The Effect of Neuroticism on Trust Perception of Time Series Visualizations

Carlota de Oliveira Lopes Dias

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

Supervisors: Prof. Sandra Pereira Gama
Prof. Daniel Jorge Viegas Gonçalves

Examination Committee

Chairperson: Prof. António Manuel Ferreira Rito da Silva
Supervisor: Prof. Sandra Pereira Gama
Member of the Committee: Prof. Carlos António Roque Martinho

November 2021

Acknowledgments

First of all, I would like to thank my dissertation supervisors Prof. Sandra Gama and Prof. Daniel Gonçalves, for their support and valuable contributions throughout this project. In particular, I would like to thank Prof. Sandra Gama for allowing me to explore this fascinating multidisciplinary theme. I would also like to thank Tomás Alves that always guided me during this project, and was always available to answer all of my questions, even the silly ones.

I would also like to acknowledge my parents and my sister for their patience and for always encouraging me to do the best work possible. Moreover, getting through my dissertation required more than working, therefore, I would like to show my gratitude to my colleagues and friends that were always there listening and making work more fun. After all, we were all going through the same experience and I am extremely glad we were.

Last but not least, I would like to thank all of the participants in my study, that voluntarily decided to help and were genuinely interested in our investigation.

To each and every one of you - Thank you.

Abstract

Trust is one of the most important factors when it comes to deciding whether the information displayed in a visualization is going to be acquired as knowledge and used by the users. There are many factors in a visualization that can affect perceived trust, such as design factors, the display of uncertainties, or even individual characteristics such as personality. However, nowadays most visualizations are still designed using a one-size-fits-all approach. Our work proposes a study to evaluate whether the neuroticism trait has an effect on data trust perception in the context of information visualization of time-oriented data. In particular, we want to evaluate whether time granularity has an impact on perceived trust. This work acknowledges that granularity may be a relevant feature in the design of time-based visualizations, since people trust more in a line chart with a larger number of data points and interact more with a line chart that they trust less. Contrarily, we did not find any relation between neuroticism and perceived trust. However, a deeper analysis concerning the facets of neuroticism and other traits from the Five Factor Model (FFM) suggests that personality characteristics might influence perceived trust and user interaction with a visualization.

Keywords

Trust perception; time granularity; time series; health emergencies; neuroticism; Five Factor Model (FFM).

Resumo

A confiança é um dos fatores mais importantes para promover o conhecimento gerado quando os participantes analisam uma visualização. Existem muitos fatores que impulsionam a confiança como: elementos de design, incertezas exibidas na visualização, ou mesmo características pessoais como a personalidade. Contudo, hoje em dia a maior parte das visualizações são desenhadas sem ter em conta as características individuais dos utilizadores. Este trabalho propõe um estudo para avaliar se a granularidade é um elemento relevante no design de visualizações sobre séries temporais. Em particular, o objetivo deste trabalho é avaliar se a granularidade das visualizações tem um impacto na confiança percecionada pelos utilizadores. Assim, conseguimos identificar que os participantes confiam mais em gráficos de linhas que contém um maior número de pontos de informação, e interagem mais com os gráficos em que confiam menos. Contrariamente, não conseguimos encontrar nenhuma relação significativa entre o neuroticismo e a confiança percecionada. Contudo, após uma análise mais profunda das facetas do neuroticismo e de outros traços do Five Factor Model (FFM) foi possível concluir que algumas características de personalidade poderão influenciar a confiança percecionada, bem como as interações dos utilizadores.

Palavras Chave

Confiança; granularidade temporal; séries temporais; emergências médicas; neuroticismo; Five Factor Model (FFM).

Contents

List of Figures

List of Tables

Acronyms

FFM	Five Factor Model
InfoVis	Information Visualization
NEO-FFI	NEO Five-Factor Inventory
IPIP	International Personality Item Pool
TVCG	IEEE Transactions on Visualization and Computer Graphics
BFI	Big Five Inventory

1

Introduction

Contents

The main purpose of a visualization is to help users discover, explain, and form decisions based on the information that is conveyed [8]. Visualizations help guide the users [9], by allowing them to interact with data and create a mental model that will help them make informed choices [10]. However, there is a lack of information concerning how in fact users make decisions using information visualization and how visualizations should be designed in order to improve the decision making process [11].

In order to explain in what consists the decision making process, Simon [12] divides it into three key stages: *intelligence*, *design* and *choice*. The final stage, *choice*, consists of the decision stage where a user has to make a decision between several alternatives using some criteria. These choices can vary in complexity level, spanning from a person deciding which car to buy to which energy plan is more friendly to the climate. It is important to mention that the decisions participants make using visualizations become more relevant when the choice at stake can not be solved by computational solutions and cannot be fully automated [11]. An example of these types of decisions are personal decisions such as deciding which real-estate property to buy, or which medical treatment to undergo. Due to the personal level involved in these type of decisions, they become vulnerable to cognitive biases and uncertainty [13, 14]. Moreover, we believe this type of decision might be influenced by other human factors such as the level of trust perception a user has in a visualization.

Humans need to calibrate trust between their own previous knowledge and the information that is displayed by a visualization in order to build knowledge [15]. Trust can be described as the process of confirming and accepting knowledge [15]. In particular, in the context of Information Visualization (InfoVis), trust perception is related to how users evaluate the quality and reliability of the visualization [16]. Even though trust might not be relevant for every interaction with InfoVis, it becomes extremely important when there is some kind of risk associated with the information [16]. Therefore, trust becomes a critical aspect of knowledge building because it allows the user to minimize the uncertainty associated with digital information, especially when the user is vulnerable to suffer a loss if he or she believes in the information displayed.

There are many factors that have an effect on the way users perceive trust in InfoVis. In particular personal characteristics such as personality can have an impact on trust perception. Personality is what makes us different from other individuals and consequently differentiates our behavioural patterns and the way we make decisions [17, 18]. Throughout history, there have been many proposed models that tried to define personality. Most of the models subdivide personality into different traits. One of the traits that is transversal to multiple personality models is neuroticism. Neurotic individuals are characterized by the tendency to feel nervous, depressive, impulsive and even have feelings of guilt, fear and anxiety [19]. These types of characteristics make neurotic individuals more suspicious of others [17] and consequently lead to a decrement in perceived trust [17, 20, 21].

Another important aspect that affects the way users perceive information is the idiom and techniques that are used to create the visualization. There are several forms of visualizing information, to choose which one better suits our needs it is essential to understand the characteristics of the data. One important characteristic to explore is whether or not the data is time-oriented [22]. Time-oriented data allows us to understand the evolution of data over time and therefore find interesting patterns [23]. Visualizations that display this type of data can sometimes be used to display information that is intrinsically connected to risk. For example, during this pandemic, we frequently observed the evolution of COVID-19 new cases throughout the days. Furthermore, we sometimes found ourselves questioning the credibility of the information we were analysing.

Consequently, our work focuses on analysing whether **time granularity and users level of neuroticism have an impact on trust perception in the context of time-oriented linear charts of healthcare information**. Therefore, during our experiment, we place participants in an uncertain scenario and prompt them to make a risky decision. In particular, we place subjects in a health emergency scenario while there is an overcrowding crisis (there is no space left to meet the timely needs of the next patient requiring emergency care [24]), and then ask them to make decisions and perform tasks based on time-oriented information visualizations with a varying time granularity factor. We opted for this topic because we believe that having to wait for medical support in the emergency department is a situation that most people are familiarized with, thus emphasising the risky nature of the decision. Moreover, the analysis of different levels of granularity might be a first step in understanding design features that might increase users perceived trust in the process of decision making [11].

1.1 Contributions

The development of our work allowed us to understand that time granularity is a relevant feature of time-based visualizations. The majority of participants trust more in line charts that presented a larger number of data points, in our particular case, in visualizations that displayed the hours in a day (24 time points) and the days in a month (30 time points). Additionally, participants interact more with the line charts they trust the least. Results also demonstrate that the time granularity of the line charts that subjects initially rely on the most in the decision task has an impact on trust perception when visualization are examined one at a time.

Regarding the effect of personality on participants perceived trust, results did not show a significant relation between neuroticism and perceived trust. However, results point to the fact that neurotic individuals are the fastest at completing tasks when analysing visualizations for the first time. Moreover, additional findings showed a statistically significant interaction between perceived trust and other facets and factors from the Five Factor Model (FFM). In particular participants with lower levels of extraversion show overall higher scores of perceived trust. These contributions are relevant for designers of visual analytic systems, particularly when studying human decision making supported by information visualization. They provide implications to understand trust perception in a health emergency scenario as well as its variation in line chart techniques differing in the time-based granularity. Furthermore, these results suggest that the adaptation of visualization to users characteristics such as personality can positively influence the way users trust and collect insights from the information displayed.

Finally, we wrote and submitted an article for IEEE Transactions on Visualization and Computer Graphics (TVCG) with the tile *The Effect of Time Granularity in the Trust Perception of Time-Series Visualization*, that focuses on the impact of time granularity on trust perception. Additionally, we are also writing another article for TVCG with the title *How Personality Affects Trust Perception of Time-Series Visualizations*, that focuses on the impact of relevant facets from the FFM on trust perception of time-oriented data.

1.2 Organization of the Document

Chapter 2 presents the background of the concepts used in the study developed. This section provides a historical view of different personality models that approach the concept of neuroticism. It describes the neuroticism trait and its facets according to the FFM and finally, it addresses the concept of trust perception.

Chapter 3 provides an analysis of the procedures and contributions of relevant studies to the present work. Therefore, it explores the work that has previously been done concerning: the effect of personality on trustworthiness, focusing on the neurotic trait, the relationship between neuroticism and InfoVis, the relation between trust perception and InfoVis, and the relation between time series and InfoVis.

Chapter 4 presents the methodology used during our study. This chapter starts by describing the research question and hypotheses that were evaluated. Additionally, it also contains an explanation of the research design and all the statistical tests used to perform the data analysis of our results. This description is followed by a detailed explanation of how the data was collected, in particular how the participants were selected and the apparatus that was used during the experiment. Finally, it contains a description of the procedure of the study and the steps taken to analyse the data concerning users' personality traits.

Chapter 5 presents the results collected from the study developed. The results are divided into three major sections. Section 5.1 shows the main results concerning the effect of granularity on users' perceived trust. Section 5.2 displays the principal results concerning the effect of neuroticism on participants perceived trust. Finally, Section 5.3 presents additional findings regarding the impact of different facets of neuroticism and distinct factors from the FFM on users perceived trust. Moreover, this chapter provides a discussion concerning the results about the research questions previously described, the design implications and the limitations of the presented work. Finally, Chapter 6 describes the conclusions taken from the present work and possible future work.

2

Background

Contents

The following section contains a description of some of the fundamental concepts of this work, including a historical view of different personality models focused on their differences and similarities. In particular, we address the neuroticism trait and its facets. Finally, we discuss the concept of trust perception.

2.1 Fundamentals of Personality

Personality is one of the most important factors in our everyday life; it is what distinguishes us from other individuals and therefore differentiates our thoughts, emotions and behaviours [18]. Being a wide-ranging topic, it has been considered relevant in various fields concerning disparate areas, between them the field of Human-Computer Interaction. One of the main challenges of this field is to identify the users' requirements so that these interactions can be suitable to their needs, and the users' experiences can be as pleasant as possible [25]. As a result, the evaluation of personality traits can help to identify users' actions, inclinations, likes, and dislikes [25], and therefore determine users' needs.

Throughout history there were a lot of proposed models that tried to define personality and identify its dimensions in a way that would prevail during adulthood and be valid across countries and cultures. For many years the two most popular models were the Three-factor model defined by Eysenck (1950) [26] and the 16-factor model defined by Cattell (1957) [27]. Later, the FFM by Costa Jr and McCrae (1992) [28] was introduced. This model divides personality into five factors: neuroticism, extraversion, agreeableness, conscientiousness and openness. The FFM represented a breakthrough and is considered until today one of the best paradigms for personality structure. However, despite its success, as expected, there were still some critics made, and different ideas continued emerging. These ideas gave origin to models such as the Alternative Five Factor Model by Zuckerman (1994) [29] and others.

All the previously mentioned models focus mainly on identifying the basic dimensions of personality and how they can be described and evaluated. The basic dimensions of personality are the most important factors that allow us to differentiate individuals based on their emotions, interpersonal relations, experiences, attitudes and motivations [30]. The main discussion among researchers has been identifying which dimensions are transversal across models, and which can more accurately describe the personality structure. When identifying similarities between the models it is possible to see that all the previously discussed models identify a series of different traits. However, the number of traits varies across models. This difference is especially relevant when looking at the 16-factor model, one of the main criticism made to this model is the fact that most studies couldn't account for the presence of all 16 traits. It is argued that the 16 traits represent narrow factors that are comprised by the broader traits defined by other models [31]. The three remaining models share another similarity, all three models agree that one of the main factors is neuroticism. Moreover, this trait is evaluated using similar scales across all models [32].

As a result, in order to be able to evaluate and identify, in real individuals, the different traits described by the different models, several questionnaires have been developed across the years such as the NEO PI-R, the Big Five Inventory (BFI) and the International Personality Item Pool (IPIP). The NEO PI-R consists of a questionnaire with 240 items to which participants have to respond using a five point Likert scale. It evaluates the five traits defined by the FFM and six additional facets for each of the traits [33]. This questionnaire can be administered to teenagers (over 17) and adults. The BFI is a 44-item questionnaire developed to assess the big five personality dimensions (defined by the FFM) when there is no need for more differentiated measurement of individual facets [34]. Finally, the IPIP was developed to be a broad-bandwidth personality inventory, whose items are available in the public domain and can be used for both scientific and commercial purposes. IPIP's item format was specifically developed to ease translation to many different languages [35].

2.2 Neuroticism

Neuroticism is considered one of the basic dimensions of personality by many models. It is characterized as the propensity to be nervous, have feelings of depression, frustration, guilt and self-consciousness. These feelings are normally related to low self-esteem, impulsiveness behaviours and ineffective coping. Contrarily, individuals with low levels of Neuroticism are described as calm, relaxed and composed [19].

According to the FFM, neuroticism can be subdivided into 6 facets: Anxiety, Angry Hostility, Depression, Self-Consciousness, Impulsiveness and Vulnerability. **Anxiety** is related to individuals that experience tension, fear and apprehension. Individuals with high anxiety might experience phobias; **Angry Hostility** is visible in individuals that are typically impatient and that easily get frustrated and angry; **Depression** is perceived as the tendency to feel sad, melancholic, lonely and even desperate; **Self-consciousness** is related to the feelings of shame and embarrassment. Individuals with high self-consciousness have the tendency to feel social anxiety and uncomfortable with their peers; **Impulsiveness** refers to the inability to resist temptations; Finally, **Vulnerability** relates to the tendency to experience stress and panic during emergencies. Vulnerable individuals find it hard to deal with tension which makes them highly dependent [36].

As presented in section 2.1, neuroticism is defined in a similar way across different models. However, it is important to highlight at least one difference. While both the Three-factor model and FFM consider angry-hostility and anxiety as facets of the Neurotic trait, the Alternative Five Factor model divides the latter two facets into two different traits: Neuroticism-Anxiety (N-Anx) and Agression-Hostility (Agg-Host) [37].

Besides the similarities and differences found across different personality models, it is also possible to find relationships between them. As we mentioned, the 16 traits in the 16 factor model can be seen as facets of the traits defined by other models. When studying the 16 traits in the perspective of FFM it was possible to find a positive relation between privateness and perfectionism relatively to Neuroticism, and a negative relation between emotional stability and Neuroticism [38]. This lack of emotional stability allows us to understand that neurotic individuals tend to be reactive and are normally affected by their feelings, which can impact the way they trust other individuals [39].

Moreover, individuals high on neuroticism can also be defined in relation to their decision-making style. There is a positive relation between neuroticism and the dependent decision-making style [40]; This style is characterized by the need to get guidance and support from others when making important decisions. In addition, due to its anxious facet, highly Neurotic individuals have the tendency to avoid risk when taking action [41].

In light of this, taking into account all the factors that define neurotic individuals, it is possible to draw a conclusion about the effects of this personality factor on trust and trustworthiness. Contrary to individuals with high Agreeableness [1], highly neurotic individuals experience social anxiety and find it difficult to relate to others which can lead to lower initial trust. Furthermore, the vulnerability facet characteristic of neurotic individuals implies great distress when dealing with uncertainty and risk. Considering that trust emerges in risky situations, vulnerability will most likely entail a decrease in perceived trust.

2.3 Trust Perception

Trust is one of the most important and central aspects of our everyday life, it has an impact on economic, social and political relations, and on the way we connect to others; it represents the willingness to be vulnerable to another [42]. To better understand this concept we have to decompose it and understand two other different concepts: trusting and trustworthiness. On the one hand, we can define trusting as the inclination of a person A to believe another individual (B) will cooperate in an action for A's benefit, and will not take advantage of A, even if he is given the chance. On the other hand, trustworthiness is the willingness of a person B to act favourably towards a person A when A has placed an implicit or explicit demand or expectation for action on B [43]. A common example that exemplifies the definition of trusting is the situation where a father allows his child to enter a swimming pool and assumes that if his child is at risk of drowning, then unknown people will help and sacrifice, at least their comfort. Oppositely, trustworthiness can be described as the same situation, where the child is in fact drowning and an unknown person is expected to do something to save the child [43]. Considering our day-to-day life it is not difficult to find ourselves in a situation where we need to trust or be trusted. In fact, there are

[1] Trust is one of the facets of Agreeableness. Individuals with high Agreeableness have the tendency to consider other individuals honest and well-intentioned [36].

many factors that trigger trust and constantly impact the way we make our decisions.

We can narrow these factors into three main aspects that drive the decision to trust another individual: attitudes toward risky prospects, trustworthiness expectations and betrayal sensitivity [44]. The first factor, **attitudes toward risky prospects**, is based on the idea that trust can be described as a decision under risk and therefore is dependent on probabilities and outcomes. The second factor, **trustworthiness expectations**, reflects the amount and validity of available information and how this information is integrated to form a judgment [44]. This factor is also related to prior trust experiences and trust cues that can be a determinant of trust behaviour. Finally, **betrayal sensitivity** explains the way individuals accept vulnerability when there is a potential betrayal by another person or source of information [44].

When it comes to InfoVis, trust is perceived in a slightly different way, it refers to the tendency of a user to rely on a visualization and to build on the information displayed [16]. In particular during our study we focused more specifically on trust perception. Trust perception explores the way users interact and evaluate the data in a visualization in order to be able to rely or not on the information displayed [16].

Trust is something that needs to be built (or rebuilt). Furthermore, trust is also characterized by the situations where it emerges. As mentioned before, trust becomes a relevant factor especially when there is some kind of risk [45] and uncertainty associated with the information. Another aspect that sparks trust when a user interacts with a visualization is transparency and its different dimensions such as accuracy, thoroughness, and disclosure [5]. Transparency is characterized as the perceived quality and quantity of intentionally shared information [46]. Finally, not every user will perceive trust in the same way, trust perception is intrinsically connected to users' individual differences and personality traits [47]. In fact, trust has been seen as a construct with a stable propensity [48, 49] and researchers believe that individual differences might be responsible for differences in trust perception between individuals. Section 3.1 explains in more detail the relation between neuroticism and trust.

3

Related Work

Contents

This section contains a description of the procedures and contributions from relevant studies to the present work concerning: the effect of personality on trust perception, with focus on the neurotic trait, the relationship between Neuroticism and InfoVis, the relation between trust and InfoVis, and the relation between time series and InfoVis.

3.1 Personality and Trust Perception

Researchers in personality have yet to find out whether personality has a direct impact on our actions, behaviours and in determining important life outcomes. However, many agree that personality impacts several aspects of our life and therefore affects the way we live and make our decisions [17]. In order to study this relationship between individual differences and trust many studies have been conducted over the years.

A first approach, used in several studies, was to take advantage of traditional games and puzzles related to the context of risk and trust perception, such as the trust game, to explore participants level of perceived trust. By conducting these games researchers were able to explore the relations between users' interactions and their respective personality scores. Alarcon et al. [50] explored the role of personality across the trust process with three different studies. For one of the studies, 66 pairs of participants were recruited, where each pair was composed of two individuals that already had a trusting relationship (family members, friends). The participants were asked to play a version of the Prisoner's Dilemma, where they had to cooperate or defect with their partner. They were asked to complete this task twice, the first time with their familiar partner and the second time with someone they did not know. Two different methods were used to assess participants' personality: the NEO Five-Factor Inventory (NEO-FFI) questionnaire to measure personality concerning the FFM, and Mayer's Davis [51] scale to measure propensity to trust. To measure trustworthiness participants were asked to classify their peers' ability, benevolence, and integrity using a seven-point Likert scale. They found out that both in unfamiliar and familiar pairs, individuals high in agreeableness and low in neuroticism tend to consider others' people actions as well-intentioned, which leads to higher initial trustworthiness. Taking into account the familiarity between participants, they were also able to find out that, for familiar pairs, previous experiences with their peers had a higher impact than the participants' beliefs about the trustworthiness of people. This leads to the conclusion that individual differences tend to be more relevant when the person to be trusted (trustee) is unknown or participants are exposed to a new environment or situation.

Another example is the study conducted by Müller and Schwieren [17] that analyzed the impact of personality factors on behaviour in the Trust Game. The Trust Game consists of an interaction between two players, a sender and a receiver. At the beginning of the game, both players have ten units of a fictional currency. The sender starts by deciding how much money he/she would like to invest, he/she can decide to keep all his money. If the sender decides to invest, the receiver will receive triple the amount invested. Finally, the receiver decides how much money he/she would like to send back to the sender (and how much to keep to himself/herself). The NEO PI-R questionnaire was used to measure participants personality scores, and the overall trust was measured by the amount each participant sent. Firstly, results showed that personality only impacted the decisions of the first player (the trustor), the decisions of the second player were better explained by the situation. Secondly, when evaluating the correlations between the amount sent and personality, results showed that individuals high on neuroticism sent the least amount of money and therefore were the ones that least trusted their peers. This conclusion is possibly explained by the fact that individuals scoring high on neuroticism tend to be anxious and therefore will avoid the risk of not getting paid back.

The two previously mentioned studies focused on evaluating trust through the analysis of participants actions towards others. However, different studies present different approaches to measure trust. Evans and Revelle [21] developed a study with the purpose of analyzing the validity of a new trust Inventory - Propensity to Trust Survey. Furthermore, they examined the relationship between trust, trustworthiness, and the Five-Factor Model. Initially, participants were asked to answer an online questionnaire that contained items taken from the IPIP. Second, participants were asked to answer 10 items from the propensity to trust survey, this survey contains a list of 40 items to measure trust and trustworthiness, also taken from the IPIP. The survey was previously rated by individuals familiar with the trust literature. The study found out that the Propensity to Trust Survey was reliable for evaluating both trust and trustworthiness. Besides that, when analyzing the correlations between the personality traits and propensity to trust it was possible to find a negative correlation between Neuroticism and trust. It was also possible to find regression models that defined trust and trustworthiness as composites of the FFM, more specifically trust was predicted by neuroticism.

Additionally, other works focused on the factors that trigger trust in order to study its relation with different personality factors. As mentioned before, a factor that affects trust are user's attitudes toward risky prospects, specifically the way people assess risk and consequently make decisions [44]. Cho et al. [20] studied the effect of personality traits on phishing susceptibility. Therefore, researchers created probability-based mathematical model using Stochastic Petri Nets to analyse this topic. In order to develop this model they investigated the correlations between users' personality, perceived risk and perceived trust. With this model, they were able to conclude that the increase of the Neuroticism probability led to an increment in perceived risk and, at the same time, a decrement in perceived trust.

Finally, recent research has begun to explore the effects of personality on user trust in AI-Enabled User Interfaces. Böckle et al. [1] explored the impact of the FFM personality traits on trust in AI-enabled user interfaces in order to define best practices and design guidelines for distinct personalities. In order to perform this experiment researchers created a survey with different storyboards covering the best practices of design guidelines proposed by Google. These practices are divided into four categories: (i) *mental model*: concerns the end users' understanding of how AI systems work, (ii) *model confidence*: address how the end-user receives an appropriate level of explanation regarding how the system works and its degree of confidence in its output, (iii) *feedback and control*: mechanisms that provide a better end-user experience by suggesting personalized content, (iv) *errors and graceful failures*: guidelines that help identify and diagnose AI context errors and communicate the way forward. Participants were asked to analyse four sets of storyboards based on each of the previously mentioned principles (example: Figure 3.1) and rate each storyboard on their perceived trust based on four factors: (i) The system is reliable, (ii) I am confident in the system, (iii) I can trust the system, (iv) I am not suspicious of the systems' intentions, actions, or outputs.

Results showed that for participants with high levels of extraversion and agreeableness there was a positive relationship with the aforementioned best practices, while for participants with high levels of openness a negative relationship was found.

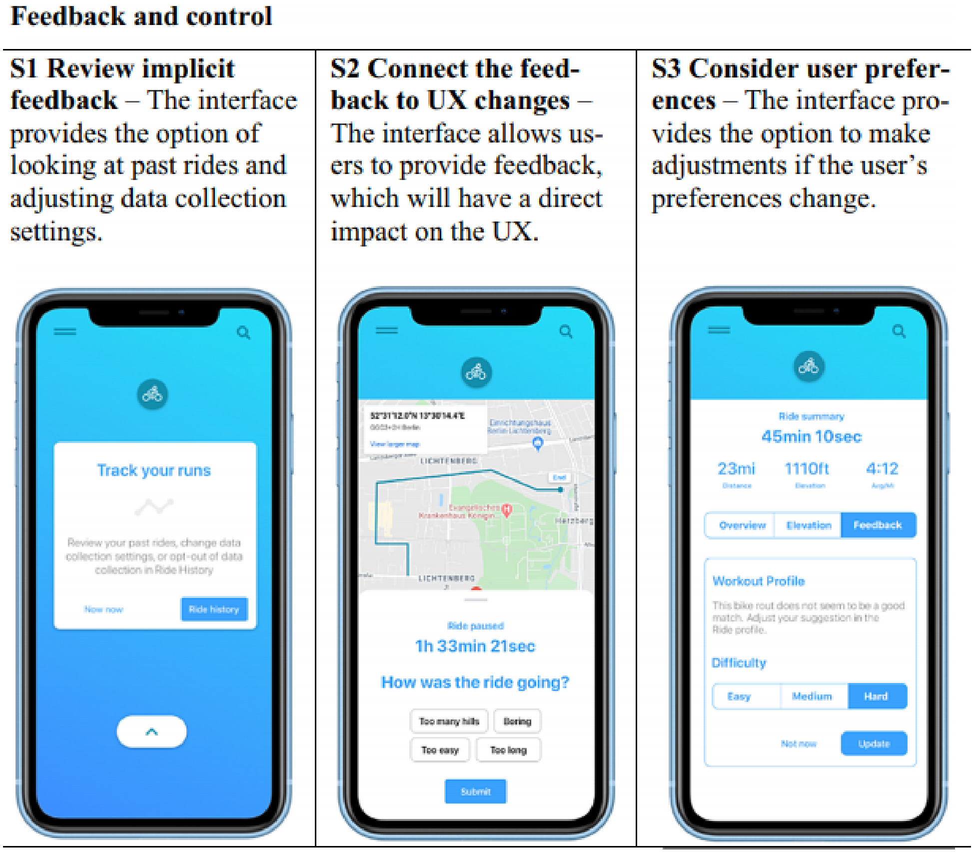


Figure 3.1: Storyboard based on feedback and Feedback and Control [1].

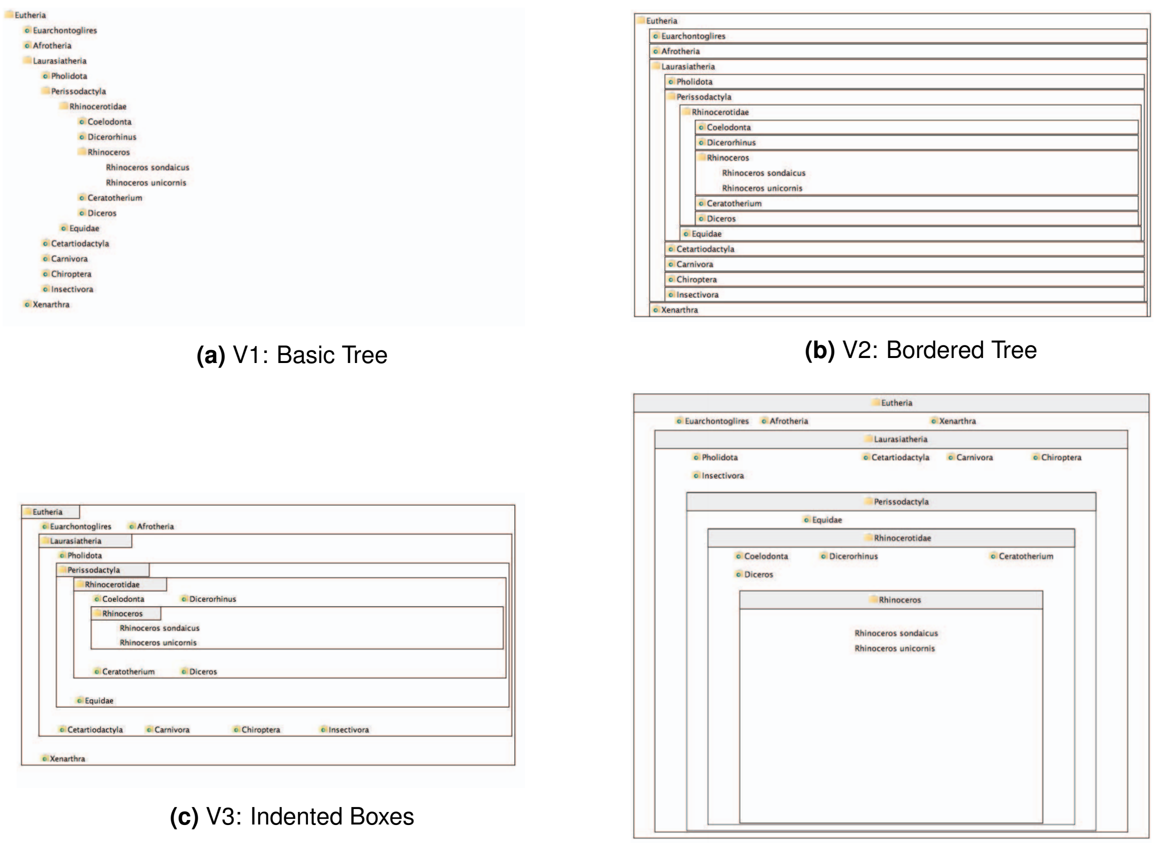
3.2 Neuroticism and Information Visualization

Visualizations are used to support the development of new information by those who see it. They can be used to build knowledge or even trigger reasoning processes and infer conclusions when interaction is used. Consequently, it is impossible to understand a visualization and its potential without knowing how the visualization's user thinks [52]. It is also important to understand the built-in learning "presets" of the user, once what is considered an intuitive visualization for one subgroup of users may be difficult to interpret by another [53].

When designing a viable and successful technological development, such as a visualization, it is necessary to consider the target users, and the differences that represent them. These differences include age, job, gender, culture, personality, among others [54]. Personality traits are individual characteristics that influence the way we think and behave. Thus, the interaction between personality traits and a visualization can largely impact important decision making processes [55], dexterity using the interface, and the way users solve problems [52]. A trait is considered impactful in a problem-solving situation, using visualizations, if it usually results in a problem-solving, decision-making or socioeconomic advancement [55]. Furthermore, personality traits seem to be more impactful when more complex tasks are evaluated [52], and in situations where it is required a higher cognitive load, such as tasks involving inference and metaphorical reasoning [52]. There are several traits to consider when evaluating personality, one of them being neuroticism. As explained in Section 2.2, neurotic individuals are characterized, among other things, by the anxiety and fear they feel.

Taking into account user's personal characteristics when designing visualizations, it is also important to consider that the design of the visualization, the balance between elements and the layout disposition may also have a significant impact on how each user perceives the information presented [2]. Ziemkiewicz et al. [2] investigated the relationship between visualization layouts and personality factors. Participants were asked to answer search and inferential questions about four different visualizations. The design used for these visualizations was developed with the purpose of increasing the containment metaphor used to illustrate the hierarchy represented, as shown in Figure 3.2. To measure personality, participants were asked to fill the Locus of Control and FFM inventories. After completing the experiment, participants answered a survey about how much they liked the visualization. Results concluded that high neuroticism was positively related to accuracy. In addition, neurotic individuals became more accurate as the views became more container like. These results may be explained by the fact that individuals who have higher scores on neuroticism are more attentive and therefore find it easier to learn from unfamiliar interfaces. Another explanation is that neurotic participants feel pressured to complete all the tasks correctly, independently of the layout of the visualizations.

Moreover, another aspect that has an impact on user experience, and on the insights users are able to collect from a visualization is interaction [56]. Interaction helps the visualization convey its message,



(a) V1: Basic Tree

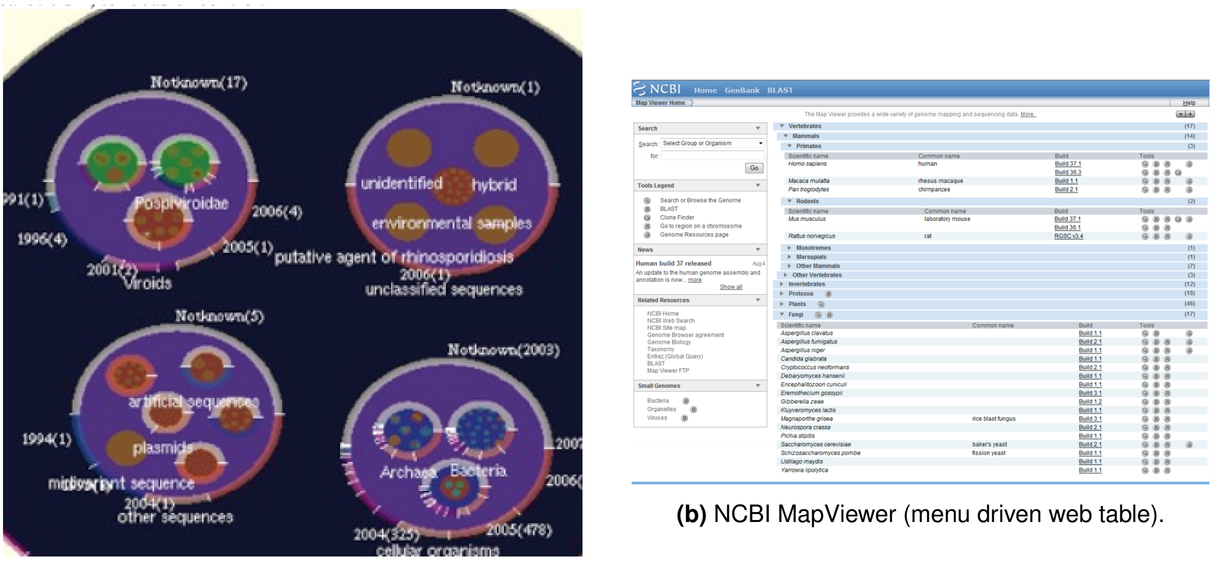
(b) V2: Bordered Tree

(c) V3: Indented Boxes

(d) V4: Nested Boxes

Figure 3.2: Four visualizations used on the study varying on the containment metaphor [2].

therefore helping users to engage and explore the data [56]. In light of this, Green and Fisher [3] explored the impact of personality features on interface interaction and learning behaviours in both an interactive visualization, and a menu-driven web table built to display genomic information as shown in Figures 3.3 a) and b) respectively. Personality was measured using three psychometric measures: Locus of Control, extraversion, and neuroticism. The latter two measures were assessed using the IPIP Mini Big Five Personality Inventory. During the experiment, two studies were performed and users conducted a series of tasks designed to test procedural performance, such as identifying a target located in the hierarchy. Finally, participants were asked to give some insights on their experience, and to classify both interfaces. Results showed that neuroticism was negatively correlated with completion times in both interfaces, meaning that higher neuroticism corresponded to faster times. Moreover, this experiment suggested that more neurotic individuals reported fewer insights than participants with a lower neuroticism score. These results indicate, once again, that highly neurotic individuals are more attentive and therefore can quickly adapt to new interfaces.



(a) The main view of GVis (interactive visualization).

(b) NCBI MapViewer (menu driven web table).

Figure 3.3: Two visualization used to display genomic information [3].

Finally, in order to create adaptive visualization it is necessary to understand the user's context and the visualization that they require [4]. Consequently, an important element to consider is the detail of information that will be displayed in a visualization. Oscar et al. [4] studied the consequences of mismatching the granularity of information presented on visualization to the user's needs. Granularity is represented by the level of detail that is presented in a visualization. Visualizations with low granularity show summarized or aggregated information, whereas visualizations with high granularity present more of the source information as presented in Figure 3.4. Participants were shown different visualization and asked to complete four "find values" tasks and four "compare values" tasks. The questions presented to the user could either match the visualization granularity or not. It was possible that users were not presented enough information to answer the questions. Personality characteristics were measured using the Ten Item Personality Inventory (TIPI). Results showed that individuals high on neuroticism were more likely to identify that a task with high need for detail could not be answered correctly when there was a mismatch with a low granularity visualization (that did not show the detail needed to answer the question). Therefore, individuals high on neuroticism were less likely fooled by mismatches in the information.

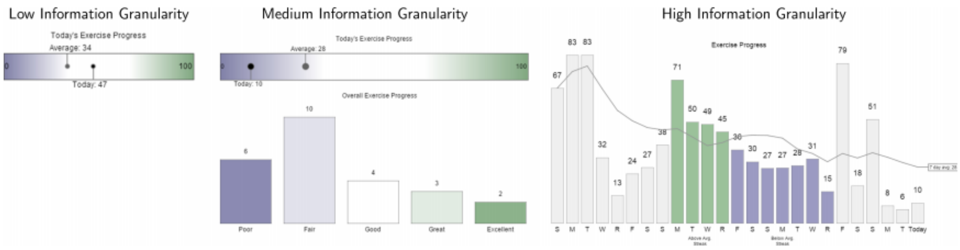


Figure 3.4: Example visualizations for different granularity levels [4].

3.3 Trust and Information Visualization

There are many important factors when a user sees and analyses a visualization. Most of the times the main purpose of a visualization is to convey the user new knowledge or information, but for that to happen users need to be able to trust the knowledge they generate through visual analytics [15]. Therefore, trust has a huge impact on InfoVis, because users depend on it to build new assumptions.

In human relations there are always two sides to relations based on trust: the trustor and the trustee. In the case of InfoVis the trustee becomes the visualization itself. Therefore, just like when analysing human relationships, in InfoVis we have two sides for trust: trustworthiness and trust perception [16].

On the one hand, trustworthiness is affected mainly by three factors: **trustworthiness of information, presence and display of uncertainties and design factors**. The first factor, **trustworthiness of information**, is affected by accuracy, currency, objectivity and completeness of the data [16]. **Uncertainty** is intrinsically connected to trust, many authors believe that, to a certain point, trust is increased when the users are informed about the presence of uncertainty in the data [6]. Finally, **design factors** increase user experience and usability, consequently activating positive feelings that trigger trust [16].

In regards to how design factor can affect trust perception, Kong et al. [57] studied whether the misalignment of a title to the visualization content had an impact on trust and recall of information. This experience took into account two different types of misaligned title: (i) a miscued slant: where the visualization emphasizes one side of the story, whereas the title addresses the other (less emphasized) side of the story, and (ii) a contradictory slant: where the information conveyed in the title is not presented at all in the visualization. Initially, participants were asked to fill out a demographical survey and report their attitudes on six controversial issues, two of which were visualization topics used in the study. Based on the participants' answers two visualization-title pairs were chosen for them to analyse. One of the titles favoured the participants' attitude toward the thematic and the other did not. After analysing the visualizations participants were asked to answer factual questions, recall the main message of the visualization, and identify whether the information was consistent with their beliefs ("Strongly biased," "Slightly biased," or "Neutral"). Finally, credibility was measured based on five factors: accuracy, fairness, trustworthiness, bias, and completeness.

Results showed that perceived credibility and impartiality of the information significantly decreased when the title was misaligned with the participant's existing attitudes in both the miscued and the contradictory conditions. Showing that participants are likely to dismiss information if it goes against their preconceived notions. Additionally, results showed that trust perception seems more dependent on the title, rather than the content of the visualization itself. This is consistent with the fact that most people recall information based on the title rather than on the visualization data. In conclusion, design factors seem to have a significant impact on the way users perceive information.

On the other hand, trust perception is affected by many factors such as **lack of clarity**, which can anger a user and set off doubts about the accuracy and certainty of the information; and **users' personal characteristics** such as prior knowledge, personality, intentions and perceived risk [16].

In order to explore the effect of some of these factors on user trust, Xiong et al. [5] conducted a study to analyze the relationship between data visualization transparency and trust. The visualization chosen to study this relationship was a map-based driving application. This visualization was chosen because most people are familiarized with it, and have experienced uncertainty with this type of application. Participants were shown two different visualizations and were told these two visualizations were a screenshot of two different driving applications. These visualizations displayed routes from several fire stations to a fire location, highlighting the fastest route in blue as shown in Figure 3.5. It is important to mention that the fastest route was always different in the two visualizations presented. Participants were then asked to put themselves in the role of a firefighter and choose which visualization to use and from which fire station to dispatch firefighters. Finally, participants were asked to rate trustworthiness, accuracy, clarity, disclosure, and thoroughness of all the visualizations they were shown using a five-point scale

To examine the effect of transparency on trust, the visualizations presented to the participants varied in the amount of information displayed by changing the number of routes presented, the number of fire stations recommended and the number of fastest routes. Results showed that participants were more likely to choose visualizations that appeared clearer, more thorough and that disclosed a greater amount of information.

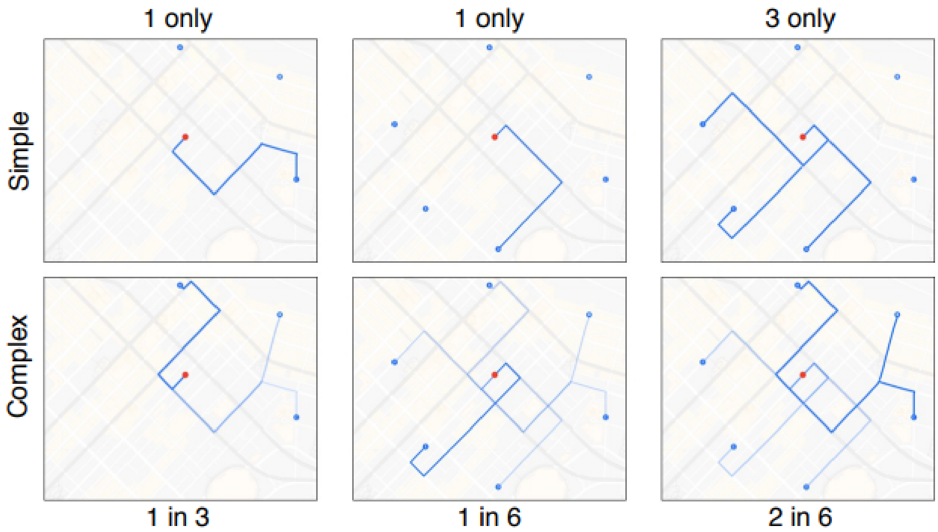


Figure 3.5: Blue dots represent the fire stations, red dots represent the fire location and the fastest routes are highlighted in blue [5].

Zhou et al. [6] investigated the effects of personality traits on user trust in human-machine collaborations, concerning the impact of uncertainty and cognitive load. To perform this experiment, they took inspiration from a case study about water pipe failure prediction. As such, participants were shown a visualization of a failure prediction model learned from historical pipe failures as shown in Figure 3.6. Participants were then asked to make a budget plan in terms of pipe length to be inspected, they could either trust the Auto Predict Assist (shown at the top right of Figure 3.6) or provide their own estimations. Each participant was asked to perform this task for three different visualizations: no uncertainty, non-overlapping uncertainty, and overlapping uncertainty as shown in Figure 3.7. For each of these three visualizations participants repeated the task four times for four different levels of cognitive load. Participants were asked to remember three, five, seven or nine random digits to intensify cognitive load. In order to measure Trust, after each decision participants had to fill a questionnaire to rate their trust using a nine-point Likert Scale. Finally, to measure personality participants filled the TIPI previously to the experiment.

It was possible to conclude that uncertainty will lead to an increase of trust, but only when the cognitive load is low. It was also possible to extract some conclusion in relation to personality, and more specifically neuroticism. Individuals high on neuroticism express higher trust when uncertainty probabilities are known.

Finally, with the purpose of facilitating interpretation and trust, Chuang et al. [58] presented some design guidelines to use in model-driven visualizations, between them: alignment and progressive disclosure. The first guideline (alignment) values the matching of the details of the visualization to the user's tasks by minimizing the information that might distract the user. The second guideline (progres-

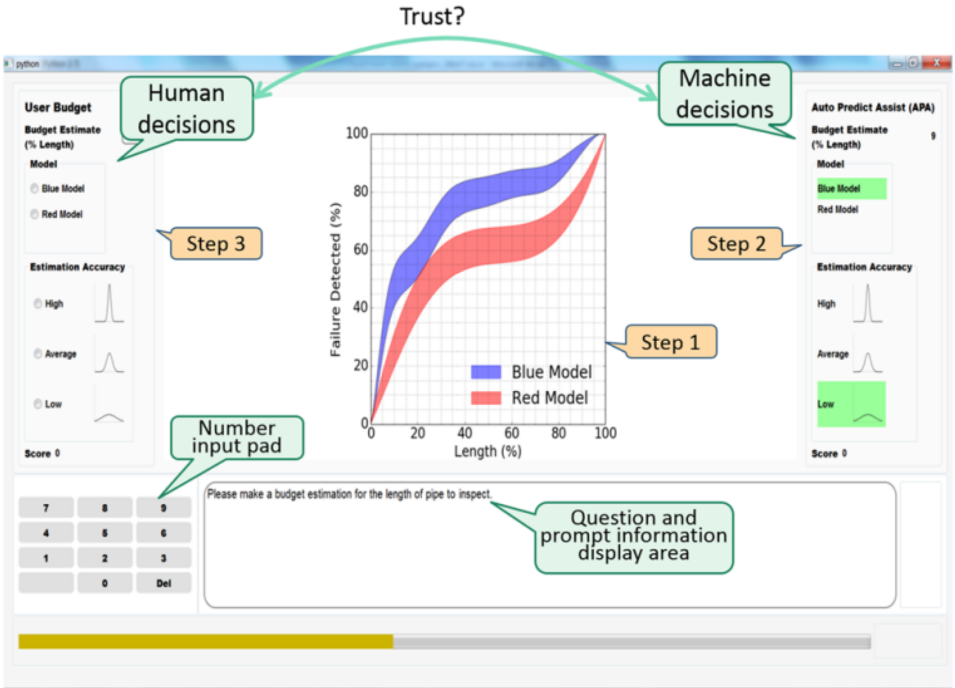


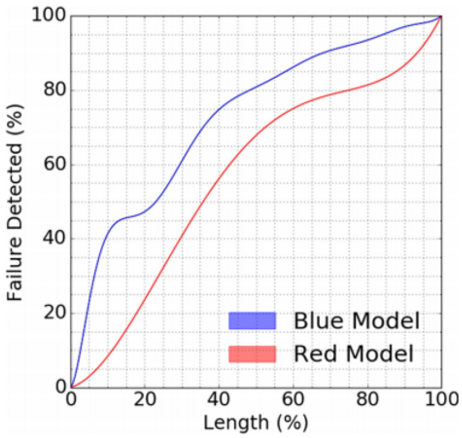
Figure 3.6: Pipe failure predictive model visualization [6].

sive disclosure) is a technique that allows users to ”drill-down from high-level overview, to intermediate abstractions, and eventually to the underlying data itself”, this allows the users to select the detail of information according to their needs.

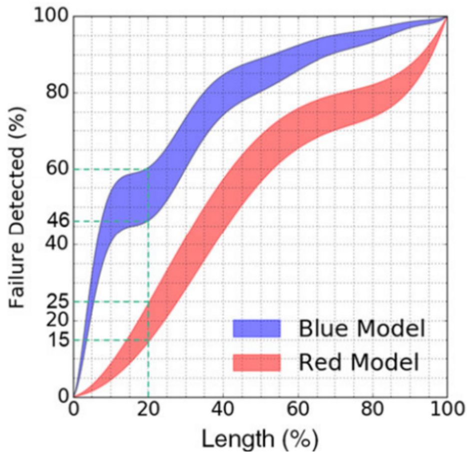
In summary, to access the quality and reliability of a visualization trust needs to be built. These process is done by analysing the data, finding its uncertainties and retaining or building knowledge from it. Nevertheless, in some cases the user is unable to become fully aware of all the properties in the data. When that happens other factors may be evaluated such as confidence in the source [16] or even personal characteristics.

3.4 Time Series and Information Visualization

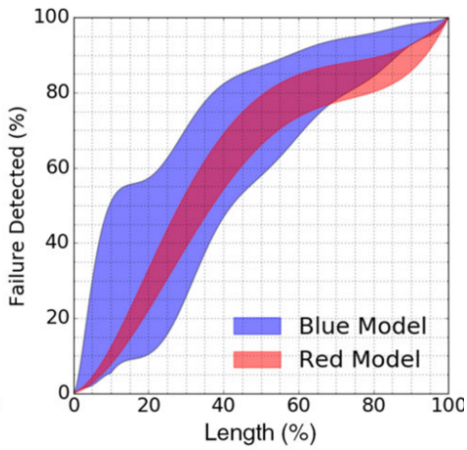
In order to design visual representations of information, it is essential to understand the characteristics of the data. One important feature to explore is whether or not the data that is going to be visualized is time-oriented [22]. There are several dimensions that characterize distinct types of time such as: linear versus cyclic time, time points versus time intervals, and order time versus branching time. Different time characteristics implicate the usage of alternative visualization techniques and idioms. Besides facilitating the visualization and analysis of time-oriented data, it is also important to use methods that parametrize the data according to users' tasks. Therefore, interaction can provide a solution that allows users to



(a) Performance curves of ML models without uncertainty.



(b) Performance curves of ML models with non-overlapping uncertainty.



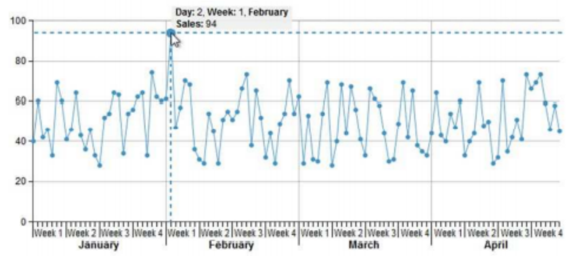
(c) Performance curves of ML models with overlapping uncertainty.

Figure 3.7: Pipe failure predictive model with different levels of uncertainty [6].

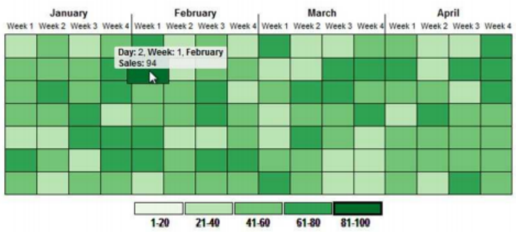
explore the data according to their needs [22].

Adnan et al. [7] performed a study to understand the effectiveness of time series visualizations when using different interaction, visual encodings, and coordinate systems. Highlighting and tooltips were used to evaluate the effectiveness of interaction on the different visualizations. For the coordinate system, both the Cartesian and polar systems were used to test the effect of alternative coordinate systems. Finally, the visual encodings techniques used were characterized as one of the three types: position, area, and color as shown in Figure 3.8. Participants were asked to perform four different tasks for each of the different visualizations. Additionally, they were asked to answer a 5-point Likert scale questionnaire that evaluated their level of confidence when performing the tasks in order to measure the impact of all the different variations. During the experiment it was registered tasks' completion times and correctness when answering to the different tasks.

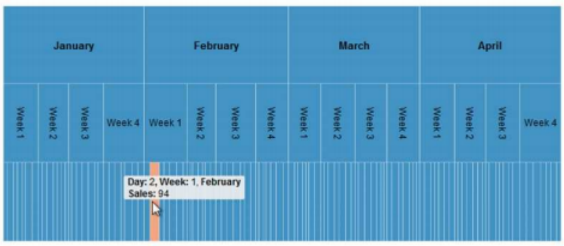
In regards to interaction, results showed that introducing interactivity enhanced the user experience without the loss of efficiency and accuracy. Concerning the impact of visual encodings, positional encodings were considered the best option to visualize time-oriented data. Finally, concerning the coordinate system it was concluded that the polar coordinate system should only be used if there are clear reasons that favor it. Otherwise, the cartesian coordinate system is the preferred option.



(a) Line chart (position).



(b) Rectangular heatmap (color).



(c) Icicle plot (area).

Figure 3.8: Visual encoding techniques [7].

Javed et al. [59] investigated the user performance for comparison, slope, and discriminating tasks using different line graph techniques involving multiple time series. The experiment conducted used four different idioms: line graphs, small multiples, horizon graphs, and braided graphs. The experiment consisted of performing three different tasks: (i) finding the time series with the highest value on a specific point in time, (ii) finding the time series with the highest increase during the entire time, (iii) finding the time series with the highest value for a point specific to each series. Besides changing the idiom, the number of series presented, and the size of the chart also changed. Each participant performed a total of 216 trials.

Results showed that shared-space techniques such as line charts and braided graphs were better for comparison tasks, once the comparison can be done in the same space. However, in shared-space techniques, the increase of time-series resulted in cluttered visualizations which led to an increase in task completion time and a decrease in accuracy. Finally, line charts and small multiples were considered the most versatile idioms for performing tasks related to the visualization of time-series data.

Finally, with the purpose of analysing the effectiveness of radial charts, Waldner et al. [60] conducted a study to evaluate the performance of radial charts when analysing daily patterns. For the purpose of this study, four different visualizations were created, two with a radial layout and two with a linear layout. The first two consisted of two juxtaposed 12-hours radial charts and one single 24 hours radial chart. The latter two consisted of two juxtaposed 12-hours linear bar chart and a 24h linear bar chart. The data used for all visualizations represented the number of traffic accidents in Staten Island from 2016. Participants were asked to perform several tasks regarding different visualizations. Firstly, participants were asked to write down everything they could observe from a specific visualization. Secondly, participants were asked to perform a series of analytic tasks such as locating time intervals and reading a value from a specific time. Finally, participants were asked to rate each visualization. Completion time and errors were also collected.

Results showed that overall linear charts had the best performance even when showing periodical daily data. This is probably due to the fact that most participants are more familiar with linear charts. Results also concluded that continuous bar charts should be preferred over separate ones. Contrarily to what was expected, most users were not able to understand, and therefore take advantage of the clock metaphor behind radial charts. This factor might also have impacted the preference of linear charts over radial charts.

3.5 Discussion

The aforementioned studies suggest that there is a negative correlation between neuroticism and trust. A possible explanation for this is the fact that the neurotic trait is often associated with anxiety and fear. Therefore, individuals with high neuroticism have difficulties dealing with risky situations and when forced to make a decision. These characteristics may lead to a decrease in trust. Different studies used different questionnaires to measure the neurotic personality trait such as the NEO-FFI, NEO PI-R and the TIPI.

Neuroticism has also been studied in the context of information visualization. Most studies concluded that neurotic individuals are more observant when performing tasks, which results in lower completion times and higher accuracy. A possible explanation for this is the fact that individuals with high neuroticism are normally associated with low stability and are used to feeling “out of control”. This feeling of anxiety perceived by neurotic individuals can be advantageous when facing unfamiliar visualizations. Another explanation is that neurotic individuals feel pressured to complete all the tasks correctly. Therefore, the effect caused by the unfamiliarity with unusual idioms or techniques is mitigated.

Concerning the effects that personality traits have on trustworthiness of information, there is still little research on the topic. To our knowledge, only Zhou et al. [6] studied the effects of personality traits on user trust in human-machine collaborations. From this study it was possible to acknowledge that neurotic individuals express higher levels of trust when uncertainty is represented. This conclusion comes in line with the previous results that suggest that neurotic individuals are more observant and feel pressured to answer correctly. This type of characteristics can cause neurotic individuals to pay more attention to details and uncertainties associated with the data. Being highly insecure, it is possible that neurotic individuals feel more comfortable when they have access to more information.

Even though most studies do not consider individual factors as a variable, some studies have already investigated what factors impact trust in the context of information visualization. From these studies, it is possible to conclude that participants tend to trust visualizations that are clearer, more exhaustive and that display more information. According to Mayr et al. [16] another factor that influences trust are design factors, yet there is still little research made in that direction. Concerning the evaluation of trust, there is still no consensual metric or system, different studies have adopted distinct approaches. Some studies measured trust by analysing the actions taken by the participants, while others used questionnaires to measure the trust perceived by the participants after completing the tasks. Xiong et al. [5] mentions that asking participants directly about their trust might not lead to accurate results. To address this problem Sacha et al. [15] suggests that trust can be measured by participants’ overall decision switching. Another proposed method is to use think-aloud evaluations and look for words such as “unsure, uncertain, maybe, perhaps”.

In this light, it is still not clear what leads a user to trust a visualization when there is not a thorough evaluation of the information presented. Most studies recognized that different participants prioritized distinct aspects when making decisions based on visualizations. These variations among participants point to the effect of psychological differences in the trust process. Consequently, there is still a considerable gap when it comes to the analysis of the implications of user traits in the perceived trust in information visualization. Research about this topic could push the development of personalized visualizations that adapt to the user needs and individual differences, and provide guidelines to design visualizations that would increase user trust.

4

Methodology

Contents

The purpose of this work is to analyse whether specific user characteristics have an impact on the way users trust and collect information from visualizations of time-oriented data. In particular, we focus on the personality trait of neuroticism.

Trust becomes more relevant in risky situations [45] (Section 2.3), such as medical emergencies. As such, we took inspiration from the emergency department overcrowding crisis. Overcrowding occurs when there is no space left to meet the timely needs of the next patient requiring emergency care [24]. More than ever the majority of admissions in hospital and emergency departments are unscheduled [24]. Indeed, by 2003 90% of large hospitals in the USA had their emergency departments operating at or over capacity [61]. In fact, there are many causes that explain this problem, among them: unnecessary visits to the ED, poor management of the surgical scheduling, increased affluence due to seasonal illness, and higher severity of illness resulting from an ageing population [24].

This problem is especially relevant because of its negative impact on the patients' safety and health. Overcrowding is associated with long waiting times, especially for patients that are not critically ill, increasing mortality, patient walkouts, and ambulance offload delays. Therefore, overcrowding results in a decrease in the quality of care. Besides that, it also harms hospitals due to financial losses [61].

In this light, our approach is to present an emergency context to each participant where they are in a position of having a health emergency and needing to decide to which hospital to go. Participants will be asked to choose a hospital having in mind that they want to be admitted as soon as possible. We formulated our study so that this choice would solely be made based on the number of patients that visited the emergency department in a certain period. We chose this approach in order to limit the variables involved. In particular, participants were presented with three visualizations where each one has a different time granularity level: day, week, and month. These three different granularities were chosen to analyse different measures of time. The day granularity allows the participants to analyse the most recent events, i.e last 24 hours. The week granularity gives users context but on a week-based scale, i.e last seven days. Finally, the month granularity will allow the users to analyse trends during a longer period, i.e last 30 days. We decided to variate between shorter and longer periods of time to analyse what users prefer to visualize when making decisions about immediate events.

4.1 Research Question and Hypotheses

The main goal of this work is to analyse **whether neuroticism has an effect on perceived trust in the context of information visualization of time-oriented data**. In particular using line charts in the context of healthcare information. Moreover, we want to analyse whether there are elements in a visualization that increase trust perception, such as the time granularity used to present time-oriented data.

Previous research showed that neurotic individuals tend to have a hard time making decisions [40], especially in risky situations [41]. We believe neurotic individuals will show lower levels of initial trust and will try to avoid any risks. Therefore, they will take more time analysing the three visualizations before choosing which one they trust the most. Furthermore, we want to explore whether neuroticism has an impact on the difference between the time taken to choose a visualization on the first time analysing the three visualizations, and the time taken to choose a visualization after analysing all three visualizations individually. This analysis will allow us to understand whether an individual extensive analysis of each visualization has an impact on the user interaction and decision making patterns. Therefore our first hypothesis is:

H1: Neuroticism positively affects the time taken to choose a visualization.

Another important goal of this experience is to explore whether people with different levels of neuroticism have different levels of perceived trust when exposed to different granularities of information. Firstly, we want to analyse whether time granularity has an effect on neurotic individuals' trust perception when analysing time series. Variations in granularity allow the display of more or less detailed information. As mention previously (Section 3.2), Oscar et al. [4] concluded that neurotic individuals were more likely to identify that a task with a high need for detail could not be answered correctly by the analysis of visualizations with low granularity. Moreover, the increase in granularity leads to greater disclosure of information. Disclosure of information is one dimension of transparency and represents one of the most important factors that influence trust. In fact, Xiong et al. [5] concluded that participants were more likely to trust visualization that disclosed more information. Therefore, we hypothesize that **H2.1: The visualization granularity positively affects trust perception**.

Moreover, we collect user personalities scores from NEO PI-R test. From this information, we want to examine whether the level of neuroticism has an impact on users' perceived trust. Since neurotic individuals tend to experience negative feelings such as anxiety and overall distress [17], neuroticism is usually associated with a decrease in trust. In light of this, we suggest that **H2.2: Neuroticism negatively affects trust perception**.

Additionally, as presented in previous sections, neurotic individuals show higher levels of trust when uncertainty is presented [6]. Moreover, highly neurotic individuals usually feel pressured to answer correctly [2]. Consequently, we believe that neurotic individuals will most likely trust visualizations that present a greater time interval and consequently will allow them to make an informed decision, taking into account past events (**H2.3: Trust perception is higher in visualizations that display a greater time interval for neurotic individuals**). In light of this our second hypothesis suggests that:

H2: There are interaction effects between neuroticism and the visualization granularity in perceived trust.

During our experiment, we also registered task completion times and hover events per data point while the participants interacted with the different visualizations. Therefore, we want to analyse whether distinct granularity levels have an impact on the way users interact with information. Additionally, we want to investigate whether users' individual characteristics such as their level of neuroticism have an impact on these factors. Hence, our third hypothesis is:

H3: Granularity and neuroticism affect user interactions with the visualisations.

Finally, users have two moments to decide on the visualization they trust the most. Therefore, we want to analyse users self-calibrated degree of confidence in a taken decision, i.e. whether performing tasks with each visualization may alter the choice that the participant made at the beginning. Moreover, we want to explore whether neuroticism has an impact on this decision switch. As mentioned before neurotic individuals are more attentive when performing tasks [2], therefore the fact that they will look closely at each granularity might affect their decision. Thus, our final hypothesis is:

H4: Neuroticism will have an impact on users' decision before and after they look at each visualization individually.

4.2 Research Design

This study leverages the variables described in Table 4.1. As mentioned, we decided to focus on one specific personality trait. Neuroticism will be evaluated in the context of the FFM and, in order to perform the statistical tests necessary to evaluate our hypotheses, participants will be divided into three neuroticism groups using the Portuguese norm [62]. This design decision is further explained in Section 4.4.

Table 4.1: Variables used in the study.

Name	Type	Dependency	Description
Neuroticism level	Ordinal	Independent	Individuals level of neuroticism (low/ medium/ high)
Time granularity	Categorical (Hierarchical)	Independent	Defines the time granularity presented in the visualization (day/ week/ month)
Time (1st attempt)	Quantitative (Ratio)	Dependent	Time took by the participants to choose between three visualizations
Time (2nd attempt)	Quantitative (Ratio)	Dependent	Time took by the participants to choose between three visualizations after analysing each of the visualizations individually
Trust	Quantitative (Ratio)	Dependent	Level of trust perceived by participants when analysing a specific visualization
Visualization chosen	Categorical	Dependent	Visualization chosen by the participants that reflects the information they trust the most
Hover per point	Quantitative (Ratio)	Dependent	Hovers events made by participant when analysing each of the different granularities
Task completion time	Quantitative (Ratio)	Dependent	Time taken by participants to choose an hospital in each granularity
Change in Decision	Categorical	Dependent	Defines whether participants changed their decision about the visualization they trust the most

Concerning time granularity, we consider three different granularities: day, week and month. As mentioned previously, these distinct granularities allows us to show shorter or larger time periods. Therefore, our participants will have access to the most recent information (daily granularity), as well as information that provides context given a larger period of time, allowing participants to understand trends during a specific time period (weekly and monthly granularities).

Regarding the idiom that will be used to illustrate the data, line charts were used for all visualizations. As presented in Section 3.4 shared space techniques are considered better for comparison tasks. Additionally, positional idioms were considered the best option to visualize time-oriented data. Finally, we do not want to explore any cyclic aspect of time. Therefore we considered that choosing an idiom with a linear approach and that most people were already familiarized with, would be the best solution for our approach [60].

The data displayed in the visualizations represents the number of patients that were admitted to emergency departments in different hospitals throughout time. As we mentioned before, we chose this type of data to provoke a feeling of emergency in the participants, and consequently trigger the need for trust. Besides that, we also want to choose an environment that most participants would be familiar with, which, in this case, is to wait for an appointment in the emergency department.

Each visualization displays the number of patients in an emergency room throughout a time period in two different hospitals, encoded by two colours randomly associated at the beginning of each experiment. This information was represented in a line chart, where the x-axis illustrated the time, the y-axis represented the number of patients, and each line presented a hospital. As such, we developed a set of visualizations (Figure 4.1) with three different time granularity values: (i) the number of patients per hour in the past day, (ii) the average number of patients per hour in the past week, and (iii) the average number of patients per hour in the past month. In particular, each visualization is created from the same database so that actual data trustworthiness is the same across all conditions. Therefore, the visualizations contain all information the user needs to evaluate the quality of the underlying information.

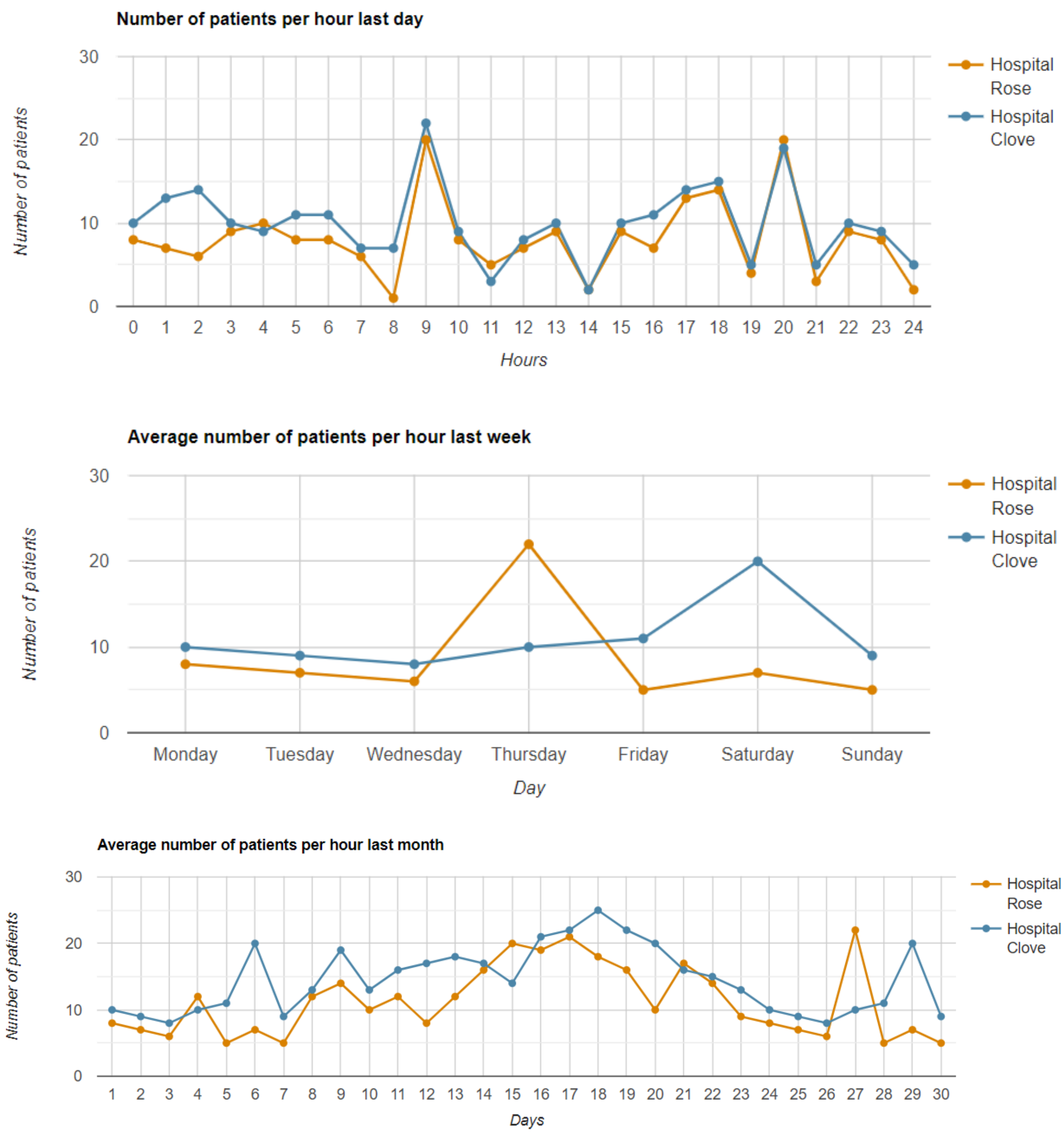


Figure 4.1: Set of visualizations with different time granularity factors.

During our study we evaluated the time participants take to choose a visualization and the number of hovers performed before and after the analysis of the three different visualizations separately. The time measured consists of the sum of the time participants take to choose a hospital and identifying which visualization they trusted the most to make that decision. Whereas, one hover is registered every time a participant hovers a specific time point. On the one hand, Alarcon et al [50] concluded that individual differences are more relevant when participants are exposed to new environments and situations. Therefore, we decided to register time when participants are exposed to three new and unfamiliar visualizations.

On the other hand, neurotic individuals are more attentive when performing tasks about visualizations [2, 3]. We believe that this characteristic may have an impact on the time taken to choose a visualization after exploring each visualization individually. Hence, we registered the time taken by participants to make a decision after they have had the opportunity to form a deeper understanding of the different visualizations. By measuring time at different moments, we are able to explore whether neuroticism has an impact on the time taken to choose a visualization considering the moment of choice (before or after analysing the three visualizations).

Besides analysing if there are changes in the time participants take to make a decision, we also want to analyse whether participants are consistent in their choices of which visualization affected their decision the most, and whether the level of neuroticism has any impact on these choice moments. In particular, we defined the visualization participants trust the most at the beginning of the experiment as the *anchor granularity*. The anchoring effect is characterized as the tendency to focus too heavily on a piece of information when making decisions [63]. In this light, we decided to explore whether the anchor granularity had an impact on participants' decisions when analysing the visualizations individually. Moreover, the fact that we ask participants to repeat the same task at the beginning and end of the experience allows us to understand the self-calibrated degree of confidence in a taken decision [64], i.e. whether performing tasks with each visualization may alter the choice that the participant made at the beginning.

Additionally, we decided to register hover events and the time participants took to chose a hospital when analysing each visualization individually. When considering neuroticism in the context of information visualization, most studies concluded that neurotic individuals show lower completion times and higher accuracy [2, 3]. Therefore, we want to examine whether these trends are verified when participants are exposed to a situation where they need to trust the visualizations.

Finally, participants were asked to evaluate their perceived trust using a five-point Likert scale. To our knowledge there is still no standard procedure to measure perceived trust, therefore to try and improve our comprehension of perceived trust we also used the think aloud method as suggested by Sacha et al. [15].

Table 4.2: Statistical tests used to evaluate the study.

Hypoteses	Tests	Statiscal test
H1: Neuroticism positively affectsthe time taken to choose a visualization.	Is the neuroticism score correlated to the time participants take to make a decision?	Spearman Correlation
	Does the neuroticism level have an impact on the time participants take to make a decision?	One-way ANOVA
	Do participants with different levels of neuroticism have distinct interaction times before and after analysing each visualization individually?	Two-way mixed ANOVA
H2: There are interaction effects between neuroticism and thevisualization granularity in perceived trust.	Are there significant differences between the level of trust perceived by participants when analysing the three different granularities?	One-way repeated measures ANOVA
	Do different levels of neuroticism display different levels of trust for distinct granularities?	Two-way mixed ANOVA
	Does granularity have an impact on users' perceived trust considering the anchor granularity?	Two-way mixed ANOVA
H3: Granularity and Neuroticism affect user's interactions with the visualisations.	Are there significant differences between the number of hover events performed in distinct granularities?	One-way repeated measures ANOVA
	Does granularity have an impact on how many times users hover a point considering the anchor granularity?	Two-way mixed ANOVA
	Are there significant differences between the time participants take to choose a hospital in distinct granularities	One-way repeated measures ANOVA
	Does granularity have an impact on the time participants take to choose a hospital considering the anchor granularity?	Two-way mixed ANOVA
H4: Neuroticism will have an impact on users' decisions before and after they look at each visualization individually	Is the visualization users trust the most consistent throughout the experience?	Chi-Square test for association
	Is the neuroticism level correlated to a change in participants decision about which visualizationthey trust the most?	Chi-Square test for association

4.2.1 Data Analysis

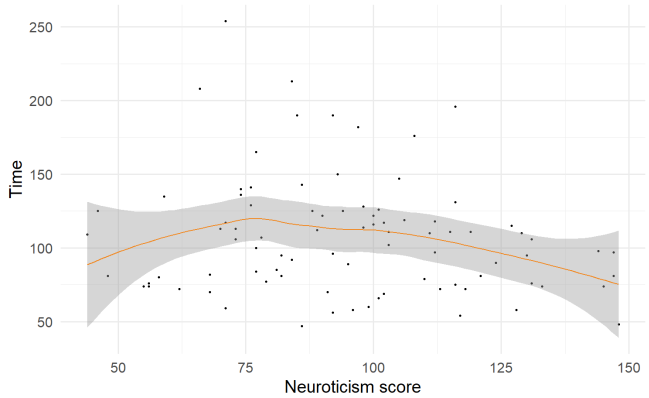
In this subsection, we discuss the different methods used to analyse our data in order to verify our hypotheses (Table 4.2). We will explain each test used and how they were used for each of the hypotheses presented in Section 4.1. All the statistical tests detailed were then evaluated using SPSS Statistics [65].

4.2.1.A Spearman's Correlation

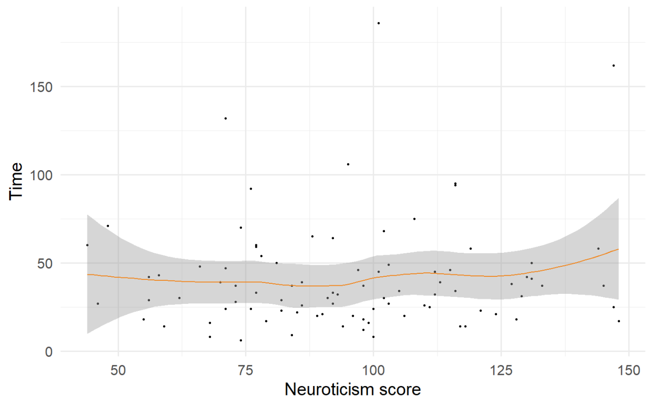
Spearman's correlation test measures the strength and direction of the association/relationship between two continuous or ordinal variables. The two variables considered must represent paired observations, and the relation between the two variables must be monotonic. This test is used to evaluate **H1**.

Using this test we are able to analyse whether there is a correlation between the neuroticism score and the time participants took to make a decision when: (i) analysing the visualizations for the first time, (ii) after analysing all the visualizations individually. Finally, we also explore whether neuroticism scores are correlated with the time difference between the initial and final moment, i.e the difference between the time participants took to make a decision when analysing the visualizations for the first time and the time participants took to make a decision after analysing all the visualizations individually.

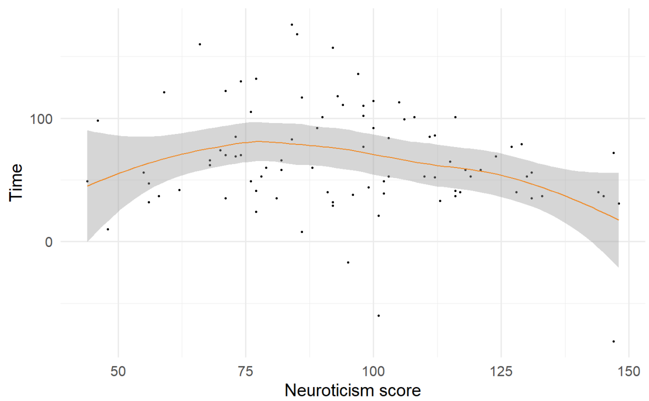
For all the three previously mentioned cases the test performed considers paired observations of two continuous variables: neuroticism scores and time, i.e. for each time measured we have a value for the neuroticism score of that participant. Moreover, we tested whether the relation between the two variables was monotonic, but the plots do not show a clear correlation (Figure 4.2), therefore we ignored this assumption.



(a) Relation between neuroticism and the initial time.



(b) Relation between neuroticism and the final time.



(c) Relation between neuroticism and the time difference.

Figure 4.2: Scatter plots of the relations between neuroticism and users' task completion times.

4.2.1.B One-way ANOVA

This test allows us to determine whether there are any statically significant differences between the means of two or more independent groups. We need to have a dependent and an independent variable, and independence of observations. Furthermore, (i) there should be no significant outliers in the groups of the independent variable, (ii) the dependent variable should be approximately normally distributed for each group of the independent variable, and (iii) there should be homogeneity of variances (i.e., the variance of the dependent variable is equal in each group of the independent variable).

This test was used to analyse **H1: Neuroticism positively affects the time taken to choose a visualization.** With this test, we assess whether there are significant differences between the time taken by participants with different levels of neuroticism. We consider three different times: (i) the time participants take to choose a chart when looking at the visualizations for the first time, (ii) the time participants take to choose a visualization after analysing all the visualizations individually, (iii) the difference between the second and first time. For the three previously mentioned cases the assumption of normality was violated. However, ANOVA is robust to deviations from normality therefore we proceeded with the test [1].

4.2.1.C Two-way Mixed ANOVA

The two-way mixed ANOVA compares the mean differences between groups taking into account two independent variables. This test is used to understand if there is an interaction between two independent variables on a dependent variable. The two-way mixed ANOVA requires that we have one dependent variable, one between-subjects factor that is categorical with two or more categories and one within-subjects factor that is categorical with two or more categories. For this test, we should also verify that there are no outliers, that the dependent variable is approximately normally distributed for each cell of the design and that there is homogeneity of variances. After performing this test, we always followed it by post-hoc Tukey's range tests, which include Bonferroni corrections. For the cases where data was not normally distributed, we ignored this hypothesis given that, as mentioned before, this test is robust to deviations from normality [2].

The two-way mixed ANOVA was used to evaluate **H1: Neuroticism positively affects the time taken to choose a visualization**, **H2: There are interaction effects between neuroticism and the visualization granularity in perceived trust**, and **H3: Granularity and neuroticism affect user interactions with the visualisations**.

Concerning **H1**, we analyse whether participants in distinct neuroticism groups had different interaction times when analysing the visualization before and after analysing each visualization individually.

[1] https://statistics.laerd.com/spss-tutorials/one-way-anova-using-spss-statistics.php
[2] https://statistics.laerd.com/spss-tutorials/mixed-anova-using-spss-statistics.php

There were no significant outliers, but the data was not always normally distributed among groups. However, as mentioned before, we ignored this assumption and ran the test anyway.

In order to investigate **H2** we tested whether different levels of neuroticism display distinctive levels of trust for distinct granularities. Moreover, we investigated whether neurotic individuals show lower levels of trust for all granularities compared to the other levels. Finally, this test allowed us to verify if participants with high levels of neuroticism trust more in the visualization that displays a larger time period.

In regards to **H2**, we explored as well whether the visualization chosen as the anchor granularity (during the first time users interacted with visualizations) had an impact on the users' trust perception. As before, we verify that the data was not always normally distributed.

Finally, concerning **H3** we started by exploring the effect of granularity on users interactions. Therefore we investigated whether the anchor granularity chose by the participants as the granularity they trusted the most had an impact on the users' interactions with different visualizations. Specifically, we investigated whether participants who trusted in distinct visualizations had different behaviours when analysing the visualizations separately. In light of this, we performed a test analysing the impact of granularity in how many times users hovered a point considering the anchor granularity, and another test analysing the impact of granularity in how much time took to choose a hospital considering the anchor granularity as the between factor. For both cases, we had to ignore the assumption that the data was normalized.

Next, we explored the effect of neuroticism on user interactions. Consequently, we performed two different tests analysing the impact of the neuroticism level on the time to choose a hospital and the number of hovers performed per granularity, considering the level of neuroticism as the between factor.

4.2.1.D One-way Repeated Measures ANOVA

The one-way repeated measures ANOVA is used to analyse whether there are statistically significant differences between the means of three or more levels of a within-subjects factor. The previously mentioned levels are related because they refer to the same cases in each level, (e.g each level has the same participants). We need a continuous dependent variable and one within-subjects factor that consists of three or more categorical levels. Moreover, it is required that there are no significant outliers, that the distribution of dependent variables in the different levels is approximately normally distributed and that the variance of differences between all combinations of levels is equal. Similarly, to the two-way mixed ANOVA, for the cases where data was not normally distributed, we ignored this hypothesis given that this test is robust to deviations from normality [3].

This test was used to investigate **H2: There are interaction effects between neuroticism and**

[3] https://statistics.laerd.com/spss-tutorials/one-way-anova-repeated-measures-using-spss-statistics.php

the visualization granularity in perceived trust and **H3: Granularity and Neuroticism affect user interactions with the visualisations**.

When analysing **H2** we focus only on one main effect: whether visualization granularity positively affects trust perception. Our objective was to analyse whether there were significant differences between the level of trust perceived by participants when analysing the three distinct granularities. In this test, we considered the dependent variable as the trust rate given by participants and the within-subject factor as the level of granularity analysed (day, week and month).

Concerning **H3**, in order to analyse the impact of different granularities on user interactions, we analysed two factors: (i) the number of hovers per point, (ii) the time participants took to choose a hospital in each granularity. The number of hovers per point consists of the total number of hover events participants performed in a visualization divided by the number of time points in that visualization, e.g the weekly granularity has five time-points.

Regarding the first factor, we performed this test to investigate whether there were significant differences between the number of hover events performed in distinct granularities. In this test, the number of hover events is the dependent variable, and the within-subject factor is the different granularities. Finally, we used the same approach to analyse the time participants took to choose a hospital in each granularity. With this test, we analysed whether the granularity had an impact on the time participants took to choose a hospital. Considering the hover analysis, we discovered some outliers. However, these appear to be genuinely unusual values. Moreover, for the two previously mentioned cases, we had to discard the assumption of normality.

4.2.1.E Chi-Square Test for Association

This test evaluates whether two categorical variables are associated. This test requires that we have two categorical variables, independence of observations and that all cells have an expected count greater than five. We used this test to evaluate **H4: Neuroticism will have an impact on users' decisions before and after they look at each visualization individually.**

Initially, we analysed if participants were consistent in their choices of which visualization affected their decision the most. Therefore we used this test to explore whether there is a relation between the visualization users trust the most at the beginning of the experiment and the visualization they report trusting the most after analysing all the visualizations individually. As mentioned, to perform this test it is required that all cells have an expected count of greater than five. However, this assumption was not met by our data. One possible solution would be to collapse some of the categories, in this case, collapse the visualizations users trust the most. Since we only have three options users could choose from (day week, or month), we decided it did not make sense to collapse these categories. Therefore, we chose to run this test anyway but keeping in mind that our results may be invalid.

We used the same approach to find out if there was any relation between the level of neuroticism and the fact that users changed or did not changed their minds about the visualization they trusted the most.

4.3 Data Collection

In this section we will describe the participants and how they were selected, the apparatus used during the experiment and finally, the procedure conducted throughout the study.

4.3.1 Participants

Participants were recruited through standard convenience sampling procedures including direct contact and through word of mouth. We conducted the study with 89 participants (38 males, 51 females) between 18 and 69 years old $(M = 27.40; SD = 12.04)$. However, we only collected the data referent to the NEO PI-R test from 88 of these participants (37 males, 51 females).

Considering our sample, we checked for outliers using the interquartile range [66], and considered outliers every record that posed as an outlier for more than 50% of the attributes we were analysing. With this strategy we did not find any outliers. Therefore, we did not discard the results from any of the participants.

4.3.2 Apparatus

We used the NEO PI-R questionnaire [67] to obtain data related to the participants' personalities. As explained in Section 2.1 the NEO PI-R questionnaire is composed of 240 items, with eight items for each facet of the five traits. Thus, there are 48 items to evaluate each trait. For this study, we used the European Portuguese version of the NEO PI-R developed by Lima and Simões [62].

At the beginning of the experiment, all of the participants verbally consented to a consent form (Appendix A) that allows us to collect their data relative to the NEO PI-R results, and the data collected during the experiment. The data collected during the experiment includes hover events, completion time, and audio records of the session. At the start of the experiment, participants had access to a Google Forms (Appendix B) that guided them through all the phases of our study.

During this study, we used line charts to display information. Therefore, we used the previously mentioned questionnaire to access the familiarity participants have with this type of idiom. During the first part of the experiment, we also used the Google Forms to inquire participants whether they were using any eye correction apparatus. This document explained as well the emergency department overcrowding crisis, which served as the theme for the visualizations, and described the data displayed in the visualizations.

We used a chronometer to measure the time users take to choose a visualization. Additionally, we used logs to register the number of hovers performed. For the second part of the experiment, we used Google Forms to access participants' perceived trust for each granularity individually. At the end of the experiment we used the previously mention questionnaire to ask participants about the last time they had gone to a hospital. We asked this question in order to discard possible biases participants might have during the study, regarding previous personal experiences. Finally, we developed the InfoVis used to display the data and used a computer to perform the experiment. Participants were also required to have a computer as the experiment was conducted in a remote setting due to COVID-19 restrictions.

4.3.3 Procedure

Due to COVID-19 restrictions, user tests were conducted through an online video conference platform, forcing the visualizations to be resized to ensure they were displayed in the same physical size, regardless of device resolution. Before the experiment, participants were asked to answer the NEO PI-R, as it is common in social science studies to collect data "in the wild" [17, 50]. The study was divided into three parts. Initially, all participants were asked to fill out a questionnaire in order to rate their familiarity with line charts. Participants rate their level of familiarity using a five-point Likert scale ranging from *"Not familiarized"* to *"Completely familiarized"*. Additionally, we asked that each subject provided demographic factors such as whether they were using any eye correction apparatus. Then, participants were asked to read a document that explained the emergency overcrowding crisis, serving as a theme for the visualizations and prompting participants to be aware of the emergency associated with this crisis while motivating trustworthiness toward this problem. We continued in a three-part test, as represented in Figure 4.3.

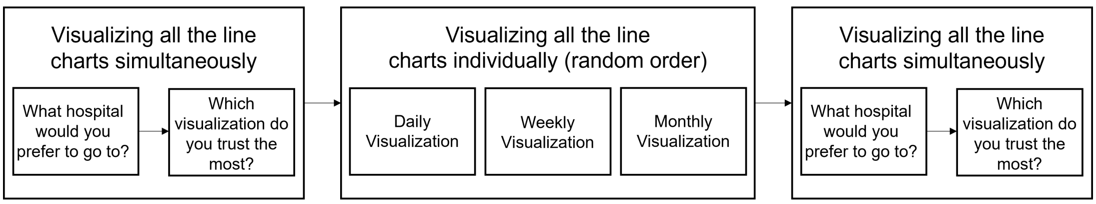


Figure 4.3: Overview of the decision processes in our experiment.

First, participants were asked to put themselves in a position of having a health emergency and needing to decide to which hospital they should go to (Figure 4.3, left). The decision was made with access to the three visualizations. We prompted subjects to decide solely based on the number of patients that visited the emergency department on the past day, week or month, considering that fewer patients would most likely lead to a shorter waiting time. Moreover, the names of the hospitals were made up and the colour encoding of each hospital was randomly assigned in order to reduce any potential biases. Additionally, the number of patients in each of the two hospitals was kept the same between the visualizations, hence prompting the subject to base their decision on the time granularity factor.

After analysing the three visualizations using a think-aloud protocol, participants were asked to choose which hospital they would prefer to go to, and in which visualization they trusted the most to make that decision. We collected both the time participants took to make a decision and registered their anchor granularity.

Next, participants were assigned a random order through which they would interact with each visualization (Figure 4.3, middle). For each visualization, participants were asked a set of three tasks in order to ensure that they acknowledged the different granularity. The tasks consisted of (i) finding the number of patients for a specific point in time, (ii) finding which hospital had the greatest growth of patients in a specific time interval, and (iii) choosing one hospital to visit in case of an emergency. After performing the tasks, users were invited to assess their perceived trust regarding the presented data using a five-point Likert scale ranging from *"I do not trust this information"* to *"I completely trust this information"*. For this part of the experiment, we collected the number of hovers that subjects triggered while interacting with each time point in the chart as well as the time they took to complete each task.

After the participants interacted with the three visualizations individually, each subject was again presented with the three visualizations and was prompted to choose which hospital they would go to and which visualisation weighted the most in their decision, similar to the first part (Figure 4.3, right). Again we collected the time participants took to make a decision and registered the visualization they trusted the most to make that decision. Finally, we asked participants what was the last time they visited a hospital for an emergency (last week, last month, last year, or never). All participants had normal or corrected-to-normal vision. In addition, the last time they visited a hospital had no effect on the dependent variables.

4.4 Clustering Study

In order to analyse our results as mentioned in Section 4.2.1, we performed several statistical tests. For some of these tests, it was necessary that neuroticism was considered as a categorical variable. Therefore, we decided to divide our participants into different levels of neuroticism. Even though the levels of neuroticism are considered an ordinal variable, it is possible to treat this variable as categorical. We can do this because we are able to consider each of the different levels (low, medium and high) as distinct nominal categories. This division was made through the exploration of two different methods that resulted in six divisions as shown in Table 4.3. Firstly, we divided our participants into two and three levels of Neuroticism according to the Portuguese Norm developed by Lima and Simões [62]. Participants were distributed according to the norm as demonstrated in Table 4.4. Moreover, participants

Table 4.3: Different approaches used to divide participants into distinct levels of neuroticism.

Technique used	Participants level	Attributes analysed
Portuguese Norm	2 (low, high)	Neuroticism score
	3 (low, medium, high)	Neuroticism score
K-means	2 (low, high)	Neuroticism score
	3 (low, medium, high)	Neroticism score
	2 (low, high)	Neuroticism facets score
	3 (low, medium, high)	Neuroticism facets score

were divided considering their gender and age group: young adults (younger than 21), and adults (older than 20).

Secondly, we divide participants into clusters using the K-means algorithm. Using this method we followed two different approaches: (i) dividing the participants according to their neuroticism level, (ii) dividing the participants according to their score on the six facets of neuroticism. For each of these approaches, we started the algorithm with 100 cluster centres. We used the elbow method to choose the number of clusters as shown in Figure 4.4. Consequently, we decided to divide the participants into two and three levels of neuroticism.

Finally, after the division was applied we followed the same methodology for all the distinct approaches. Firstly, we analysed whether the distribution was balanced between the different levels of Neuroticism. In particular, taking into account the division made according to the Portuguese distribution's percentiles we expected that the division of participants into three levels would consist in 25% of participants on the lower lever 50% on the medium level and 25% on the higher level. Additionally, we expected that the division of participants into two levels would be composed of 50% of participants with a low level of neuroticism, and 50% with high neuroticism scores. Furthermore, we analyzed whether the different levels were significantly different from each other, i.e if participants with different levels of neuroticism had significantly different scores. In order to verify if the levels were significantly different, we ran a one-way ANOVA.

Table 4.4: Levels of neuroticism.

Neuroticism level	Neuroticism Percentile
low	[0%-25%]
medium	]25%-75%[
high	[75%-100%]

Concerning the first division, which consisted in dividing the participant into three levels using the Portuguese norm, we verified that we have 24 participants with a high level of neuroticism, 35 participants with a medium level of neuroticism and 29 participants with a low level of neuroticism. These

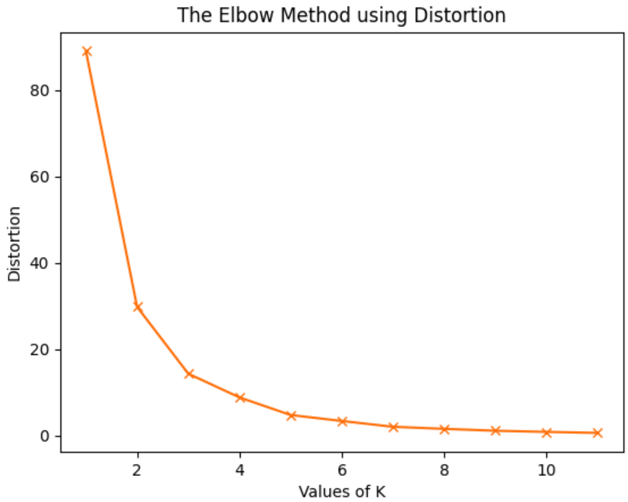


Figure 4.4: The elbow method using distortion.

results allow us to conclude that participants are balanced between levels. Additionally, we verified that participants were significantly different from each other in the distinct facets of neuroticism. The only exception was regarding participants with high and medium levels of neuroticism who were not significantly different when analysing the facet Impulsiveness $(p = 0.351)$. Figure 4.5 shows the scores of participants in different neuroticism levels across the distinct facets of Neuroticism.

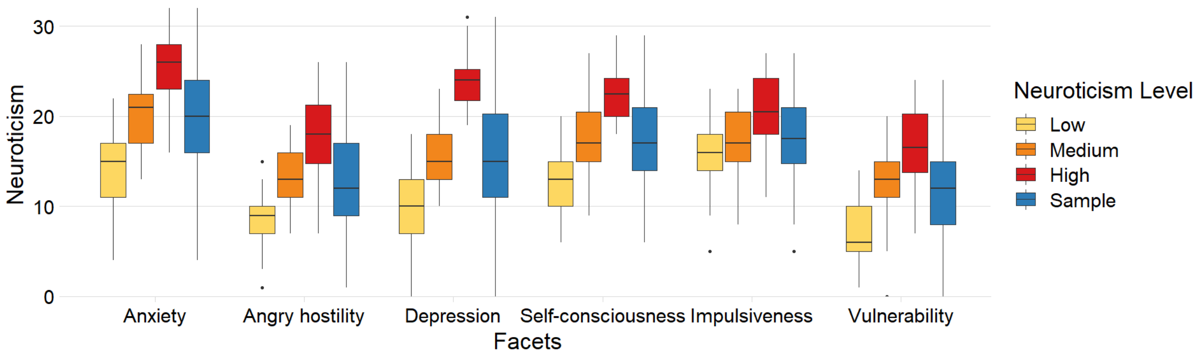


Figure 4.5: Distribution of levels of neuroticism (divided by the portuguese Norm) per facet.

Regarding the second division, that consisted in dividing participants into two levels of neuroticism using the Portuguese norm, we obtained 51 participants with a low level of neuroticism, and 37 participants with a high level of neuroticism. Additionally, we verified that participants were significantly different using a one-way ANOVA (Figue 4.6). Results showed that participants in different levels of neuroticism were different for all facets of neuroticism.

Regarding the second method (K-means), we divided participants into three different neuroticism

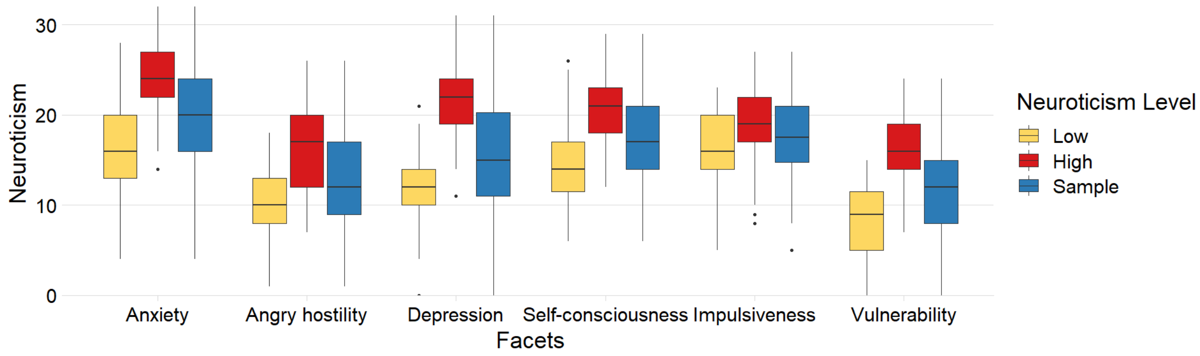


Figure 4.6: Distribution of levels of neuroticism (divided by the portuguese Norm) per facet.

levels taking into account the neuroticism score and concluded that participants were balanced between clusters. We had 30 participants in the first cluster, 36 participants in the second cluster and 22 participants in the third cluster. Furthermore, results showed that overall participants were different for all facets except Impulsiveness. Participants with low and medium levels of neuroticism were not significantly distinct for this facet $p = 0.130$ (Figure 4.7).

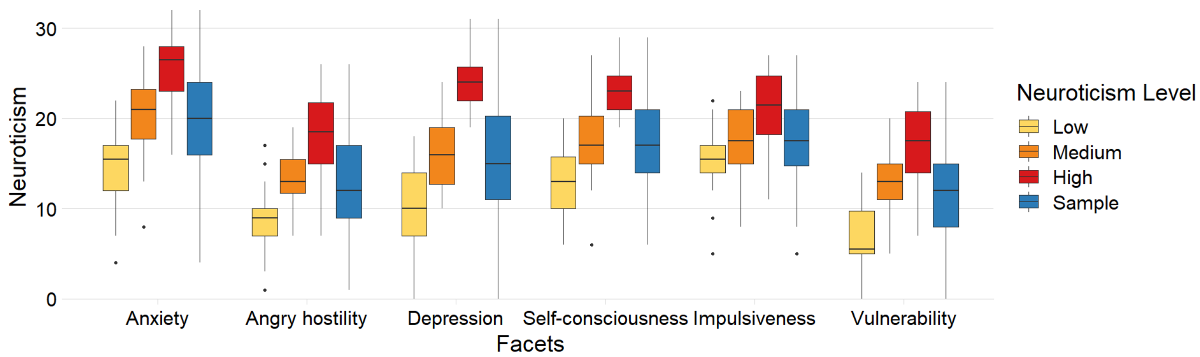


Figure 4.7: Distribution of levels of neuroticism per facet using cluster distribution.

In regards to the second division, which consisted in dividing the participants into two levels of neuroticism taking into account the neuroticism score, we obtained one cluster with 43 participants and another with 45 participants. Furthermore, results from the one-way ANOVA showed that all participants were significantly different. In Figure 4.8 we can observe the scores of participants divided according to this approach in all the different neuroticism facets.

Finally, when applying K-means algorithm to divide the participants considering the 6 facets of neuroticism, we started by dividing them into three levels of neuroticism. We this approach we obtained 45 participants in the first cluster, 24 participants in the second cluster and 19 participants in the third cluster. Concerning the results from the One-way ANOVA participants were overall different. However, when looking into the impulsiveness facet none of the levels was significantly different from each other.

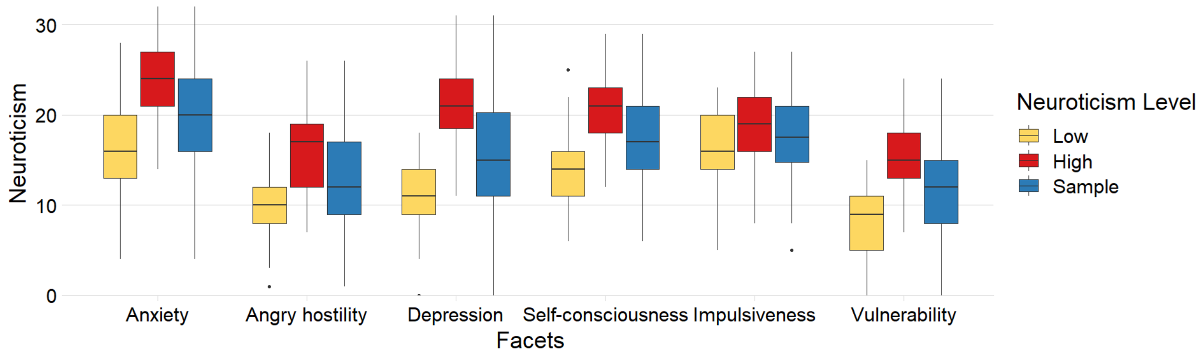


Figure 4.8: Distribution of levels of neuroticism per facet using cluster distribution.

Figure 4.9 shows the score of the participants in the different levels across all facets, allowing us to see that the average scores of impulsiveness are similar across levels.

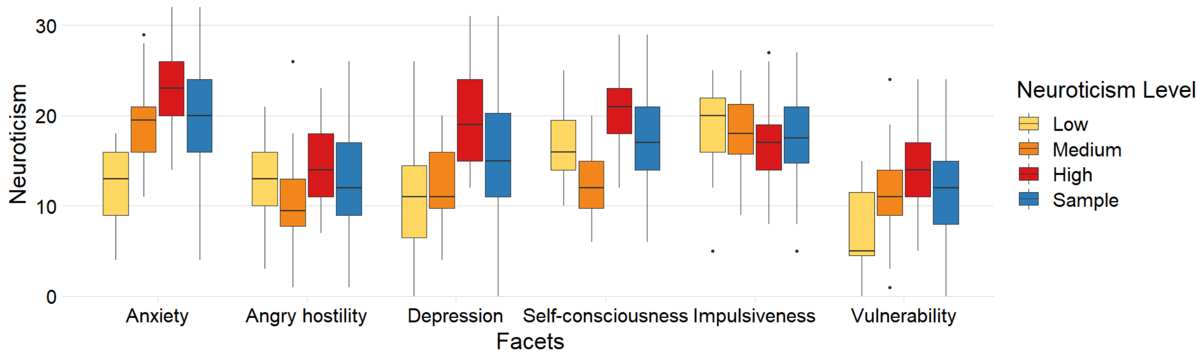


Figure 4.9: Distribution of levels of neuroticism per facet using cluster distribution.

Finally, when dividing participants into two levels, we obtained 70 participants in a cluster and 18 in the other. These numbers led us to conclude that this approach does not provide a balanced solution. In light of this, we started by excluding the option of dividing participants using clusters according to the neuroticism facets, once this option was the one that provided the worst results. Moreover, with this approach, there was no guarantee that participants were significantly different from each other.

In the end, we chose to use the Portuguese norm distribution to divide participants into three levels of neuroticism. This approach comes as an advantage because it divides participants not according to our sample, unlike the cluster approach, but taking into account the average scores of a large sample from the Portuguese population. Once our sample is composed of only Portuguese participants this seemed like the best option. Moreover, we decided to divide the participants into three levels so that we could better understand the differences among different levels of neuroticism. We believed that it is advantageous to have a medium level of neuroticism so that we can explore results beyond that of the extreme scores.

5

Results

Contents

In this chapter we discuss our results. We divided our analysis into four parts. Initially, we present our results considering only the effect of the visualizations' granularity. Next, we follow a similar analysis by considering this time the personality factors from our participants. These two sections present our major findings and answer to the hypothesis presented in Section 4.1. Then, we present results that go beyond our previously formulated hypotheses. Finally, we discuss our results and describe the main insights we obtained. Data are mean $\pm$ standard error, unless otherwise stated.

5.1 The Effect of Granularity

We started by focusing only on the aspect of how different granularities affect perceived trust. This approach allows us to understand the impact of the distinct time granularities on perceived trust of time-based visualizations. In particular, we study trust perception through two different scopes: we present the results related to perceived trust, and the results referent to user interaction.

5.1.1 Trust Perception

We started by running a one-way repeated measures ANOVA to study whether the degree of trust perception was influenced by the granularity when participants analysed the different visualizations separately. Results did not show a statistically significant interaction between granularity and perceived trust, $F(1.636, 142.309) = 2.777$, $p = .076$, partial $\eta^2 = .031$ (Figure 5.1). All distributions look similar with a positive mean rating, yet results suggest that subjects assessed their trust perception with lower grades in the weekly granularity $(4.057 \pm .092)$ compared to the daily $(4.136 \pm .097)$ and the monthly $(4.216 \pm .089)$ granularity values. In particular, a pairwise comparison reports a statistically significant increase of $.159$ $(95\%$ CI, $.002$ to $.316)$ points from the weekly to the monthly granularity, $p = .046$. These results support **H2** that suggest that **there are interaction effects between neuroticism and the visualization granularity in perceived trust**. In particular, **high levels of granularity will have a positive impact on the trust rates given by participants (H2.1).**

Next, we ran a two-way mixed ANOVA to study whether the degree of trust perception per granularity was influenced by the visualization that subjects weighed the most in their first decision: anchor granularity (Figure 5.2). **There was a statistically significant interaction between the visualization chosen and the granularity on trust perception**, $F(3.372, 143.290) = 4.496$, $p = .003$, partial $\eta^2 = .096$. In a pairwise comparison, we can observe that, for subjects that chose the monthly visualization as the anchor granularity, there were statistically significant increases of $.280$ $(95\%$ CI, $.029$ to $.531)$ points from the daily to the monthly granularity, $p = .023$, and of $.340$ $(95\%$ CI, $.144$ to $.536)$ points from the weekly to the monthly granularity, $p < .001$. These results suggest that people who trust more in a monthly granularity $(4.300 \pm .119)$ are more consistent assessing their perceived trust in the remaining granu-

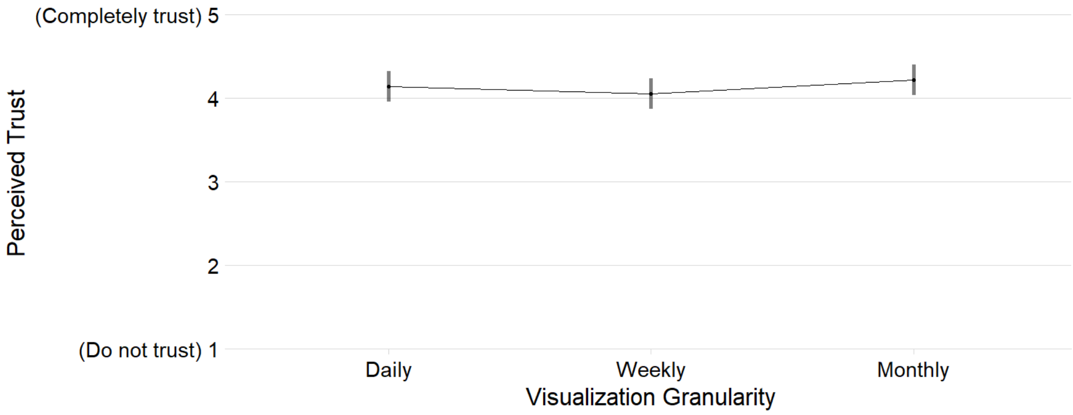


Figure 5.1: Estimated marginal means of trust perception based on the granularity of the assessed visualization.

larity options (Daily: $4.020 \pm .129$; Weekly: $3.960 \pm .122$). Similarly, people who weighted more on the daily granularity in their decision also trusted more in this visualization $(4.417 \pm .263)$ compared to the weekly $(4.250 \pm .250)$ and monthly $(4.000 \pm .242)$ granularity values. In contrast, people who favored the weekly granularity had similar values across their trust perception for the daily $(4.231 \pm .179)$, weekly $(4.154 \pm .170)$, and monthly $(4.154 \pm .165)$ visualizations.

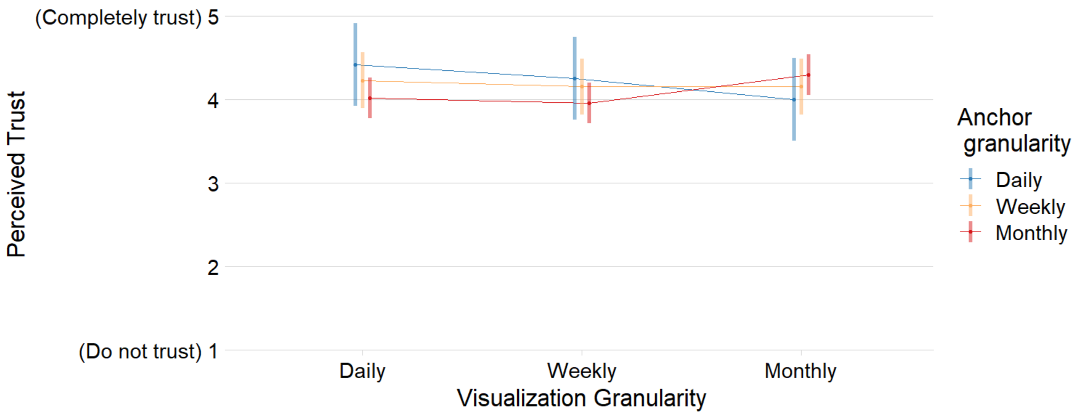


Figure 5.2: Estimated marginal means of trust perception based on the anchor granularity and the granularity of the assessed visualization.

Moreover, we ran a chi-square test of independence between the visualization that weighted the most in both decision moments. There was a statistically significant association between the chosen visualizations, $\chi^2(4) = 59.359, p < .001$. The association was moderately strong [68], Cramer's V $= 0.581$ (Table 5.1). Overall, **these results suggest that participants were consistent in their choices, reflecting a strong self-calibrated degree of confidence in a taken decision**, except for people who initially weighted more on a weekly granularity, who were more likely to change their decision after a careful analysis of the data. Moreover, the monthly granularity was the most relied on in both decision moments. As such, our next step is to find whether the interaction data reflects the lower trust perception of the weekly granularity as well as how the granularity which was most relied on affected how people rated each granularity.

Table 5.1: Crosstabulation of the visualization that weighted the most in each decision moment. Adjusted residuals appear in parentheses next to the observed frequencies.

		Granularity Chosen After			
		Daily	*Weekly*	*Monthly*	Total
Granularity Chosen Before	*Daily*	$8\ (5.5)$	$0\ (-2.0)$	$4\ (-2.3)$	12
	Weekly	$3\ (-0.6)$	$15\ (5.3)$	$8\ (-4.2)$	26
	Monthly	$2\ (-3.3)$	$4\ (-3.6)$	$44\ (5.4)$	51
	Total	13	19	57	89

5.1.2 Interaction Data

As we mentioned, we collected both the time in seconds users took to decide between the two hospitals as well as how many times they hovered a point in the line charts. **The visualization granularity showed statistically significant effects both in the time users took to decide**, $F(2, 174) = 8.401$, $p < .001$, partial $\eta^2 = .088$, **as well as in how many hovers per point subjects did**, $F(1.438, 125.117) = 22.948$, $p < .001$, partial $\eta^2 = .209$. Regarding the time to decide, participants took less time to choose an hospital when analysing the daily granularity ($6.852 \pm .558$), compared to the monthly ($9.557 \pm .933$) and the weekly (11.102 ± 1.119) granularity versions (Figure 5.3). More precisely, a pairwise comparison reported statistically significant increases of 4.250 (95% CI, 1.419 to 7.081) seconds from the daily to the weekly granularity, $p = .001$, and of 2.705 (95% CI, $.441$ to 4.968) seconds from the daily to the monthly granularity, $p = .013$. These results suggest that participants find it significantly easier to make a decision when observing the daily granularity.

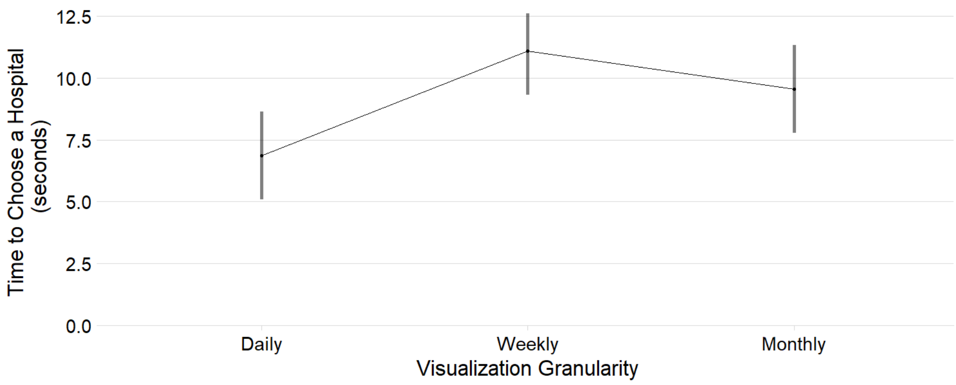


Figure 5.3: Estimated marginal means of time to choose a hospital based on the granularity of the assessed visualization.

For the hovers, **results showed that participants performed more hover events per points when they analysed the weekly granularity** $(0.747 \pm .096)$, followed by the daily $(.307 \pm .038)$ and then the monthly granularity $(0.300 \pm .043)$ (Figure 5.4). In particular, a pairwise comparison reports statistically significant increases of $.427$ (95% CI, $.220$ to $.634$) hovers per point from the daily to the weekly granularity, $p < 0.001$, and of $.434$ (95% CI, $.231$ to $.637$) from the monthly to the weekly granularity, $p < 0.001$. Therefore, users interacted more with the weekly granularity, which was the one they rated with a lower trust perception.

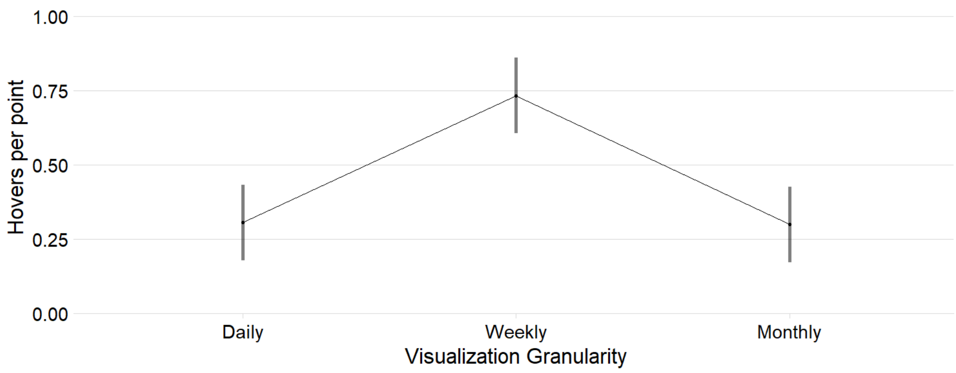


Figure 5.4: Estimated marginal means of hovers per points based on the granularity of the assessed visualization.

These results confirm **H3** suggesting that **granularity and neuroticism affect user interactions with the visualisations**. Specifically, that granularity has an impact on users interactions with the visualizations. In fact, results suggest that participants take more time making decisions when visualizing information with a lower granularity. Moreover, participants interact more with this type of visualizations.

Similarly to the trust perception analysis, we decided to verify whether the anchor granularity had an effect on the time to decide on an hospital and the number of hovers per granularity. **We found that the visualization that subjects relied the most in their decision significantly affected the time that people took to choose an hospital when they interacted with the visualizations one at a time** (Figure 5.5), $F(3.709, 157.626) = 2.679$, $p = .038$, partial $\eta^2 = .059$. In particular, a pairwise comparison showed that subjects that weighted the weekly visualization the most had statistically significant increases of 5.385 (95% CI, 1.256 to 9.514) seconds from the daily to the monthly granularity, $p = .006$. Additionally, subjects that weighted the monthly visualization the most in their decision showed statistically significant increases of 5.360 (95% CI, 1.602 to 9.118) seconds from the daily to the weekly granularity, $p = .002$, and of 3.820 (95% CI, 0.593 to 7.047) seconds from the monthly to the weekly granularity, $p = .015$. These results show that people who weight either the daily (11.000 ± 3.003) or the monthly (12.740 ± 1.471) granularity the most take more time to choose an hospital when analysing the weekly granularity. Nevertheless, only the people who chose the daily granularity actually take less time to choose an hospital when they analyse their chosen visualization (5.917 ± 1.519). This trend is not present for people that chose the other granularity values; people who focused either on the weekly (6.269 ± 1.032) or monthly $(7.380 \pm .744)$ granularity are faster in a visualization with a daily granularity.

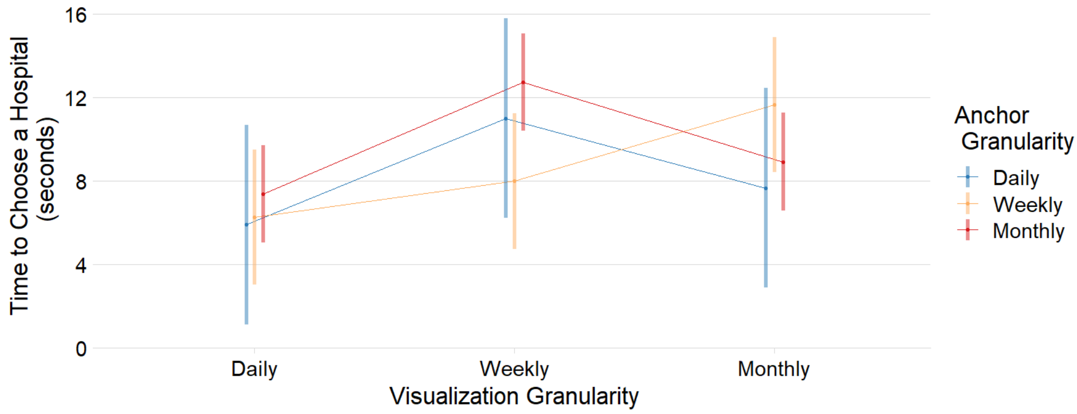


Figure 5.5: Estimated marginal means of time to choose an hospital based on the anchor granularity and the granularity of the assessed visualization.

Finally, **we found that the visualization that subjects relied the most in their decision did not significantly affect the number of hovers per point performed by the participants when interacting with the visualizations individually** (Figure 5.6), $F(2.863, 121.677) = 0.282$, $p = .829$, partial $\eta^2 = .007$. Even though there was not a significant interaction, a pairwise comparison showed that subjects that weighted the weekly visualization the most had statistically significant increases of 0.437 (95% CI, $.051$ to 0.822) hovers per points from the daily to the weekly granularity, $p = .021$. Moreover, subjects that weighted the monthly visualization the most in their decision showed statistically significant increases of 0.439 (95% CI, 0.161 to 0.717) hovers per points from the daily to the weekly granularity, $p = .001$, and of 0.496 (95% CI, 0.224 to 0.767) hovers per points from the monthly to the weekly granularity, $p < .001$.

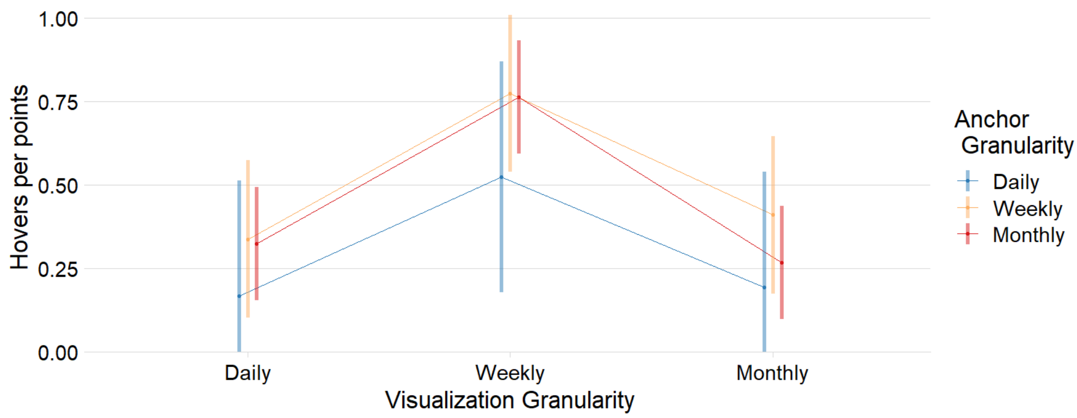


Figure 5.6: Estimated marginal means of hovers per points based on the anchor granularity and the granularity of the assessed visualization.

5.2 The Effect of Neuroticism

Our second approach was to study the impact of neuroticism and its facets on perceived trust when analysing time-series visualizations. In particular, we present our results according to two different perspectives: the impact of neuroticism on users perceived trust and the impact of neuroticism on user interactions with the visualizations.

5.2.1 Trust Perception

We started by running a two-way mixed ANOVA to study whether the degree of trust perception per granularity was influenced by the level of neuroticism of the participants. Results **did not show a statistically significant interaction between granularity and perceived trust considering the neuroticism level** (Figure 5.7), $F(3.381, 138.066) = 2.275$, $p = .078$, partial $\eta^2 = .051$. All distributions look similar with a positive mean rating. More precisely results showed that participants with a high level of neuroticism trusted more on the daily granularity $(4.375 \pm .184)$ followed by monthly granularity $(4.333 \pm .172)$ and the weekly granularity $(4.083 \pm .177)$. Concerning participants with a medium level of neuroticism, results demonstrated that trust levels were lower for the daily granularity $(3.914 \pm .152)$ followed by the weekly granularity $(3.971 \pm .147)$ and higher for the monthly granularity $(4.200 \pm .142)$. Finally, participants with a low level of neuroticism showed similar levels of trust across all different granularities (daily: $4.207 \pm .167$, weekly: $4.138 \pm .161$, monthly: $4.138 \pm .157$).

These results suggest that participants with high levels of neuroticism do not attribute lower trust scores. In reality, for both the daily and monthly granularity these participants are the ones that give higher trust scores. Moreover, a pairwise comparison showed that for participants with high levels of neuroticism there was statistically significant increase of $.292$ (95% CI, $.046$ to $.538$) points from the weekly to the daily granularity, $p = .014$. Overall, these findings do not support **H2** since there is not a clear relationship between neuroticism and perceived trust.

Furthermore, these **results contradict our assumption that participants with high levels of neuroticism would attribute lower scores of perceived trust (H2.2)**. This can be related to the fact that neurotic individuals might want to be complacent towards others and therefore attribute higher scores to the visualizations. Finally, concerning **H2.3: Trust perception is higher in visualizations that display a greater time interval for neurotic individuals** our results show that individuals with high scores of neuroticism give higher scores of trust to the daily and monthly granularities. These **results suggest that participants with high neuroticism trust more on visualizations that present more time points**. Therefore, it is possible to conclude that trust increases with disclosure of information rather than with time granularity.

Additionally, we ran a chi-square test of independence between the neuroticism level and a variable

Figure 5.7: Estimated marginal means of perceived trust based on the level of neuroticism and the granularity of the assessed visualization.

that defined whether or not the participants changed decision from the beginning to the end of the experiment (Table 5.2). There was not statistically significant association $\chi^2(2) = 0.552, p = .759$. In fact, results showed that **independently of level of neuroticism the majority of participants did not change their decision from the beginning to the end of the experiment**. Therefore, these findings do not support **H4: Neuroticism will have an impact on users decision before and after they look at each visualization individually.**

Table 5.2: Crosstabulation between the neuroticism level and the change of decision. Adjusted residuals appear in parentheses next to the observed frequencies.

		Changed Decision		
		Yes	*No*	Total
Neuroticism Level	*Low*	$23(0.5)$	$6(-0.5)$	29
	Medium	$27(0.2)$	$8(-0.2)$	35
	High	$17(-0.7)$	$7(0.7)$	24
	Total	67	21	88

5.2.2 Interaction Data

As mentioned before, during this experiment we registered the time participants took to choose a hospital to go to, and the number of hovers per point when users analysed a specific visualization. Taking into account these results, in this section we present the findings about the impact of neuroticism on user interactions and how this interaction may relate to perceived trust.

Initially, we started by analysing whether there was a relationship between neuroticism and the time participants took to choose a visualization the first time they analysed the three visualizations simultaneously. We ran a Spearman correlation test to analyse this hypothesis. **Results showed that there**

was no statistically significant correlation between neuroticism and the time participants took to make a decision, $r_s(86) = -.150, p = .117$.

We use the same approach to find whether there was a relation between the neuroticism score and the time participants took to make a decision after analysing all the visualizations individually. Again, we did not find any statistically significant relation, $r_s(86) = .032, p = .767$. Finally, we explored whether there was any relationship between the neuroticism score and the time difference between the two moments previously analysed. The results from the Spearman correlation test did not show any statistically significant correlation $r_s(86) = -.164, p = .128$.

Even though the previous experiences concluded that there was not any statistical correlation between neuroticism and the time to make a decision, we decided to explore whether there was any correlation between the participants' level of neuroticism (low, medium and high) and the time participants took. We started by analysing whether the neuroticism level had an impact on the time participants took to choose a hospital and identifying which visualization they trust the most when exposed to the visualizations for the first time. We ran a one-way ANOVA to perform this analysis. There was homogeneity of variances, as assessed by Levene's test for equality of variances $(p = .258)$. **Participants with high neuroticism were the fastest** (92.580 ± 30.780)**, followed by participants with a medium level of neuroticism** (115.70 ± 46.553) **and participants with a low level of neuroticism** (111.490 ± 37.641), but the differences between these groups was not significant, $F(2) = 2.473$, $p = .090$. These results are in line with literature that suggests that participants with high neuroticism complete tasks faster [3].

We followed the same approach to verify whether there was a relationship between neuroticism and the time participants took to choose a visualization after they analyse all the visualizations individually. There was homogeneity of variances, as assessed by Levene's test for equality of variances $(p = .969)$. In this case, the fastest participants were the ones with a medium level of neuroticism (38.660 ± 32.867), followed by low neuroticism (41.620 ± 27.076), and high neuroticism (43.830 ± 32.904). However, the difference between groups was not significant $F(2) = .205$, $p = .815$.

Moreover, we investigated whether neuroticism had an impact on the time difference between the two previously discussed moments, running a one way ANOVA. There was almost homogeneity of variances, as assessed by Levene's test for equality of variances $(p = .045)$. On the one hand, participants that had a higher difference in time were the ones with medium and low neuroticism (low: 73.550 ± 41.593, medium: 72.830 ± 47.548). On the other hand, participants with high neuroticism were the ones that least differed between the two times (48.75 ± 33.372). Results showed that there is no statically significant relation between groups $F(2) = 2.930$, $p = .059$.

Finally, we analysed whether the neuroticism level affected the time participants took to make a decision per moment of decision. During our experience we have two decision moments: (i) when participants analyse the visualization for the first time, (ii) when participants analyse the visualization after exploring each visualization individually. Results show that there was not a statistically significant interaction $F(2.000, 85.000) = 2.930$, $p = .059$, partial $\eta^2 = .064$ (Figure 5.8). **However they suggest that participants with high neuroticism show the smaller difference in time between the two moments.**

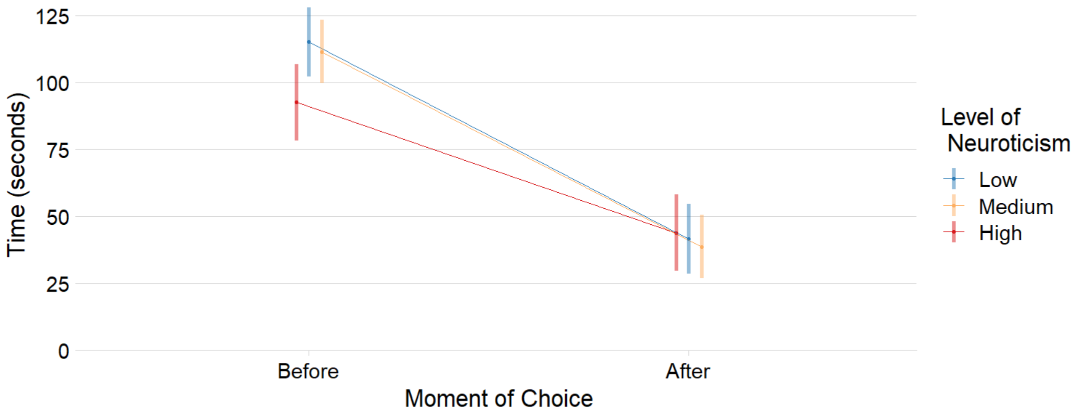


Figure 5.8: Estimated marginal means of time to choose a visualization and a hospital to go to based on the level of neuroticism and the moment of choice.

These results **do not support our assumptions for H1: Neuroticism positively affects the time taken to choose a visualization**. In fact, neurotic individuals are the fastest when analysing new visualizations. However, against our expectation after analysing all the visualizations individually, neurotic individuals are the ones who take more time choosing a visualization. Contrarily to the other participants, neurotic individuals have more difficulties making a decision after a careful analysis of the visualizations. This might mean that neurotic individuals start second guessing their choices after a detailed analysis.

Afterwards we investigated the impact of neuroticism on users' interactions when analysing each visualization individually. Firstly, we performed a two-way mixed ANOVA to analyse whether the level of neuroticism had an impact on the time participants took to choose an hospital per granularity (Figure 5.9). Results did not show a statistically significant relation, $F(4.000, 170.000) = .293$, $p = .882$, partial $\eta^2 = .007$.

Results suggest that overall participants took more time deciding one hospital when analysing the weekly granularity independently of the level of neuroticism. Furthermore, in a pairwise comparison we noticed that, for subjects who score medium on neuroticism, there were statistically significant increases of 3.686 (95% CI, $.076$ to 7.295) seconds from the daily to the monthly granularity, $p = .044$. Additionally, results showed that participants with low neuroticism were the fastest in all granularities (daily: $5.172 \pm .956$, weekly: 10.138 ± 1.963 , monthly: 7.897 ± 1.608) followed by participants with high neuroticism (daily: 7.125 ± 1.051, weekly: 10.500 ± 2.158 , monthly: 8.375 ± 1.768), and participants with a medium level of neuroticism (daily: $8.057 \pm .870$, weekly: 12.314 ± 1.787 , monthly: 11.743 ± 1.464). Finally, it is also important to mention that overall all the participants were faster when analysing the daily granularity.

These results go against most studies that suggest that neurotic individuals are faster when performing tasks. However, as mentioned before it is possible that a deeper analysis of the visualizations confuses participants with high neuroticism values, and causes them to take more time deciding.

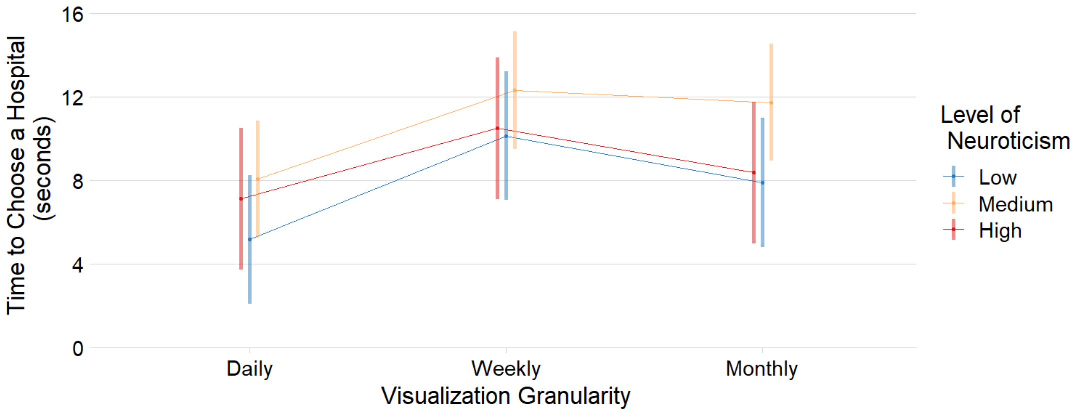


Figure 5.9: Estimated marginal means of time to choose an hospital based on the level of neuroticism and the granularity of the assessed visualization.

We used the same method in order to analyse whether the level of neuroticism had an impact on the number of hovers per points users made per granularity (Figure 5.10). Results did not show a statistically significant interaction, $F(2.885, 122.615) = 1.665$, $p = .160$, partial $\eta^2 = .038$.

Results point to the fact that participants perform more hovers in the weekly granularity independently of the level of neuroticism (high: $.970 \pm .183$, medium: $.727 \pm .152$, low: $.547 \pm .166$). They also suggest that highly neurotic individuals perform the most hovers, followed by participants with a medium level of neuroticism and participants with a low level of neuroticism. Even though overall participants with a high level of neuroticism perform more hovers, there is an exception for the monthly granularity where participants with a medium level of neuroticism perform more hovers.

Additionally, a pairwise comparison showed that for participants with high neuroticism there are statistically significant increases of 0.621 (95% CI, $.225$ to 1.018) hovers per points from the daily to the weekly granularity, $p = .001$, and of 0.736 (95% CI, $.353$ to 1.118) hovers per points from the monthly to the weekly granularity, $p < .001$. The same happened for participants with medium levels of neuroticism who showed significant increases of 0.372 (95% CI, $.043$ to $.700$) hovers per points from the daily to the weekly granularity, $p = .021$, and of 0.336 (95% CI, $.019$ to $.653$) hovers per points from the monthly to the weekly granularity, $p = .034$.

These results show that participants with high levels of neuroticism perform significantly more hovers in the weekly granularity that is also the granularity they trust the least. This might mean that this individuals tend to analyse in more detail information they trust the least.

Overall, looking at **H3** our results we conclude that**there is no significant interaction between the neuroticism level and users' interaction with the visualizations**. Contrarily, users tend to have the same behaviours when analysing the visualizations individually.

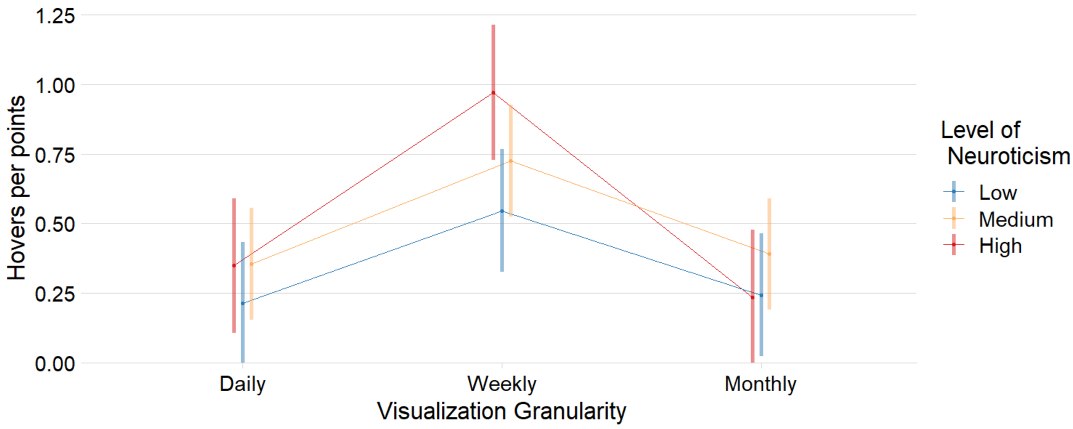


Figure 5.10: Estimated marginal means of number of hovers based on the level of neuroticism and the granularity of the assessed visualization.

5.3 Additional Findings

In order to scrutinize our findings, we decided to analyse whether any external artifacts had an effect on our study. Since we ask participants to choose between two data trends in each time granularity of the line charts, we need to check whether the decision of the subject was affected either by the hour, weekday, or day of the month when the test was conducted.

In particular, we started by analysing whether the hour, weekday, or day had any impact on the anchor granularity. We ran chi-square tests of independence for each of the time dimensions (hour, weekday, day of the month). Results showed that neither the hour $(\chi^2(22) = 15.718, p = .830)$, weekday $(\chi^2(10) = 6.196, p = .799)$, or day of the month $(\chi^2(36) = 33.056, p = .609)$ had an impact on the decision made by the participants.

Afterwards, we ran one-way ANOVAs in order to analyse whether the hour, weekday, or day when participants conducted the study had any impact on the time participants took to choose an hospital and to decide in which visualization they trusted the most, during the first time participants analysed the visualization. Again, results showed that neither the hour $(F(5, 83) = .443, p = .817)$, the weekday $(F(11, 77) = 1.453, p = .167)$, or the day of the month $(F(18, 70) = .680, p = .819)$ had a statically significant impact on the time taken by participants. Finally, we conducted one-way ANOVAs to analyse the impact of the moment participants conducted the study on the trust perception each of the granularity values. Results showed that neither the hour, the weekday, or the day of the month had any impact on trust perception level for each granularity visualization. These results lead us to believe that the time when the subjects conducted our study did not affect their decision making and trust perception.

Regarding visual acuity, all participants had normal or corrected-to-normal vision. In addition, the last time they visited an hospital had no effect on the dependent variables.

Furthermore, our previous results, contrarily to what we expected, did not show any significant relation between neuroticism and users' perceived trust. Therefore, we decided to further explore this relationship through the evaluation of the impact of different FFM facets on perceived trust.

5.3.1 Neuroticism Facets

Our first approach was to analyse in more detail the six different neuroticism facets. Table 5.3 displays the results from the tests performed to assess the impact of the six different facets of neuroticism on users perceived trust.

Table 5.3: Effect of neuroticism facets on perceived trust

Facet	df	dferror	F	p
Anxiety	3.283	139.510	.371	.792
Angry Hostility	3.256	138.377	2.645	.045
Depression	3.218	136.760	1.142	.336
Self-Consciousness	3.250	138.126	.326	.860
Impulsiveness	3.273	139.105	1.283	.282
Vulnerability	3.252	138.199	.673	.581

We started by running a two-way mixed ANOVA to study the impact of anxiety on users perceived trust across different granularities. Results did not show a statistically significant interaction, $F(3.283, 139.510) = .371$, $p = .792$, partial $\eta^2 = .009$. However, looking more closely at the results, it is possible to see that **there is a clear difference between the values of trust obtained by participants with high anxiety compared to participants with medium and, low anxiety** (Figure 5.11). In fact, looking into a pairwise comparison results showed that in the daily granularity there were statistically significant increases of $.625$ (95% CI, $.078$ to 1.172) points from participants with medium anxiety to participants with high anxiety, $p = .019$. Moreover, in the the weekly granularity there were significant increases of $.556$ (95% CI, $.038$ to 1.175) points from participants with medium anxiety to participants with high anxiety, $p = .031$. In fact participants with high anxiety had the highest score of trust for all the different granularities (daily: $4.500 \pm .161$, weekly: $4.400 \pm .152$, monthly: $4.533 \pm .149$). We believe these results might be justified by the fact that individuals with high anxiety might feel the need to be complacent and therefore attribute higher scores of trust to validate our visualizations.

Afterwards, we ran a two-way mixed ANOVA to study the impact of hostility on perceived trust across granularities. Results showed a statiscally significant interaction, $F(3.256, 138.377) = 2.645$, $p = .047$, partial $\eta^2 = .059$ (Figure 5.12). Participants with high levels of angry hostility have tendency to feel frustrated and angry. On the other hand, participants with low levels will almost never get angry or

Figure 5.11: Estimated marginal means of preceived trust based the level of anxiety and the granularity of the assessed visualization.

fight with others. Results showed that participants with medium scores in this facet had the highest levels of trust for all granularities (daily: $4.343 \pm .152$, weekly: $4.143 \pm .146$, monthly: $4.343 \pm .142$). Participants with low levels of angry hostility, on the other hand, attributed similar levels of trust to all the distinct granularities. However, contrarily to the majority of the participants these participants trusted more on the weekly granularity (daily: $4.094 \pm .159$, weekly: $4.125 \pm .152$, monthly: $4.094 \pm .149$). Finally, participants with high scores show the lower trust scores for the daily and weekly granularity (daily: $3.857 \pm .197$, weekly: $3.810 \pm .188$), in contrast looking into a pairwise comparison there are significant increases of $.429$ (95% CI, $.117$ to $.740$) points from the weekly granularity to the monthly granularity, $p = .004$.

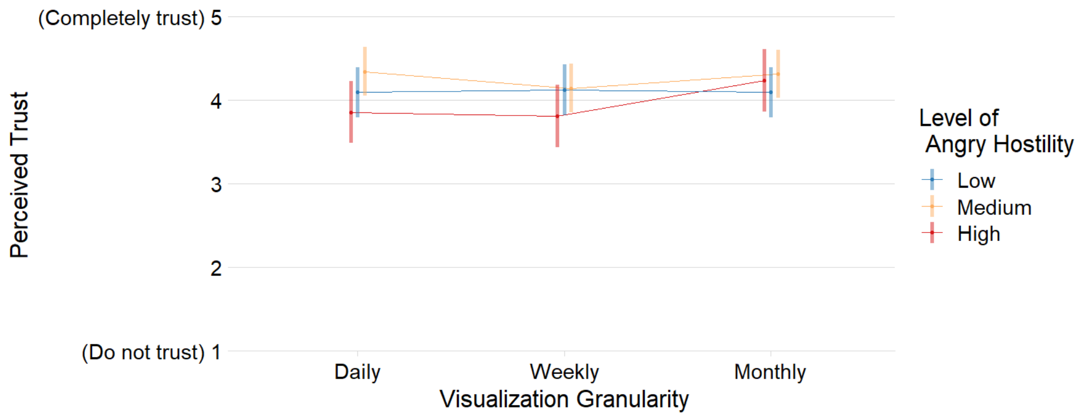


Figure 5.12: Estimated marginal means of perceived trust based the level of angry hostility and the granularity of the assessed visualization.

Finally, we discovered that **anxiety has an impact on the number of times participants hovered a visualization**. We investigated this affect by running a two-way mixed ANOVA. Results showed a statically significant relation, $F(2.942, 125.033) = 3.492$, $p = .009$, partial $\eta^2 = .076$, between participants' anxiety level and the number of hovers per point in each granularity (Figure 5.13). A pairwise comparison, showed that for participants with high anxiety there are statistically significant increases of $.677$ (95% CI, $.334$ to 1.020) hovers per point from the daily to the weekly granularity, $p < 0.001$, and of $.705$ (95% CI, $.368$ to 1.043) hovers per point from the monthly to the weekly granularity, $p < 0.001$. The same trend is repeated for participants with low anxiety levels where there significant increases of $.505$ (95% CI, $.137$ to $.873$) hovers per point from the daily to the weekly granularity, $p = 0.004$, and of $.436$ (95% CI, $.074$ to $.799$) hovers per point from the monthly to the weekly granularity, $p = 0.013$. Moreover participants with high anxiety have the highest number of hovers for the weekly granularity $(.981 \pm .161)$. This trend is not repeated for participants with medium anxiety levels who perform approximately the same number of hover for all granularities. These results suggest that participants on extreme levels of anxiety find it more difficult to make decisions. Moreover, **for participants with high anxiety the difficulty in analysing the weekly granularity becomes even more relevant**.

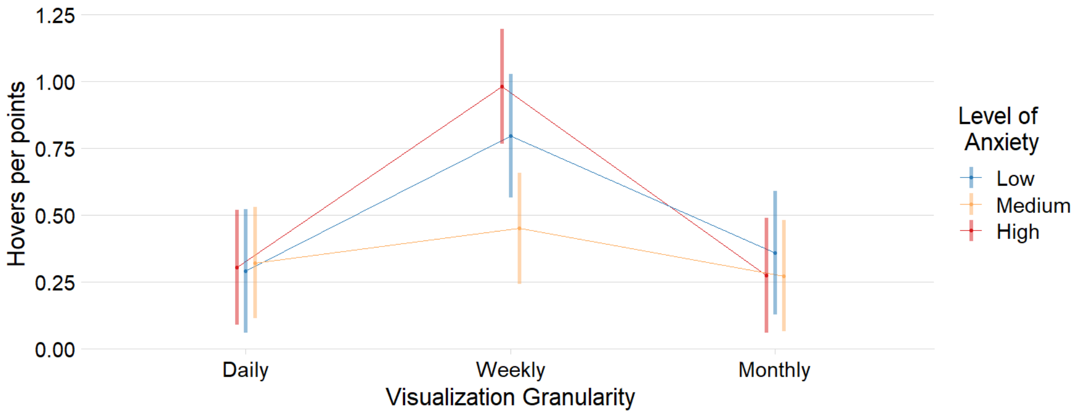


Figure 5.13: Estimated marginal means of hovers per points based on the level of anxiety and the granularity of the assessed visualization.

5.3.2 Other FFM traits their Facets

Our second approach was to analyse facets from the FFM that might have an impact on perceived trust. In light of this we analysed the impact of these facets on users perceived trust, as well as their impact on users interaction behaviours.

Our first conclusion was that **there is a significant relation between the participants' level of Compliance and perceived trust** (Figue 5.14). Individuals with high scores of compliance are characterized by easily accepting others opinions without questioning. We investigated this facet because we believe participants with high scores of compliance will more likely accept the data as truthful and trust the information provided. We conducted a two-way mixed ANOVA to investigate this relationship, results show a significant relation , $F(3.334, 141.687) = 2.880$, $p = .033$, partial $\eta^2 = .063$. Results show that participants with high level of compliance have higher perceived trust scores for both the daily and weekly granularity $(daily : 4.179 \pm .174, weekly : 4.143 \pm .165)$, whereas participants with low compliance present the lowest scores $(daily : 4.000 \pm .238, weekly : 4.000 \pm .225)$. This trend is reverted for the monthly granularity where participants with low scores of compliance present the highest levels of perceived trust $(4.533 \pm .215)$. In fact, pairwise comparison show that there are significant increases of $.533$ (95% CI, $.066$ to 1.001) points from the daily to the monthly granularity, $p = 0.020$, as well as increases of $.533$ (95% CI, $.164$ to $.903$) points from the weekly to the monthly granularity, $p = 0.002$.

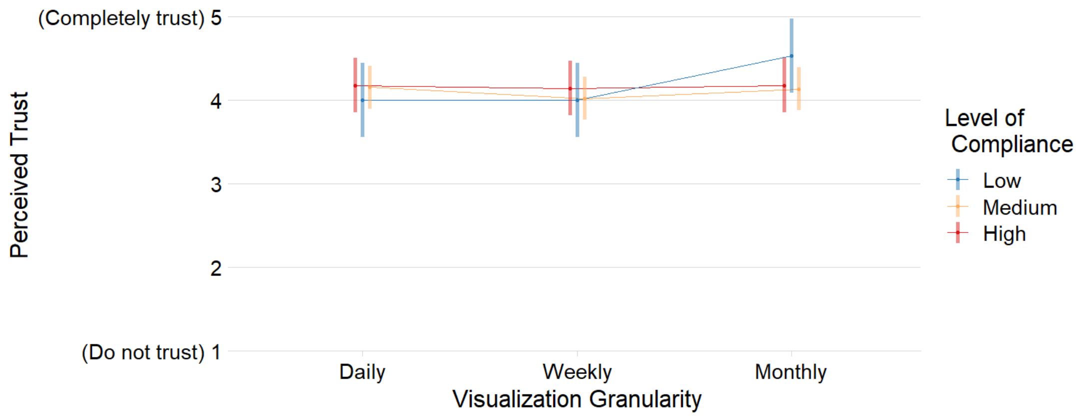


Figure 5.14: Estimated marginal means of perceived trust based the level of compliance and the granularity of the assessed visualization.

These results suggest that **participants with low scores of compliance present higher scores of perceived trust when analysing the visualization they trust the most** (overall the monthly visualization was the one participants most trusted). Participants with high scores of compliance present similar trust levels for all the distinct granularities. These results come in line with the fact that participants with high scores of compliance easily accept others' opinions, contrarily to individuals with low scores of compliance who do not have a problem in showing a different opinion [67]. Moreover our research showed that some facets had an impact on the time participants took to choose an hospital and identifying the visualizations they trust the most, before and after analysing each visualization individually.

Initially, we analysed the impact of the trust facet on the time participants took to choose a hospital and identifying which visualization they trusted the most considering the moment of the experiment (before or after analysing the visualizations individually). We ran a two-way mixed ANOVA and results showed that there was statistical significant interaction $F(2.000, 85.000) = 3.387$, $p = .038$, partial $\eta^2 = .074$ (Figure 5.15). As expected, the participants took a shorter amount of time when answering to the questions after analysing all the visualizations separately. However, **even though participants with low trust are the fastest during the first part of the experiment** (91.462 ± 10.987)**, they are**

the ones who take the longest during the last part of the experiment (51.231 ± 8.558)**, similarly to participants with high neuroticism**. Equivalently, we believe that a closer analysis of the visualizations make participants with low trust doubt their initial assumptions.

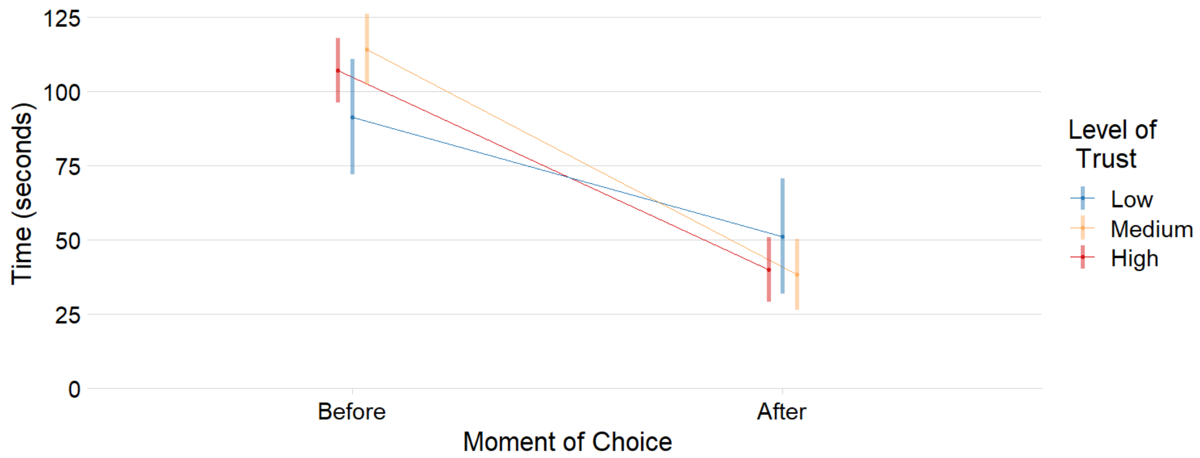


Figure 5.15: Estimated marginal means of time to choose an hospital based on the level of trust and the moment of choice.

Additionally, we also found that **deliberation had an impact on the visualization chosen by the participants as the one they trust the most**, $\chi^2(4) = 10.685$, $p < .030$. The association was moderately strong (Cohen, 1988), Cramer's $V = .246$. Results suggest that **contrarily to the majority of participants, users with low scores of deliberation trusted the most in the weekly granularity** (Table 5.4). Individuals with high scores of deliberation are characterized by the tendency to think carefully and plan things out before acting or make decisions [67]. Therefore, we believe that individuals who have low scores of deliberation prefer visualization that have less information and allow them to make quicker decisions.

Table 5.4: Crosstabulation of the anchor granularity and the level of deliberation of the participants. Adjusted residuals appear in parentheses next to the observed frequencies.

		Granularity Chosen Before			
		Daily	*Weekly*	*Monthly*	Total
Level of Deliberation	*Low*	$2.6\ (-)$	$11\ (3.1)$	$7\ (-2.0)$	19
	Medium	$8\ (1.5)$	$8\ (-1.9)$	$25\ (0.7)$	41
	High	$3.8\ (-)$	$7\ (-.6)$	$18\ (1.0)$	28
	Total	12	26	50	89

Afterwards, we analysed whether there were other factors from the FFM that had an effect on users' trust perception. From these results we **were able to find interesting relationships between extraversion and participants trust level**. Extraversion reflects the quantity and intensity of interpersonal relationships, and their environment, activity level and need for stimulation.

We started by running a two-way mixed ANOVA to understand the effect of the extraversion level on participants perceived trust. Results did not show a significant relation, $F(3.268, 138.911) = .818$, $p = .495$, partial $\eta^2 = .019$. However, taking a look at Figure 5.16 it is possible to see that distinct levels of extraversion have distinct levels of trust. In fact, looking into a pairwise comparison in daily granularities there significant increases of $.741$ (95% CI, $.064$ to 1.419) points from participants with medium extraversion to participants with low extraversion, $p = 0.027$. Moreover, in the weekly granularity there are significant increases of $.723$ (95% CI, $.086$ to 1.360) points from participants with medium extraversion to participants with low extraversion, $p = 0.020$. We believe this might be related to the fact that participants with low extraversion have a tendency to be shy [67, 69], and might not want to give their honest negative opinions out loud.

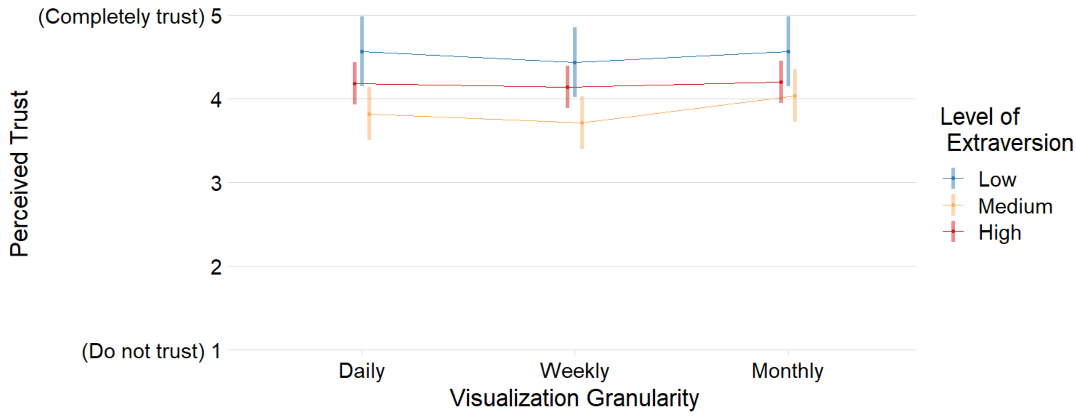


Figure 5.16: Estimated marginal means of perceived trust based on the level of extraversion and the granularity of the assessed visualization.

After, we perform a two-way mixed ANOVA to understand the impact of extraversion on time users take to make a decision. Results did not show a statistical significant interaction, $F(3.714, 157.842) = 2.316$, $p = .064$, partial $\eta^2 = .052$. Nevertheless, looking at Figure 5.17 it is possible to conclude that participants with low extraversion present a different behaviour than the majority of our participants.

In fact looking into a pairwise comparison it is possible to see that for the daily granularity there are significant increases of 3.665 (95% CI, $.017$ to 7.313) seconds from participants with low extraversion to participants with high extraversion, $p = 0.049$. These results point to the fact, that contrarily to the majority of the participants, participants with low extraversion take less time analysing visualizations that show a wider range of time. Opposed to other participants, **the time participants with low extraversion take to choose a hospital seems to be related with the time interval displayed, rather than to the number of time points presented.**

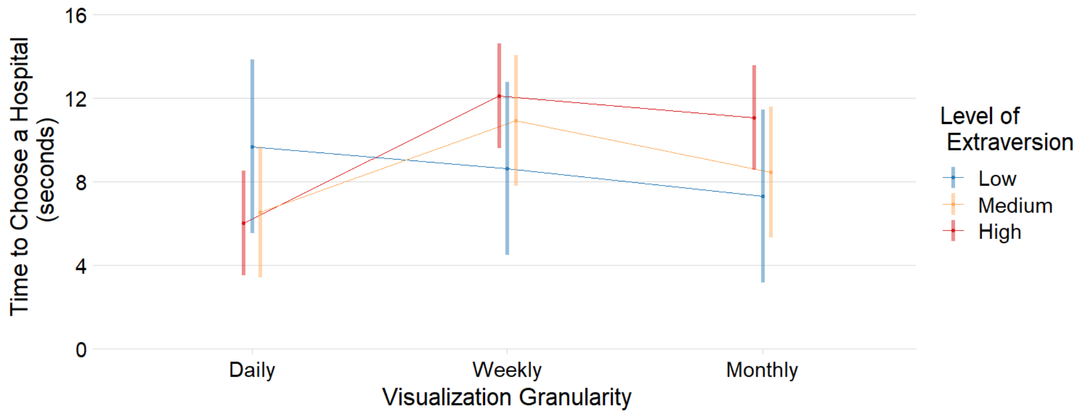


Figure 5.17: Estimated marginal means of time to choose an hospital based on the level of extraversion and the granularity of the assessed visualization

5.4 Discussion

Our results shed a new light on the understanding of the effects of time granularity on trust perception in the context of InfoVis.

5.4.1 Answering the Research Question

Even though there was not a significant relationship between granularity and the trust perceived by the participants, they were more likely to attribute a lower score of perceived trust to the weekly granularity. These results were corroborated by the larger number of hover events per point and more time to decide between hospitals while interacting with the weekly granularity visualization. In particular, we noticed that, while analysing the weekly granularity, people paid more attention to the peaks than the evolution of the data. In opposition to this trend, results showed that overall participants made a decision faster when using the daily granularity. Moreover, results showed that participants' interactions were affected by the anchor granularity. In particular, results showed that overall the visualization participants trusted the most when analysing each visualization individually had the same granularity as their respective anchor granularity. These findings suggest an anchoring effect [63] caused by the participants initial choice of which visualization they trusted the most.

Contrarily to what we expected, neuroticism did not seem to have an impact on users' perceived trust. Actually, independently of the neuroticism level, the number of hovers registered and the time taken to choose a hospital was higher when participants visualized the weekly granularity. Additionally, participants with high neuroticism scores were the ones that show higher mean scores of trust, which contradicts our assumption that neurotic individuals would find it difficult to trust our visualizations due to feels of anxiety and distress. In fact, neurotic individuals were the ones that attributed higher scores of trust, and were also the fastest in choosing the visualization they trust the most and which hospital to go, when analysing the visualizations for the first time. These results come in line with literature that suggests that these individuals are faster under pressure [3]. Moreover, neurotic individuals might feel pressure to answer to the tasks correctly [2], and might not be comfortable in giving a negative evaluation.

Moreover, we verified that highly neurotic individuals were the ones that had the highest difference of time to chose a hospital and identifying the visualization they trust the most, between the beginning and end of the experiment. We believe that neurotic individuals were the ones that least benefited from a deeper analysis of each of the visualizations individually. In fact, when analysing the visualizations individually, neurotic individuals were no longer the fastest. These results point to the fact that neurotic individuals quickly adapt to new interfaces and have better performance results when analysing visualizations for the first time [2]. However, a deeper analysis might make them question their previous answers.

Finally, we analysed the self-calibrated degree of confidence in a taken decision by firstly asking the participants to state which granularity they would find more reliable to decide between hospitals. Overall, the participants, independently of their neuroticism level, were coherent and chose the same visualization at the beginning and end of the experiment. The monthly granularity was the most chosen level when subjects were asked which visualization granularity was more relevant to decide between the two hospitals. This finding is in line with Xiong et al. [5], since participants relied more on a visualization that showed a larger amount of information. Additionally, participants who initially trusted the most in the monthly granularity, also perceived it with higher trust compared to the remaining granularity options. In contrast, participants who initially weighed the weekly visualization the most were the ones more likely to change their decision about which granularity weighted the most on their decision after a careful analysis of the data.

In conclusion, our results show that disclosure of information positively affects trust perception. Contrarily to what we expected, the number of points presented in a visualization has a higher impact on users perceived trust than the time interval presented. As mentioned by Xiong et al. [5] users prefer visualisations that disclose more information, in this case, the daily and monthly visualizations. Moreover, literature suggests that there is a negative impact between neuroticism and trust [17, 20, 21] in human interactions. However, our results did not show a statistically significant interaction between neuroticism and trust perception.

5.4.2 Design Implications

In this light, we were able to recognise some patterns related to the analysis of time-series visualizations. **Overall participants seem to trust more in visualizations which disclosure more information.**

In particular, the monthly granularity, which displayed the most information (30 points), was the one participants trusted the most. This finding was exacerbated through the think-aloud protocol, since participants mentioned that they felt more confident predicting future events when they were able to analyse data from a larger period of time. As mentioned before these results come in line with Xiong et al. [5].

Additionally, we were also able to understand that **participants appeared to attribute more relevance to the maximum values when compared to the overall variation of the data when less information was displayed** such as in the weekly granularity (7 points). Regarding interaction data, we were able to notice that **people interact more with a visualization that they trust the least.** This relationship may be leveraged to adapt the granularity version of visualization when the system detects a large amount of interaction data, as it may indicate that users are not trusting in what they are seeing.

Concerning personality factors, neurotic individuals show lower completion times when exposed to the visualizations for the first time. However, it might not be helpful to show more detailed information to neurotic individuals in emergency situations. As mention by [6], neurotic individuals appear to be the ones that would perform better when conducting predictive decision-making tasks under highly critical situations once these individuals quickly adapt to new environments. Contrarily, individuals with medium and low levels of neuroticism show great benefits from looking at each visualization individually before having to make a decision.

Moreover, the analysis of other traits and facets from FFM allowed us recognize other important interaction effects that can be considered in the design of new visualizations. Firstly, results concluded that participants with low extraversion scores showed a decrease in the time to choose an hospital with the increase of the time granularity displayed in the visualization. Therefore, we believe that for these participants the time interval presented is more relevant the number of time-points displayed in a visualization. Consequently, visualizations with greater time intervals may help these participants in their decision making process.

Secondly, results showed that participants with high anxiety levels performed considerably more hovers, particularly when visualizing the weekly granularity. This might be related to the fact that participant with high anxiety have a tendency to feel apprehensive [62] and therefore prefer to explore the data carefully before making any decision. In this light, the use of interaction seems to be relevant and should be take into consideration in the decision making process of individuals with high anxiety.

5.4.3 Limitations

There are some important factors that may explain the lack of significance observed in the effect of granularity in perceived trust. The assessment of perceived trust through a Likert scale may have confused participants, as they have assessed their perceived trust quite similarly in each granularity. Additionally, we tried to invoke a feeling of risk so that participants felt the need to trust the information presented. However, it would not be ethical to submit participants to any form of real risk. Therefore, the fact that we relied on participants previous experience to imagine the purposed scenario might not have been enough to trigger trust variations.

Moreover, results showed longer completion times and an increase in hover events performed when participants analysed the weekly granularity. Taking a closer look at the weekly line chart (Figure 4.1), the low amount of data points may have led participants to have an exacerbated perception of the peaks. Notice that the monthly granularity (Figure 4.1) also has the same peaks in the last seven days, yet the amount of data points reduces the area they cover, hence their reduced impact. As such, this design may have led to additional user interaction when participants explored this visualization, so future experiences should use randomly generated datasets. Different data encodings could also be considered.

Furthermore, results showed that participants trusted more on visualizations where a larger number of time points were represented, and contrarily to what was expected there was not a relationship between trust and time granularity. Consequently, in order to analyse the impact of time granularity we believe that all the three different granularities could have had the same number of time points. Therefore, eliminating the bias associated with the quantity of information. In addition different time granularities could also be examined.

6

Conclusions

Contents

The purpose of this work was to analyse whether time granularity and level of neuroticism had an impact on trust perception in the context of time-oriented linear charts concerning healthcare information. Consequently, we started by developing a collection of three visualizations that displayed the number of patients in two different hospitals using three distinct time granularities (day, week and month). Using this approach we wanted to explore whether the visualization of distinct time intervals, that enabled the participants to see more or less information throughout time, had an effect on the way they trusted and perceived the information displayed. Additionally, we collected personality characteristics using the NEO PI-R questionnaire. Using this data, we divided participants into three levels of neuroticism and analysed the impact of the level of neuroticism on perceived trust. During the experiment, we asked participants to choose which visualization they trusted the most to choose a hospital to go to, and also to evaluate each of the visualizations in terms of their perceived trust.

Results showed that granularity is a relevant feature to explore in order to increase participants perceived trust in time-series visualizations. In particular, results showed that participants trust more in visualizations where more information is presented. However, the relation between trust and granularity is related to the disclosure of information rather than the time interval presented, i.e participants trust more in visualizations that have more time points (day, month). Moreover, we found that participants interact more with the visualizations they trust the least. This finding gives us a crucial insight about participants' interaction patterns in emergency situations. Furthermore, these results can help find potential problems in the visualization design that lead to a decrease in perceived trust.

Regarding the effect of neuroticism on perceived trust, we did not find any statistically significant relation. Nevertheless, looking into the neuroticism facets we discovered interesting conclusions concerning anxiety and angry hostility. First, participants with high anxiety showed higher levels of trust. We believe these results are related to the fact that participants with high anxiety may feel the need to be complacent, and might be intimidated to attribute lower scores of trust. Therefore, these conclusions might motivate the investigation of new and improved methodologies to evaluate trust. Moreover, these participants interacted significantly more than others with the weekly granularity. Second, we were also able to find a significant relationship between participants' level of angry hostility and perceived trust.

In light of this, we believe that the way visualizations are designed has an impact on users perceived trust and can affect the users' decision making process. Moreover, even though there was not a statistically significant relation between neuroticism and perceived trust, the analysis of other facets and factors from the FFM suggest that there are benefits from creating visualizations that adapt to participants' characteristics such as personality.

6.1 Future Work

Our work allowed us to understand that the visualizations' design has an impact on users perceived trust. Future work may leverage from the usage of temporal aggregation in the visualization design. An example of this is the use of moving average in the finance domain, where the moving average sums up the data points of a financial security over a specific time period and divided the total by the number of data points to obtain the average. In this case, the moving average is continually recalculated based on the latest price in data[1]. This approach allows data to be smoother and less unpredictable [70]. Consequently, we could use this approach to study whether the usage of a different visualization approach in the different granularity visualizations affects the way users see and perceive information.

Moreover, we found some limitations concerning the way we evaluated trust. Results showed that participants were confused when asked to evaluate their levels of perceived trust. Therefore, we believe that we could further analyse users perceived trust by evaluating the different dimensions of trustworthiness, such as accuracy, coverage, up-to-dateness and objectivity. Additionally, we would like to explore how the design of distinct visualization impacts users' perceived trust regarding other thematics besides healthcare and how trust is influenced by other factors of the FFM.

Finally, literature suggests that uncertainty plays an important role in trust building [6, 15, 16], since trust increases when users are aware of the presence of uncertainty in data [15]. Consequently, the inclusion of this element in visualizations should also be explored when designing visualizations to increase users' perceived trust in information.

[1] https://corporatefinanceinstitute.com/resources/knowledge/other/moving-average/

Bibliography

[1] M. Böckle, K. Yeboah-Antwi, and I. Kouris, "Can you trust the black box? the effect of personality traits on trust in ai-enabled user interfaces," in *International Conference on Human-Computer Interaction*. Springer, 2021, pp. 3–20.

[2] C. Ziemkiewicz, A. Ottley, R. J. Crouser, A. R. Yauilla, S. L. Su, W. Ribarsky, and R. Chang, "How visualization layout relates to locus of control and other personality factors," *IEEE transactions on visualization and computer graphics*, vol. 19, no. 7, pp. 1109–1121, 2012.

[3] T. M. Green and B. Fisher, "Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction," in *2010 IEEE Symposium on Visual Analytics Science and Technology*. IEEE, 2010, pp. 203–210.

[4] N. Oscar, S. Mejía, R. Metoyer, and K. Hooker, "Towards personalized visualization: Information granularity, situation, and personality," in *Proceedings of the 2017 Conference on Designing Interactive Systems*, 2017, pp. 811–819.

[5] C. Xiong, L. Padilla, K. Grayson, and S. Franconeri, "Examining the components of trust in map-based visualizations," 2019.

[6] J. Zhou, S. Luo, and F. Chen, "Effects of personality traits on user trust in human–machine collaborations," *Journal on Multimodal User Interfaces*, pp. 1–14, 2020.

[7] M. Adnan, M. Just, and L. Baillie, "Investigating time series visualisations to improve the user experience," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 5444–5455.

[8] M. Card, *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[9] C. Tominski and H. Schumann, *Interactive Visual Data Analysis*. CRC Press, 2020.

[10] R. Spence, *Information visualization*. Springer, 2001, vol. 1.

[11] E. Dimara and J. Stasko, “A critical reflection on visualization research: Where do decision making tasks hide?” *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[12] H. A. Simon, “The new science of management decision.” 1960.

[13] E. Dimara, S. Franconeri, C. Plaisant, A. Bezerianos, and P. Dragicevic, “A task-based taxonomy of cognitive biases for information visualization,” *IEEE transactions on visualization and computer graphics*, vol. 26, no. 2, pp. 1413–1432, 2018.

[14] A. Kale, M. Kay, and J. Hullman, “Visual reasoning strategies for effect size judgments and decisions,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 272–282, 2020.

[15] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim, “The role of uncertainty, awareness, and trust in visual analytics,” *IEEE transactions on visualization and computer graphics*, vol. 22, no. 1, pp. 240–249, 2015.

[16] E. Mayr, N. Hynek, S. Salisu, and F. Windhager, “Trust in information visualization,” in *EuroVis Workshop on Trustworthy Visualization (TrustVis), Robert Kosara, Kai Lawonn, Lars Linsen, and Noeska Smit (Eds.). The Eurographics Association. DOI: http://dx. doi. org/10.2312/trvis*, vol. 20191187, 2019.

[17] J. Müller and C. Schwieren, “Big five personality factors in the trust game,” *Journal of Business Economics*, vol. 90, no. 1, pp. 37–55, 2020.

[18] D. C. Funder, *The personality puzzle.* WW Norton & Co, 1997.

[19] R. R. McCrae and P. T. Costa, “Validation of the five-factor model of personality across instruments and observers.” *Journal of personality and social psychology*, vol. 52, no. 1, p. 81, 1987.

[20] J.-H. Cho, H. Cam, and A. Oltramari, “Effect of personality traits on trust and risk to phishing vulnerability: Modeling and analysis,” in *2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*. IEEE, 2016, pp. 7–13.

[21] A. M. Evans and W. Revelle, “Survey and behavioral measurements of interpersonal trust,” *Journal of Research in Personality*, vol. 42, no. 6, pp. 1585–1593, 2008.

[22] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski, “Visual methods for analyzing time-oriented data,” *IEEE transactions on visualization and computer graphics*, vol. 14, no. 1, pp. 47–60, 2007.

[23] W. Aigner, S. Miksch, W. Muller, H. Schumann, and C. Tominski, "Visualizing time-oriented data—a systematic view," *Computers & Graphics*, vol. 31, no. 3, pp. 401–409, 2007.

[24] R. Salway, R. Valenzuela, J. Shoenberger, W. Mallon, and A. Viccellio, "Emergency department (ed) overcrowding: evidence-based answers to frequently asked questions," *Revista Médica Clínica Las Condes*, vol. 28, no. 2, pp. 213–219, 2017.

[25] R. Tadeusiewicz and A. Horzyk, "Man-machine interaction improvement by means of automatic human personality identification," in *IFIP International Conference on Computer Information Systems and Industrial Management*. Springer, 2015, pp. 278–289.

[26] H. J. Eysenck, *Dimensions of personality*. Transaction Publishers, 1950, vol. 5.

[27] R. B. Cattell, "Personality and motivation structure and measurement." 1957.

[28] P. T. Costa Jr and R. R. McCrae, "Four ways five factors are basic," *Personality and individual differences*, vol. 13, no. 6, pp. 653–665, 1992.

[29] M. Zuckerman, "An alternative five-factor model for personality," *The developing structure of temperament and personality from infancy to adulthood*, pp. 53–68, 1994.

[30] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *Journal of personality*, vol. 60, no. 2, pp. 175–215, 1992.

[31] M. Zuckerman, D. M. Kuhlman, J. Joireman, P. Teta, and M. Kraft, "A comparison of three structural models for personality: the big three, the big five, and the alternative five." *Journal of personality and social psychology*, vol. 65, no. 4, p. 757, 1993.

[32] S. Aziz and C. J. Jackson, "A comparison between three and five factor models of pakistani personality data," *Personality and Individual Differences*, vol. 31, no. 8, pp. 1311–1319, 2001.

[33] P. T. Costa Jr and R. R. McCrae, *The Revised NEO Personality Inventory (NEO-PI-R).* Sage Publications, Inc, 2008.

[34] O. P. John, E. M. Donahue, and R. L. Kentle, "Big five inventory," *Journal of Personality and Social Psychology*, 1991.

[35] L. R. Goldberg *et al.*, "A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models," *Personality psychology in Europe*, vol. 7, no. 1, pp. 7–28, 1999.

[36] P. T. Costa and R. R. MacCrae, *NEO PI-R: inventário de personalidade NEO revisto: manual profissional*, 2000.

[37] M. Zuckerman, “Zuckerman-kuhlman personality questionnaire (zkpq): An alternative five-factorial model,” *Big five assessment*, pp. 377–396, 2002.

[38] A. Byravan and N. V. Ramanaiah, “Structure of the 16 pf fifth edition from the perspective of the five-factor model,” *Psychological Reports*, vol. 76, no. 2, pp. 555–560, 1995.

[39] H. E. Cattell and A. D. Mead, “The sixteen personality factor questionnaire (16pf).” 2008.

[40] N. Bayram and M. Aydemir, “Decision-making styles and personality traits,” *Int J Recent Adv Organizational Behaviour Decision Sci*, vol. 3, pp. 905–15, 2017.

[41] J. Müller and C. Schwieren, “What can the big five personality factors contribute to explain small-scale economic behavior?” 2012.

[42] R. C. Mayer, J. H. Davis, and F. D. Schoorman, “An integrative model of organizational trust,” *Academy of management review*, vol. 20, no. 3, pp. 709–734, 1995.

[43] A. Ben-Ner and F. Halldorsson, “Trusting and trustworthiness: What are they, how to measure them, and what affects them,” *Journal of Economic Psychology*, vol. 31, no. 1, pp. 64–79, 2010.

[44] I. Thielmann and B. E. Hilbig, “Trust: An integrative review from a person–situation perspective,” *Review of General Psychology*, vol. 19, no. 3, pp. 249–277, 2015.

[45] S. Marsh and M. R. Dibben, “The role of trust in information science and technology,” *Annual Review of Information Science and Technology (ARIST)*, vol. 37, pp. 465–98, 2003.

[46] A. K. Schnackenberg and E. C. Tomlinson, “Organizational transparency: A new perspective on managing trust in organization-stakeholder relationships,” *Journal of Management*, vol. 42, no. 7, pp. 1784–1810, 2016.

[47] J. T. Cacioppo and R. E. Petty, “The need for cognition.” *Journal of personality and social psychology*, vol. 42, no. 1, p. 116, 1982.

[48] E. M. Uslaner, “The moral foundations of trust,” *Available at SSRN 824504*, 2002.

[49] Uslaner, “Trust as an alternative to risk,” *Public Choice*, vol. 157, no. 3-4, pp. 629–639, 2013.

[50] G. M. Alarcon, J. B. Lyons, J. C. Christensen, M. A. Bowers, S. L. Klosterman, and A. Capiola, “The role of propensity to trust and the five factor model across the trust process,” *Journal of Research in Personality*, vol. 75, pp. 69–82, 2018.

[51] R. C. Mayer and J. H. Davis, “The effect of the performance appraisal system on trust for management: A field quasi-experiment.” *Journal of applied psychology*, vol. 84, no. 1, p. 123, 1999.

[52] C. Ziemkiewicz, A. Ottley, R. J. Crouser, K. Chauncey, S. L. Su, and R. Chang, "Understanding visualization by understanding individual users," *IEEE computer graphics and applications*, vol. 32, no. 6, pp. 88–94, 2012.

[53] T. M. Green, D. H. Jeong, and B. Fisher, "Using personality factors to predict interface learning performance," in *2010 43rd Hawaii International Conference on System Sciences*. IEEE, 2010, pp. 1–10.

[54] T. Alves, J. Natálio, J. Henriques-Calado, and S. Gama, "Incorporating personality in user interface design: A review," *Personality and Individual Differences*, vol. 155, p. 109709, 2020.

[55] Z. Liu, R. J. Crouser, and A. Ottley, "Survey on individual differences in visualization," *arXiv preprint arXiv:2002.07950*, 2020.

[56] M. Sorapure, "Text, image, data, interaction: Understanding information visualization," *Computers and Composition*, vol. 54, p. 102519, 2019.

[57] H.-K. Kong, Z. Liu, and K. Karahalios, "Trust and recall of information across varying degrees of title-visualization misalignment," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–13.

[58] J. Chuang, D. Ramage, C. Manning, and J. Heer, "Interpretation and trust: Designing model-driven visualizations for text analysis," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 443–452.

[59] W. Javed, B. McDonnel, and N. Elmqvist, "Graphical perception of multiple time series," *IEEE transactions on visualization and computer graphics*, vol. 16, no. 6, pp. 927–934, 2010.

[60] M. Waldner, A. Diehl, D. Gračanin, R. Splechtna, C. Delrieux, and K. Matković, "A comparison of radial and linear charts for visualizing daily patterns," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 1, pp. 1033–1042, 2019.

[61] J. A. Acuna, J. L. Zayas-Castro, and H. Charkhgard, "Ambulance allocation optimization model for the overcrowding problem in us emergency departments: A case study in florida," *Socio-Economic Planning Sciences*, vol. 71, p. 100747, 2020.

[62] M. Lima and A. Simões, "Neo-pi-r manual profissional," *Lisboa: CEGOC*, 2000.

[63] I. Cho, R. Wesslen, A. Karduni, S. Santhanam, S. Shaikh, and W. Dou, "The anchoring effect in decision-making with visual analytics," in *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 2017, pp. 116–126.

[64] A. Dasgupta, J.-Y. Lee, R. Wilson, R. A. Lafrance, N. Cramer, K. Cook, and S. Payne, "Familiarity vs trust: A comparative study of domain scientists' trust in visual analytics and conventional analysis methods," *IEEE transactions on visualization and computer graphics*, vol. 23, no. 1, pp. 271–280, 2016.

[65] IBM Corp., "Ibm spss statistics for windows," Armonk, NY: IBM Corp. [Online]. Available: https://hadoop.apache.org

[66] J. Han, J. Pei, and M. Kamber, *Data mining: concepts and techniques*. Elsevier, 2011.

[67] P. Costa and R. McRae, "Revised neo personality ınventory (neo-pir) and neo five-factor ınventory (neo-ffi)," 1992.

[68] J. Cohen, *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

[69] M. M. Kwiatkowska and R. Rogoza, "A modest proposal to link shyness and modesty: Investigating the relation within the framework of big five personality traits," *Personality and Individual Differences*, vol. 149, pp. 8–13, 2019.

[70] A. Raudys, V. Lenčiauskas, and E. Malčius, "Moving averages for financial data smoothing," in *International Conference on Information and Software Technologies*. Springer, 2013, pp. 34–45.

A

Informed Consent

Consentimento Informado

Objetivo

Estamos a conduzir um estudo observacional sobre como indivíduos interagem com visualizações de informação. Iremos também validar se fatores de personalidade modelam este comportamento. O objetivo desta sessão é executar um conjunto de tarefas com diferentes visualizações e recolher as suas observações sobre as mesmas.

Características da sessão

A duração desta sessão não deve ser maior que 30 minutos e terá lugar na plataforma de videoconferência Zoom. Não identificámos riscos que não sejam os da vida quotidiana.

Tratamento dos dados pessoais recolhidos durante a sessão

Na sessão serão gravados os seus dados referentes a (i) *interações com interfaces (e.g. movimentos de rato e cliques)*, (ii) *respostas a questionários*, (iii) *áudio*, e (iv) *gravação de tela*. Todos os dados recolhidos serão mantidos em sigilo. Os dados mencionados em (i) e (ii) poderão também ser utilizados para apresentação ou exibição de resultados, devidamente pseudonimizados, em publicações científicas, conferências ou eventos semelhantes. Os dados mencionados em (iii) e (iv) serão exclusivamente usados para ajudar a interpretação da experiência. A gravação de tela não inclui a face do participante. Assim, nenhum destes dados será divulgado em publicações científicas, conferências ou eventos semelhantes.

Estes dados vão ser armazenados em unidades de armazenamento externas privadas a cargo do responsável pelo tratamento de dados. De forma a preservar a pseudo-anonimidade dos seus dados, ser-lhe-á atribuído um identificador numérico único. Os seus dados de contacto e os seus dados da experiência serão guardados em unidades de armazenamento externas privadas diferentes de forma a manter a confidencialidade dos mesmos. Os dados pseudonimizados da experiência (não incluem identificador) serão analisados, exclusivamente, pelos membros da equipa de investigação. Para além destes dados, vamos também usar os seus dados referentes aos questionários de personalidade que preencheu numa fase anterior. O seu tratamento será igual aos que recolhemos nesta sessão.

Os seus direitos

A sua participação é voluntária e livre, sendo que tem o direito de desistir a qualquer momento sem qualquer prejuízo pessoal. Caso tal aconteça, os dados relativos à sua experiência serão removidos e destruídos. Tem igualmente o direito de solicitar ao responsável pelo tratamento acesso aos dados pessoais que lhe digam respeito, bem como os direitos de rectificação, apagamento, limitação e oposição do tratamento, incluindo o direito de retirar consentimento em qualquer altura, sem prejuízo da licitude do tratamento eventual e previamente consentido. Tem igualmente o direito de apresentar uma reclamação à CNPD (Comissão Nacional de Proteção de Dados). Todos os dados serão destruídos ao fim de três anos desde a data desta sessão, de acordo com a Lei de Proteção de Dados Portuguesa. Por último, tem também o direito de saber as entidades a quem possam os dados ser comunicados e possibilidade da transferência dos dados para países terceiros (fora do Espaço Económico Europeu).

Se tiver alguma questão, sinta-se à vontade para a colocar. Para participar nesta experiência, pedimos-lhe que leia o consentimento informado e caso concorde em participar de acordo com os termos abaixo, pedimos-lhe que assine o formulário no local indicado.

---

1 - Li e compreendi o significado deste estudo. Tive a oportunidade de colocar questões, caso necessário, e recolher as respetivas respostas.

2 - Compreendo que a participação neste estudo é voluntária e que posso desistir a qualquer momento, sem apresentar qualquer explicação. Caso tal aconteça, não serei alvo de qualquer penalização e os dados relativos à minha experiência serão removidos e destruídos.

3 - Autorizo a gravação dos dados durante a sessão.

4 - Autorizo o processamento dos dados no âmbito deste projeto para fins de análise, investigação e disseminação de resultados em publicações científicas ou conferências na área do projeto, pelos investigadores deste projeto.

5 - Compreendi que os dados recolhidos neste estudo serão utilizados como mencionado anteriormente.

6 - Autorizo novamente o processamento dos meus dados demográficos e de personalidade recolhidos anteriormente.

7 - De acordo com o descrito acima, autorizo a minha participação neste estudo e aceito as suas condições.

Obrigado pela sua colaboração!

____
(participante)

____
(investigador responsável)

____
(data)

Ao participante será entregue uma cópia assinada deste formulário.

Equipa

Sandra Gama (Investigador responsável)
sandra.gama@tecnico.ulisboa.pt

Tomás Alves (Responsável pelo tratamento de dados)
tomas.alves@tecnico.ulisboa.pt

Tatiana Nunes (Encarregado de proteção de dados)
dpo@inesc-id.pt

Daniel Gonçalves
daniel.goncalves@inesc-id.pt

Carlota Dias
carlota.lopes.dias@tecnico.ulisboa.pt

Alexandra Maroco
alexandra.maroco@tecnico.ulisboa.pt

Ricardo Velhinho
ricardo.velhinho@tecnico.ulisboa.pt

Graphics and Interaction - INESC-ID
Instituto Superior Técnico
R. Alves Redol 9,
1000-029 Lisboa, Portugal

Joana Henriques-Calado
jhcalado@psicologia.ulisboa.pt

CICPSI, Faculdade de Psicologia,
Cidade Universitária,
Alameda da Universidade,
1649-004 Lisboa, Portugal

B

Study Questionnaire

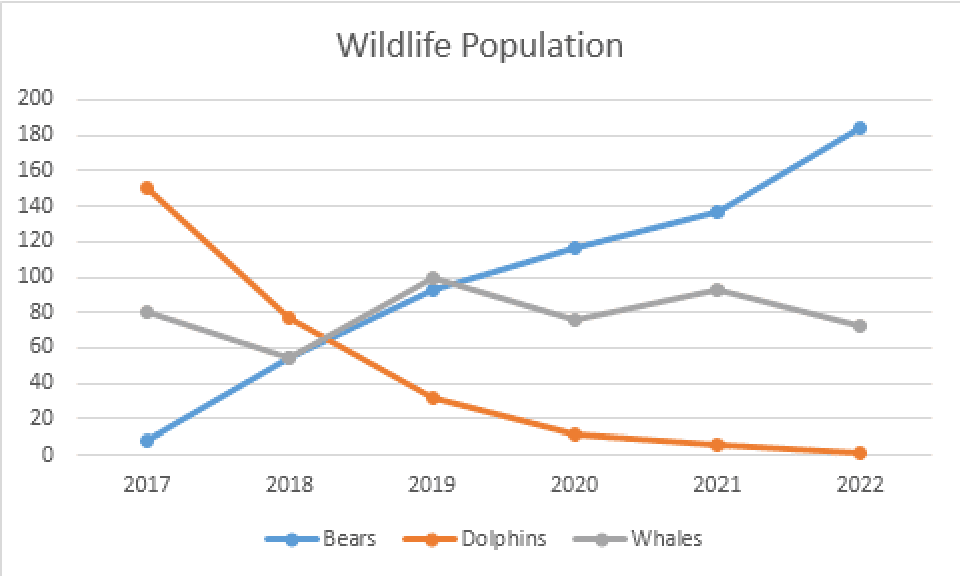
Teste com Utilizadores

* Required

1. Insira o seu id de participante *

Questionário de familiaridade

2. Qual a sua familiaridade com este tipo de visualizações? *



Mark only one oval.

	1	2	3	4	5	
Não familiarizado	◯	◯	◯	◯	◯	Completamente familiarizado

3. Está a usar óculos ou alguma tipo de correção visual? *

Mark only one oval.

◯ Sim

◯ Não

Introdução

Nos últimos anos tem sido experienciado um aumento exponencial do número de pacientes que visita diariamente diferentes departamentos de emergência espalhados pelos hospitais por todo o mundo. Atualmente, a sobrelotação destes departamentos é um problema transversal a muitos países, entre eles Portugal. A sobrelotação acontece quando não existe espaço para que os pacientes que necessitam de tratamento rapidamente o possam receber. Quando uma situação urgente é adiada devido ao congestionamento existente, então estamos perante um caso de sobrelotação.

Mais do que nunca, deparamo-nos com um aumento no número de hospitais que lida diariamente com departamentos de emergência a funcionar com mais do que a sua capacidade máxima. Esta crise pode ser explicada por diversos fatores, entre eles: as visitas desnecessárias às urgências, a má gestão da calendarização das intervenções cirúrgicas, o aumento de afluência em alturas específicas do ano devido a doenças sazonais, e o aumento da severidade das doenças resultantes de uma população cada vez mais envelhecida.

Este problema é considerado especialmente relevante devido ao impacto negativo que tem na segurança e saúde dos pacientes. A sobrelotação está associada ao aumento dos tempos de espera, especialmente para pacientes que não se encontram em estado crítico, ao aumento da mortalidade, ao aumento do número de pacientes que acabam por sair das urgências sem terem sido atendidos, e finalmente ao aumento dos tempos de espera por ambulâncias. Assim, é possível concluir que a superlotação leva a uma diminuição notável da qualidade dos serviços de saúde. Além do mais, esta crise afeta também os próprios hospitais devido aos gastos financeiros associados a esta situação.

Tendo em conta tudo isto, pedimos agora que se imagine na posição de estar perante uma urgência de saúde e ter de decidir a que hospital quer ir, de modo a poder ser assistido. Para realizar esta decisão terá acesso a diversas visualizações, onde poderá ver o número de pacientes no departamento de emergência de dois hospitais distintos, no último dia, semana e mês. Apesar de existirem diversos aspetos que poderiam ser considerados aquando desta tomada de decisão, pedimos que considere apenas o número de pacientes. Tendo em conta que um menor número de pacientes será equivalente a um menor tempo de espera.

1ª Visualização

4. Qual a visualização que está a observar? *

Mark only one oval.

- ◯ Número de pacientes no último dia
- ◯ Número de pacientes na última semana
- ◯ Número de pacientes no último mês

5. Quanto confia nos dados que lhe estão a ser apresentados ? *

Mark only one oval.

	1	2	3	4	5	
Não confio	⬭	⬭	⬭	⬭	⬭	Confio totalmente

2ª Visualização

6. Qual a visualização que está a observar? *

Mark only one oval.

- ⬭ Número de pacientes no último dia
- ⬭ Número de pacientes na última semana
- ⬭ Número de pacientes no último mês

7. Quanto confia nos dados que lhe estão a ser apresentados ? *

Mark only one oval.

	1	2	3	4	5	
Não confio	⬭	⬭	⬭	⬭	⬭	Confio totalmente

3ª Visualização

8. Qual a visualização que está a observar? *

Mark only one oval.

- ⬭ Número de pacientes no último dia
- ⬭ Número de pacientes na última semana
- ⬭ Número de pacientes no último mês

9. Quanto confia nos dados que lhe estão a ser apresentados ? *

Mark only one oval.

	1	2	3	4	5	
Não confio	⬭	⬭	⬭	⬭	⬭	Confio totalmente

Untitled Section

10. Quando foi a última vez que foi a um hospital num contexto de urgência médica? *

Mark only one oval.

- ⬭ Na última semana
- ⬭ No último mês
- ⬭ No último ano/s
- ⬭ Nunca