model an individual's emotional behaviour. For the context of this thesis the important factor is not the nature of emotion itself, but how it can be measured and used in benefit to the user of a specific system. After extensive research one system seemed particularly adapted to this hypothesis, having actually been built in order to allow experimentation in the field of visual stimulation in order to create an emotional response on the subject.

International Affective Picture System (IAPS) is a system currently being developed to provide ratings for a large set of emotionally-evocative, internationally-accessible, colour pictures that includes contents across a wide range of categories. The IAPS is product and property of the NIMH Centre for Emotion and Attention (CSEA) at the University of Florida. The associated model also allows for the definition of any emotion in a three dimensional space, having in account values in three different parameters to define specific emotions. Its goal is to provide standardized materials that are available to researchers in the study of emotion and attention. The existence of these collections of normatively rated affective stimuli should allow better experimental control in the selection of emotional stimuli, facilitate the comparison of results across different studies conducted in the same or different laboratory, and encourage allowed exact replications within and across research labs who are assessing basic and applied problems in Psychological science.

This system was developed based on the assumption that emotion can be defined by a co-incidence of values (Bradley and Lang, 1954) on a number of different strategic dimensions. This view is founded in Osgood's (Osgood, Suci, and Tanenbaum, 1957) ground-breaking work with the semantic differential, in which factor analysis conducted on a wide variety of verbal judgments indicated that the variance in emotional assessments were accounted for by three major dimensions. Dimensional views of emotion have already been advocated by a large number of theorists through the years, by the likes of Mehrabian and Russell in 1974, Tellegen in 1985 or even going back to Wundt in 1898.

The IAPS is then a collection of normatively rated affective stimuli, considering three dimensions: affective valence, ranging from pleasant to unpleasant, arousal, ranging from calm to excited, and dominance, ranging from controlled to in-control. These three dimensions, rated in a scale of points from 1 to 9, allow for the measuring of emotion in a more complex level than mere substantive denominations of "sad" or "afraid", since these are

often highly subjective and simply not very scientific at all. This scale can be used to establish a three dimensional model of a subject's emotional reactions to the stimuli produced by the researcher, a model that can afterwards be applied in the creation of personalized responsive systems that can benefit from the new knowledge of the users state of mind at any given moment, knowledge gained through the users actions and patterns of reaction. This way there is both no need to comment on the users emotion using substantives nor to inquire about the meaning and purpose behind the actions the user takes. The machine knows the user has a certain emotional state in that particular moment due to that particular interaction, and reacts to it accordingly without any need or intention of knowing what caused said reaction. The training method of exposing the subject to a collection of images is exactly that, a training method, unavoidable but temporary. The creation of the model necessitates knowledge of the causes of the emotional variations, and the image set of the IAPS model allows for that knowledge to be unobtrusive and standard. The user's habits will not be monitored, only their reaction to the aforementioned pictures, and even so all monitoring of content will end after the model is created. The use of this model allows for a dissociation between content and reaction to the content, necessary to prove the point of the detection of emotions through means that do not blatantly break the trust and privacy of the user. As such, these methods are invaluable if implemented in everyday products, even if the training phase can be a little overbearing for the casual consumer.

<div align="right">*3*</div>

HUMAN COMPUTER INTERACTION

Human Computer Interaction (HCI) is an area of research and practice that emerged in the early 1980s, and since it has high relevance in the context of the work done in this dissertation it shall be explored a bit more here, by exposing a view of its importance and evolution through the times. It is seen as a specialty area in computer science embracing cognitive science and human factors engineering, something unusual at a time where computation was seen as a dark art accessible only to the technically gifted. It has since been expanding rapidly and steadily for three decades, attracting professionals from many other disciplines and incorporating diverse concepts and approaches. To a considerable extent, HCI now aggregates a collection of semi-autonomous fields of research and practice in informatics, a form of viewing computation with man at its centre, an anthropomorphic approach that sees the computer as being developed and created to suit the necessities of human society and individuals instead of simply to accelerate existing processes or production needs (Carroll, 1997). However, the continuing synthesis of disparate conceptions and approaches to science and practice in HCI has produced a dramatic example of how different epistemologies and paradigms can be reconciled and integrated in a vibrant and productive intellectual project.

For a long time after the creation of the computer, the only humans who interacted with computers were information technology professionals and dedicated hobbyists, who either needed the machines to conduct their specialized work or found them fascinatingly distant and yearned for more knowledge. All of this changed drastically with the emergence of personal computing, in the late seventies. Personal computing, including both personal software (productivity applications, such as text editors and spreadsheets, and interactive computer games) and personal computer platforms (operating systems, programming languages, and hardware), made everyone in the world a potential computer user, and made apparent to all that the computer as a tool was not ready to be used by the masses, since its obvious usability deficiencies and hard learning curve showed that what was beginning to be viewed as a useful tool open to all was in fact a specialty one, made only for those who studied the field of computing itself. This challenge presented itself at an opportune time,

since only then where many field and experts considering throwing themselves into work in this new field of computing for masses. Thus, at just the point when personal computing presented the practical need for HCI, cognitive science presented people, concepts, skills, and a vision for addressing such needs through an ambitious synthesis of science and engineering. As such HCI was born as a cognitive and computational science, presenting an unusual cooperation of what was before as separate as math and poetry. John M. Carrol, a specialist and founding father of practical psychology as a pillar of design for computing, had a much defined look into what HCI was and what it could evolve into. "HCI is a science of design. It seeks to understand and support human beings interacting with and through technology. Much of the structure of this interaction derives from the technology, and many of the interventions must be made through the design of technology. HCI is not merely applied psychology. It has guided and developed the basic science as much as it has taken direction from it. It illustrates possibilities of psychology as a design science." (Carroll, 1997).

With the necessity for HCI proved and the ability to make it happen suddenly obtained, the scientific community focused its efforts on making commercial products ready for mass consumption that were usable by the typical user of the time. HCI therefore opened the way to true personal computing through the creation of standards for many tools we still use to this day, like word processors and such. This was possible through advances in areas adjacent to HCI that many times overlapped with its principal components, like human factors engineering and the creation of heavy documentation for the tools being developed. Human factors had begun the develop empirical and analytical steps and techniques for the evaluation of the interactions between humans and their tools and machines, and moved to study and define contexts in which humans interacted with systems in a way that operators were able to exert greater problem solving potential than before. Documentation was also adapted to this new principle, changing from merely being technical logs to a full manual of instructions for any and all users, employing methods of media production only used until then in literature, news and advertising. They needed to be useful and usable, in order to both accommodate the product to eh user and diminish the learning curve of using the machine itself.

At the heart of HCI sits its core principle: usability. Everything done in the context of applying social and human sciences to computation and technical ones was done in order to add more usability to that which was being developed, whatever it was. This stepped through understand the users mind, the way society dealt with machines and tools, the social and economic context of the users themselves, the capabilities of the machines in relation to what was expected of them by the users, among many thing. Usability now

often subsumes qualities like fun, well being, collective efficacy, aesthetic tension, enhanced creativity, flow, support for human development, and others. A more dynamic view of usability is one of a programmatic objective that should and will continue to develop as our ability to reach further toward it improves. Usability engineering became the basis for HCI going forward. It emphasizes the involvement of stakeholders in the process of design itself, not only after the design has been finished and produced. Iterative techniques are necessary, since hypothesis needed to be proposed and evaluated an errors corrected, features added and removed. There were three key notions in this philosophy of design. First, it was proposed that iterative development be managed according to explicit and measurable objectives, specifications that needed to be attained in order for the project to succeed. In designing a program, one would iteratively design and redesign, prototype and evaluate, include the intended users of the product in the process, and produce specific metrics and objectives to be met in the small term. Usability specifications are now standard practice in HCI development. The second key notion in usability engineering was a call to broaden the practical scope of design itself, subverting the all-scientific approach to the development process. A variety of approaches and techniques for user participation were developed, to facilitate collaboration between users, who bring expertise on the work situation the tool or program proposes to assist, and developers, who bring the ability to put into action and create what was found necessary through discussion. This again emphasizes the involvement of users before the product is finished, in order to assess how they would react to it and why that is essential to the design of the product itself. The third key notion in usability engineering was cost effectiveness. It is expensive to carry out many iteration cycles of development and error correction, while disregarding the product of work done until then. Developers need to employ efficient methods throughout and to know when they have reached a point where it simply is not sustainable to keep iterating. As we can see with these principles, HCI helped open the door for the art of iterative design, a technique so ingrained in computer science and engineering these days one wonders how we could ever have lived without it.

The desktop model, full of selectable little icons, is one of the design achievements of HCI that most people are aware of. However, the initial conception of HCI has become more broaden, more open to other possibilities of interaction that do not depend on designing an arrangement of graphical elements on a screen alone. Mental model theories began to be utilized to explain perception of objects, and information processing psychological theories and models were used to model the behaviour and cognition capabilities of the individual users, through their interactions with the graphical elements of design and with the main methods of introducing input to the computer, like keyboards and such. This is obviously where HCI establishes a direct relation with the work done in the context of the disserta-

tion, since the basis for the tool to be developed is the modulation of the user's emotional reactions and their interpretation through causal input methods. Through theories and applications such as these, the advances in all fields related to this unifying area of study have evolved in some way or another. Codifying and using relatively austere models made it clear what richer views of people and interaction could be articulated and what they could help in the development of a program or product, while at the same time personal devices became portals for interaction in the social and physical world, requiring richer theoretical frameworks for analysis and design due to their newfound relevance.

One aspect of HCI that confuses many is its relation with design itself. Although it was always talked about a guideline to designers, it was firstly highly differentiated and many boundaries were set in order to differentiate the areas and simply separating the areas as different aspects of the professional expertise necessary for the field. However, through the times many designers were assimilated into HCI communities, due to the interconnected nature of computation and design experienced by them. Initially seen as a simple acceptance of techniques and methods that came from a field more artistic and less technical than their own, the advances in technology that allowed for better graphical integration of the design elements allowed for the notion that these elements were not something separate that was useful but something essential to the experience itself. Somewhat ironically, designers were welcomed into the HCI community just in time to help remake it as a design discipline. A large part of this transformation was the creation of design disciplines and issues that did not exist before. As such, no one can accuse HCI of resting after the work is done, since the basis for this field of work is the subversion of accepted principles. Conceptions of how underlying science informs and is informed by the worlds of practice and activity have evolved continually in HCI since its humble beginnings of allowing casual users to perform minor computation feats. Throughout the development of HCI, paradigm-changing scientific and epistemological revisions were deliberately embraced by a field that was, by any measure, succeeding intellectually and practically. This resulted in a very complex field of study that to this day keeps a very important position in the priorities of developing any given product in the area of computation and technology.

All in all the future looks promising for the defenders of HCI. The role of cognitive and social sciences in more technical fields like computer science has always been reduced to a mere afterthought, but these days things are changing. No longer can experts ignore the implications and possibilities of interacting and connecting to users in a more deep, personal level, a level that surpasses the definition of computers as mere machines or tools but turns them into reals parts of human society, cogs in the big machine, essential not only to the functioning of our society but also to the functioning of people's minds, for the leading

of their daily lives and routines. This will impact society in an inexorable way. As Carrol puts it, "it is exciting to see that the emerging role of social and cognitive science in computer science and the computer industry is far more diverse, pervasive, and critical than imagined in the 1970s. As it has turned out, that role was not to support a received view of design but to help overturn it and help clarify the real nature of design. Nor was that role to recast psychological laws as human factors guidelines; it was indeed to play a part in driving the evolution of social and cognitive science, and the recognition that computers can be deliberately designed to facilitate human activity and experience only when social and cognitive requirements drive the design process throughout. There is unprecedented potential for interdisciplinary synergy here. Social science has always borne the vision of what human society might become, but it has typically lacked the means to be constructive. Computer science - quite the converse - cannot avoid causing substantial social restructuring. An integrated and effective HCI can be a turning point in both disciplines and, perhaps, in human history." (Carroll, 1997).

As such, advances in human computer interaction lead to unavoidable advances in the structure of modern society, since our dependence on computers to perform everyday tasks and to even perform the acts of governance and maintenance necessary to sustain our society mean that when our way of interacting with these machines we prize so much changes, so will the way we interact with other people and with society itself. With machines having more relevance every day, with the internet taking its place as the main method of communicating and with personal computation devices becoming not only commonplace but essential to life itself in the minds of many individuals, the way the users interact with these newfound pillars of human society will clearly shape the direction we evolve as a society. Evolutions in technology always provoke great changes in our societal structure, and this new emphasis on different ways to interact with ever smarter machines that are prepared to anticipate our needs or even our thoughts will turn this century into a new era of constant revolution and reinvention, where systems that are able to interact with users in interesting new ways and connect with them through the use of both logic and emotion will pave the way for the rest of us. We can only make predictions on how this will affect us as a species, and if these changes to society will lead to improvements to the lives of many or simply to a loss of that which truly makes us human like many seem to fear. Even though these fears exist the choice on whether to walk the road to this uncertain future has already been made, although rather quietly, and the only path is forward.

4

RELATED PROJECTS AND ANALYSIS

Not much has been done in regards to practical application of emotion in engineering tools. Although the IAPS developers intended for it to be used in a research point of view, it still has not been applied to its full potential, at least in a consistent and investigate manner. Even though emotional integration in tools and services is something that is rising up in attractiveness these days, the idea of using non-intrusive methods for measuring biological phenomena on users is nothing new. Since the concerns of lack of privacy and storage of user information surfaced a lot of investigators tried to obtain different methods to deal with the issue of adapting to the user without compromising the user's rights and identity. A difficult endeavour, only possible if attained though the examination of the interactions the user takes part in rather than the examination of the user's characteristics themselves. A lot of work has been done in the field of fatigue, especially by ISLab investigators, in using non-intrusive methods to evaluate and measure levels of fatigue the user possesses and adapting to these levels (Novais et al., 2012). Success in this work was what led to the idea that such methods could be used in other type of measurements, and in the creation of models more deep and intertwined with the user than evaluation of their level of stress.

As such, and even though other works do exist, I would like to give a higher degree of attention to two works in particular than both highly motivated and served as base for this dissertation. Both works are adequately documented in the bibliography. The first being the paper **"An application to enrich the study of Auditory Emotion Recognition"**, by Renato Rodrigues, Augusto J. Mendes, Davide Carneiro, Maria Amorim, Ana P. Pinheiro and Paulo Novais (Rodrigues et al.). This work focuses heavily on auditory recognition of emotion, which differs heavily in both method and form from the subject of this dissertation, but is highly similar in its motivation and approach to emotion as an essential factor in both human and human computer interaction. It is stated that the ability to convey, as well as to accurately and rapidly decode, emotions is fundamental for the success of communication and interactions, something that is also one of the basic principles behind this dissertation. This work introduces a different way to assess auditory emotional recognition, in the form of a tool, one that can be used both in research and clinical settings, focused

on a mobile competition device model. While directly focused on emotional recognition through voice, the developed system can be easily changed for other types on voice based detection. The developed tool was also used in conjunction with Psychology experts, in order to test its efficiency. The framework present for the project depicted in this paper is therefore very similar to the one being described in this dissertation, even though the way the emotion is being measured and the way the emotional model is being constructed differ.

The second work directly related to this dissertation is the one that effectively justifies the existence of the dissertation itself, since they are connected at a very basic level. The paper **"An environment for studying visual emotion perception"**, by Davide Carneiro, Hélder Rocha and Paulo Novais, another work developed in the context of ISLab research (Carneiro et al.). This paper constructs the whole basis for the work of this dissertation, since its main objective and methods are the same. The team behind the project proposed the creation of a tool that can be used to measure emotion through the same application of visual stimuli. This paper explores the best architecture for such a tool, and the ways it could be used. It therefore serves as a basis from which to further build on the hypothesis, yet unproven, that such a system would be effective on creating a model of emotion useful in other kinds of circumstances that not simply tests with the tool. This paper goes more into the architecture of the system itself than on its uses, and one of the ways this dissertation seeks to build upon this predecessor of sorts is the creation of a better measurement model, as well as a tool that is more easily applicable to other investigative projects.

As such it can be said that even if there isn't much work done in this area of study and even though there aren't many related projects, the previously developed projects that serve as the basis for this dissertation are solid and comprised of adequate academic work, providing a strong pillar of support over which to innovate on. The purpose of this dissertation is therefore to use ideas explored in these papers and expand them in order to make them available to use in different contexts, as well as add features and a new vision of emotional integration and testing when creating a model that can be used as a way to better connect man and machine.

TOOL REQUIREMENTS

The usage of requirement engineering has long been a staple of good and responsible development in the field of informatics (Pohl, 2010). This kind of engineering is a field of its own, with a deep crevice of characteristics and formalization of the way one documents and plans a tool or product. As such I will not be constructing a full-fledged requirements document, but will only be using some of the conveniences and rules of requirements engineering in documenting the mains requirements necessary for the development of the projected tool.

These shall be divided in functional and non-functional requirements, seeing as these are the main two groups where every requirement can be encapsulated without much error. By describing the tool to be developed in such a fashion I intend to focus development in the main features and necessities of the program, and ensure the presence of all the main and necessary parts of the tool. Changes made in development will also be catalogued in the requirements themselves, in order to maintain consistency and to allow the perception of the whole of the systems capabilities in this chapter.

5.1 FUNCTIONAL REQUIREMENTS

Functional requirements represent the backbone of any system. These are the requirements that describe the main functionalities of the system, and are therefore essential for usage. These requirements are the features that the tool must present tot eh user in order to ensure the success of this study. As this is not a product to be made available for purchase client satisfaction will not be quantified and indicated, as is custom in this format. All requirements shall be numbered, allowing for simple reference procedures, and priority shall not be expressed since only necessary requirements will be listed.

Requisite Nº1
Description: The tool is able to create new tests based on available images;
Reason: In order to allow researchers to create their own custom made tests based on IAPS;

Criteria: Being able to select and go through the created tests;

Requisite Nº2
Description: The tool is able to list available images during test creation;
Reason: In order to allow researchers to choose the images they want on their tests;
Criteria: Being able to select and go through the available images;

Requisite Nº3
Description: The tool is able to list available tests ready for usage;
Reason: In order to allow users to go through the test and store data;
Criteria: Being able to select and go through the created tests;

Requisite Nº4
Description: The tool is able to show images during the test taking procedures;
Reason: In order to allow users to react to the image and store relevant data;
Criteria: Being able to see images in the test taking procedures;

Requisite Nº5
Description: The tool is able to record text written by the user about the image;
Reason: In order to allow users to react to the image and store relevant data;
Criteria: Being able to write text in the test taking procedures;

Requisite Nº6
Description: The tool is able to record the three emotion parameters as specified by the user about the image;
Reason: In order to allow users to react to the image and store relevant data;
Criteria: Being able to indicate emotional parameter levels in the test taking procedures;

Requisite Nº7
Description: The tool is able to record time stamps of the user taking the test;
Reason: In order to allow study of the emotional responses;
Criteria: Being able to verify the time data in the database;

Requisite Nº8
Description: The tool is able to record non-intrusive keyboard and mouse data during testing;
Reason: In order to allow study of the emotional responses;
Criteria: Being able to verify the mouse and keyboard test data in the database;

Requisite Nº9
Description: The tool is able to save all data relating to the test taking in an external database;
Reason: In order to allow study of the emotional responses;
Criteria: Being able to verify the extensive test data in the database;

Requisite Nº10
Description: The tool can display usage instructions detailing all its main features;
Reason: In order to allow users without further training to use the tool;
Criteria: Being able to read the instructions using the tool;

5.2 NON-FUNCTIONAL REQUIREMENTS

Non-functional requirements represent the external aspects of the program, things that cannot be functionally measured or perceived in a technical way but are as important as the remaining requisites to the usability of the program. These include graphical, usability and legal requirements, among many other subcategories. However, this tool is not a product, and as such not much emphasis will be spent on attractiveness or other qualities necessary for a well successful product to sell. As they are not quantifiable there will not be any criteria to verify their implementation. Still, as before all requirements shall be numbered, allowing for simple reference procedures. All requirements shall be numbered, allowing for simple reference procedures, and priority shall not be expressed since only necessary requirements will be listed.

Requisite Nº1
Description: The tool instructions and all text must be in English;
Reason: English is the most widely spread language, opening the tool up to a majority of users;

Requisite Nº2
Description: The tool must have a simple and instinctive graphic user interface;
Reason: To allow for all users to instinctively understand how to interact with the tool;

Requisite Nº3
Description: The tool must be able to work with different image folder and databases;
Reason: To allow for usage of the tool with different projects and research purposes;

Requisite Nº4
Description: The tool must use only proprietary or open source code;
Reason: To allow for free distribution, usage and avoid copyright infringement;

Requisite Nº5
Description: The tool must be able to work in different computers after configuration;
Reason: To allow for usage of the tool by different people;

Requisite Nº6
Description: A user must not be able to access or change any data collected during testing;
Reason: To allow for unadulterated study of the collected data;

<div style="text-align: right">

6

</div>

DEVELOPMENT

This chapter contains the documentation of the development of the tool itself, complete with the main decisions behind the development process, the description and definition of the developed application, the user interface design and process of though and the database functioning and structure. As such it is expected that in this chapter the whole reasoning behind the developed tool is present and presented to the reader.

6.1 TOOL DESCRIPTION

The developed tool was nicknamed "Horus" in the name of the Egyptian god whose all-seeing eyes could see even a person's soul, relating to the intended purpose of measuring emotion through interaction. The tool's purpose is not to create models of emotion itself, but to collect data with which to build those models, as well as validate the data collected in order to ensure that a model construct with this data would be sufficiently exact. As such, this tool was developed using Microsoft Visual Studio, with a basic Windows Presentation Form template in C#, utilizing a Mongo Database connection to store and read tests and data. These technologies were chosen for their ease of use and their wide adoptability, allowing for the tools to be used in systems with only some easily installed support software.

The tool consists on a window based program that allows for two basic modes, namely test creation and test taking, envisioned according to the requisites to accommodate researchers and test takers' needs. The tool possesses a simple interface with clearly defined buttons that can be used to proceed to the subsequent window. When using the creation part of the tool, chosen through the main window, the user can select multiple images from a list based of a local image folder, where images can be dropped off as necessary. After selecting the chosen images and typing a name for the test, the test is created in the database and the user is returned to the main window. When using the test taking part of the tool, also accessible through the main window, the user can then choose a test to take from a list of previously created tests, taken from the connected database. After selecting the test

one wishes to take the user goes through a series of windows with images selected for the test, and is asked to describe the image presented and evaluate the user's own emotional reaction to the image using the three dimensions of emotion defined in IAPS methodology. This information is saved each time the user presses the next button to proceed to the next picture, until the end of the test. Unbeknownst to the user there are other types of data being saved, relating to the users interaction with the system, which will be later explained in detail. After finishing the test the user is thanked for the participation and sent once again to the main window.

As such, the developed tool can be described as a simple interface of visual stimuli testing, with a focus on allowing for easy and simple creation of tests and the even simpler taking of such tests. The data being recorded is not shown to the user, and is saved in the database as individual test pieces related to the image shown and the test itself, allowing for simple access for posterior data parsing and analysis. The main objective of this application is therefore the collection of data rather than the treatment of data, which could be implemented in later versions, as well as the visualisation of said collected data. As it stands the tool is built for test subjects with no experience in the fields of informatics, psychology of neuroscience, as well for researchers that want a simples yet elegant way of collecting data. This tool is very dependent of the IAPS system of image evaluation with which it was built, and is therefore unable to be used with studies that do not take this system in mind.

6.2 USER INTERFACE

As this is a tool to be used by any researcher, knowledge of informatics non withstanding, user interface is an integral part of the experience of the application. Therefore the choice was made to use a simple style, based on basic Widows Presentation Form applications, with a series of windows and buttons that directly take the user to the desired feature (Nathan, 2014). We will therefore go through all the main windows that constitute the user interface for this tool. As the application starts, the user is greeted by the main window of the program, which contains three buttons that represent the three essential functionalities of the application, which are "Learn", where you can read instructions on how the use the tool, "Create", where you can create your own tests based on the IAPS image database, and "Take", where you can go through the tests previously created and obtain test data.

Once one selects the "Learn" button, a window will be shown where the instructions for the usage of the tool can be found. The instructions are divided between instructions for researchers and for test takers, which are loosely connected to the "Create" and "Take"
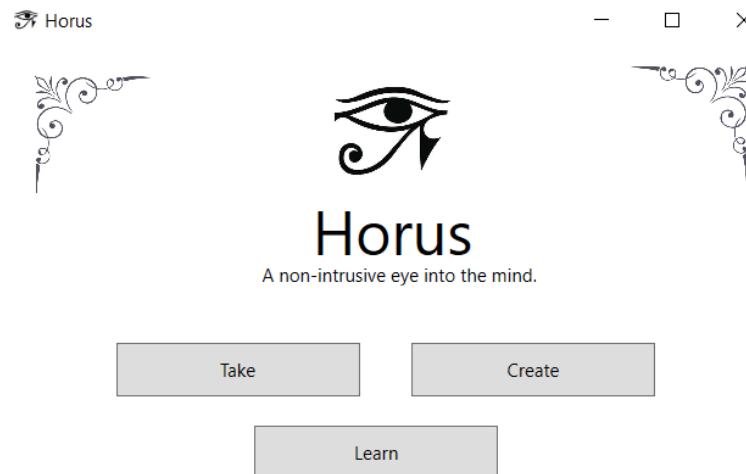
Figure 1: Main window of the tool

functions. This window is a simple if somewhat inelegant way of conveying key information to users, in a way that can be accessed even if there is no human contributor talking about the tool to the potential user. These instructions are crude and to the point, serving
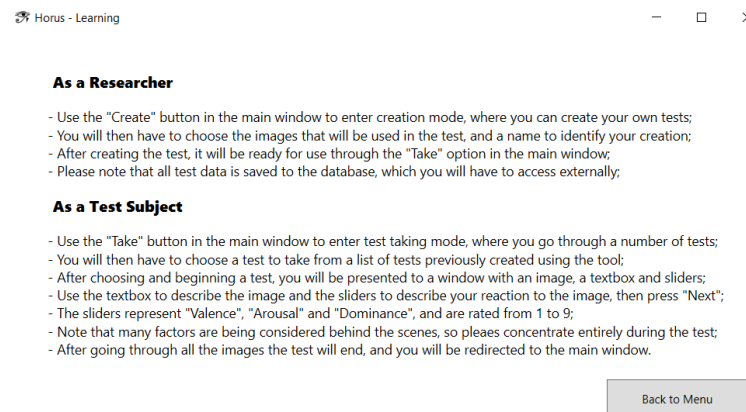

Figure 2: Instructions window

nothing more than to create a written version of what someone explaining the tool would say. Further development of this feature would probably imply a mock test, or a window where one could have a go into a model test in order to understand how it works. In any way, this feature is essential in order to ease the introduction of the tool, as time measuring is integral part of the data collecting and first time observation could influence the time a user takes to do anything in a test.

The "Create" window, encompassing the test creation process, was quite simplified, considering it represents one of the most important features of the application. In order to

create a test the user has simply to select the images of the IAPS database necessary for the test, represented by their IAPS identification number. This list of images is located in a predefined folder that can only be changed by editing the File Path variable in the source code of the Creation Window. Further implementation could include a file choosing option, as well as a format choosing option, since only images in jpg are considered. After selecting the images the user needs to introduce a test name, that will serve to identify the test in both the database and the tool interface, and press the create button, following which the test will be created and added to the available test pool.
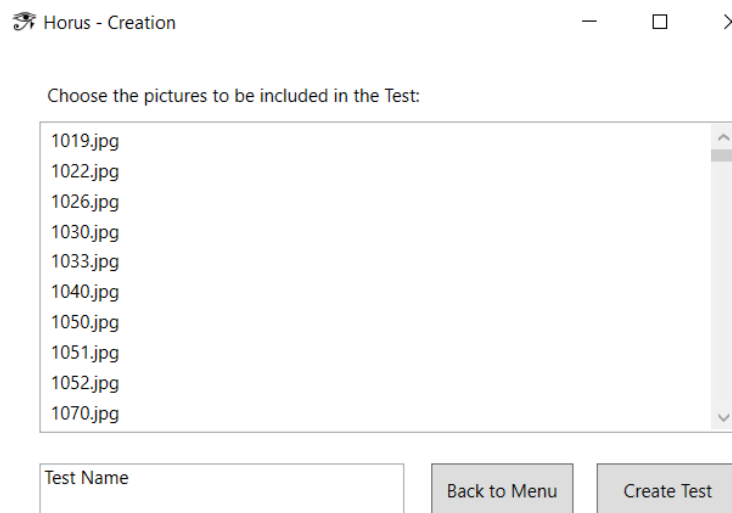

Figure 3: Test creation window

Going back to the main menu, the only option still unexplored is the test taking one. Once the user selects the "Take" option a windows containing the available tests list is displayed. These tests are the ones created using the tool or by directly inserting information in the appropriate section of the database. The user can choose only one form the list, which will then start once the take test button is pressed. This button should only be pressed once the test subject is ready to take the test, since time measurement and data collection will begin immediately after.

The test taking window is arguably the most complex of the set, even though most of its complexity is hidden from the user. There is a progress bar at the top, indicating the progress of the test through an evolving green bar that starts empty and ends fully filled. This bar varies its rate of filling depending on the number of images the test contains. Under the bar there is the image, the main attraction of the window. The image occupies a central location in order to immediately drawn in the users' attention, but is not too big in order to allow for simple usage and detection of the other interactive test components.
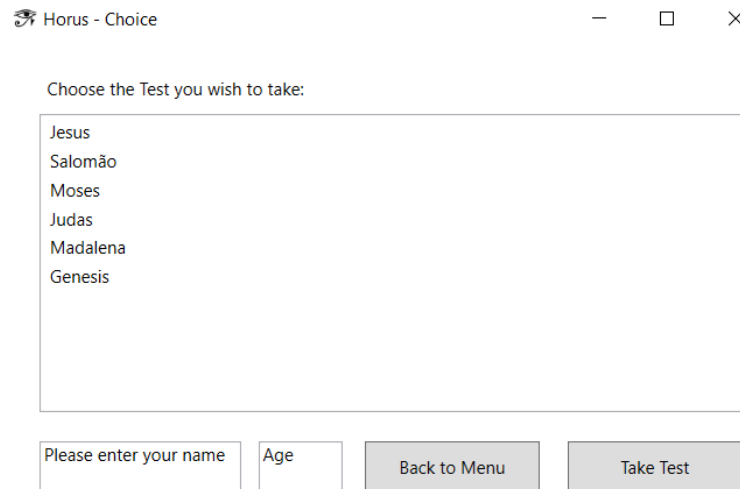
Figure 4: Choosing a test window

Under the image is a text box that is used to collect the users' description of the image, in order to collect keyboard data and text patterns. Finally there are three sliders that each represent one of the three dimensions of the emotion model being used, namely Valence, Control and Dominance. The user is asked to adapt these sliders to the way they are feeling, in order to compare it with the emotion that is predicted to be caused by the image. After clicking in the next button, a new image will appear and all values will be reset, until the test is over and all data is saved.
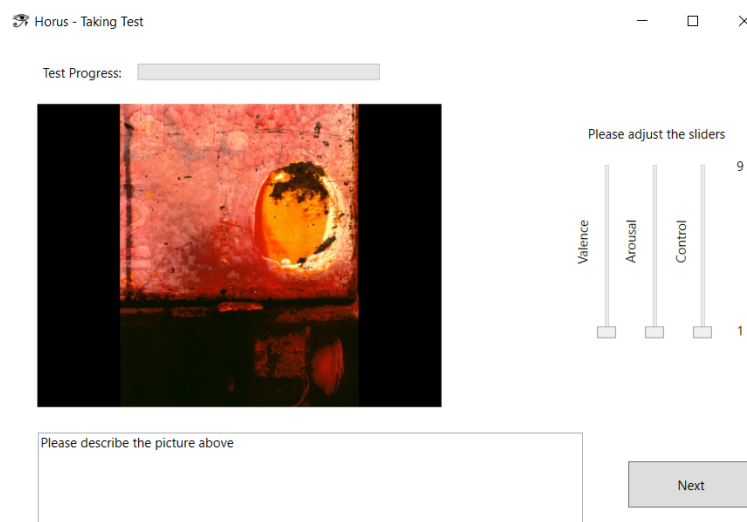

Figure 5: Test window

These were therefore the main windows and user interface details of the application, which

however simple should be and are now detailed and specified in order to support a solid documentation of the tool.

6.3 TEST FORMAT

The test format itself is quite simple, consisting of a series of images being shown to the user in quick succession. The number of images and the particulars of which one depend only on the will of the test creator, and can be manipulated to achieve different goals and study different aspects of human emotion, according to the IAPS three dimensional view. The user must describe the image in the text box in a quick but expressive manner, and evaluate the way they feel about the image using the three dimensional sliders.The user taking the test must be exposed to a quick explanation of the test and the actions they are expected to perform, in order to ease their introduction to the test and diminish the tool their learning and habituation take on the values collected in the first images. The user must concentrate on the test and must be warned to focus on the images and on what they make them feel, for constant distraction and wandering of the mind might diminish the effects the images have.

The objective of the test is always to measure the way the user interacts with the tool itself, through a number of non intrusive methods not immediately obvious to the blind eye. As such what the user writes in the text box is not nearly as important as the way the user writes it. In this same vein the levels assigned by the user in the sliders are not nearly as important as the way they move the mouse, the speed they use in their interactions, the time they take to complete each section of the test. This means that the basic explanation given to the user must be a slight misdirection, to allow for unadulterated data to be collected. The user must believe that what they are writing and evaluating is meaningful for the experience, if not only to ensure their focus on the task at hand. The warning at the beginning of the test detailing the possible perturbing nature of some images and the thank you at the end are methods to both be polite and detail the periods when the user is allowed to focus or relax.

6.4 DATA COLLECTED

There are many types or data being collected during the test. This data is what allows the study of the user's interactions, and obtaining it is therefore the main goal of the tool itself. Each of the types of information listed below are collected for each of the images in a test, being directly related to the window they are collected from. Each time the user presses the "next" button the window's information is stored, and all of the variable created to

store the data below are reset and readied for new insertions. As such these values are directly connected to the images, which is of course the point of this constructions since the relations between the values and the psychiatric evaluation of the images are what will be used to obtain conclusions that are hopefully useful. The most useful of these values are the ones that are derived from Keyboard or Mouse Dynamics, that allow for behavioural analysis. The types or data are the following:

- **Text:** Although the main objective of the test is not to evaluate what the user writes, since that would be a bit too intrusive according to the motives behind this whole venture, a lot of data can be collected through text processing that does not include the comprehension of the message in the text. As such, and in order to allow for different types of study, the text the test taker writes in the text box is saved in its entirety;

- **Valence Slider:** The first slider refers to Valence, and goes from unpleasant (1) to pleasant (9). This dimension encompasses how much the user likes the displayed image, and is the easiest to assess, be it by the investigator or by the test taker himself. The evaluation the user makes is not the most relevant, but this value is nevertheless stored each time;

- **Arousal Slider:** The second slider refers to Arousal, and goes from calm (1) to excited (9). This dimension encompasses the level of energy the image causes on the user. Arousal is classically associated with eroticism, but is in this case not limited by this view, being more of a evaluation of attractiveness of the image relating to the user personally. This is again not too relevant to the test but always saved;

- **Dominance Slider:** The third and final slider refers to Dominance or Control, and goes from controlled (1) to in control (9). This dimension encompasses the level of stress and discomfort the image causes, since human beings feel in control when they are comfortable with their surroundings or present situation and under stress when they are not. This data is saved with the rest of the sliders;

- **Time Elapsed:** This data contains the time between the moment a new image window is loaded until the next button is pressed, encompassing the entirety of the time, in seconds, the user spends on a particular image. This information is crucial since it is simple to make connections between the time a user spends in a particular type of images with particular 3D emotion values compared to others with different parameters;

- **Time Between Key Presses:** This data is stored as a list of millisecond periods, representing the time spent between key presses on the keyboard by the user. The time

between these keyboard presses can be representative of the state of mind of the test taker, and as such is valuable information in the field of keyboard dynamics;

- **Time Between Mouse Clicks:** Similarly to the previous one, this type of data is composed of a list of millisecond periods between left mouse button clicks, which are usually connected to relevant actions when in the context of a computer application. This time between clicks is also extremely relevant in the vein of mouse dynamics;

- **Mouse Click Duration:** Still in the way of mouse dynamics, this list of millisecond values represent the time the left mouse button is pressed down. This period begins when the button is pressed down and ends when it is released upwards. It can also be relevant to ascertain acuteness of though and emotion;

- **Key Press Duration:** Same as the previous one but in the area of keyboard dynamics, this list of millisecond values represent the time between the moment one key of the keyboard is released and the moment a new one is pressed, being directly related to the behaviour of the individual writing on the text box;

- **Number of Backspace Presses:** This single integer value represents the number of grammatical mistakes or errors done by the user during the typing process. These mistakes can have many causes, from the usual misspelling of a word to the rethinking of a whole sentence, which gives the possibility of creating a connection between the error and the image;

- **Number of Characters in Text:** Easily obtained by counting the number of letters and symbols in the text, this integer value can be representative of the length of the message written and therefore of its potential as a carrier of emotional burden;

- **Time Spent Typing:** This value comes from a timer that starts when the test taker begins typing and stops when the last character is inputted, representing the total time spent on the text box itself, which may vary greatly from picture to picture;

- **Mouse Movement Speed:** This final mouse dynamic indicator is a list of pixels per millisecond values measured every fifty milliseconds by calculating the location of the mouse pointer and the time it took to get there from the previous location on the screen. These speed values might vary greatly, especially when the user is thinking, but may give important information about the user's readiness and assertiveness.

Each of these items are essential for the understanding of the interactions between the test subjects and the machine providing the test. The expectation is that through the study of this data some kind of pattern and model can be established that can allow for the creation of future ways to understand the users emotion through nothing more than their interactions with the peripherals and the computer's software itself.

6.5 DATABASE FORMAT

The database, powered by *MongoDB* technology, was named "horus" according to the name of the tool itself. Using a local connection easily adaptable to remote through the connection function in the database functions file, the tool establishes a connection with the database in order to save and load data directly to and from the created tool. Although initially the plan was to create a remote database in order to allow for multiple devices to use the developed software, the lack of accessible hardware forced the database to become limited to the one computer i possess. As such this connection was hard coded, although in a simples enough way that can be modified in a matter of seconds by any knowledgeable developer.

The data saved to the database can be divided in two simple categories, namely Tests and Test Entries. As in any *MongoDB* database, the basis of data organisation are *BsonDocuments*. Each document contains a series of fields of data, and those fields can vary greatly. Each test window with a specific image has the data it generates in its own document, saved whenever the next button is pressed. Along with all the types of data described in the previous chapter, each *"testentry"* document also includes the identification of the user, namely *"UserId"*, the identification of the test, namely *"TestId"*, the age of the user requested on the test options window, namely *"UserAge"*, and finally the identification of the image on screen, namely *"ImageId"*. While these additions are not that valuable in the sense of being useful variables in understanding the subjects interactions, they are essential to identify the subjects themselves and to create basic statistic interpretations of the data. Presented below is the function utilised to save each new test entry document into the database.

```
public static void SaveTestEntry ( string UserId , string TestId , int
    UserAge , string ImageId , string Text , double SliderA , double
    SliderB , double SliderC , double TimeElapsed , List < double >
    BetweenPressesTime , List < double > BetweenClicksTime , List < double >
    ClickDuration , List < double > PressTime , int BackNum , int CharNum ,
    int TypeDuration , List < double > MouseSpeed) {
        var collection = db.GetCollection <BsonDocument >(" testentries ")
            ;
        BsonDocument testentry = new BsonDocument ();
        testentry.Add (new BsonElement ("UserId", UserId ));
        testentry.Add (new BsonElement ("TestId", TestId ));
        testentry.Add (new BsonElement ("UserAge", UserAge ));
        testentry.Add (new BsonElement ("ImageId", ImageId ));
        testentry.Add (new BsonElement ("Text", Text ));
        testentry.Add (new BsonElement ("SliderA", SliderA ));
        testentry.Add (new BsonElement ("SliderB", SliderB ));
        testentry.Add (new BsonElement ("SliderC", SliderC ));
        testentry.Add (new BsonElement ("TimeElapsed", TimeElapsed ));
```

```
        testentry.Add(new BsonElement("BetweenPressesTime", new
            BsonArray(BetweenPressesTime)));
        testentry.Add(new BsonElement("BetweenClicksTime", new
            BsonArray(BetweenClicksTime)));
        testentry.Add(new BsonElement("ClickDuration", new BsonArray(
            ClickDuration)));
        testentry.Add(new BsonElement("PressTime", new BsonArray(
            PressTime)));
        testentry.Add(new BsonElement("BackNum", BackNum));
        testentry.Add(new BsonElement("CharNum", CharNum));
        testentry.Add(new BsonElement("TypeDuration", TypeDuration));
        testentry.Add(new BsonElement("MouseSpeed", new BsonArray(
            MouseSpeed)));
        collection.InsertOne(testentry);
    }
```

Listing 6.1: Saving Test Entry document function

The second type of data entry in the database is the Test itself, a simple insert with two fields, namely *"ImageList"*, a list of the names of the images included in the test, and *"Test-Name"*, the name that the creator choose for the test. These fields are merely utilitarian, as the image list is necessary to search for the images in the indicated file so they can be shown to the user and the name of the test is essential to identify the test entries that come from each particular tests, serving a similar purpose as a primary key. Below the function used to insert new tests into the database can be observed in full.

```
    public static void CreateTest( String TestName, List<string> Images )
    {
        var collection = db.GetCollection<BsonDocument>("tests");
        BsonDocument test = new BsonDocument();
        test.Add(new BsonElement("ImageList", new BsonArray(Images)));
        test.Add(new BsonElement("TestName", TestName));
        collection.InsertOne(test);
    }
```

Listing 6.2: Saving Test document function

These two functions are the only way the application interacts directly with the database in a storage manner. The tool also searches for the tests in order to present them in the test options window, but as that is a simple search of all the test names it will not be specifies here. As such we can summarise that the database is constituted only by two types of documents, tests and test entries, each with the types of data explored above. A new type of entry will be created for data analysis, but that operation will be explored in a later chapter.

TESTS AND DATA COLLECTION

Having completed the development and the documentation of the tool itself it was time to put it to use. As such a well defined testing phase was necessary. In this chapter the decisions and process relating to the tests will be explored, since data collection is an essential part of this project, relating to the fourth main objective of this thesis.

7.1 PLANS AND REASONING

The main objective of the testing phase was to obtain sufficient data to articulate an hypothesis regarding the effectiveness of the tool and specifically of the usage of this kind of data collection in the identification of the emotions of users. In order to do so it was necessary to gather a study group large enough to produce a relevant amount of data. It was determined that at least twenty test subjects were necessary in order to generate enough data to create an effective study. It also was estimated that the test would have to include at least twenty images in order to create a strong enough impression on the user so that their emotional output fluctuated enough to detect. These two necessities were estimates reached after trial and error experimentation. As such a test was created that had this number of pictures, ready for usage by any willing test subject, named "Genesis" in going with the biblical theme present in the test naming conventions to this point.

However, since many aspects of emotion could be studied through the developed tool it was decided that the dimension of Valence would be the focus of the study, since Valence, or pleasantness in other words, is the dimension easier to influence through visual stimulation and therefore the one that would most likely create interesting results in this planned test. In order to obtain strong data regarding the Valence dimension two sets of images were chosen from the IAPS database, namely the ten images with the highest Valence values and the ten images with the lowest Valence values. The ten pictures with high values were to be shown first in order to ease the user into the test process, and the low value pictures were to be shown last, in order to create a lasting impression on the user strong enough to be expressed in their usage of the device. We can observe below the specifications of the

images chosen through this process and that constitute the "Genesis" test.

IAPS	_id	aromn_all	arosd_all	om1mn_a	om1sd_all	om2mn_a	om2sd_al	set_all	valmn_all	valsd_all	ï»¿desc
1710	5a0d8c52€	5.41	2.34	6.55	1.98	.	.	3\r	8.34	1.12	Puppies
1750	5a0d8c52€	4.1	2.31	6.15	2.01	.	.	3\r	8.28	1.07	Bunnies
5833	5a0d8c52€	5.71	2.66	.	.	6.97	1.93	16\r	8.22	1.08	Beach
1460	5a0d8c52€	4.31	2.63	6	2	.	.	5\r	8.21	1.21	Kitten
2050	5a0d8c52€	4.57	2.53	7.71	2.53	.	.	1\r	8.2	1.31	Baby
1440	5a0d8c52€	4.61	2.54	6.05	2.38	.	.	6\r	8.19	1.53	Seal
2040	5a0d8c52€	4.64	2.54	7.33	2.54	.	.	1\r	8.17	1.6	Baby
2070	5a0d8c52€	4.51	2.74	7.14	2.1	.	.	4\r	8.17	1.46	Baby
8190	5a0d8c52€	6.28	2.57	6.14	2.74	.	.	5\r	8.1	1.39	Skier
2080	5a0d8c52€	4.7	2.59	7.08	1.96	.	.	4\r	8.09	1.47	Babies
3053	5a0d8c52€	6.91	2.57	2.33	1.94	2.17	1.59	7\r	1.31	0.97	BurnVictim
3102	5a0d8c52€	6.58	2.69	2.18	1.9	2.42	1.56	7\r	1.4	1.14	BurnVictim
3000	5a0d8c52€	7.26	2.1	2.99	2.1	.	.	1\r	1.45	1.2	Mutilation
3064	5a0d8c52€	6.41	2.62	2.63	2.04	.	.	10\r	1.45	0.97	Mutilation
3170	5a0d8c52€	7.21	1.99	2.7	1.89	.	.	3\r	1.46	1.01	BabyTumor
3080	5a0d8c52€	7.22	1.97	2.85	2.1	.	.	6\r	1.48	0.95	Mutilation
3063	5a0d8c52€	6.35	2.6	2.7	2.02	.	.	10\r	1.49	0.96	Mutilation
3131	5a0d8c52€	6.61	2.34	2.73	1.93	.	.	17\r	1.51	0.97	Mutilation
9410	5a0d8c52€	7.07	2.06	2.81	1.99	.	.	4\r	1.51	1.15	Soldier
3015	5a0d8c52€	5.9	2.82	2.84	2.13	.	.	11\r	1.52	0.95	Accident

Figure 6: Images chosen for the Genesis test

The difficulty with this endeavour was the fact that the test itself had to be taken in a single computer possessing the database and the developed tool, due to the lack of more hardware capable of supporting the necessary specifications. This, along with the fact that going through the test takes from fifteen to thirty minutes makes the data gathering rather difficult in the small time-frame available. The test also demands an environment that allows for concentration in order to ensure that the user is not too distracted by outside stimuli do devalue the data collected. All these factors aligned made it so that the data gathering process took a lot more time than desired, since the conjunction of elements necessary to produce a valid test were many times hard to find or create. It was still possible to obtain the necessary data, with the process detailed in the section below.

7.2 TESTING PROCESS

The testing process was quite simple, but highly thought out in order to both standardise the test experience and improve the quality of the test data itself. As such the general procedure was similar for all the test subjects, with the exception of the test venue itself. The hardware was not a variable, since all the tests were executed in my personal computer, assuring a stable executable environment. As for the steps in the test process, each test went by the following sequence:

- **Setup:** The computer was connected to a power source and positioned in a stable surface, usually a table. The first step was to execute a batch file starting the database in the command line. After the database was ready a shortcut to the release version of the developed application was used to start it up. After ensuring that the tool and the database were connected the set up process was complete, only needing the user to kick start the process;

- **Explanation:** After the set up process the user was sat comfortably in front of the computer and provided with a simple voiced explanation about the tool and what was expected of them. The explanation revolved around the three dimensional sliders, the text box and the tasks the user needed to perform. This was necessary in order to ensure the user understood the tool before interacting with it. Almost no understanding about what was actually being evaluated and saved in the test was implanted in the user, and no preparation for the images was allowed. The disclaimer about possible disturbing imagery was the only previous data about the images the user got;

- **Test:** The test itself begins when the user clicks the "ok" button in the disclaimer window. However the process begins beforehand, when the user selects a test to take and enters his or her name and age. After the test begins, the user looks at the image displayed in the screen and writes its description, followed by the usage of the sliders to describe the user's feeling towards it. The user sometimes asks about the sliders, since their nature can be elusive and an unfocused user can be forgetful. After reaching the eleventh image the user the "bad" images begin, and the user tends to have a negative reaction towards it. The user was reassured in this case and asked to continue even if they find the images disturbing, since their reaction is the purpose of the study. After finishing the twenty images the user is thanked for their part in the test, both by the researcher and by the application itself.

- **Post-Test Video:** After finishing the test a small one minute internet video is introduced to the user. The video, based on the internet *Youtube* website, consists of a photographer trying to get ten baby kittens together for a photograph, and is considered highly pleasing. This video's purpose is twofold: firstly to reassure the user and allow for him or her to relax, since the previous ten images were highly disturbing; secondly to verify if the user can smile and emit positive reactions to this video after being shown the disturbing imagery, in order to verify if such interactions have a long term effect or not. When the video ends the user is thanked once again and dispensed.

- **Data Verification:** After the test was finished and the user dispensed from their strenuous task the data was verified using the "Studio 3T" *MongoDB* interface program. This was done to certify the data and ensure that the test went smoothly, on order to

not waste any data through errors and mistakes relating to the database connection itself. This verification also allowed for some early conclusions, since some of the data stored could be compared to the previous tests data in order to form basic conclusions useful during the more in depth study of the data.

The specified test sequence was verified in every test by every test subject. All test subjects are noted by name in this dissertation's acknowledgements, but no data was used in an individual sense, in the way that no one but myself has access to individual results and that all explorations and study processes executed on the data itself looked into the origin of the data itself. The results are therefore independent of the subjects themselves, something that is essential in any study that wishes to protect its participants privacy and that is doubly essential in the context of this dissertation, since the whole point of the thesis is the understanding of people's emotions through a non intrusive method.

7.3   COLLECTED DATA CONSIDERATIONS

The test data collected amounts to **four hundred and eighty (480)** test entry documents, which originates from tests taken by **twenty three (23)** volunteers. Each of these documents involves thousands of records, especially in the mouse speed and time between clicks and presses categories. The test data itself will not be displayed nor included here since it is identified mainly by name of the user, and since the whole point of the project is to avoid being intrusive the users personal information will not be displayed here. All data analysis will be displayed in the next chapter in a collective manner, with no discrimination of any kind of identification. Statistical treatment of the data is essential due to the fact that it protects anonymity of the subjects and allows for conclusions about the subjects' interaction in a simple visual manner.

Even so there are some details that must be explained about the data. Firstly, even though there are 480 entries some of these ended up being discarded in the early analysis since they referred to images where the user skipped the writing and the slider usage due to double click or mistake in the usage of the next button, not being representative of the user's state. Even though these mistakes could be linked to the users emotional state they are too sporadic and uneven to use as an evaluative tool. Another detail worth mentioning is the elimination of zero values during this early analysis, since although they can be representative they are not good for statistical purposes and are in some cases too prevalent due to user inactivity, as is case in the mouse movement speed, where the user obviously does not use the mouse when writing. These and more early data treatment will be better described in the following chapter of analysis and discussion of the obtained data.

ANALYSIS AND DISCUSSION

In this chapter the data collected in the form of the test entries saved after the volunteers took the "Genesis" test prepared to evaluate valence are treated and analysed. There is also a discussion about the effectiveness of the methods used, as well as of the effectiveness of the developed tool itself in obtaining data usable in the study of emotion in human computer interactions.

8.1 EARLY TREATMENT

Before a statistical treatment and subsequent definite conclusions could be made it was necessary to correct some initial issues with the data, as well as prepare it for analysis. With a total of 23 test being done the amount of document entries in the database amounted to 480, a respectable if somewhat small amount. The existence of 18 variables per document, some with hundreds of entry, make the data a bit larger, but still fragile in its ability to support decisive conclusions. This was exacerbated by some issues. The first issue was that some of the users had skipped some of the images during their tests, as stated before. Errors due to quick double clicks in the next button or due to forgetfulness of some of the actions necessary during the testing contributed to this issue. This caused that from the total of 480 test entries only 464 remained usable, due to the necessity of eliminating 16 database entries. These entries had near zero values and no text, due to the immediate skipping of the image on display. As such the pool of test data diminished slightly, which was unfortunate if unavoidable. The second issue was that a lot of zero values were present in some of the variables, due to actions too quick to measure, to measurement errors or simply due to inactivity, fruit of distraction or of focus on other parts of the test. This was dealt with later and not directly on the database.

The first step in studying the data was readying it for the study itself. The used test, "Genesis", was built from the ground up in order to study the effects of Valence variations in the interactions of the test subjects. As such it was necessary to connect the test entries

obtained to the existing values of valence attributed by the official International Affective Picture System (IAPS) subjects evaluation, present in the "AllSubjects_1-20" file obtained with the database itself when requested to the laboratory that created and supports the system. A small script was created in order to give origin to a new collection, one that joined the information given by the laboratory and the data obtained from the tests in the same spot, in order to allow for study based on the valence levels attributed by the high amount of subjects used to categorise the IAPS image collection. The script utilized to fuse the two collections can be observed below.

```
db.getCollection("testentries").aggregate(
 // Pipeline
 [
   // Stage 1
   {
     $addFields: {
         "ImageId": {$substr : ["$ImageId", 0, {$indexOfBytes:  ["
             $ImageId", "."]}]}
     }
   },
   // Stage 2
   {
     $lookup: // Equality Match
     {
         from: "AllSubjects_1-20",
         localField: "ImageId",
         foreignField: "IAPS",
         as: "IAPS_Info"
     }},
   // Stage 3
   {
     $unwind: {
         path : "$IAPS_Info",
         includeArrayIndex : "arrayIndex", // optional
         preserveNullAndEmptyArrays : false // optional
     }
   },
   // Stage 4
   {
     $out: "dados_joined"
   },
 ]
);
```

Listing 8.1: Joining IAPS entries

After creating a collection where the data was all joined came the time to begin data treatment. In order to do so a platform and appropriate tools had to be chosen. Due to the enormous variety of options and high entry toll of most of them the chosen platform was *RStudio* with its highly mathematical R programming language, a choice also motivated by its simple integration with *MongoDB* databases that would allow for seamless integration. R is a language that is widely used by data scientists and miners in order to design statistical or math based software, and is described as a free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. As such, and with this early treatment of the data complete, the data was migrated into the new platform in order to give it a mathematical evaluation using R.

8.2 GENERAL DATA ANALYSIS

The first step in integrating the data into the new environment was importing it from the database. As such it was necessary to import a dedicated library and connect the database to the studio environment. Afterwards two data collections were created, one with the data entries regarding images with high (above 5) Valence values and one regarding images with low (under 5) Valence values. This was done to differentiate the data and to allow for later considerations about a specific group to be taken, as patterns could hopefully be identified relating to high or low valences in these test entries. The code used can be observed below.

```
#Conection
library(mongolite)
collection = mongo(collection = "dados_joined", db = "horus")
dados <- collection$find('{}')
valence_low <- dados[dados$IAPS_Info$valmn < 5, ]
valence_high <- dados[dados$IAPS_Info$valmn >= 5, ]
```

Listing 8.2: Importing data and generating groups

After the operation i noticed some Type Duration values were negative, due to some software error during testing. As such these were manually eliminated in the database, after which the data was reintroduced in the study environment. This made it so that the number of documents went from 464 to 461. After importing the data and creating the two collections of different Valence values it was time to analyse each relevant variable in a statistical manner. First however it was necessary to elaborate some auxiliary functions in order to both eliminate bad or irrelevant values in the data and ignore the significant amount of outliers that plague the tests. This had to be made in order to allow for the analysis of the data without considering values like zeros or "Not a Number" values that occur occasionally and to reduce the amount of error in the graphical analysis, since these values

and outliers, which are extreme discrepancies in the data due to error or any unimportant factor, can create a disturbance in the limits of the values of these variables. As such, the two functions created for these purposes can be observed below.

```
#Remove Outliers Function
remove_outliers <- function(x, na.rm = TRUE, ...)
{
    qnt <- quantile(x, probs=c(.25, .75), na.rm = na.rm, ...)
    H <- 1.5 * IQR(x, na.rm = na.rm)
    y <- x
    y[x < (qnt[1] - H)] <- NA
    y[x > (qnt[2] + H)] <- NA
    y[!is.na(y)]
}
```

Listing 8.3: Function to Remove Outliers

```
#Filter Function
filtra <- function(input)
{
  input <- input[!is.na(input)]
  input <- as.numeric(input)
  input <- input[input != 0]
  input
}
```

Listing 8.4: Function to Filter bad values in the Data

The other auxiliary function that was necessary for a good analysis was the "unlist" one, a function already defined in the basic RStudio environment. This function was necessary in order to allow for the usage of the data in lists, as is the case in most of the variables. As such it was applied in every variable that exists as a list. With these three described the auxiliary functions were accounted for, and specific analysis could be done. The first analysis done was a general summary of statistical information in the form of data like mean and median of each and every relevant data type, specifically the nine before mentioned measured dynamic values. the standard deviation and the KS Test were also calculated for each of these. The deviation is a simple tool to understand and quantify the amount of variation or dispersion of the values present in the data set, since a low standard deviation indicates that the data points tend to be close to the mean of the set, while a high deviation shows that the existing data is spread out over a wider range of values. The Kolmogorov–Smirnov test, or KS test for short, was used in its two-sample facet, since it is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples presented, in this case the high and low Valence collections. The test checks if two

data sets come from the same distribution, and the important thing is to check its p-value, since one can reject the null hypothesis that the two samples were drawn from the same distribution if the p-value is less than the established significance level, in this case 0.05. These tests were performed for each and every of the nine main data types, for each of the two Valence collections, in order to gain a general understanding of the state of the data collected through the testing phase. The functions utilised to obtain these values are listed below for one of the data types, being used in the others in its adapted state.

```
# Character Number
ks.test(remove_outliers(valence_low$CharNum), remove_outliers(
    valence_high$CharNum))
summary(remove_outliers(valence_low$CharNum))
sd(remove_outliers(valence_low$CharNum))
summary(remove_outliers(valence_high$CharNum))
sd(remove_outliers(valence_high$CharNum))
```

Listing 8.5: General Statistical Analysis

After executing this code, adapted for each and every data type, the data was compiled onto a table in order to allow for simpler analysis and comparison between values. The table contains the whole data given by the general summary, as well and the standard deviation and the p-value from the KS test. With this visualisation of the data we can already obtain some useful information about how the indicators are related, and take some conclusions about the relevance of said data types. The table can be examined below, with each data type indicated on the left, divided between High or Low Valence, each statistical indicator indicated on the top, with the final one being the KS test.

		Min	1st Qu.	Median	Mean	3rd Qu.	Max	S. Dev.	Ks Test p-value
Character Number	High Valence	4	14	21	26.14	34	79	17.21531	0.02884
	Low Valence	3	15.75	26	30.66	38.25	91	19.6656	
Backspace Presses	High Valence	0	0	1	2.048	3	10	2.543394	0.05069
	Low Valence	0	0	1	3.074	5	16	4.165282	
Time Elapsed	High Valence	10.92	24.01	34.34	38.56	49.67	93.22	18.39447	0.7668
	Low Valence	7.689	22.632	33.159	37.907	48.268	85.456	19.66392	
Type Duration	High Valence	532	4458	10562	12715	19079	42862	10083.26	< 2.2E-16
	Low Valence	263	5082	11058	14514	21770	50125	11800.24	
Click Duration	High Valence	1	132	133	421.8	617.5	2075	499.1917	6.00E-15
	Low Valence	11	121	133	235.5	181	1092	244.888	
Time Between Presses	High Valence	1	104	152	183.3	232	639	130.4182	2.89E-15
	Low Valence	1	72.25	143	167.88	219	595	124.3557	
Time Between Clicks	High Valence	11	425	1058	1439	1882	6689	1399.354	0.007886
	Low Valence	1	328	1080	1733	2100	9825	1990.433	
Press Time	High Valence	14	70	104	92.27	107	147	28.11155	0.05288
	Low Valence	13	70	105	92.62	107	147	29.43688	
Mouse Speed	High Valence	-0.0000071	0.03125	0.0662913	0.13517	0.1803281	0.7048763	0.1559297	< 2.2E-16
	Low Valence	0.01266	0.03226	0.08224	0.17407	0.23423	0.90253	0.2054756	

Figure 7: General Statistics Table

One of the things that are immediately noticeable is the negative value of the High Valence collection **Mouse Speed** data type Minimum. It is the only negative value of the whole data set, and was not manually eliminated since it is so infinitesimally small it is almost zero, and a great amount of Mouse Speed values are zeros. Therefore the analysis suffers almost no impact. Since it does not create significant error in the analysis the document was maintained due to the already small data pool, in order to not diminish it any further.

When it comes to the values of standard deviation, one can attain some conclusions on the dispersion of the data. First of all, in seven of nine cases the deviation is higher in the Low Valence collection that in the High Valence collections, something that cannot be ignored. The lowest deviations can be found in the **Mouse Speed** values and the **Backspace Presses** values, although these can be justified due to the concrete nature of the values themselves, since the first is calculated in milliseconds and the second has a very small maximum and therefore a very small margin to verify divergence. The high values like the **Type Duration** can also be justified due to the nature of the data type, that represents time spent writing and therefore can be quite extensive. One of the only conclusions that can be stated relative to the data types is that some data types like **Click Duration** and **Time Between Clicks** present significantly large differences between the two Valence collections, according to their means and standard deviation.

When it comes to the KS test, results are simpler to assess. Since the established significance level is 0.05 it can immediately be noted which data types possess p-values under this level. Once again seven in nine cases have a p-value well under the level, although not the same seven as before. Only **Time Elapsed**, **Backspace Presses** and **Press Time** have a p-value higher than the level, and in the case of the last two the difference is very small. **Click Duration** and **Time Between Presses** have values well under the level, with ridiculously small values, and **Type Duration** and **Mouse Speed** have values so low that go beyond the minimum value calculated by the program, namely 2.2E-16. Although it has been stated that this type of scientific method of testing hypotheses by statistical analysis stands on a flimsy foundation due to the scientific community's over reliance on the p-value level of 0.05, it was still the best option to evaluate the data (Goodman, 1954). As such these results tell us that the seven variables under the established level suffer significant differences between the two Valence collections, although the level of difference is not equal to the difference in p-value. Still, we can already tell that the three variables that have a higher p-value will most likely not be useful in the detection of emotions, since the difference between collections is so small. This goes hand in hand with the results obtained through the study of the rest of the elements in the table, even if not completely (C., 2011). The detection of seven possible variables that provide relevant data is however encouraging,

even if insufficient.

The information about the data is relevant and useful for taking conclusions, but it is still not enough to relate the variations in emotion to specific fluctuations of data types. As such, and having done a general analysis, a more specific analysis individual to each data type was necessary. Therefore that was an obvious next step.

8.3 INDIVIDUAL DATA ANALYSIS

The statistical analysis done before was useful but not determinant. Conclusions made about this matter need to be built on a strong foundation, a foundation built on individual analysis focused on each and every data type, in order to find if one or more of these are relevant in the detection and modulation of emotion in computer systems. As such a graphical comparison of the data types was necessary, since a graphical analysis can be critical to conclusions on comparisons of data collections. The two types of graphs utilised were the Box Plot, or Box and Whisker Diagram, since it displays the distribution quite well with an attention to relevant outliers for as many collections as needed, and the Density Plot combination of graph, with the addition of lines for the collection not initially used. These choices were influenced by the necessity to observe the differences between data points on the same screen at the same time and for the high amount of easily obtainable information on the density and distribution of the data presented on these graphs. The code used for the graphics and for the accessory necessities is displayed below for one of the variables, and used in its adapted form for the rest.

```
# ClickDuration
  boxplot(remove_outliers(filtra(unlist(valence_low$ClickDuration)))
      , remove_outliers(filtra(unlist(valence_high$ClickDuration))),
      ylab="ClickDuration", xlab="Valence", names=c("Low", "High"),
      outline=FALSE)
  plot(density(remove_outliers(filtra(unlist(valence_low$
      ClickDuration)))))
  lines(density(remove_outliers(filtra(unlist(valence_high$
      ClickDuration)))), lty=2)
  legend(x=800, y=0.015,c("Low Valence","High Valence"), lty=c(1,2))
```

Listing 8.6: Individual Graphical Analysis

The usage of the "filtra" function is restricted to variables that are lists with high amount of values, since these tend to exhibit too many values could perturb the analysis. The "unlist" function is also used to work with these lists, since it unpacks them. The first function, "box-

plot", creates the Box Plot with the two collections. The "plot" and "lines" functions create a density plot that represents one of the collections with continuous lines and the other with discontinuous ones, identified with the "legend" function. The parameter "outline=FALSE" is used whenever the amount of outliers is too much to be useful and therefore distorts the graph too much. Whenever it is used it will be stated as so.

8.3.1 *Backspace Presses*

Beginning with the **Backspace Presses** data type, the number of times the we can see in the Box Plot below that the lower quartiles of the two collections are identical, which makes sense since erasing the line of instructions on the data box was the first step of any test taker. However, the higher two quartiles are somewhat different, with the Low Valence collection presenting a higher amount of values, with a few relevant outliers, enough to state that with Low Valence images the users tend to use the backspace key more.
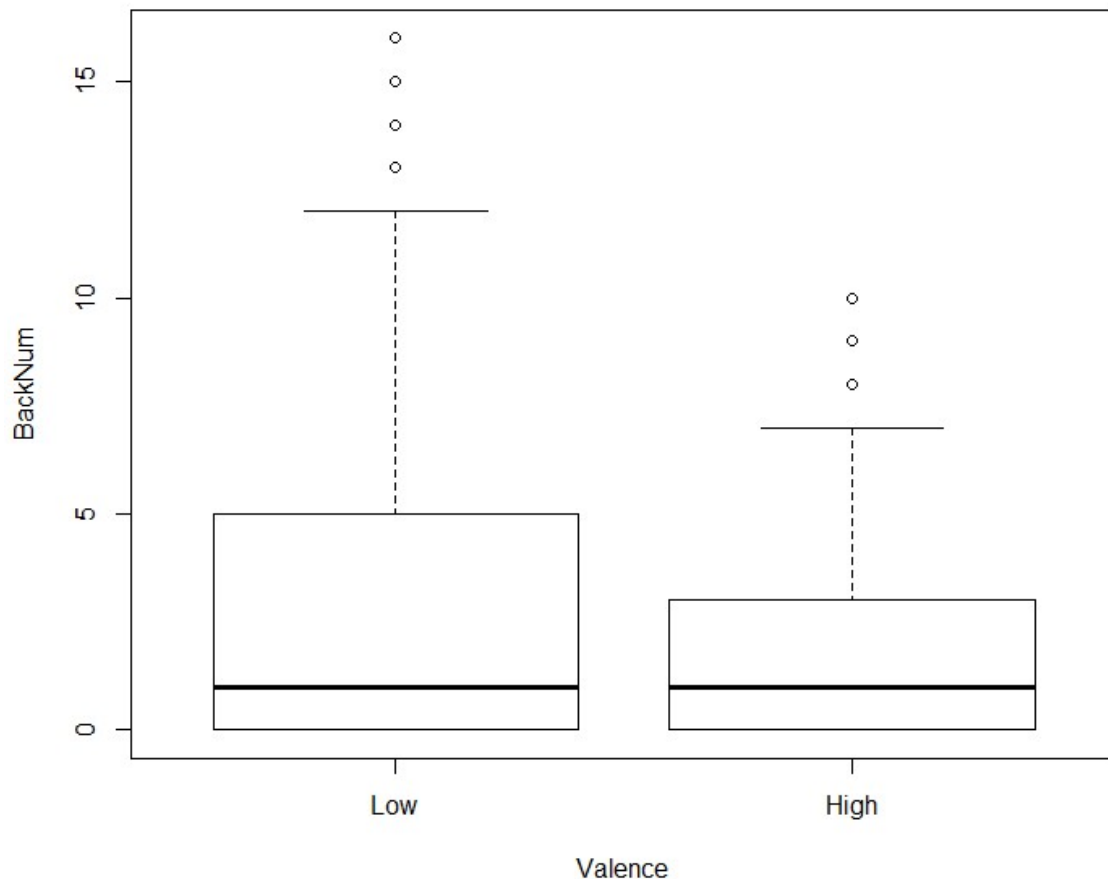
Figure 8: Backspace Presses Box Plot

Following up with the Density Plot for the same indicator we can deduce that even if the number of backspace presses is higher with Low Valence images this difference is not too high. However it is safe to assume that most test subjects do not commit as much mistakes with High Valence images as they do with Low ones, which can be connected to the negative emotions the Low ones tend to cause. People tend to type more nervously and therefore have a higher probability of committing mistakes when exposed to disturbing imagery, an hypothesis proved here. The plot can be seen below.
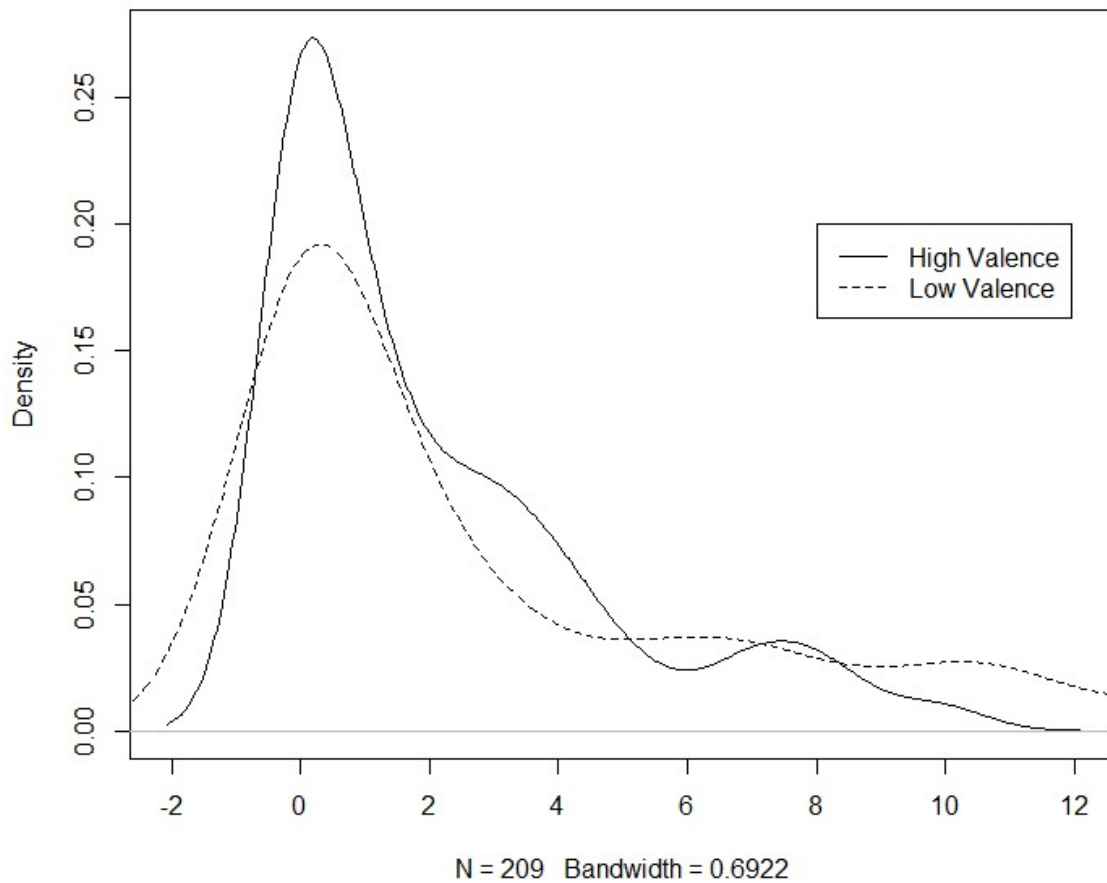
Figure 9: Backspace Presses Density Plot

8.3.2 *Time Between Clicks*

Moving onto the **Time Between Clicks** data type, we can see the Box Plot below. This indicator represents the time a user take between clicks on the left mouse button, which in turn can be understood as actual actions. The outliers were disabled on this graph since the enormous amount made visualisation of the quartile distribution difficult. The fact that

this was necessary is in itself a bad sign, since the amount of valuable data is small and the distribution was too wide for this type of graph. We can however see that the distribution we are left with is very similar, with only a slightly higher amount of values present in the Low Valence collection. The lower halves of the collections are essential the same, something that is to be expected due to the fact that clicks tend to be faster when an action has been decided by the user and lower when the user is undecided. We can deduce that there is a difference in values, but not much, and maybe not enough to detect emotional change. However the difference is still there, and therefore cannot and should not be ignored.
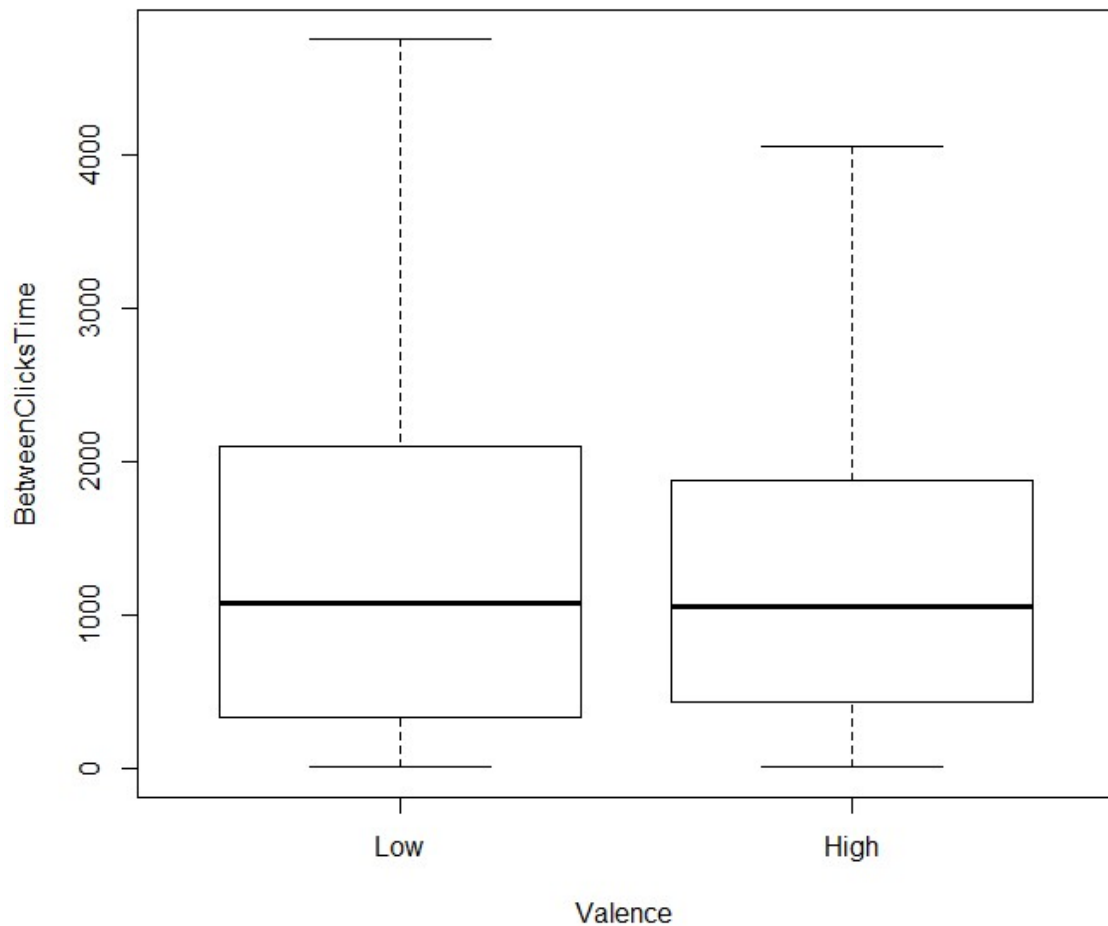
Figure 10: Time Between Clicks Box Plot

In the indicator's Density Plot graph we can observe that both collections are almost matched, with the exception of some Low Valence values that experience a small low in the second peak of the High Valence values. The differences between the rest of the graphs are however too small, and this small discrepancy cannot be safely attributed to an emotional change in the subject. The nature of the indicator itself is an issues here, since even if filtering a majority of unwanted values a user tends to spend a lot of time without clicking

when thinking about what to do and especially when writing. As such this indicator has a low value in the context of this study, even if with more data this could possibly change. The differences are there and the p-value confirms them, but the amount of data and the divergence between collections is simply too small to make a strong case based on the effects of Valence variation on this indicator. The plot is displayed below with both collection's lines clearly defined.
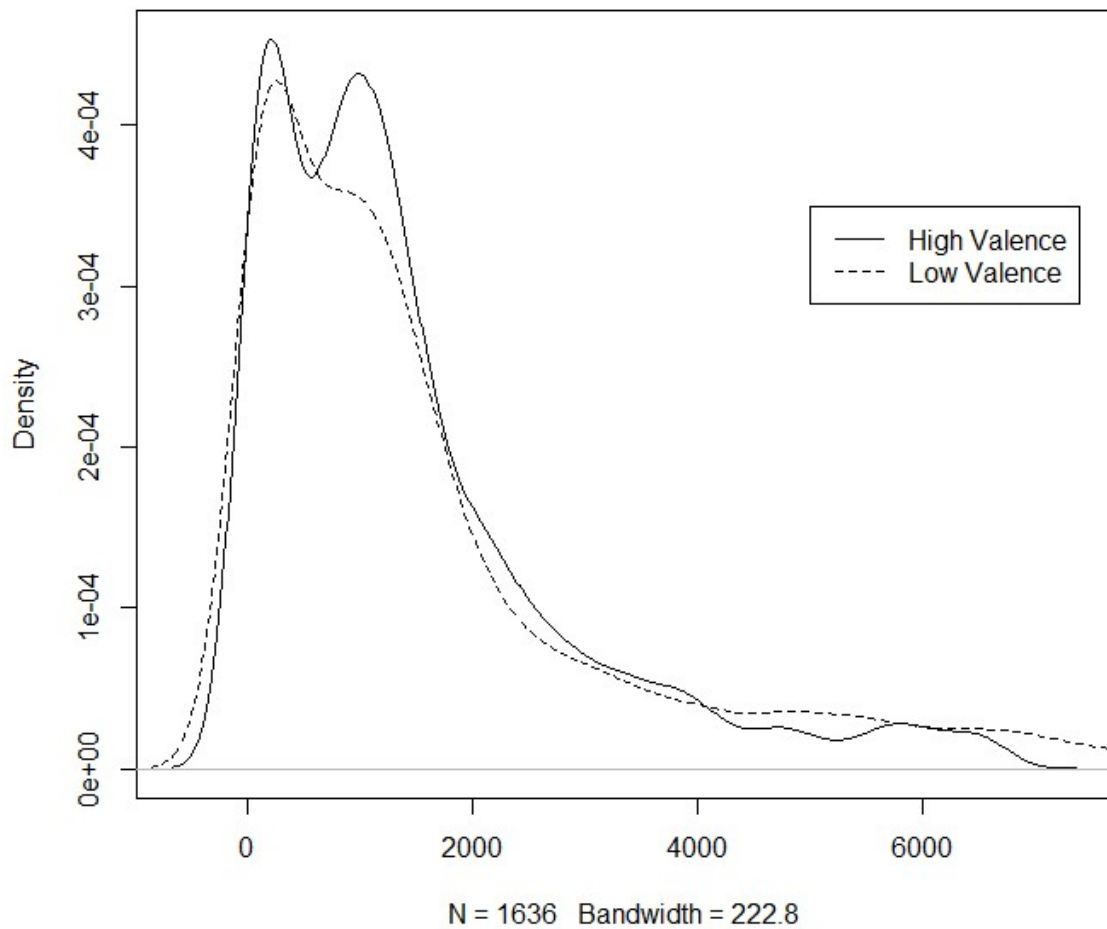

Figure 11: Time Between Clicks Density Plot

8.3.3 *Time Between Presses*

The next data type to be analysed is the **Time Between Presses**. The Box Plot found below can allow for some conclusions on the disparities between the two Valence collections, which are minor. At a first glance it can be said that the Low Valence collection has a majority of the values slightly under the High Valence collection, but this difference is min-

imal. The fact that both collections have the same minimum and almost the same median makes differences difficult to affirm. The Low Valence collection also spots values over the maximum of its counterpart, but again in a very small amount of change. This graph doesn't give us many reasons to believe this indicator can be useful enough in the detection of Valence changes in the user, even if the p-value would say so. The graph was also presented without outliers due to the great amount of variety they present, something that is inevitable when dealing with large amounts of calculations, but still somewhat negative in the scheme of things, since the amount of useful data decreases and any hypothesis formed is weakened.
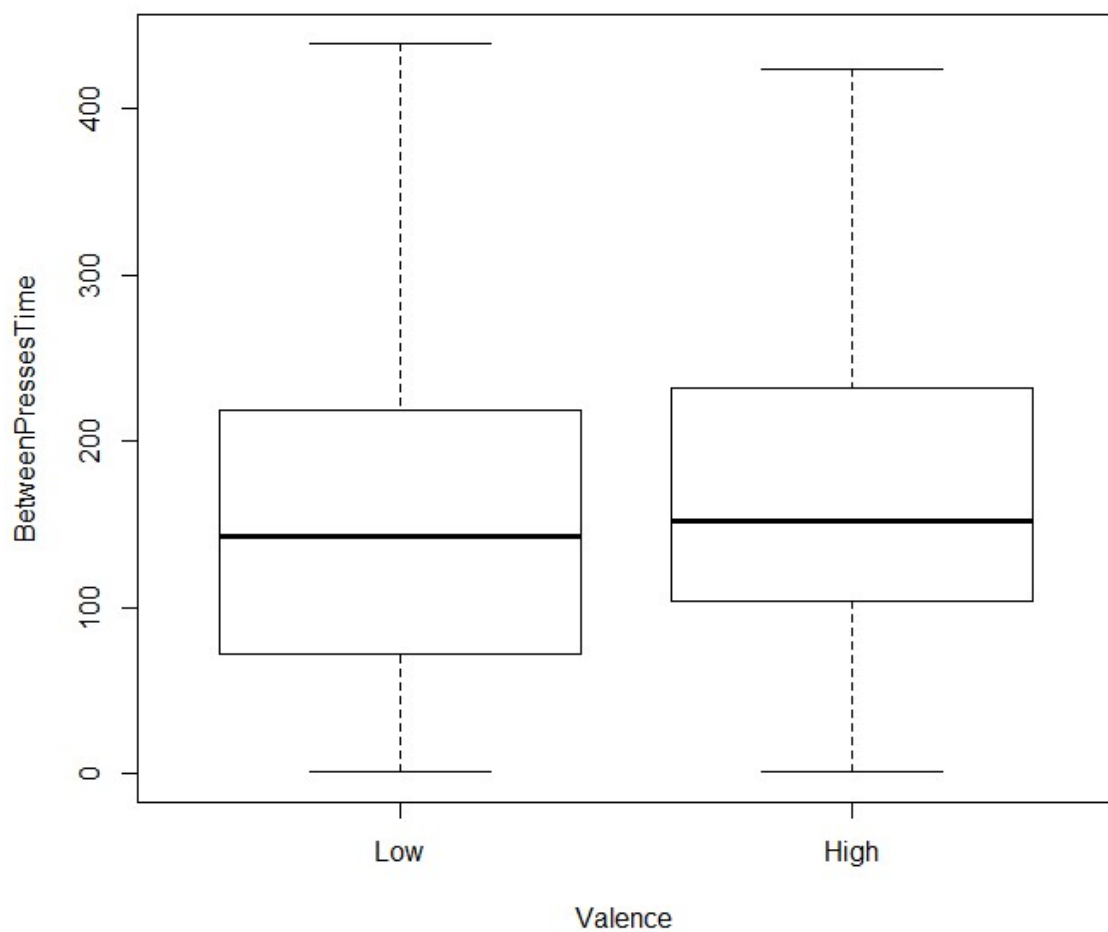
Figure 12: Time Between Presses Box Plot

Examining the Density Plot presented below we can observe that in fact the two collections present very similar lines, with a slight difference around the zero value of time between presses, which is higher in the Low Valence collection than in the High Valence one. It could be said that when exposed to disturbing imagery a majority of people write faster

in order to get it over with and move on, taking the image out of their field of view. Even though a lot of the highs in the early parts of the graph can be naturally attributed to the fact that most people type rather efficiently and fast in this modern day and age, the differences in the Valence collections give us enough ground to state that when exposed to things they don't like, a significant amount of people spend less time between key presses. The graph can be observed below.
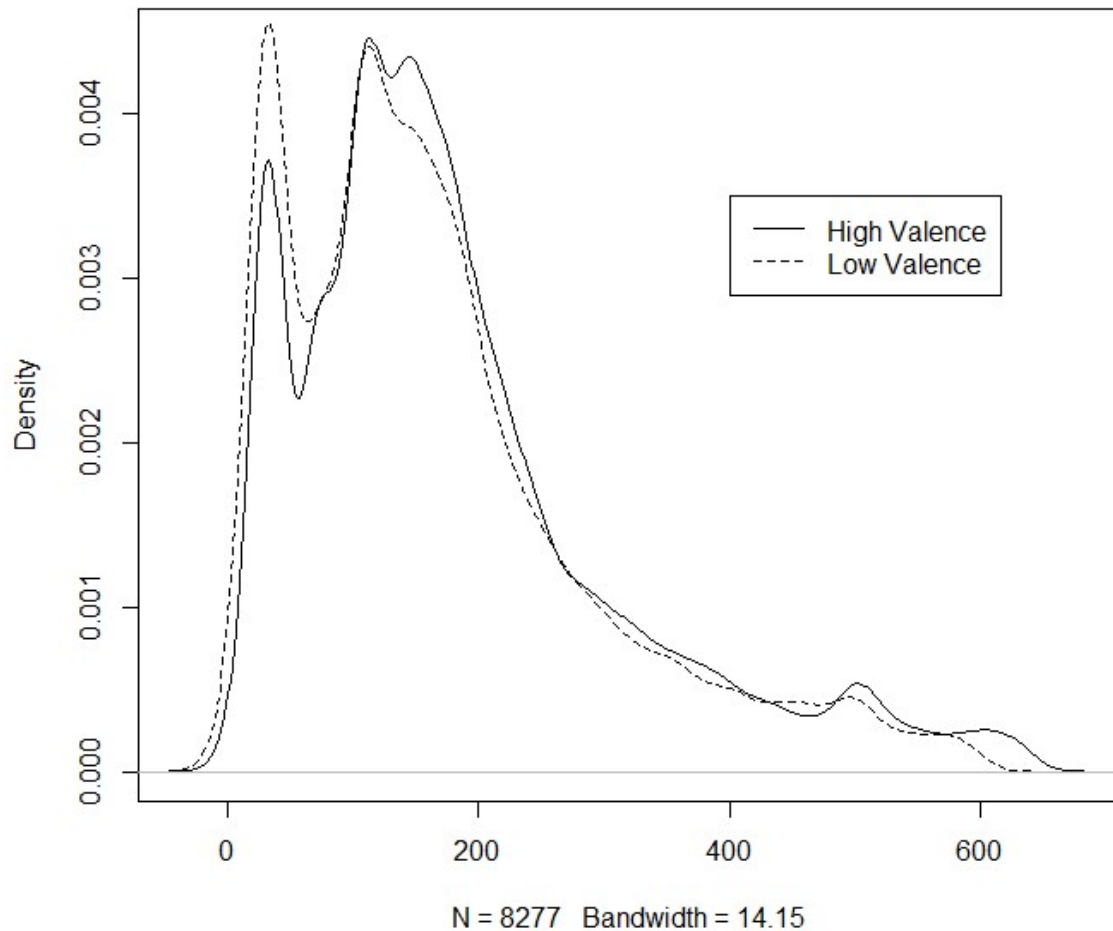

Figure 13: Time Between Presses Density Plot

8.3.4 *Character Number*

Now beginning with the **Character Number** data type, we can observe its Box Plot below. This indicator is a difficult one to rely on, since the amount of text a person writes depends on much more than his or hers emotional state. The level of culture one has, the amount of books one has read, the literary production practice one has cultivated, these are all

more relevant to the amount of text written. But we can still affirm that emotional state is relevant, as can be seen in the graph. Even with high amount of outliers we can observe a pattern that when the Valence is Low the amount of characters written rises in general. The distribution tends to go into higher values with the last three quartiles in the Low Valence collection, compared to the High Valence one.
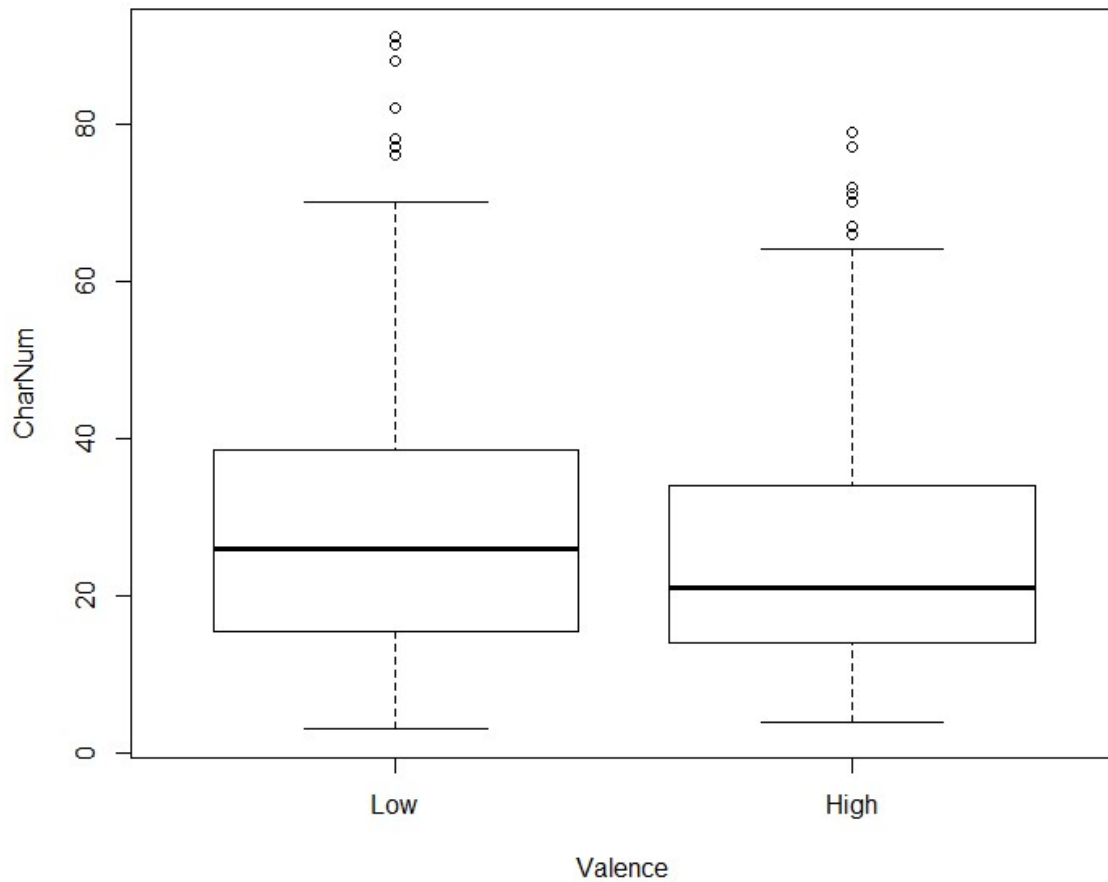

Figure 14: Character Number Box Plot

Following with the Density Plot graph, presented below, there are some more observations we can take, the first of which is that in the fact the amount of characters written seems higher in the Low Valence curve, even if most of the values in general are concentrated in the low spectrum of the graph. There are however three distinct areas where the Low Valence curve is higher than the High Valence one, which proves that this hypothesis is highly likely to be true. The difficulties of using this data type for emotional detection are however maintained, due to the amount of factors that can influence it without having nothing to do with the image displayed on screen. Even so, if we discard these factors and

only look at the presented data, this data type can be considered relevant in the detection of emotional Valence changes.
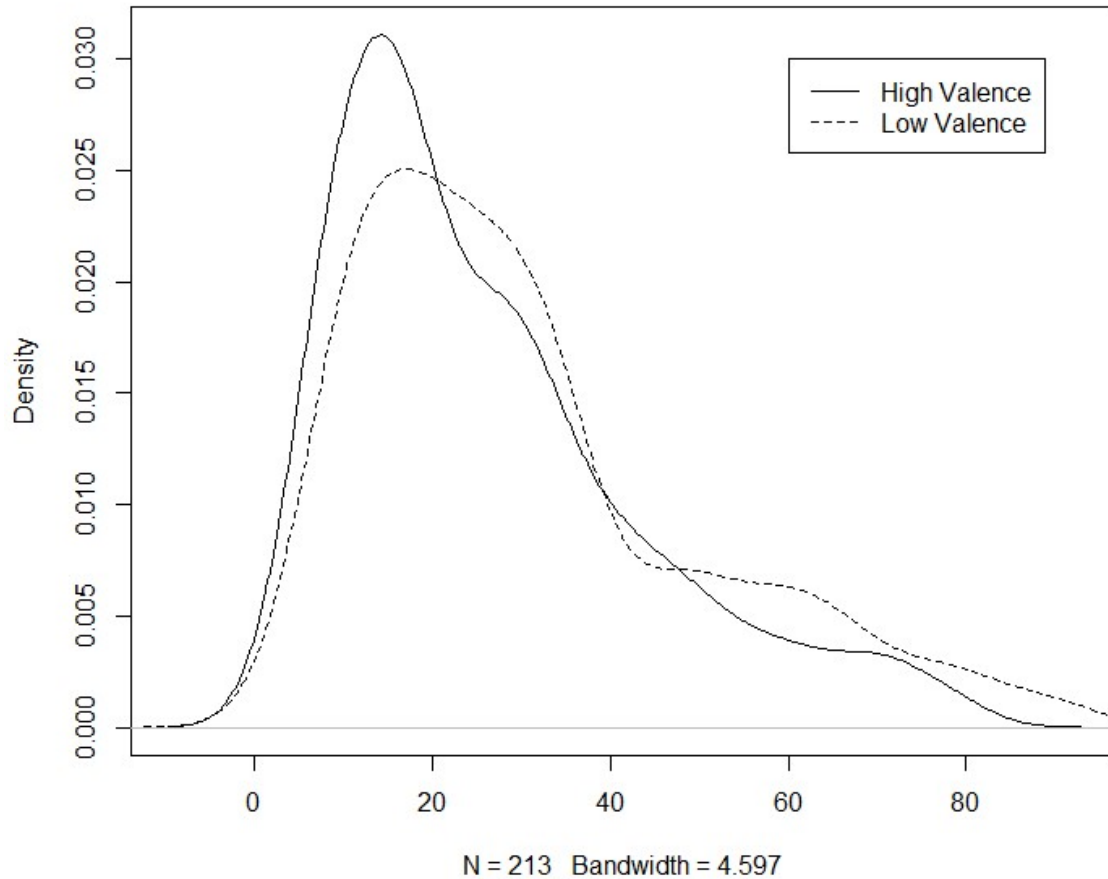

Figure 15: Character Number Density Plot

8.3.5 *Click Duration*

Now looking at the **Click Duration** data type, with the Box Plot presented below, e can finally observe obvious and drastic differences between collections. The Low Valence collection has a distribution focused only on small values, with the High Valence collection being more widely distributed but with most values far above the ones found in its counterpart. The time spent during each click, meaning between pressing and releasing the left mouse button, can be influenced by our amount of peace of mind, or stress, since a common hypothesis is that when under pressure a person is less likely to keep the sufficient pressure on the mouse necessary for it to be kept pressed for high amounts of time. The amount of values and differences between these collections seem to prove this hypothesis

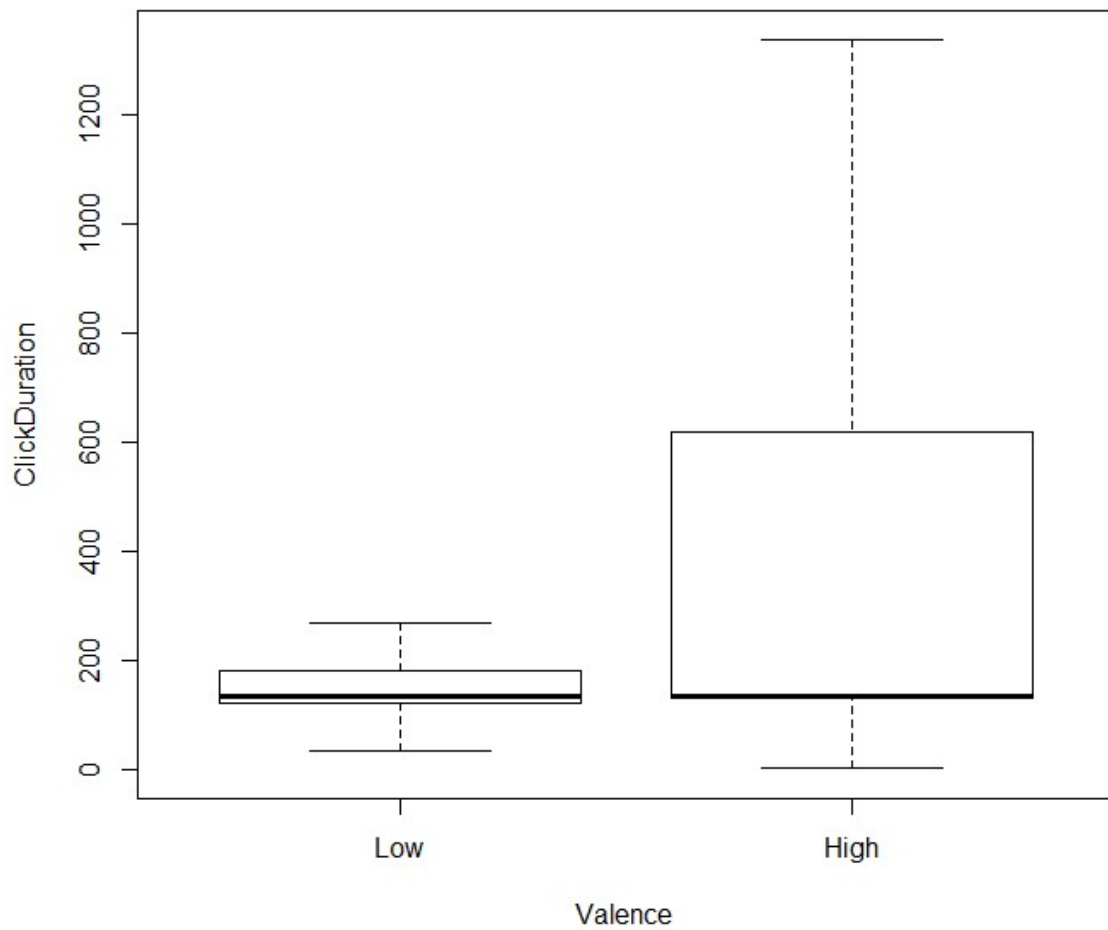or at least to highly tend towards a positive response to it.

Figure 16: Click Duration Box Plot

Looking a the Density Plot below we can confirm the early observations on the previous graph. There is a complete lack of values in the 1000 to 2000 millisecond range in the Low Valence collection, while a a stable amount of High Valence values can be observed in that range. The amount of Low Valence values in the lower range of the graph is so high that the top of the line is not covered in the window. Even though it is unusual this window size was chosen as such in order to permit the visualisation of the High Valence values in the higher values range, something that would not be noticeable if the graph was oriented toward the Low Valence ones or if the top values of the window were adapted to fit the top of the Low Valence line. There is an obvious difference difficult to ignore, that can safely be assumed as a tool to differentiate Valence. This difference was however in all likelihood exacerbated by the nature of the test itself. The existence of the three dimensional sliders ensures that many users will spend more time than usual with the left mouse button

pressed down, in order to move them. Even so the data has value due to two main factors. First, that many users did not in fact drag the sliders to arrange the values, but click repeatedly in the level they wished to take the slider to, something that should lower the values, not bring them higher. The second is that when exposed to disturbing images a user many times tends to end the task quickly, and this includes moving the sliders, meaning that even if the inclusion of the sliders can inflate the values, this inflation is not enough to devalue the importance of the data, making this an important data type.
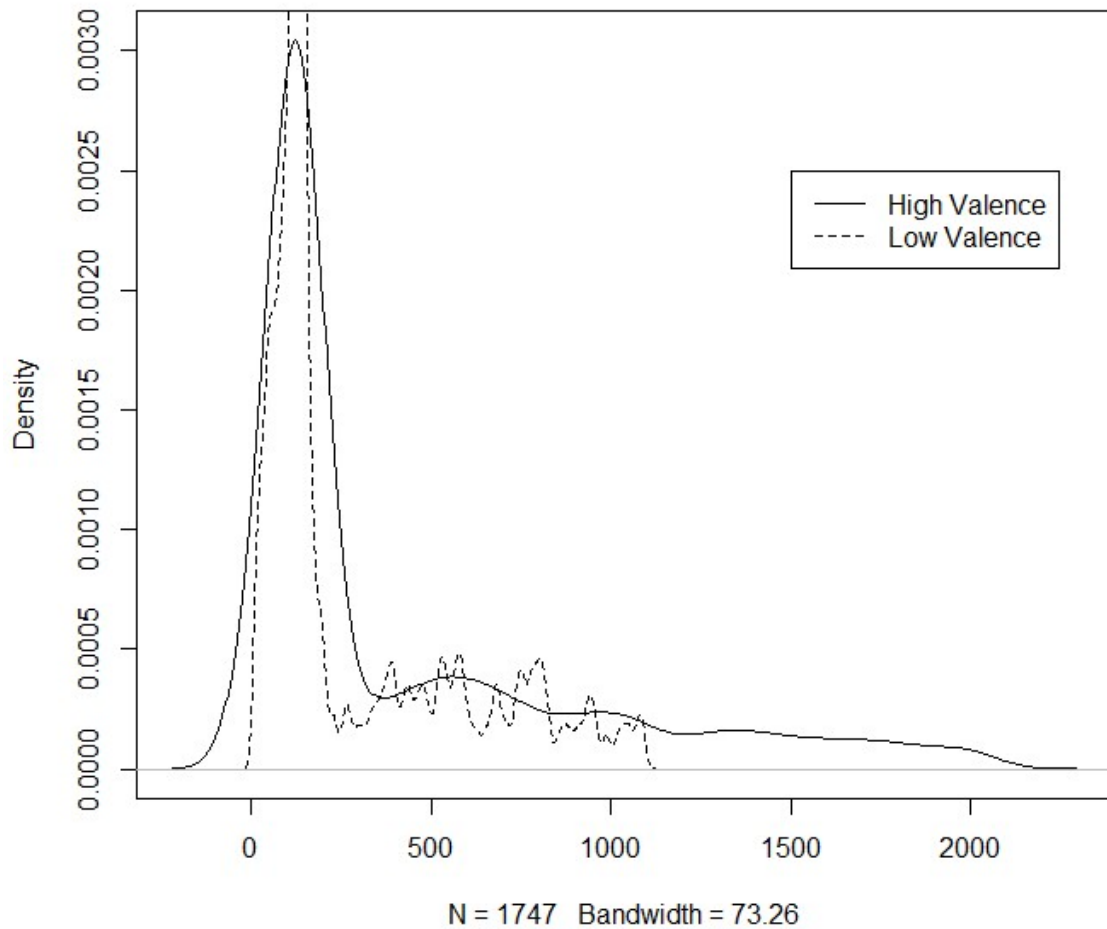
Figure 17: Click Duration Density Plot

8.3.6 *Mouse Speed*

Moving onto the **Mouse Speed** data type, we can observe its Box Plot on the image below. This indicator, representing how fast the mouse moves in a small millisecond interval, is simultaneously one of the most well grounded indicators due to the sheer amount of values

and the nature of mouse movement and what it implies. Immediately we can observe some significant differences between collections. The Low Valence collection's last two quartiles are noticeably higher than the High Valence ones. It is important to note that most of the values in this data type were zeros, since there is a significant amount of time when the user has stopped using the mouse to write or to decide what action to take, and generally stands still to decide when it comes to mouse movement. These values were however filtered in order no to disturb conclusions, since speed can not be negative or zero once movement is indeed verified. The presented graph has therefore no outliers. As such we can conclude that in general values of mouse movement speed are higher in the Low Valence collection. This can support the hypothesis that people move and act faster when they do not want to spend more time in the presence of images that they do not enjoy.
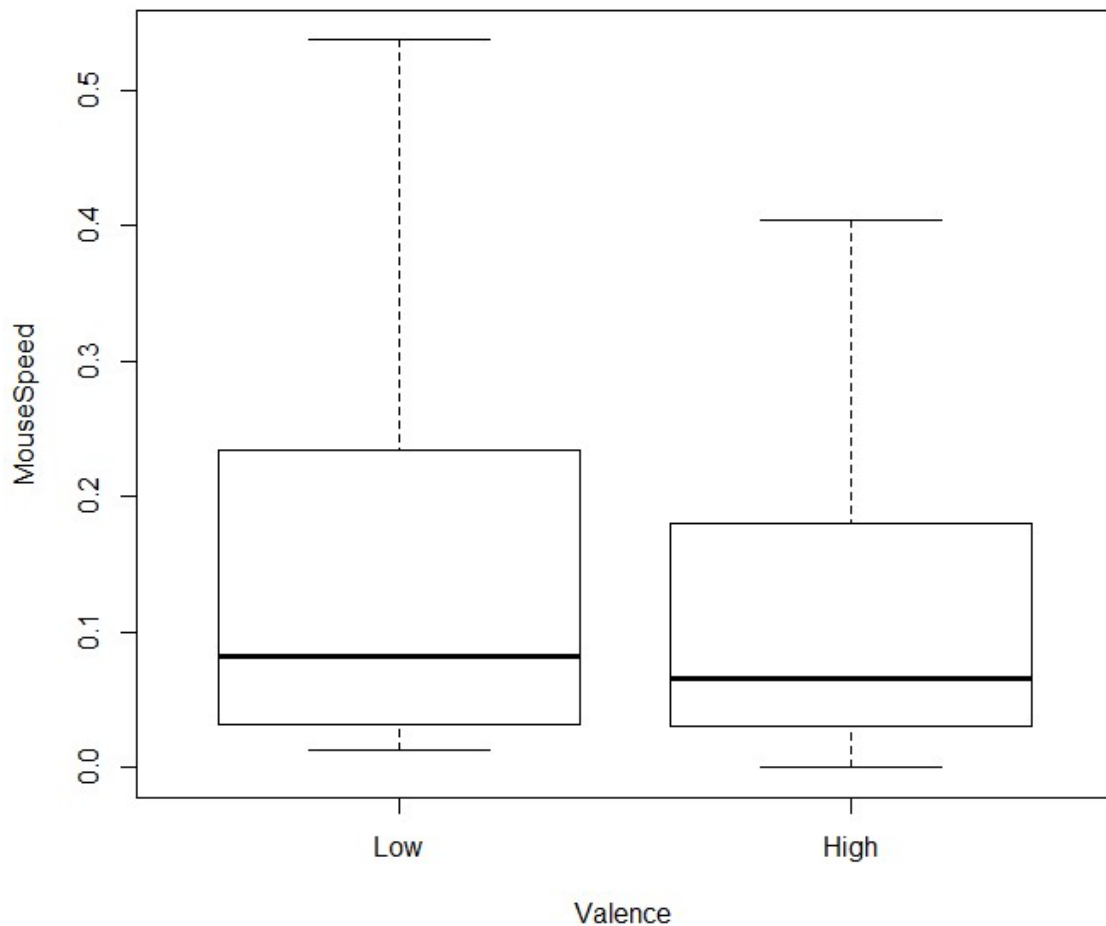

Figure 18: Mouse Speed Box Plot

Looking at the Density Plot graph presented below we can in fact confirm that there is a defined difference between density in the lower spectrum of the graph. The difference between the collections is compensated in the other areas of the graph, where there is a

slight but concrete and constant difference. It is also noticeable that the line of Low Valence continues beyond the limit of the High Valence line, increasing the amount of high values, With by far the largest data sample of any of the data types, this information assures us that Low Valence images are connected to higher mouse speed movement values, confirming the hypothesis created with other previous graphical analysis. It can therefore be stated that a high enough amount of users tend to perform mouse movements faster when presented with an emotional state of Low Valence, even if in this case it is probably caused by the desire to stop the emotional stress to which they were exposed.
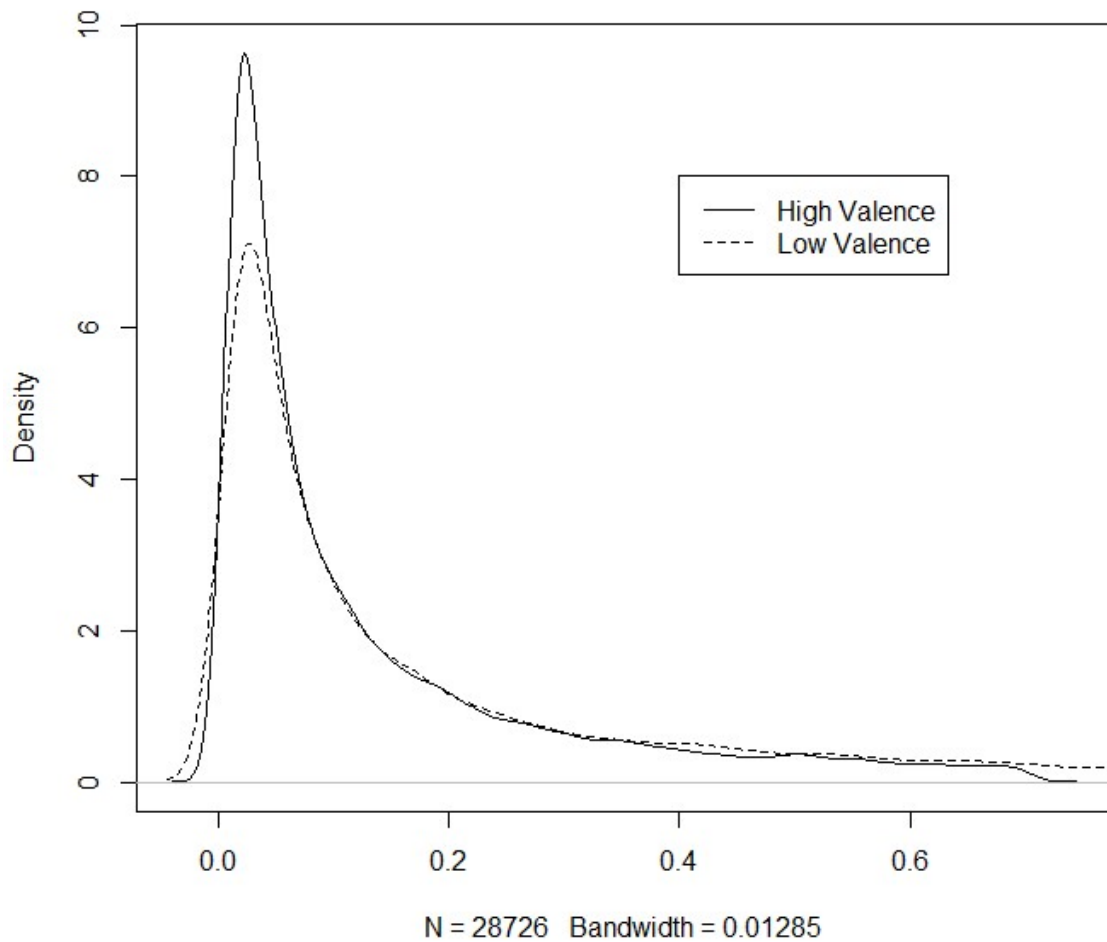

Figure 19: Mouse Speed Density Plot

8.3.7   *Press Time*

Continuing on, we arrive at the **Press Time** data type, with its Box Plot below. This indicator represents the time a keyboard key is maintained pressed down, namely the time

between the activation and release of said key. At a first glance we can notice that the two collections are almost exactly similar, with no discernible differences. A small difference can be noticed on the High Valence third quartile, but it is rather minuscule, and therefore not very important. From this graph we can deduce that this form of keyboard dynamic may very well have suffered no effect by the variation in Valence. Not much more relevant information can be retrieved from the observation of the graph, besides the conclusion that the amount of time a person spends pressing a keyboard key tends to be focused between very small values, with only ever the small variation, since the median lines are almost at the top of the boxes presented in the graph.
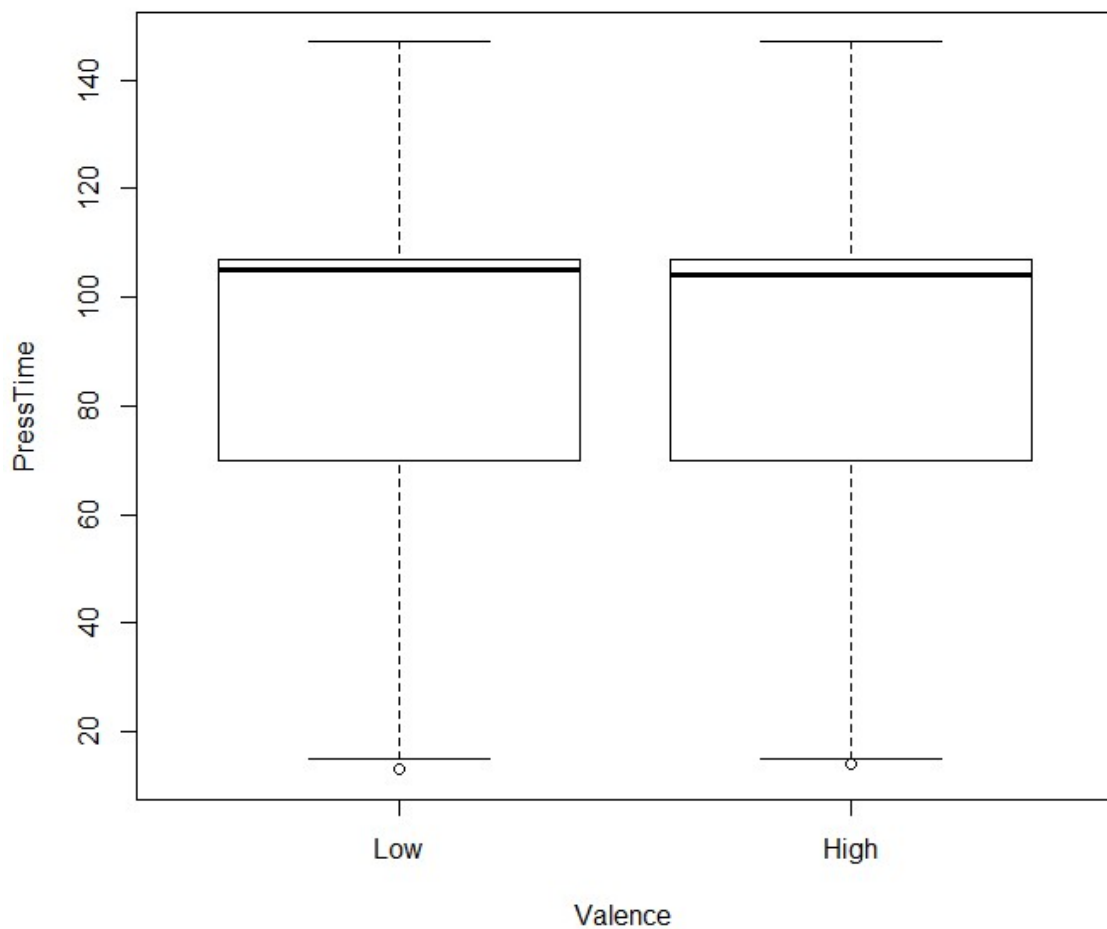

Figure 20: Press Time Box Plot

Looking at the Density Plot presented below we can come to some conclusions. There seems to be an unusual concentration of values in four different areas, with barely any values in between, if any. Secondly the two lines are almost coincidental, but with some small differences. In three of the main heights of the graph the Low Valence collection stands

above the High Valence one, with the exception of the second "mountain" of the graph, where the High Valence values are higher. This however does not seem to be enough to create a stable and grounded connection between the indicator and the Valence variations.
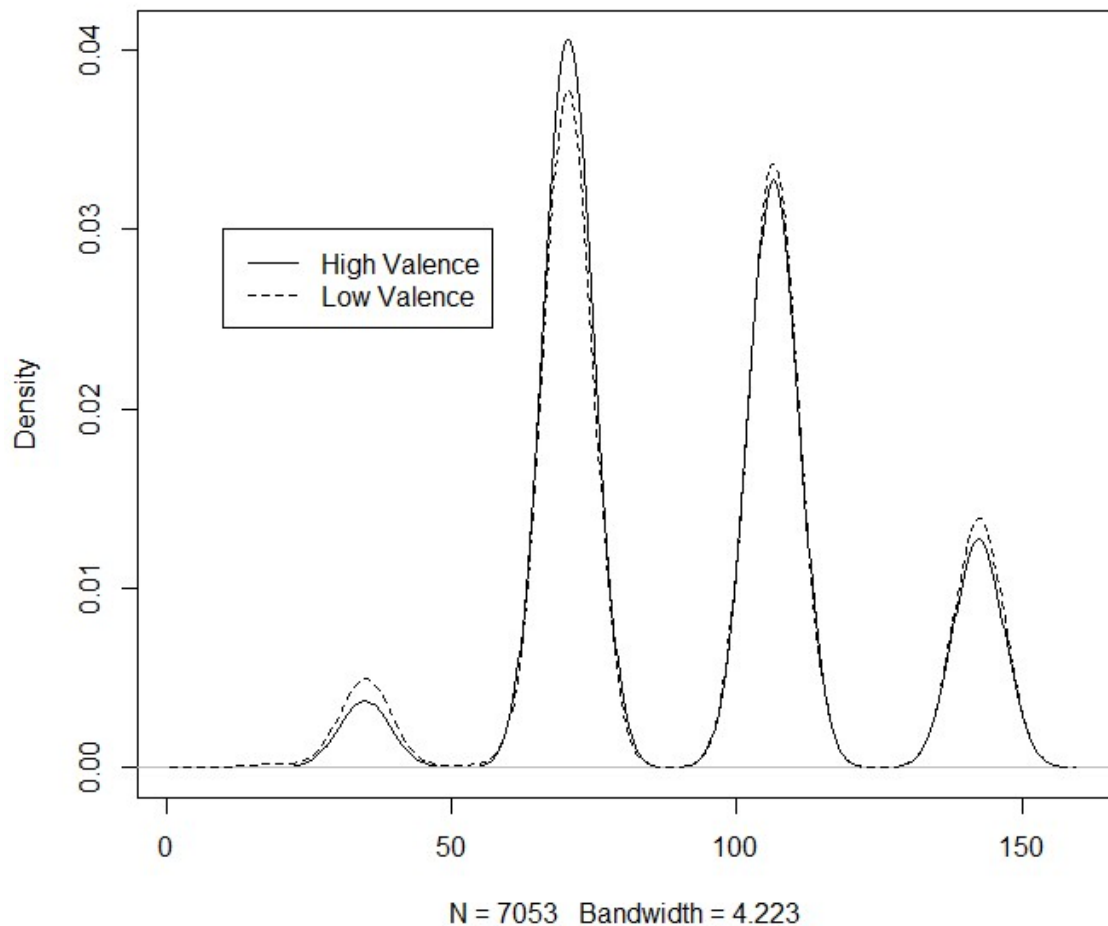

Figure 21: Press Time Density Plot

8.3.8   *Time Elapsed*

The next data type we analyse is the **Time Elapsed**, namely the time spent in a test window, corresponding to a single image and related tasks. This time period is measured here in seconds, opposed to the milliseconds used on most other time based indicators in this study. Immediately we can observe that even if the graph's collections are very similar, the Low Valence collection is slightly moved down comparing to the High Valence one. This suggests that in general the time spent on Low Valence images is shorter than on the High Valence ones, even if not by much. With the collections this similar it is however difficult to

base a system's evaluation of emotional Valence based on the time spent in an image. As such more information is necessary, information that may reside in depths of the following graph, in order to take decisive conclusion about this indicator.
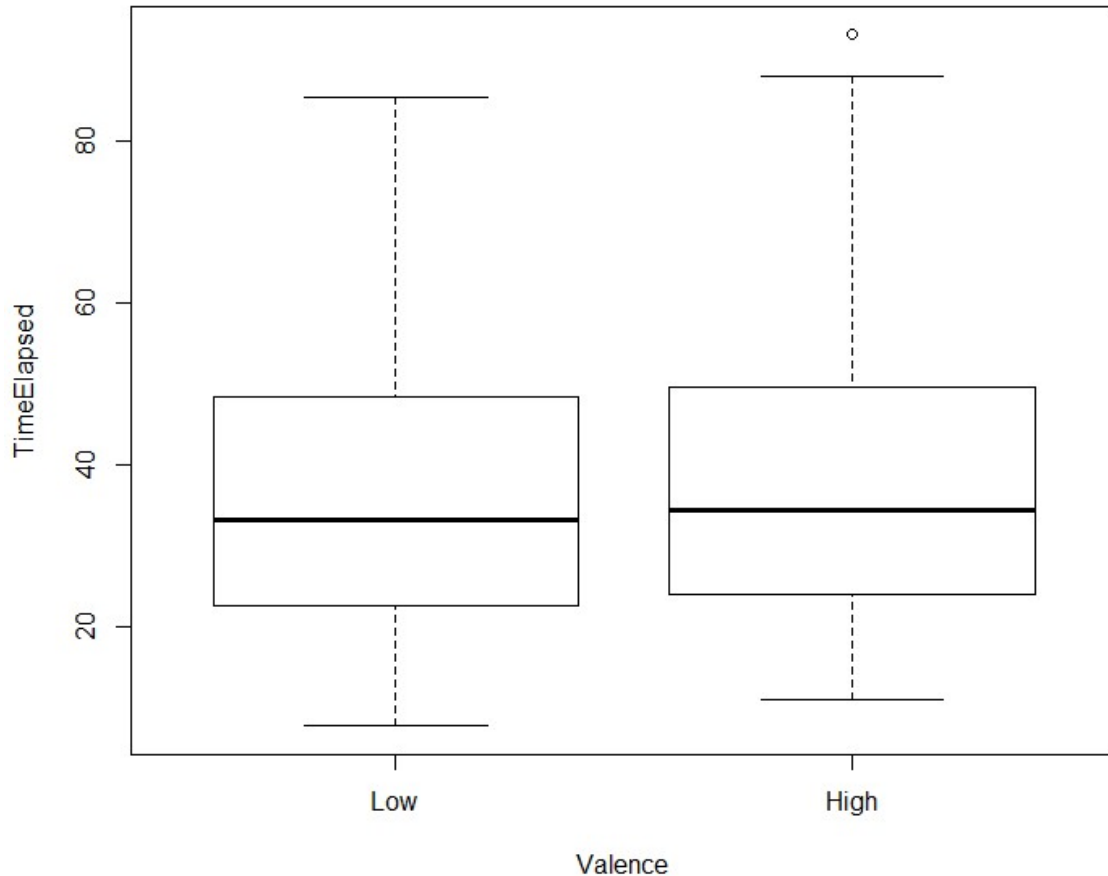

Figure 22: Time Elapsed Box Plot

Looking at the corresponding Density Plot graph, we can verify the previous conclusions, but with some new added information. The lines are effectively similar, with the Low Valence values slightly lower than the High Valence ones in general, as seen by the dislocation of the curve to the left. There is however a small high in high values in the Low Valence collection compared to the High in the high end of the graph, denoting a larger concentration than expected of 80 second time periods in Low Valence images. There is a possibility that this is an outlier, a high point created due to distractions, due to the obvious turn to left of the Low Valence line. However it is not possibly to assess if that is the case, and as such this occurrence throws doubt into the hypothesis that in general users spend less time in the Low Valence images. A higher amount of users and therefore test data cold be instrumental in confirming if this strange high point is an isolated incident or a verified trend. The value of this indicator as an evaluating tool is therefore unconfirmed, even if

it goes into the hypothesis that users tend to want to run quickly through the disturbing imagery that causes their Valence values to lower.
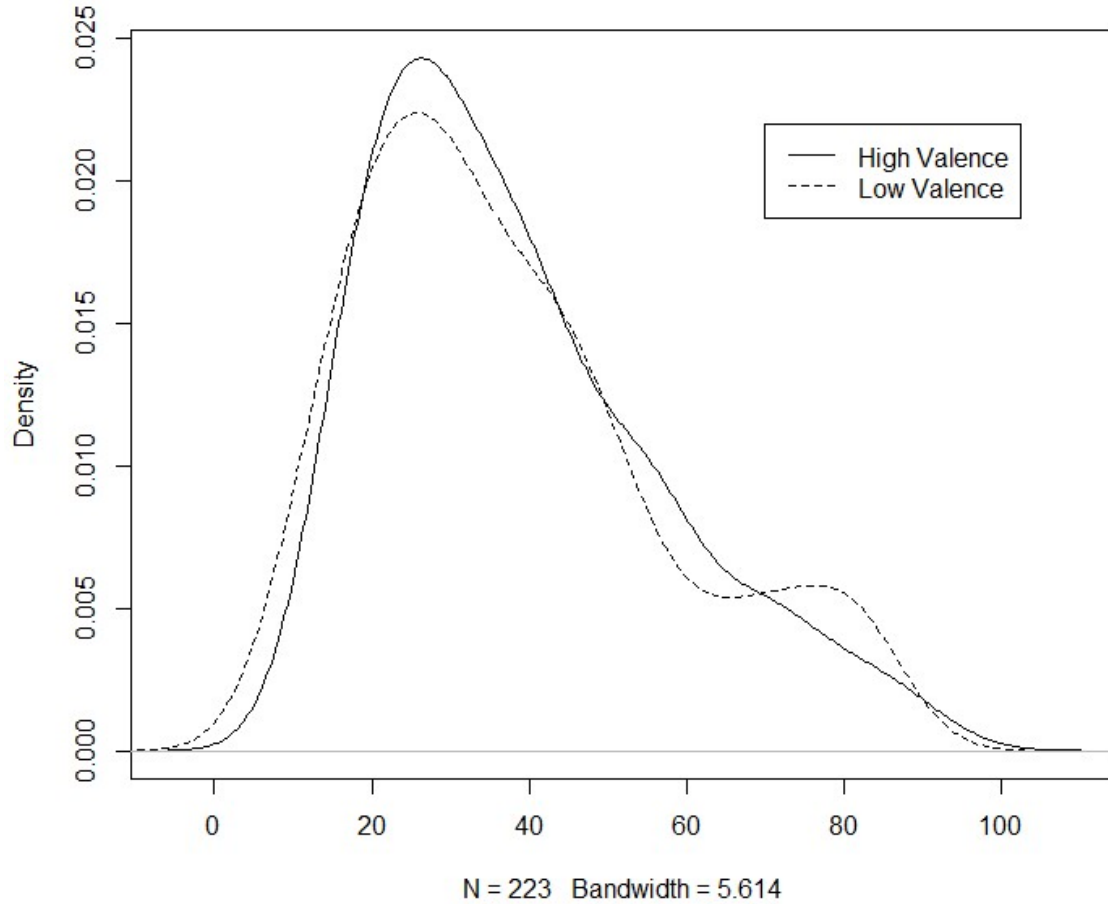

Figure 23: Time Elapsed Density Plot

8.3.9   *Type Duration*

Heading into our last data type we finally reach the **Type Duration** indicator. The Box Plot presented below has its first half almost identical, with both median standing comfortably slightly above 1000. The two latter quartiles of the collections, however, seem to exhibit significant differences. The first is that the Low Valence collection seems to show higher values in general, as well as a more spread out distribution of those values. Even the outliers, which have not been removed, respect this kind of observation. At first the conclusion that a lot of users spend more time writing when exposed to Low Valence images may seem contrary to the hypothesis that they want to get the test over with faster when faced with them, but it may not be so. The fact that they are mentally and emotionally affected causes

people to have more difficult to decide what to write and how to describe what they are seeing. This also goes hand in hand with the previous character number high observed in the Low Valence collection.
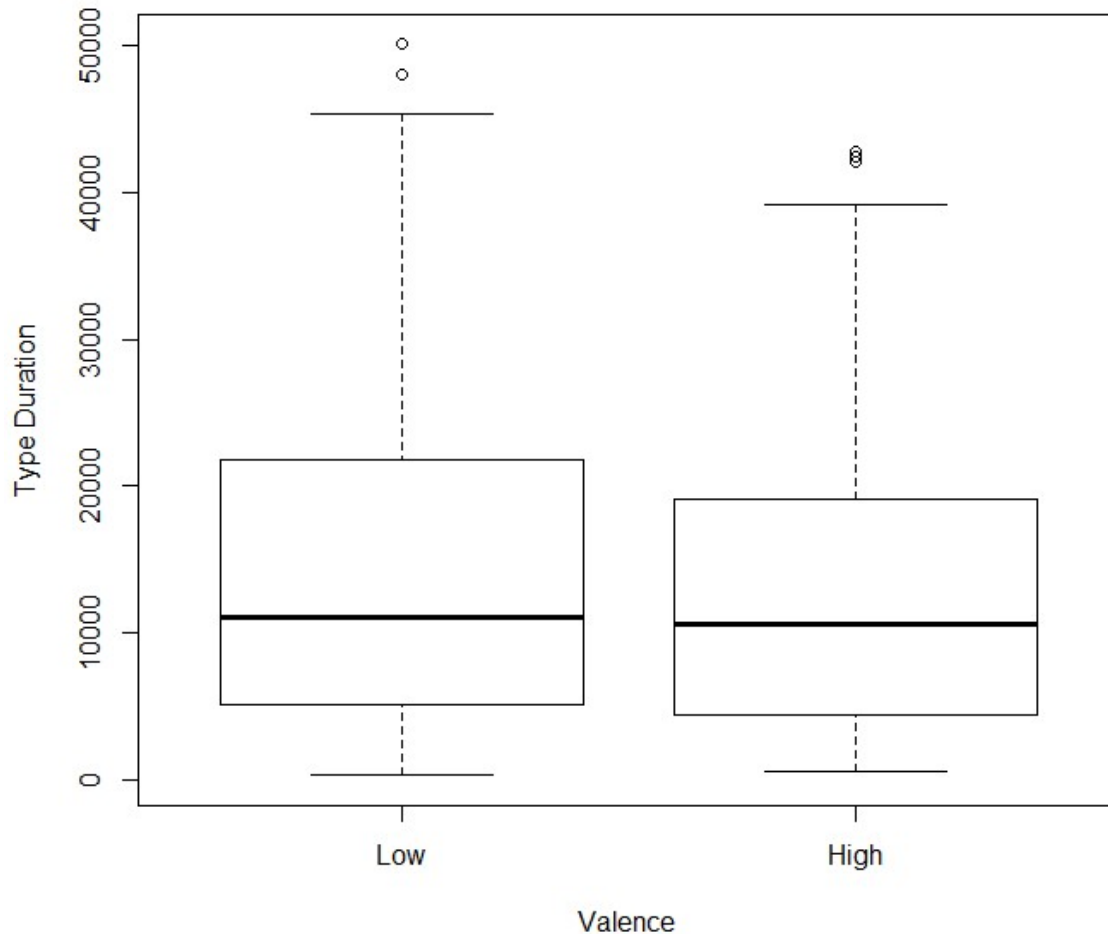

Figure 24: Type Duration Box Plot

Looking at the final graph displayed below, the Density Plot graph for this indicator, we can obtain new information. The values are in fact higher in the Low Valence collection, and their distribution is also higher, but the amount of high values is not very high. The low values are also mostly the same, exhibiting no significant differences. The values suggest that users react very differently to these variations, since the values are very spread out, but the graphs point out that the concentration of high values is distinctly located on the Low Valence collection. Once again we come to the conclusion that even if the data points toward a high concentration of values in the low spectrum of the graph it is on the high values that we can observe a difference between collections. As such, and with this, we finish this individual graphical analysis, with the last Density Plot graph presented below.
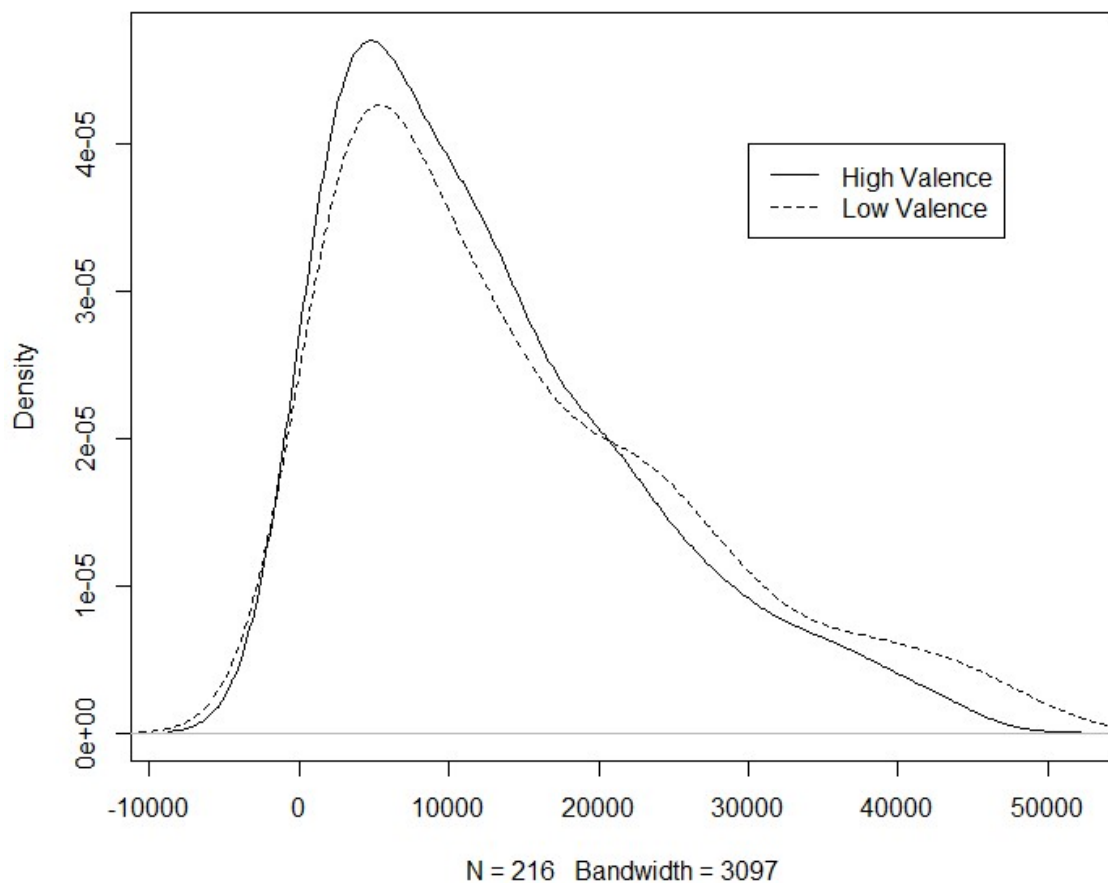
Figure 25: Type Duration Density Plot

8.4 DISCUSSION

After analysing the data in the general and individual senses it is time to focus on the questions raised in the beginning of this dissertation, and on the discussion about the efficacy of the solutions presented. This discussion will be focused firstly on the performance and helpfulness of the developed tool, and secondly on the results on the attempt to obtain relevant data capable of creating a conclusion about the effects of Valence changes on test subjects and general computer users.

8.4.1 *Tool Effectiveness*

Starting with the discussion on the developed tool, the Horus software, the question being posed is if the application does what it is intended to do, and if it does it well. The effectiveness of the tool is important since one of the goals of the whole dissertation and associated

project was its development, in order to help in the collection of data for researchers of emotions, be it psychologists or engineers. As a tool it is fully functional and capable of obtaining the data necessary for a study, but is still highly limited. The local and offline nature of the tool make it difficult do distribute and to promote use out of this project, and its dependency of the International Affective Picture System (IAPS) image database means that the local databases used in the project would need to be transformed into online remote ones in order to allow for use outside of the installed hardware. This transformation would be simple to implement due to the structure of both the code and the tool, only limited by available hardware, not being a hurdle particularly hard to surpass. Besides the localisation issue the tool has some small adjustments that would most definitely need to be made in order to allow for usage outside of extremely controlled environments, be it in terms of hardware, associate software, or personnel.

After dealing with these issues however, and besides its functionality as a piece of software, we have to discuss its abilities to collect relevant data and to be versatile enough to serve many studies and purposes. In that aspect the tool is considered a success, since the focus on the test of Valence done in this dissertation proves that the tool can be adapted to study different aspects of emotion through the choice of images to be included in the test. In this way, the full functionality of the test creation and test taking parts of the tool are a testament to its possible use in the future. The data gathered through the tests is varied and independent of much of the structure, meaning that if a developer wished to introduce some other timer or data gathering structure inside the tool's code it would be possible and rather simple, after reading the development report part of this dissertation. The relevance of the indicators being calculated by the tool is still in discussion, and its variety of use makes it so that to take definite conclusions about its importance in the field of emotion detection is difficult. However, their presence and the possibility to use them is confirmed, and as that was what the tool was projected to achieve, this in itself is a success.

It is estimated though that due to the nature of the tool, being non-intrusive in its collection of information on the user, a higher amount of test subjects would be necessary to definitely prove the influence or the lack of thereof of the user's emotional state in their routinely use and interaction with their personal devices. The tool's effectiveness in collecting the data necessary to take these conclusions is however proven, as it allows for the creation and utilisation of tests made with IAPS and subsequent collection of the data in those tests. As such, this objective has been reached in a successful manner, even if the tool must be adjusted for further use by different investigators ans teams in future studies.

8.4.2 *Valence Detection Effectiveness*

The experiment created with the developed tool had as its main goal to understand if any of the collected data types or variables could be used to detect emotional Valence changes in users without knowing what they are doing of being exposed to. The non intrusive angle of the tool and of the project is therefore both its greatest asset and its biggest hurdle. The test was projected to be made with limited subjects and with limited time, in an environment where most users present similar characteristics, ans these must be taken in consideration. All users were Portuguese citizens, with ages ranging from eighteen to fifty one years, and with considerable amounts of education. Their medium income privileged status as citizens of a democratic nation currently at peace means they have less resistance to disturbing imagery than someone used to it due to circumstance, necessity or other. These results are therefore fruit of the social and economic environment where the subjects reside, and would be subject to significant change if the environment was a drastically different one. Even so, the data collected from the subjects interactions was unadulterated and representative of their group, due to their willingness to comply with the test's necessities and to their status. No data treatment or examination was done to individual subjects, maintaining their rights and privacy. As such, and after a general and individual study of the variables in the collected data in regards to its difference in Valence, there are some conclusions that can be made.

The evaluation of the variables has been done in three levels, namely Low, Medium and High relevance. From the nine variables considered relevant, three were proven mostly unchanged in the face of Valence changes, namely **Time Between Presses**, **Time Between Clicks** and **Press Time**. The analysis of these three indicators observed very little change, especially in the case of the third one, and due to their nature the creation of a model using these values would be difficult. The third one of these also possessed a p-value higher than the expected level, which confirms its difficulty in establishing useful patterns. However, the small differences between Valence collections in these indicators might be enhanced if more test subjects were to be used in the test collection. Even so, these three are categorised as having Low relevance to the modelling of Valence emotional models. From the remaining six, four of the indicators were considered to have Medium relevance, namely **Backspace Presses**, **Character Number**, **Time Elapsed**, and **Type Duration**. Two of these, the first and the third, also had high p-values, and it is true that there are not many differences in an initial analysis, but both the amount of mistakes during writing and the amount of time spent in an image tend to vary sufficiently, after the graphical analysis. Even so, they might not be enough to confirm their usefulness, even if there is data to suggest it. A higher amount of tests is mandatory in order to confirm their significance. As for the

last two, namely **Mouse Speed** and **Click Duration**, they were evaluated as having High relevance, since there are very significant changes in the two collections, be it after the statistical analysis be it in the graphical individual one. As such, we finish the analysis having confidence that two of the variables are highly likely to help detect changes in Valence, that four have a distinct possibility of being helpful in this endeavour and that three are most likely not as useful as the others. All these evaluations would benefit immensely from a second round of testing with a larger amount of data available, as stated multiple times in the dissertation, but as it stands these are the results presented here through the analysis of the available data.

As a formulated hypothesis i earlier suggested that users tend to want to go through the disturbing images at a faster rate but have difficulty in writing down their thoughts because of the nature of the images. Even though there are significant outliers there is enough data to support said hypothesis, even if this may only apply to situations where visual stimuli is presented. keeping in mind that all these results come from exposing the users to visual stimuli, that the users interactions were highly limited and monitored and that the amount of users themselves were rather small, the results are satisfactory when it comes in discovering patterns in interactions. The users emotional change, caused by the variation in Valence of the images, has indeed affected the way they interacted with the tool and with the computer, and has allowed for the identification of a group of variables that might help in the adaptation of systems to the emotional state of their users. As such, and even if the experiment was not a resounding success, it was successful in creating both a proof of concept and a starting point that can be used to continue studying the effects of emotions on interactions. The particulars of the Valence changes have also been sufficiently analysed to affirm that the method is sufficiently effective, even if not as much as one would want.

9

CONCLUSIONS AND FUTURE WORK

In this last chapter there are three main things to do, namely state a summary of the dissertation and the project that allows for the simple comprehension of the work done, an analysis of the completion and final state of the goals and objectives of the dissertation, step by step, and a projection of future work to be done in the project and in the field of research itself in order to further develop the ideas explored in the document as a whole.

9.0.1 *Overview*

This project and dissertation had a series of objectives centered around the idea that emotion and emotional states are something so relevant in the way people interact with each other and the world that a machine capable of understanding these emotional states and reacting to them could be highly useful to its user. The way people interact with computers and other information systems change with a great many factors, and the hypothesis that emotion was one of these factors was obvious and yet unproven. However, the nature of emotion means that it is in itself very difficult to measure and understand. A concept that is extremely abstract and yet has such effect on human existence is difficult to detect and quantify due to the obvious amount of subjectivity involved. As such it was necessary to obtain a way to better understand this phenomenon in order to take a more scientific approach to it. Psychology was an obvious choice, and through its study a possibility of using an already developed system of defining emotional states arose.

This system needed a platform that could adopt its theoretic principles and implement them in computational language. As such a tool was created and a way of testing if the system worked was devised. The main focus of such tool was to detect the way the emotional state of the user influenced their interaction by causing intentional emotional fluctuation. By comparing the data with the expected emotion, the hypothesis was finally proven and the beginning of a evaluative model was discovered. An essential part of this whole endeavour was the way the data was collected. Society has unconsciously given up privacy for ease of use, and in order to restore a sense of self preserving identity and safe space in

the digital world any study into a user done by a machine would have to be non-intrusive, especially so since emotions are something especially intimate and therefore any intrusion in their emotional space would be especially grave. Such were the means used in the developed tool, and through these the data collected was clean of any identification of the user. The identification of patterns in the data help prove the hypothesis of emotional influence in human computer interaction, as well as formulate a possible interpretation of the way these interactions work and could be used for future study. As such, and even though the data was not as useful as desired, the project was all in all a success.

9.0.2 *Completion of Objectives*

This project and dissertation had as its main objectives the following: The better understanding of emotion and of its importance for intelligent systems and specifically in human-computer interaction, the creation of a different method of measuring emotion through non-intrusive means, the development of a tool that has the ability to study emotional changes in an individual using it, and the exploration and treatment of data obtained through such tool, in order to judge the effectiveness of the method. After tackling these objectives I can safely say that they were all successfully fulfilled, all in their own way and while dealing with predicted and unforeseen circumstances and obstacles. The different nature of each of these objectives, from more theoretical to more practical, creates the necessity to review them one at a time, in order to obtain a full understanding of the degree of completion of each one.

Starting with **"The better understanding of emotion and of its importance for intelligent systems and specifically in human-computer interaction"**, it was instantly clear that an in depth research into the various concepts and areas related to this understanding was necessary. As such a comprehensive study of emotion and theories relating to it belonging to the field of psychology was made, as well as a study in the history and basis of all work done around interactions between human and machines. This research allowed for the projection of a way to intertwine the field of computation and psychology in order to better understand the phenomenon on a practical level. The study of human computer interaction and the role that emotion has and could have in the way people use and interact with technology in a daily basis has therefore been very productive in both validating the use of emotion in intelligent systems and demonstrating the value of studying ways to implement that use. As such, this step was complete in full, since the knowledge gained was both compiled in this document and allowed for the construction of the structure over which the next objective was founded.

This opened the door to the completion of the next objective, namely **"The creation of a different method of measuring emotion through non-intrusive means"**. Finding the International Affective Picture System (IAPS) and understanding its uses was key in developing a method to both measure emotion and do it in a way that respected user privacy to a maximum. The method was created through the idealisation of a tool that could read some essential inputs done by the user, and was based in simple notions of human computer interaction that have been explored ever since the filed came to be. However, the innovation in the method is found in the joining of both visual stimulation and the acute observation of the use of peripherals during the changes incurred by the user caused by the exposition to said stimuli. As such, and with the help of IAPS as a theoretical basis, the idea to use non-intrusive methods to identify emotion, namely keyboard and mouse dynamics, was validated as a valuable and possible answer to the dilemma, be it in its moral and intellectual facets be it its practical application. Respecting privacy is essential if we are to better integrate technology into our societies, and doing so while adapting systems to their users' necessities is key in the vision of a better connected and more civil future. The method was therefore created, and all that was necessary was to both implement and evaluate it.

As such it was necessary to develop a tool as idealised in order to implement the method devised before, consisting of the objective worded as **"The development of a tool that has the ability to study emotional changes in an individual using it"**. The tool was firstly designed having in account both the specificity of what was being tested and the usefulness that it could provide to any researcher wishing to test similar concepts. With a high amount of reliance on the main tenets of the project and on the specifications obtained through the use of requirements engineering, the tool was therefore developed using available technology that could be easily included in any system and utilised by a variety of people. The designed tool was implemented successfully, according to the requisites and necessities recognised in early planning. It is completely functional and it was capable of collecting data in the test phase, being utilised by a variety of users in a great variety of environments. As such an objective was once again completed as well as possible, considering the time and restraints put on the project and the great obstacle that was not knowing exactly what was going to be found by the usage of the tool, if anything at all.

The final objective, and the one that is able to connect the whole project together and validate the previous objectives, was **"The exploration and treatment of data obtained through such tool, in order to judge the effectiveness of the method"**. The only way to collect data was to attempt a thorough testing phase, even with limited hardware, personnel and voluntary test takers. Even so the tool was able to be used as a way to collect enough data to explore and analyse in search of patterns and ways to understand emo-

tional responses. The data collected was used to identify patterns and some indicators that could be used in the creation of models that are able to identify emotional Valence. This data was also used to validate an hypothesis on the Valence variations using non-intrusive methods, something that was essential in the context of what were the goals of the project from the beginning. The method was validated, even if with some questions still standing due to the lack of sufficient data. The focus on a mixture of psychology and engineering as a way to identify and evaluate the methods and the data was essential in its success, and this association of different areas of study is something that must be more explored in the future, especially considering the way informatics engineering is sometimes not very available to do so. As such this final objective was successfully attained, completing the list of goals proposed in the beginning of this dissertation.

9.0.3 *Future Work*

A project as this is always limited by its scope and experimental nature. As such it is very important to take that what was learned and obtained through the project and extend it in order to reach the ever distant goal of improving our computer systems. It is important to note that even if the data is enough to consider these hypothesis and methods as valid, and to take conclusion on the subject, the amount of test data is not enough to safely implement a system that depends on it, and therefore more in-depth testing with a larger pool of test subjects is required. With this second round of tests the new hypothesis could be proven either right or wrong, and new aspects and patterns could be found. New types of test subjects, should also be found, in order to not overspecialise the system by only exposing it to users in a certain social and economic range. The data types and indicators that were identified as having high relevance to the detection of emotion might also not be as useful in other dimensions of emotion that not Valence, and therefore testing dedicated to its other two dimensions is mandatory in order to have a decisive conclusion on the validity of the method as a whole. The creation of models of emotion is only possible if all dimensions are considered, and reactive systems based on such models should be able to understand a persons emotion better than the uni-dimensional perspective presented here. The developed tool would also benefit of more dedicated development and of an increase in functionality, in order to allow for more variety of users, tests and research projects to make use of it. Its structure is solid but it is not a valid product, and is therefore only useful in a very restrictive and specific academic area, something that could be easily changed for the better with enough development time. Even so, this dissertation is concluded with the certainty that a small but relevant step was given in the direction of better human computer interaction based on emotions, even with the amount of future work required to finally reach this so desired goal.

BIBLIOGRAPHY

J. Beck. Hard feelings: Science's struggle to define emotions. https://www.theatlantic.com/health/archive/2015/02/hard-feelings-sciences-struggle-to-define-emotions/385711/. Online, 22/12/2017.

M. M. Bradley and P. J. Lang. The international affective picture system (iaps) in the study of emotion and attention. *Handbook of Emotion Elicitation and Assessment*, pages 29–46, 1954.

M. M. Bradley and P. J. Lang. Motivation and emotion. *Handbook of Psychophysiology (2rd Edition)*, pages 581–607, 2006.

Lopes R. H. C. *International Encyclopedia of Statistical Science*. Springer, Berlin, Heidelberg, 2011.

D. Carneiro, H. Rocha, and P. Novais. An environment for studying visual emotion perception.

D. Carneiro, J.C. Castillo, P. Novais, A. Fernandez-Caballero, and J Neves. Multi-modal behavioural analysis for non-invasive stress detection. *Expert Systems with Applications*, 2012.

J. M. Carroll. The encyclopedia of human-computer interaction, 2nd ed. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/human-computer-interaction-brief-intro. Online, 8/12/2017.

J.M Carroll. Human-computer interaction: Psychology as a science of design. *Annual Review of Psychology, Co-published (slightly revised) in International Journal of Human-Computer Studies*, (46):501–522, 1997.

K. Cherry. Overview of the 6 major theories of emotion. https://www.verywell.com/theories-of-emotion-2795717. Online, 19/12/2017.

J. Collis and R Hussey. *Business Research. A Practical Guide for Undergraduate and Graduate Students*. Palgrave Macmillan, 2 edition, 2003.

L. A. Goodman. Kolmogorov-smirnov tests for psychological research. *Psychological Bulletin*, (51(2)):160–168, 1954.

M. Goossens, S. Rahtz, and F. Mittelbach. *The LaTeX Graphics Companion*. Addison-Wesley, 1997. ISBN 0-201-85469-4.

P. J. Lang, M. M. Bradley, and B. N. Cuthbert. Emotion, attention, and the startle reflex. *Psychological Review*, 97:377–395, 1990.

D. G. Myers. *Theories of Emotion*. Worth Publishers, New York, NY, USA, 7 edition, 2004.

A. Nathan. *WPF 4.5 Unleashed*. Pearson Education, 1 edition, 2014.

P. Novais, D. Carneiro, M. Gomes, and J. Neves. Non-invasive estimation of stress in conflict resolution environments, in advances on practical applications of agents and multi-agent systems. *10th International Conference on Practical Applications of Agents and Multi-Agent Systems*, 2012.

P. Novais, D. Carneiro, and H Rocha. An environment for studying visual emotion perception. 2013.

K. Pohl. *Requirements Engineering: Fundamentals, Principles, and Techniques*. Springer, 2010.

R. Rodrigues, A. Mendes, D. Carneiro, M. Amorim, A. Pinheiro, and P. Novais. An application to enrich the study of auditory emotion recognition.