TÉCNICO LISBOA

Studying the Distinction Cognitive Bias in Information Visualization

Vasco Carvalho Ferreira Pires

Thesis to obtain the Master of Science Degree in

Computer Science and Engineering

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

Examination Committee

Chairperson: Prof. Diogo Manuel Ribeiro Ferreira
Supervisor: Prof. Sandra Pereira Gama
Member of the Committee: Prof. Catarina Pinto Moreira

October 2022

This work was created using LaTeX typesetting language
in the Overleaf environment (www.overleaf.com).

Acknowledgments

I will start by thanking my dissertation supervisors Prof. Sandra Gama, Prof. Daniel Gonçalves and Tomás Alves for the instrumental guidance throughout this entire process. Your insights and above all your patience have been invaluable, both to this work and to my development.

Reaching this milestone is only possible with all the support of the special people in my life. Thank you. Life with all of you is an endless source of teachings for me in so many ways. I would especially like to thank my family for their boundless love, care and unconditional encouragement through this journey. For always teaching me to do my best and reminding me that that is enough.

Thank you specifically to all the people I am grateful to have shared the highs and lows of the last few years with. To Cate, for caring for me, not allowing me to doubt myself and always reiterating my ability and worth. For sharing a home with me in an intensive month that felt more like a year. To Dani, for your enthusiasm, for motivating me to try new things and particularly for making it your duty to recruit people for this experiment. To Leão, for simply getting me, deeply and entirely, and pushing my creativity and self-development. To Xico, for teaching me to put things in perspective and that life doesn't need to be so serious by turning nightmares into humour. All of you have contributed to the person I have become and I am forever grateful to have navigated this roller-coaster ride with you.

Last but not least, thank you to everyone who agreed to be part of the experiments. This work would not exist without all of your contributions.

Abstract

Distinction Bias is described as the tendency to view two options as more dissimilar when evaluating them simultaneously than when evaluating them separately. Distinction is included in a group of systematic errors of human cognition denominated cognitive biases, impacting our judgments and decisions. While information visualization often supports complex thinking under uncertainty, visualizations have been observed to be sensitive to different cognitive styles and heuristics utilized in the decision-making process. However, the study of cognitive biases in the field of information visualization remains largely unexplored, with Distinction Bias standing as a likely relevant yet not discussed topic. Moreover, individual differences such as personality have an effect on how we perceive and process information and therefore also play a role when it comes to human-machine systems. This work proposes a study focused on Distinction Bias in the context of information visualization, and how a personality trait, neuroticism, may affect it. Our study acknowledges and validates the relevance of studying Distinction Bias, specifically in the information visualization context. Contrarily, our results do not exhibit correlations between neuroticism and the effects of Distinction Bias. However, further analysis raised some interest in the influence of personality characteristics on user interaction metrics.

Keywords

Distinction Bias; Cognitive Bias; Human-Computer Interaction; Information Visualization; Personality; Neuroticism.

Resumo

O Viés de Distinção é descrito como a tendência de considerar duas opções como mais divergentes ao avaliá-las simultaneamente do que avaliando-as separadamente. O Viés de Distinção está incluído num conjunto de erros sistemáticos da cognição humana denominados vieses cognitivos, impactando os nossos julgamentos e decisões. Embora visualização de informação suporte pensamento complexo sob incerteza, observou-se que as visualizações são sensíveis a diferentes estilos cognitivos e heurísticas utilizadas no processo de decisão. Contudo, o estudo de vieses cognitivos no campo da visualização de informação permanece em grande parte inexplorado, sendo o Viés de Distinção definido como um tópico presumidamente relevante, mas ainda não discutido. Adicionalmente, características pessoais como a personalidade afetam a forma como percebemos e processamos informação e, portanto, também desempenham um papel quando se trata de sistemas pessoa-máquina. Este trabalho propõe um estudo focado no Viés de Distinção no contexto de visualização de informação, e como um traço de personalidade, o neuroticismo, pode afetá-lo. O nosso estudo reconhece e valida a relevância de estudar o Viés de Distinção, especificamente no contexto de visualização de informação. Contrariamente, os nossos resultados não apresentam correlações entre neuroticismo e os efeitos do Viés de Distinção. No entanto, uma análise aprofundada levantou algum interesse na influência de características de personalidade em métricas de interação pessoa-máquina.

Palavras Chave

Viés de Distinção; Viés Cognitivo; Interação Pessoa-Máquina; Visualização de Informação; Personalidade; Neuroticismo.

Contents

List of Figures

List of Tables

Acronyms

ANOVA	Analysis of Variance
FFM	Five-Factor Model
HCI	Human-Computer Interaction
InfoVis	Information Visualization
JDM	Judgments and Decision-making
JE	Joint Evaluation
SE	Separate Evaluation

1

Introduction

Contents

Nowadays people have to deal with increasingly growing amounts of information, produced and collected at incredible rates, while simultaneously having to make progressively more complex and critical decisions [2, 7]. Visual analytics systems turn the overload of information into an opportunity for decision-makers to gain insight and take action into solving problems, blending sophisticated data analytics with interactive visual interfaces such as visualizations [3,4]. Information Visualization (InfoVis) provides users with graphical representations that allow them to efficiently explore, analyse and communicate patterns of bodies of data with various degrees of complexity, combining the strengths of humans and machines in human-in-the-loop systems [3]. In this sense, in order to design effective visualizations we must consider the process of human reasoning, judgement and decision-making, as well as its limitations [1].

Human cognition deals with finite resources such as working memory capacity and attention [8,9] and thus, in order to make decisions quickly, with limited information and resources, the human mind makes use of heuristics, rules of thumb and approximations, even when we are not consciously aware of these strategies. The imperfections of these strategies manifest themselves as cognitive biases [1, 2]. Among them and the focus of this work is Distinction Bias, described as the tendency to view two options as more dissimilar when evaluating them simultaneously than when evaluating them separately, influencing our predictions and choices [10].

Distinction bias may affect many areas of our lives. Take as an example visual analytics systems intending to compare products we are looking to buy. While evaluating options side by side, Distinction Bias may cause people to become hypersensitive to small differences that in practise will not return the additional happiness that was hoped for. For example, it can cause people to go over budget when shopping for goods like a new car, house or television. While we might have been thrilled with a cheaper model when viewing it on its own, when viewing it in comparison to another model, it may seem lackluster [10]. Or, when faced with options involving a trade-off along two attributes, such as a house with more square footage but without a characteristic people find valuable, such as an outside space, and one with less square footage but containing a backyard space, are people able to choose the option that will bring them the greater overall happiness? [10]

While visualization tools aim to support judgments and decisions in the context of problem-solving, the topic of cognitive biases and its impact on how people use these tools is still relatively unexplored. As for Distinction Bias specifically, while considered potentially relevant, it still had not been discussed in the field [1]. With this work we intend to provide a first step into the exploration of Distinction Bias in visual analytics systems powered by InfoVis, exploring how the joint evaluation of options affects people's expected utility.

Furthermore, with InfoVis often supporting complex thinking under uncertainty, visualizations may be sensitive to different cognitive styles in the decision-making process [9]. Specifically, individual differences, such as personality, have been shown to have an effect on how humans perceive and process information and therefore may impact the way we live and make our decisions [11]. Trait-based personality theory defines personality as traits that predict an individual's behavior, defining a trait as a distinguishing and relatively stable personal characteristic, manifested in consistent and enduring ways of reacting to the environment [12, 13]. Throughout history many personality models set on trait theory were proposed, being the Five-Factor Model (FFM) one of the most widely accepted at present times. Throughout most models, a personality trait that prevails is neuroticism. Highly neurotic individuals are characterized by a tendency to feel worried, nervous, depressed, self-conscious and to overall more easily experience negative emotions [14, 15]. The topic of personality has been considered relevant in several fields, among them the field of Human-Computer-Interaction [16] and research has demonstrated how distinct personality types and cognitive abilities can make a difference in problem-solving and behavioral patterns [17], with the neurotic trait being observed to play a part in participant patterns when interacting with visualizations [4, 6, 18]. With existing Distinction Bias literature focusing on happiness forecasting as one's demonstration of expected utility [10], the relevancy of studying the influence of personality in a bias-prone context was identified and therefore additionally explored in this work, specifically the possible influence of the neurotic trait in the context of our study.

1.1 Objectives

The main focus of this study is to understand the effects of Distinction Bias in the context of Information Visualization, as well as its possible relationship with a personality trait, neuroticism. In order to achieve this goal, several intermediate steps were defined to progress through the course of this study:

- Examination of available Distinction Bias literature as well as prior overall Cognitive Bias and Personality literature in the context of InfoVis;
- Craft and development of the visualizations to be used in the study;
- User testing and collection of user personality data;
- Data analysis and investigation of results in order to verify hypotheses.

1.2 Document Structure

Chapter 2 presents the theoretical knowledge of the relevant concepts for this study, from an overview of the concept of cognitive bias, followed by a description of the bias in the focus for this work, Distinction

Bias, to a brief history and introduction to the concepts of personality and trait theory and an overview on the main concepts of the specific trait we will analyze, neuroticism.

Chapter 3 provides an analysis of existing literature contributions that pose relevance to our work, namely the work focusing on both cognitive biases and personality in the field of InfoVis, discussing the main takeaways and limitations that set the foundations for our study. Chapter 4 defines the methodology and approach to our study, presenting the visualizations utilized and describing the tasks performed in the study, as well as the gathered measures. It will follow onto the statement of our research questions and hypothesis, followed by a description of the data collection and procedure process and an explanation of the data analysis process. Chapter 5 presents the results obtained from our data analysis, as well as a discussion of our findings.

At last, Chapter 6 denotes the conclusions drawn from the present work as well as some identified limitations and future opportunities to extend the work.

2

Background

Contents

This section provides the fundamental theoretical background related to this work, namely the concepts of cognitive biases, specifically the Distinction Bias, as well as the fundamental aspects of personality, comparing various models, with particular focus on the neuroticism trait, both relevant for a thorough understanding of the proposed work.

2.1 Cognitive Biases

When we make judgments and decisions we like to consider ourselves capable of objectively and logically absorbing and evaluating the information available. The study of Judgments and Decision-making (JDM) revolves around the comparison of judgments and decisions to standards, opening up the possibility to assess their value as better or worse. These standards stem from normative models of judgment and decision-making, which are set on the assumptions that people follow principles of rationality, have a fixed set of preferences and attempt to maximize their benefit when making decisions [1, 19]. However, evidence of systematic deviations from these models were identified by Tversky and Kahneman [20] and denominated cognitive biases.

Tversky and Kahneman presented specific mechanisms that people rely on to simplify the decision-making process. In order to make quick judgments in our everyday lives, the human mind unconsciously makes use of fast thinking habits and strategies that aid the process of decision-making [9]. These heuristics and rules of thumb can simplify complex problems and provide effective judgments and decisions, given that humans have limited time and cognitive abilities to make truly rational decisions [8, 9]. Fortunately, many times making use of the fast and automatic heuristic strategies can lead to satisfactory results. However, it's not always the case that these strategies result in optimal judgments. Some can routinely lead to decisions that deviate from optimality [9, 21], leading to what we refer to as cognitive biases.

A cognitive bias was defined by Pohl [22] as a cognitive phenomenon involving a reliable deviation from reality that is predictable and systematic. Additionally, the author explains that a person subject to a cognitive bias is unaware of it, as the phenomenon occurs involuntarily, being convinced that their judgments and decisions are unbiased, even though the effects often persist when subjects are informed on how to overcome them, being difficult or even impossible to avoid. As a last touchstone, cognitive biases differ from the regular course of information processing, in contrast to more mundane forms of human error, such as misunderstanding or misremembering [1, 22].

In recent years there has been a noticeable interest in cognitive biases. Some widely known examples are confirmation bias, occurring when people tend to seek and favor information that confirms their own previous assumptions, and availability bias, according to which people tend to estimate the likelihood of an event by how easy it is for them to recall similar events [1, 21].

2.1.1 Distinction Bias

Distinction Bias describes a cognitive bias in decision-making in which choices and predictions are affected due to the joint evaluation of the options. Essentially the Distinction Bias consists of a situation in which two options appear to be more dissimilar when we examine them together, as opposed to separately, thus having the evaluation mode we find ourselves in interfere with judgments and decisions, leading us to fail to make accurate predictions and, consequently, optimal choices [10].

The evaluation mode is a contextual feature in decision-making. The Joint Evaluation (JE) mode refers to the situation in which options are evaluated simultaneously [23]. Hsee and Zhang [10] detail that when people make utility and affective predictions for multiple options they tend to be in JE mode, being usually faced with several alternatives that entail different values of an attribute, this way being able to differentiate them and compare the utility and desirability of each of them. The Separate Evaluation (SE) mode refers to the situation in which one option is evaluated in isolation [23]. When experiencing or predicting an experience in SE mode, people are presented with a single event to evaluate, this way not engaging in comparisons with alternative options [10].

People often make judgements and decisions regarding available alternatives based on predictions of utility or happiness as one's manifestation of utility [24]. Examples can range from deciding to get a bigger or smaller quantity of a dessert, paying more for a larger television or opting for smaller and cheaper one, getting a less interesting yet higher salary job or even being the one hiring and deciding what will be the preferred trade-offs among several attribute values of the possible candidates.

In the process of examining several alternatives together, in JE mode, the differences between some attribute values may appear more salient, often resulting in different evaluations than when only one of the options is examined in isolation, in SE mode.

Specifically, Hsee and Zhang [10] corroborated through a set of experiments that when people in JE mode predict the affective state they would get from various outcomes, juxtaposing attribute values relative to one another, they tend to assess increasingly higher levels of happiness for both increasing values of a desirable attribute or decreasing values of an undesirable attribute. For example, if someone is a book author and is trying to sell their book, a higher number of book sales would be seen as desirable and when comparing and forecasting the happiness experienced from selling 100 or 200 books in JE mode this person will likely predict a higher level of happiness for an outcome in which they sell 200 books than for the one in which they sell 100. Similarly, people in JE mode also tend to assess decreasingly lower levels of happiness for both increasing values of an undesirable attribute or decreasing values of a desirable attribute.

In contrast, people in SE mode often tend to not have a precise idea of how quantitatively good or bad an attribute value is by itself, though they may be able to evaluate it as qualitatively positive or negative in relation to a baseline or reference point.

Attribute values are considered solely quantitatively different if they are both situated on the same side of a reference point and qualitatively different if they involve different valences or one is above the reference point and one below. This is independent from the attribute value itself being qualitative or quantitative, as the unit of analysis is the value difference. For example, the difference between reading a list containing 10 negative words or a list containing 20 negative words is only quantitative yet the difference between reading a list of 10 positive words and reading a list of 10 negative words is qualitative. As such, when people find themselves in SE mode, the happiness predicted or experienced for the several alternatives was not observed to increase or decrease in a linear proportion with increasing values of a desirable or undesirable attribute, respectively, instead depending mostly on whether the options are qualitatively considered positive or negative.

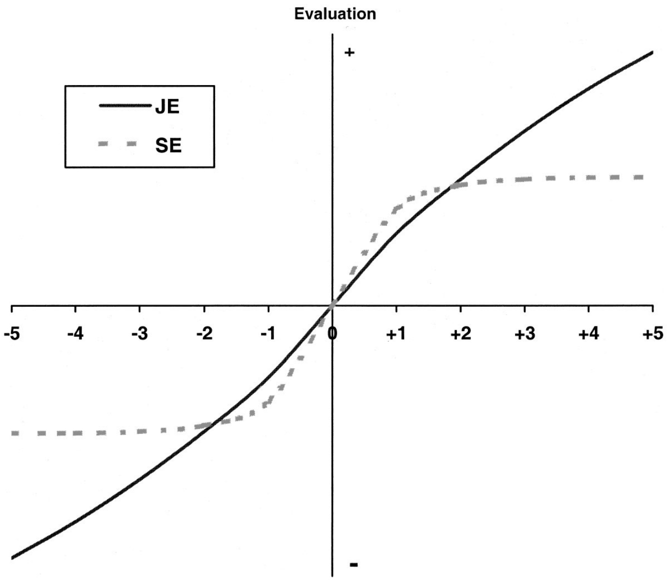


Figure 2.1: Joint-evaluation (JE) curve and separate-evaluation (SE) curve of a hypothetical attribute [10].

Thus, Distinction Bias occurs as people make overpredictions of their affective state when comparing an option's attribute value with another. In this sense, depending on the evaluation mode people find themselves in, the evaluation function of an attribute will be consistently different, as represented in Figure 2.1. We can observe that JE mode's evaluation function is relatively steep and smooth, as people perceive alternative values relative to one another. In comparison, SE mode's evaluation function is steep when close to the baseline value but somewhat flat elsewhere, as in SE mode people are mostly merely able to distinguish if attribute values are either above or below a baseline. This way people in SE mode tend to predict or experience contrasting levels of utility or happiness for qualitatively different options, yet similar levels when alternatives differ only quantitatively.

In conclusion, Hsee and Zhang [10] corroborate that Distinction Bias poses as the result of people in JE mode having the tendency to overpredict the difference in utility and affect regarding attribute value differences if these are quantitative, such as the difference in happiness between being able to sell 100 or 200 copies of our newly published book, but not if they are qualitative, such as the difference in happiness between selling 0 books and 100 books, in relation to people's experienced or predicted utility in SE mode.

2.2 Fundamentals of Personality

Personality can be defined as the combination of patterns of behavior, emotion, motivation and thought that define each individual. It is what makes each of us unique and builds our distinct ways of feeling, thinking and behaving [25, 26].

Across research a frequently shared assumption is that an individual's personality begins with biologically innate components, both those shared with others and those that are distinct because of heredity or other influences, that over the life course are channelled by the influence of many factors, including family experience, culture and other experiences [13]. Although it is not considered to be entirely rigid, it is described as a set of enduring characteristics, being generally resistant to sudden changes and considered to be fairly stable and predictable throughout life [13].

The topic of personality has been considered relevant in several fields, among them the field of Human-Computer Interaction (HCI). When it comes to human-in-the-loop systems personality can play a considerable role, as classifying human personality can aid the recognition of individual needs and behaviours, therefore opening the possibilities for system interactions to become more suitable to the users and positively impact user satisfaction [16].

Throughout history diverse approaches to the scientific psychology field have contested, each being developed over time with several theorist and researcher contributions. The trait approach stands as the largest and most dominant in contemporary personality psychology [26].

The trait approach has its focus on the ways that people differ psychologically and how these differences might be conceptualized, measured and followed over time [26]. Trait-based personality theories define personality as traits that predict an individual's behavior. A trait can be defined as a distinguishing and relatively stable personal characteristic, manifested in consistent and enduring ways of reacting to our environment [12, 13].

Gordon Allport was an early pioneer in the study of traits in the 1930s, reflecting on the significance of heredity and environment in personality. Allport believed that our uniqueness as individuals stems from the interaction of our genetics with our social environment. In this sense heredity equips the personality with the resources, such as physique, intelligence and temperament, being then shaped, expanded or limited by our environment's conditions [12, 13, 25].

Several important efforts have been made and a wide variety of personality models were developed over the years attempting to discover which are the truly essential traits and how they could be organized. Cattell [27] identified sixteen factors or dimensions of personality, in a model denominated the Sixteen Personality Factor. Eysenck [28] identified only three factors in the Three-Factor Model. For many years these models remained the most popular. More recently, prominent psychologists such as McCrae and Costa [14] and Goldberg [29] settled on five recurring dimensions with the Five-Factor model. This is the most widely accepted framework in the present time [26]. HEXACO, a six-factor model of personality, later builds on the FFM, adding an additional dimension [30]. The Five-Factor model organizes personality traits along five dimensions, often referred to as Big Five or the acronym OCEAN, being them Openness to experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism. Traits are scored along a continuum in which, for each dimension, one's scores can be considered high or low. These domains are found to be stable during the course of life and to predict emotions and behaviors in various situations, influencing many aspects of human behavior across many cultures [26]. Each of the five broad personality domains encompasses six specific facets [15].

2.2.1 Neuroticism

Neuroticism is considered a fundamental domain of personality, being recognized since the beginning of basic science personality research [31]. Among various trait approach personality models emerging throughout the years, neuroticism continues to stand as a generally recognized dimension, even its scale being highly correlated across many models' assessment questionnaires [32]. As with the remaining personality traits, neuroticism is generally assessed as a continuous dimension. Individuals with low scores of neuroticism are described as calm, relaxed and even-tempered [14]. Individuals with high scores of neuroticism are characterized by feelings related with worry, nervousness, frustration, self-consciousness, with tendency to be impulsive and have ineffective coping mechanisms [14, 15].

The central aspect of the neurotic dimension is the tendency to easily experience negative emotions [14, 15], even being associated with the development of common mental disorders including anxiety, mood and substance use disorders [33].

According to the FFM, the neurotic dimension includes six specific facet-level traits, each being a representation of a distinct aspect encapsulated by the dimension. These are:

- Anxiety (N1), measuring the tendency for an individual to feel apprehensive, tense, fearful and worried. Individuals with low anxiety scores, on the contrary, tend to feel more calm, relaxed and stable;

- Angry Hostility (N2), measuring the tendency for an individual to fell anger, frustration and be generally hot-tempered. Individuals with low angry hostility scores tend to be friendlier and good-tempered;

- Depression (N3), measuring the tendency for an individual to feel hopeless, guilty, sad melancholic, alone and desperate. Individuals with low depression scores tend to rarely experience the above-mentioned emotions;
- Self-Consciousness (N4), measuring the tendency for an individual to feel shame, embarrassment and consider themselves inferior, often experiencing social anxiety. Individuals with low self-consciousness scores tend to feel more socially adequate and safe;
- Impulsiveness (N5), measuring the tendency for an individual to lack both control and the ability to resist temptations. Individuals with low impulsiveness scores tend to more easily resist temptations and have high tolerance to frustration; and
- Vulnerability (N6), measuring the tendency for an individual to panic in emergencies and being incapable of dealing with tension. Individuals with low vulnerability scores tend to remain calm in emergencies and be more competent. [15].

3

Related Work

Contents

This section describes several contributions from relevant studies as well as conclusions to the work proposed on the areas of cognitive biases, specifically the Distinction Bias, information visualization and personality, with a focus on the trait of neuroticism. First we explore the relevant work for cognitive biases in the context of information visualization and subsequently address the existing literature on the relations of personality with information visualization. We conclude with a discussion of the findings posing as significant for our work.

3.1 Cognitive Biases and Information Visualization

In order to grasp the ways in which visualization tools support judgements and decisions, we need to understand and consider the limitations around visual data analysis [1].

Visualization design presupposes a series of trade-offs. In order to reach effective visualizations designers must take into account three types of limitations, being these computers, displays and humans [34]. On the human side, memory and attention constitute finite resources [8, 9, 34]. As discussed prior, in order to make decisions quickly with limited resources the human mind makes use of heuristics, rules of thumb and approximations, oftentimes resulting in imperfections known as cognitive biases, which constitute a crucial human limitation that visualization researches must be aware of [1, 2]. Thus, the process of decision making can be hindered by human biases and uncertainty in a plethora of scenarios, from deciding which real-estate property to buy, which medical treatment to undergo or who to vote in national elections [35, 36].

While visualizations aim to support judgments and decisions, relatively little is still known regarding how cognitive biases affect how people use visual analytical tools, as only recently the interest grew in the topic of both decision-making [37–40] and cognitive biases [3, 41] such as the availability bias [42], the attraction effect [43, 44] and the anchoring effect [45] within the visualization field [1].

Despite growing interest, there is still limited research and empirical work on biases in the field of visualization, with few works focusing on a specific bias. Furthermore most of the studied tasks in the existing bias literature make use of textual data representations or oftentimes even no data at all, making the link between cognitive biases and information visualization still a largely unexplored area [1].

In order to help bridge this gap by identifying and defining biases in the scope of information visualization, several taxonomies and frameworks have been crafted with the aim to both categorize and define biases, as well as guide the development of new studies.

A task-based cognitive bias taxonomy was developed by Dimara et al. [1], specifically targeted to information visualization researchers. Instead of, as the majority of existing taxonomies, putting the focus of the categorization process on the psychological reasoning and cognitive mechanisms of why each bias occurs, defined as explanatory, this taxonomy categorizes biases based on experimental tasks in which they have been observed and measured, in order to help visualization researchers identify biases that may affect visualization tasks. This taxonomy identifies 154 different biases and sorts them into seven different categories as shown in Figure 3.1, according to the types of tasks participants had performed when each bias was measured, being them Estimation, Decision, Hypothesis Assessment, Causal Attribution, recall, Opinion Reporting and Other. Further subcategories, denominated flavors, were also developed in a more intuitive way as an additional contribution to help readers identify connections between biases.

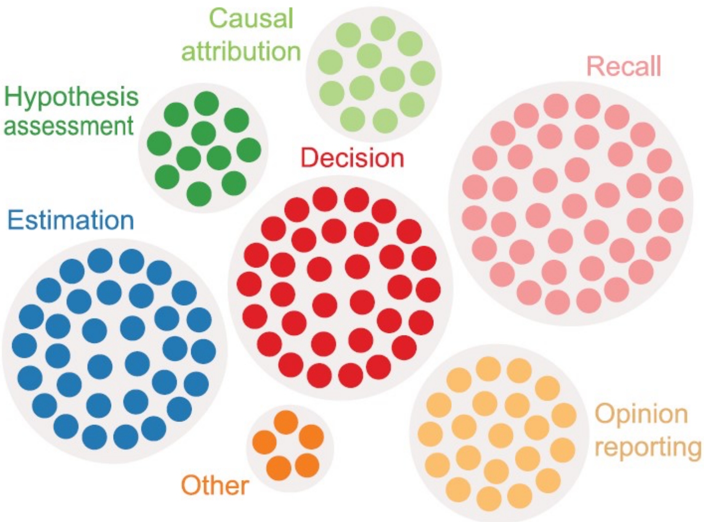


Figure 3.1: Overview of 154 cognitive biases organized by experimental task. Dots represent the different cognitive biases [1].

Decision tasks refer to tasks considered by psychology as choice studies, in which participants are "required to exhibit a preference for one of the several stimuli or make a different prescribed response to each of them" [46] and group 31 of the identified biases, including Distinction Bias. The Distinction Bias is also sub-categorized with the flavor of Baseline, as cognition is biased by a comparison with a perceived baseline, and is described by the authors as likely relevant to the field of visualization research but not yet discussed.

Valdez et al. [2] proposed a lightweight framework aiming to help organize and guide upcoming research questions around biases in visualization and visual analysis. The framework operates on providing a frame of reference when investigating a bias. In this frame of reference, different biases may occur on different levels as well as in between them. In order to investigate biases on different levels, different methods and methodologies are required.

The authors contrast their framework against the Cognitive Bias Codex [47], working orthogonally to its "cause-strategy" logic and instead referring to the different levels of cognitive processing, also taking inspiration from Norman's Human Action Cycle [48], proposing a three-tier model, as demonstrated in

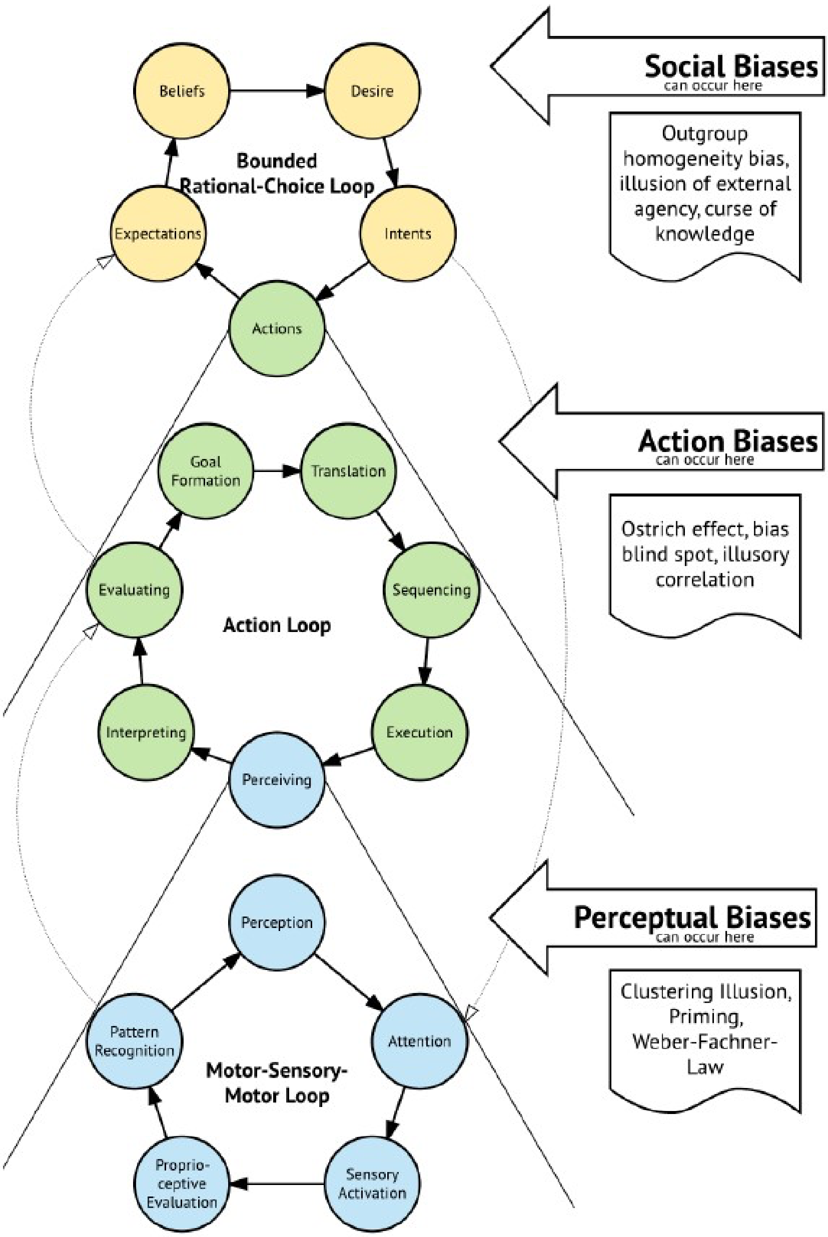


Figure 3.2: Layered closed-loop perception, action and choice model. Since no hard boundaries exist between layers, cross-talk is part of the closed loop model (see exemplary dashed arrows) [2].

Figure 3.2, additionally suggesting possible methods for the study of biases at each of the levels:

- Perceptual Biases, referring to biases occurring on a perceptual level, being based on the motor-sensory-motor loop. Perceptual biases can oftentimes be measured effectively using psychophysics methods;
- Action Biases, referring to biases made in decision making, when perception is adequately mapped to a mental representation yet interpretation or evaluation is distorted, being based on the human-action loop. Action biases measuring methods tend to be far more diverse and tailored to the bias in question;
- Social Biases, referring to biases which affect judgement on a social level, being based on the bounded rational-choice loop. As for Action biases, the methods for measuring Social bias are diverse and even more dependent on the individual bias.

Wall et al. [3] established a conceptual framework for considering bias assessment through human-in-the-loop systems and laid theoretical foundations for bias measurement. The authors describe methods for real-time detection of potentially biased analysis behavior through the measurement of user interaction sequences, proposing six preliminary metrics to systematically detect and quantify bias, these being data point coverage, data point distribution, attribute coverage, attribute distribution, attribute weight coverage and attribute weight distribution, shown in Figure 3.3.

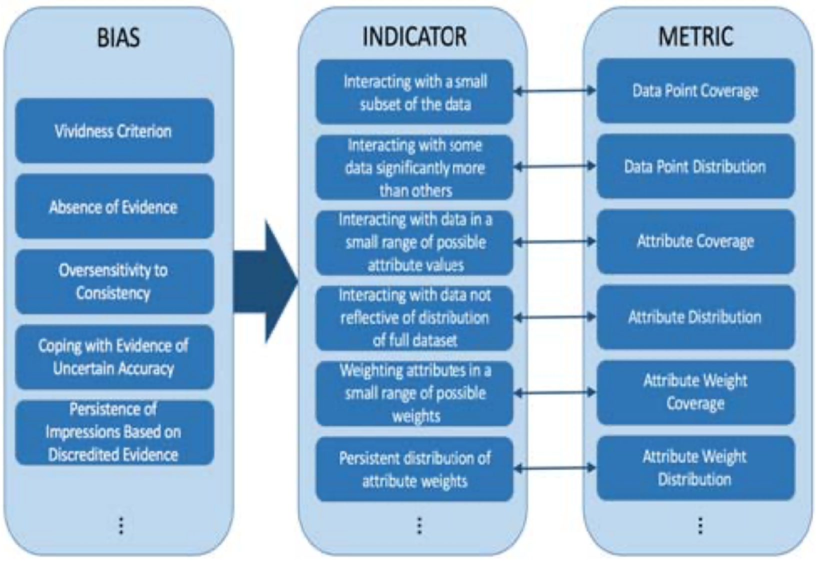


Figure 3.3: Cognitive biases result in behavioral indicators that are measurable by the proposed metrics [3].

When compared to baselines, these six metrics allow for the assessment of meaningful deviations that may reflect cognitive biases, being that each type of bias may impact people's behavior in predictable ways. When a user is biased, we expect to find some patterns in their interactions, denominated as behavioral indicators of bias, yet detecting a particular indicator does not necessarily point to a specific cognitive bias that may have caused the behavioral response. Additionally Wall et al. discuss the ways in which the proposed metrics, constituting an approach to real-time user state assessment, can be used by visual analytic systems with the aim to mitigate negative effects of cognitive biases through the use of three strategies: by making users aware of biased processes by providing them with feedback, by providing feedback to the machines or to external agents.

3.2 Personality and Information Visualization

Information visualization plays an important role in facilitating analytical reasoning by assisting users on drawing insights from abstract data. With visualization achieving larger importance, being used as cognitive aid to solve increasingly more complex and difficult tasks, rises the need to understand how different users think and how to apply visualizations given their individual differences [5]. Individual differences refer to an individual's stable tendencies in the responses to specific stimuli or scenarios in ways that can be predicted [49] and furthermore, personality traits can be defined as the individual differences in people's thinking and behaving characteristics [12].

In the last decades computational sciences have begun to recognize the role that individual differences can have when it comes to the interaction with human-machine systems [17] and research has demonstrated how distinct personality types and cognitive abilities can make a difference in task-solving and behavioral patterns [50–54], as well as on how people perceive design efficiency [55] and accept technologies [56–58].

The interest on extending these findings to the data visualization field is present [5,17,59] and a group of surging research corroborate that personality traits can demonstrate impact on task performance [4, 60], patterns of usage [6, 61] and user satisfaction [62], suggesting that individual characteristics, as much as the visual design, determine visualizations' value and potential.

Green and Fisher [4] explored the impact of individual differences in personality factors on interface interaction and learning performance behaviors. Two studies were conducted, comparing procedural learning behaviors in two interfaces of genomic relationships, an interactive visualization and a menu-driven web table shown in Figure 3.4 and 3.5, respectively. The experiments aimed to explore the impact of personality on the interaction with the interface and user learning performance, resorting to search and inference tasks and collecting measures of accuracy, completion times and insights from the user on their experience, in the end having the participants also classify each interface. The researched focused on Five-Factor Model traits, namely extraversion and neuroticism assessed using the IPIP Mini Big Five Personality Inventory, as well as Locus of Control, assessed using the Internal-External Locus of

Control Inventory. Results demonstrated that all three measures predicted completion times, specifically detecting that highly neurotic participants presented faster task completion times, showing a negative correlation between neuroticism scores and completion time. Additionally, results demonstrated that personality factors also predicted the number of insights participants reported while completing the tasks in each interface, detecting that highly neurotic participants reported fewer insights when compared to participants with lower neuroticism scores.

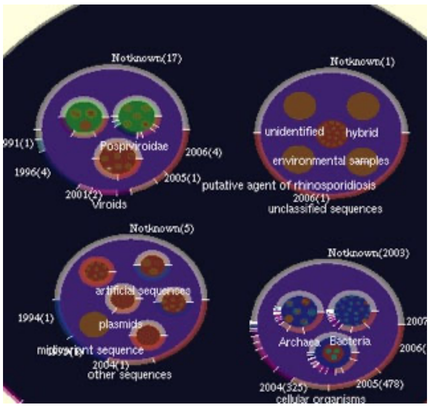


Figure 3.4: The main view of GVis Interface [4].

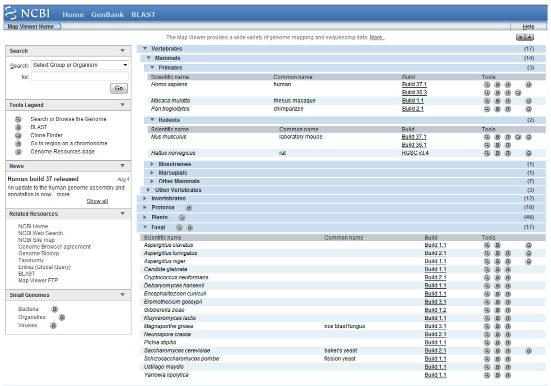
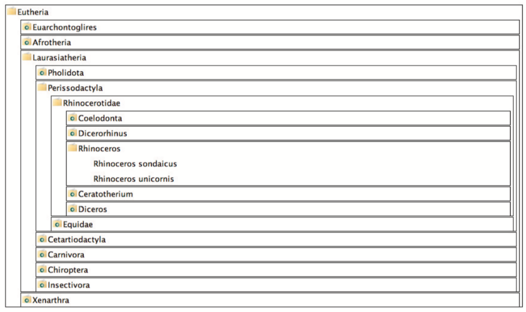


Figure 3.5: NCBI MapViewer Interface [4].

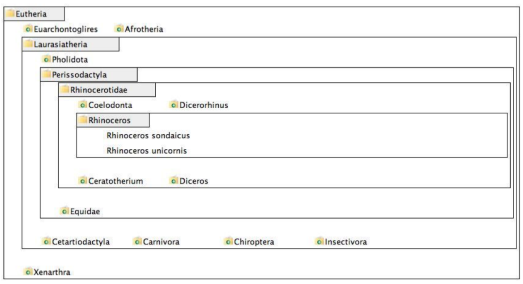
Ziemkiewicz et al. [18] conducted a user study with four visualizations presenting hierarchies between elements, gradually shifting from a list metaphor to a containment metaphor, shown in Figure 3.6. The research compares participants' speed, accuracy and layout preference throughout search and inference tasks with their Locus of Control, extraversion and neuroticism scores, assessed through the Locus of Control and Big Five Personality inventories, aiming to analyse the relationship between visualization layouts and personality factors and extending the aforementioned Green et al. [4] research. The results showed that highly neurotic participants demonstrated higher accuracy in tasks, setting a positive correlation between neuroticism scores and accuracy, being increasingly more accurate as layouts shifted from more list-like to more container-like, with participants with lower neuroticism scores demonstrating the exact reverse. The authors discuss that their findings may be due to neurotic participants possibly pressuring themselves to perform well and being more attentive, consequently more easily grasping unfamiliar layouts. Furthermore, the study corroborated that visualization layout is the key variable that determines the interaction between Locus of Control and compatibility with different system designs.

Brown et al. [6] analysed the ways in which participant interaction data may provide information about themselves and their personality to the machines, conducting an experiment using a spatial visualization in which participants were instructed to complete a visual search task consisting in finding Waldo, as shown in Figure 3.7. For this experiment the authors studied the five dimensions from the Five-Factor

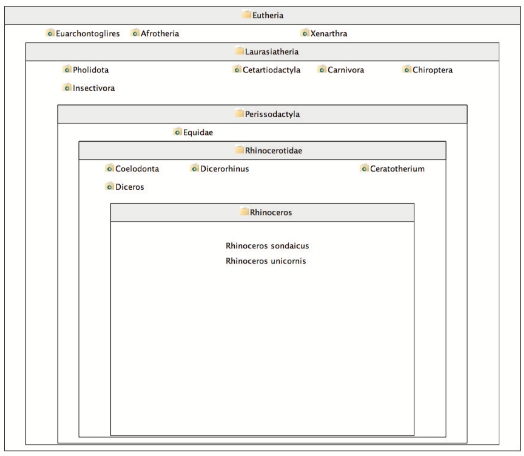
(a) V1: Basic Tree.



(b) V2: Bordered Tree.



(c) V3: Indented Boxes.



(d) V4: Nested Boxes.

Figure 3.6: The four visualizations used in Ziemkiewicz et al. [5].

Model as well as Locus of Control, both assessed through a twenty-seven-question survey, with five Locus of Control questions and twenty Five-Factor Model questions. Throughout the task interaction data such as time, mouse clicks and mouse activity were retrieved and used through machine learning techniques in order to predict participants' task performance as well as their personality factors. Overall, findings showed that interactions can provide information to the computer about its human collaborator, predicting users' task performance and personality traits, and suggest a foundation for realizing mixed-initiative visual analytic systems.

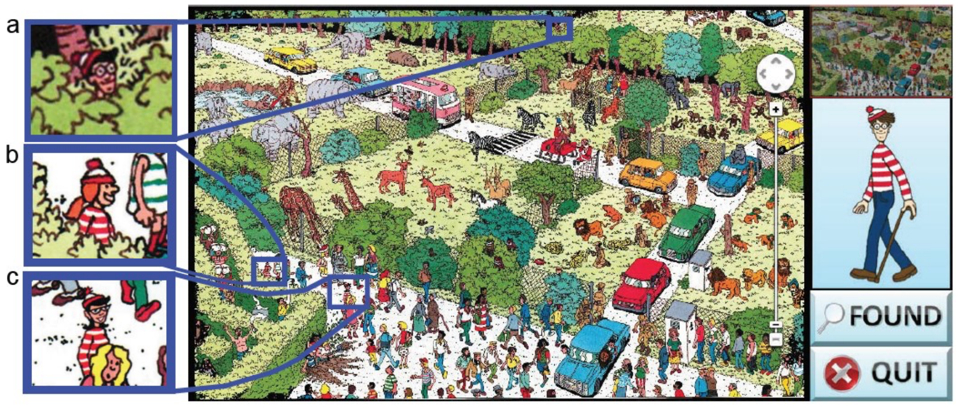


Figure 3.7: The interface from the study conducted by Brown et al. [6] in which participants found Waldo while mouse interactions were recorded.

3.3 Discussion

Visualizations combine the complementary strengths of both humans and machines in systems that involve a human-in-the-loop. The aforementioned studies pose as a significant foundation for our work, analyzing how both individual differences and human biases might affect user interaction with visual analytic tools.

In general, there seems to be a substantial lack of discussion focused on how users make decisions with data presented in a visualization or even how to design visual analytic tools taking into account decision tasks [35] as well as little consideration given to the ways inherent cognitive biases might shape the visual analytic process [3]. In regards to works concerning Distinction Bias in the InfoVis context, Dimara et al. [1] describe the relation of this cognitive bias with visualization research as "not discussed in visualization but likely relevant". However, the small body of literature surrounding the study of specific biases in the context of visualizations appears to advocate for an approach of first verifying the persistence of the studied bias effects in visualizations, basing experiments on previous psychology studies that found clear results, later generalizing the procedure to explore aspects we can deem as relevant in the visualization context [43].

The aforementioned studies surrounding personality psychology in regards to information visualization launched pertinent findings around neuroticism in the visualization context. Ziemkiewicz et al. [18] observed highly neurotic users to have higher accuracy rates, specially when dealing with tasks revolving around unfamiliar layouts. The authors point possible explanations as neurotic individuals putting more pressure on themselves to perform well on tasks, as opposed to abandoning them due to unfamiliar visualizations. Furthermore, neurotic individuals were observed to take less time in search and inference tasks in the experiment conducted by Green and Fisher [4]. Ziemkiewicz et al. [18] as well as Green and Fisher [4] suggest that these results can be explained by a higher attention of neurotic individuals when dealing with some problem-solving tasks, resulting in a better adaptation. Additionally, Brown et al. [6] observed how personality traits can be correlated with mouse activity, by having machine learning techniques being able to make use of task interaction data to predict, with variable accuracy, participant performance and personality traits.

It is also important to note that Neurotic individuals have a tendency to easily experience negative emotions [15, 18] tending to also be more pessimistic [15, 63], an important concept to take into account as Distinction bias literature deals with utility and happiness forecasting.

Many researchers agree that personality impacts several aspects of our life and therefore affects the way we live and make our decisions [11]. This could suggest correlations between certain traits and behavior that can be considered prone to a specific bias. However, information and empirical work around Distinction Bias is still scarce, and possible correlations with personality psychology are still undetermined. In the context of our work we can hypothesize how personality psychology findings regarding neuroticism may correlate with the different evaluation modes of the Distinction Bias context.

4

Methodology

Contents

Cognitive biases have been gaining relevance in the field of information visualization due to their importance when it comes to decision support systems. Even so, as discussed in Section 3.1, studies involving empirical work on the link between cognitive biases and visualizations are still rare, especially ones that focus on a specific bias.

Distinction Bias stands as a cognitive bias that, although unexplored in the information visualization field, is seen as likely relevant [1]. The purpose of our research began with exploring whether the Distinction Bias effects are still present when information is presented through InfoVis graphical encodings since, to the best of our knowledge, Distinction Bias studies are few and appear to not be linked to Information Visualization in any way.

As stated in Section 2.1.1, Distinction Bias describes the way people's utility and affective predictions for several alternatives can vary depending on the evaluation mode they find themselves in, having these predictions then possibly lead to sub-optimal decisions [10].

Across two studies Hsee and Zhang [10] tested and supported this new source of misprediction. Each study started by asking participants to consider they would find themselves in a given hypothetical situation. Taking it into account, participants were then asked to consider either several alternative outcomes for this situation, for those in JE mode conditions, or a single one, for those in SE mode conditions. Some of these possible outcomes would differ quantitatively while others qualitatively. Participants in JE mode conditions were instructed to predict their happiness levels for each individual outcome. Participants in SE mode, faced with only one outcome, would have to, in some situations, predict their affect while in others actually experience it and then assess their happiness as well. It is worth noting that both SE mode predictors and experiencers arrived at similar results, supporting that Distinction Bias effects apply both to predictions and actual experiences.

Comparing the results, Hsee and Zhang [10] studies demonstrated how people in JE mode tend to overpredict experiential differences between quantitative attribute differences but are not as likely to do so when these differences are only qualitative, when compared to people in SE mode. Essentially this means that, for example, someone in JE mode would be likely to overpredict the difference in happiness between earning $60,000 and $70,000 a year, yet less likely to overpredict the difference in happiness between working an interesting job and a tedious job [10].

In order to reach the goal of our study, we took Hsee and Zhang [10] studies and findings as inspiration, crafting the intended visualizations and tasks as models that aim to replicate and test the effects of this bias in an InfoVis context. With this work we aim to shed light on the Distinction Bias and study its possible relevance in the field of visualizations, being an opportunity to explore a little discussed yet ripe topic.

Likewise, as individual differences have an effect on how we perceive and process information and therefore may impact the way we live and make our decisions [11], we also found pertinent to incorporate the study of personality and its effects on bias-prone scenarios in our work, additionally focusing on the impacts of a specific personality trait, neuroticism.

Our study is comprised of two experiments, each composed of a set of tasks. The first experiment aims to adapt the existing Distinction Bias literature studies into the information visualization realm in order to understand the ways that the identified effects may transfer to visualizations. For each task participants are asked to examine information regarding hypothetical outcomes of a given life scenario, presented through a visualization. The second experiment further expands our exploration by introducing variation of the amount of outcomes being presented for each task, as well as the scale in which they are displayed.

This section introduces the methodology for the study. We start by laying down the thought process behind the visualizations, moving on to describing the tasks that were performed in the study, as well as the required measures to be gathered. We then go into detail on the research questions and hypothesis, followed by a description of our data collection process and an explanation of the data analysis.

4.1 Visualization

Taking into consideration the gap of research regarding Distinction Bias within the InfoVis context [1] together with, to the best of our knowledge, the absence of prior work specifically linking visualization idioms to the study of Distinction Bias, for this work we opted to make use of a bar chart idiom to encode the data being presented in our visualization, as illustrated in Figures 4.1 and 4.2.

The choice of the idiom was motivated by its simplicity and the likelihood that bar charts are a well-known idiom to participants [34], as well as its suiting core, standing as an idiom that is defined as fitting for encoding quantitative value attributes with one spatial position channel [34]. Furthermore, the aligned spacial position channel is described as the most effective for ordered attributes and found to consistently offer good perceptual precision compared to other idioms [34,64]. Additionally, bar charts were observed to have similar results compared to showing data to users without the use of visualizations and when compared to other well-known visualization idioms, were shown to provide the most relevant findings in the InfoVis context [40]. Bar charts are also reported as being well suited for the abstract tasks of looking up and comparing individual values which, in addition to the dimension of our datasets being small enough not to run into issues regarding scalability, make it a good fitting idiom choice for the prepared tasks [34].

In order to study the Distinction Bias effects, the layouts for our first experiment simulate each of the two evaluation modes: SE mode and JE mode. In this sense, for a given user, through the course of several tasks, the presented visualization is comprised of a bar chart that either displays one single bar as in Figure 4.1, encoding a value referent to a specific outcome of a given scenario, constituting the SE mode, or several as in Figure 4.2, constituting the JE mode.

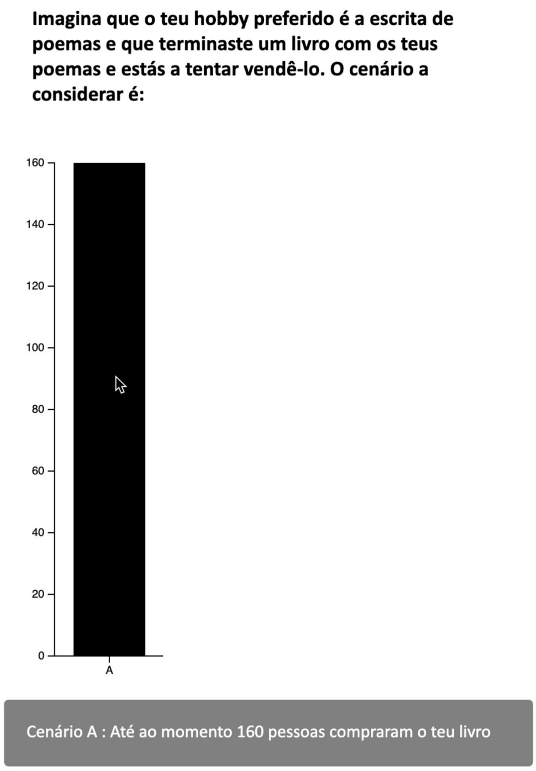


Figure 4.1: Visualization layout for an SE mode task, presenting only 1 outcome. The mouse is hovering over the bar revealing more information on the selected scenario.

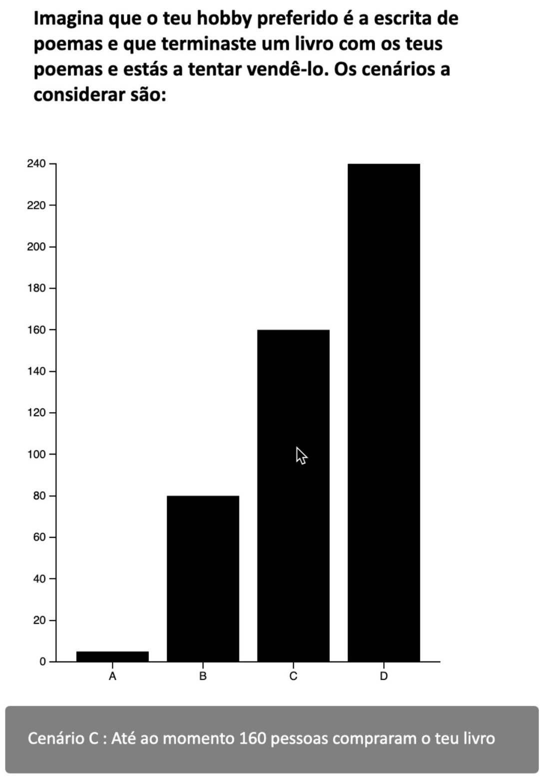


Figure 4.2: Visualization layout for a JE mode task, presenting 4 different outcomes. The mouse is hovering over a bar revealing more information on the selected scenario.

We opted to not introduce color to our bar chart layout, as we are dealing with a bias in which cognition is influenced by comparison, and in this regard the monotone layout allows us to work with the largest color contrast levels available.

The crafted visualization reveals a different bar chart for each of the tasks, titled according to the task's specific hypothetical scenario being studied. Each of the bars in the x-axis represents a possible outcome to the presented scenario, identified with letters on the x-axis. A full overview of tasks' scenarios and respective outcomes can be found in Section 4.1.1.

The length of the bars encodes the attribute values referent to each outcome presented. The presented outcomes may differ from each other quantitatively as well as qualitatively. The valence of each outcome is depicted by representing its attribute value as either positive or negative on the y-axis, as exemplified in Figure 4.3. The y-axis represents the scale of quantitative values being depicted. The y-axis range is adjusted to the values being encoded at each given moment throughout the first experiment. Hovering the mouse over a bar reveals a text box containing the value encoded, as well as more context regarding the selected outcome, as also depicted in Figure 4.3.

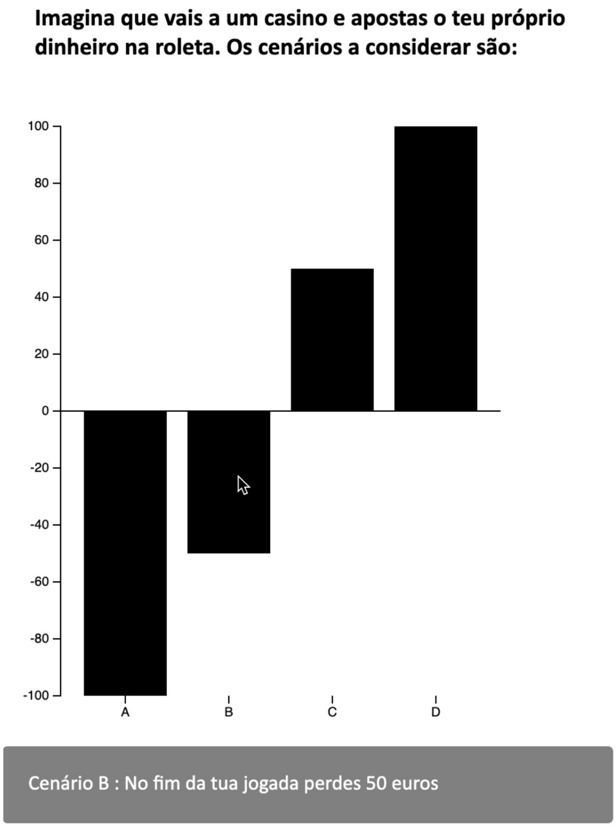
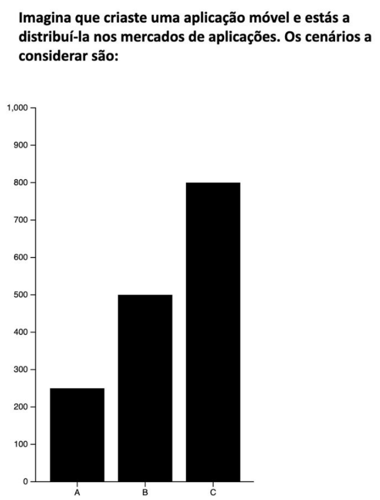
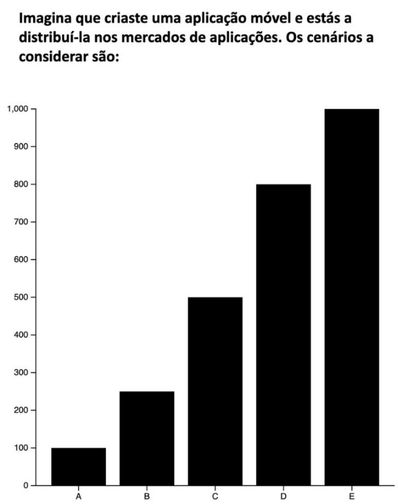


Figure 4.3: Visualization layout for a task exemplifying how valence is encoded. Outcomes A and B differ only quantitatively, representing a different quantity of money lost from a casino bet, while outcomes B and C differ only qualitatively, representing the same amount of money either lost or gained from a casino bet.

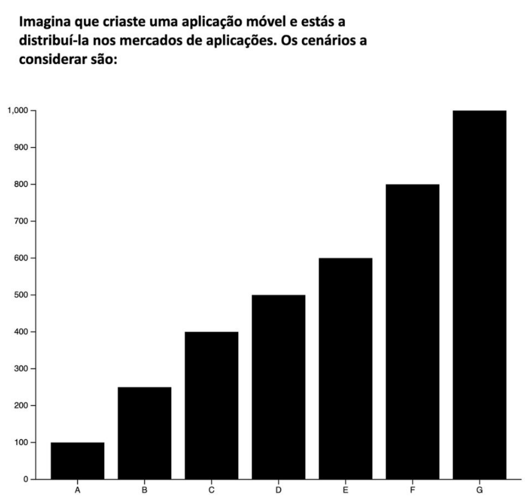
With our second experiment we further explore how the Distinction Bias manifests itself in JE mode, in relation to the number of presented outcomes that differ only quantitatively, as these are the experiment scenarios that tend to lead to biased predictions. As all participant conditions take place in JE mode for the second experiment, each are presented with several outcomes for each of the tasks. Yet, between participants, the number of bars displayed will vary, as illustrated in Figure 4.4, being each participant presented with either 3, 5 or 7 different outcomes for the entirety of tasks in the second experiment, allowing us to explore possible effects between the number of presented outcomes and the Distinction Bias effects.



(a) Visualization layout presenting 3 outcomes.



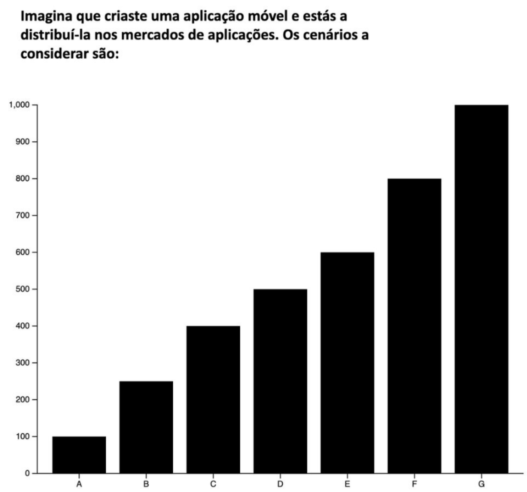
(b) Visualization layout presenting 5 outcomes.



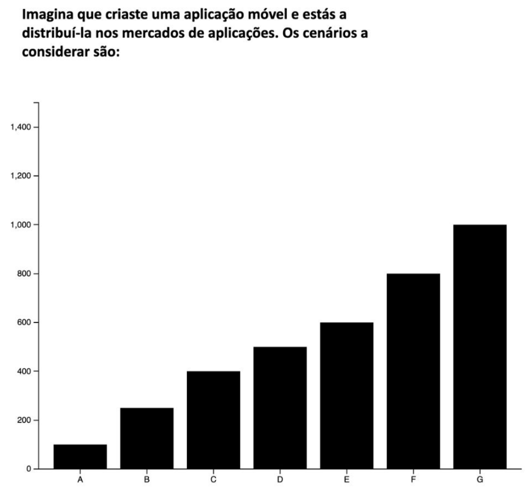
(c) Visualization layout presenting 7 outcomes.

Figure 4.4: Visualization layout presenting 3, 5 and 7 outcomes, respectively.

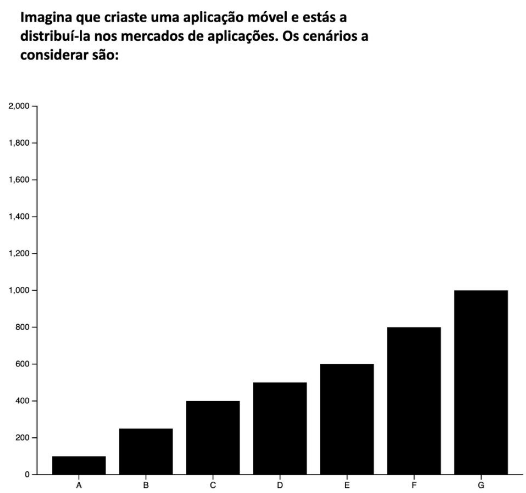
Additionally, with our second experiment we also explore the possibility that the scale in which the data is displayed may influence biased behavior in our study. We found this factor to be pertinent as we are exploring a bias in which cognition is influenced by comparison not only of several options but also with perceived baselines.



(a) Visualization layout featuring the smaller axis range level.



(b) Visualization layout featuring the medium axis range level.



(c) Visualization layout featuring the larger axis range level.

Figure 4.5: Visualization layout presenting increasing ranges for the vertical axis for a specific task.

Therefore, in order to study the repercussions of this variation of scale, the bar chart visualization counts with a vertical axis maximum value that will also vary between three possible levels, the maximum attribute value of the presented task and two progressively larger values picked arbitrarily for each task, as illustrated in Figure 4.5. A single vertical axis condition level was assigned to each participant session, consistent throughout the entirety of the second experiment tasks of that session.

4.1.1 Tasks

For each of our two experiments participants are presented information through the visualizations detailed in the previous section and for each task, as in the Hsee and Zhang [10] studies, are asked to imagine a hypothetical scenario, having to forecast their happiness levels for the different possible outcomes, rating each item from 1 (extremely sad) to 9 (extremely happy) as assessed in prior literature [10], as demonstrated in Figure 4.6.

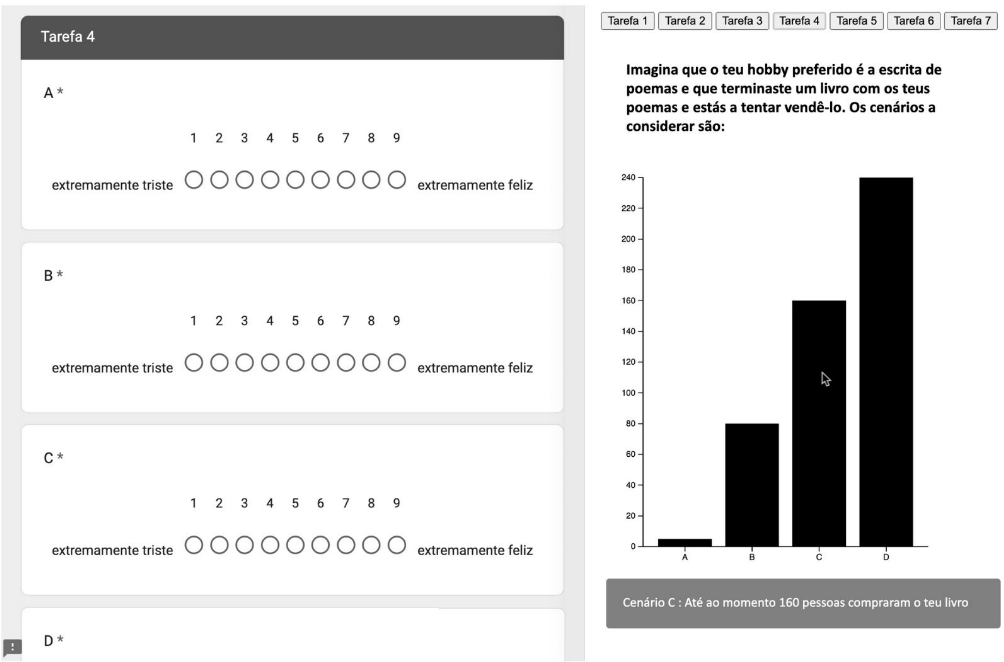


Figure 4.6: Experiment layout for a task. Participants interact with the bar chart visualization and forecast their happiness for each one of the presented outcomes on the form.

Our two experiments consist of a total of 7 tasks. As both experiments are held sequentially with the same participants, task order for each of the two experiments was randomized and tasks were then labeled sequentially for each participant session so as to not confuse participants or unnecessarily bring awareness to this fact. The first experiment is composed of 4 tasks, labeled from 1 to 4 and the second experiment adds 3 tasks, labeled from 5 to 7. The task sequence for a participant session is illustrated in Figure 4.7.

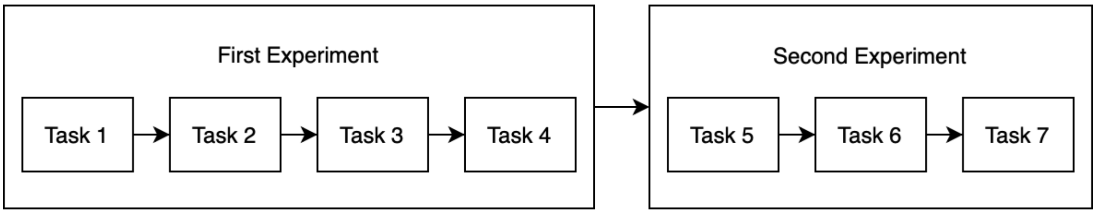


Figure 4.7: Diagram of the task flow for a participant session.

Furthermore for both experiments, participants not being presented with the entirety of the prepared outcomes have their presented outcomes re-labeled with sequential letters.

For our first experiment, in order to better understand the effects of Distinction Bias in the context of visualizations, we made use of data heavily inspired by the outcomes tested in the original Hsee and Zhang [10] misprediction studies. Furthermore, throughout Hsee and Zhang [10]'s discussion of findings, some additional concepts were briefly brought to light, which we also took as inspiration to design some additional tasks. In the end, we arrived at four different tasks, each requiring participants to make predictions of their affective states for hypothetical outcomes:

- **Task 1.1.** Participants are asked to imagine that their favorite hobby is writing poems and that they have compiled a book of their poems and are trying to sell it. The outcomes to consider are:
  - **A.** So far 5 people have bought your book.
  - **B.** So far 80 people have bought your book.
  - **C.** So far 160 people have bought your book.
  - **D.** So far 240 people have bought your book.
- **Task 1.2.** Participants are asked to imagine that they are requested to read a list of words. The outcomes to consider are:
  - **A.** Read a list of 25 negative words, such as hatred and loss.
  - **B.** Read a list of 10 negative words, such as hatred and loss.
  - **C.** Read a list of 10 positive words, such as love and win.
  - **D.** Read a list of 25 positive words, such as love and win.
- **Task 1.3.** Participants are asked to imagine that they went to a casino and gambled their own money. The outcomes to consider are:
  - **A.** You lost 100 euros on your bet.
  - **B.** You lost 50 euros on your bet.

- **C.** You earned 50 euros on your bet.
- **D.** You earned 100 euros on your bet.

- **Task 1.4.** Participants are asked to imagine that they are enrolled in a class required to finish their desired degree. Evaluation is composed by several assignments, being the final grade obtained by the average of all assignments and they have just handed in the first one and received the following classification:
  - **A.** Your assignment obtained a classification of 4 (out of 20).
  - **B.** Your assignment obtained a classification of 8 (out of 20).
  - **C.** Your assignment obtained a classification of 12 (out of 20).
  - **D.** Your assignment obtained a classification of 16 (out of 20).

It is important to note that the compared attribute in Task 1.4, an assignment grade, is presumed by the original studies authors [10] to be easier to evaluate independently, which in theory may mean that, when presented with this task, participants may be able to predict their affective state in a less biased way, making it a compelling additional situation to test.

Additionally, in order to further explore the Distinction Bias effects in InfoVis, we utilized our second experiment as a way to expand our exploratory analysis and examine how the number of presented alternatives as well as the scale in which they are displayed may affect JE mode predictions. Similarly to the first experiment's JE mode condition, participants were asked to complete several tasks for which they compared and examined several hypothetical outcomes, assessing affective state predictions for each one.

Yet, for each of the tasks in the second experiment, a participant can be presented with either three, five or seven different outcomes for the same task. In each of these tasks three outcomes serve as an anchor to understand whether more options affect user assessment, and only these were later studied. This means that, when the condition only presents three options to the participant, only the anchor options are presented in the bar chart. However, for the five or seven options conditions, we leverage the anchors and include distractor quantitatively different outcomes. Participants faced with five outcomes were presented with the three outcomes that were studied along with an outcome quantitatively larger than any other and an outcome quantitatively smaller than any other. Participants faced with seven outcomes were presented with the five previously described outcomes as well as two additional outcomes quantitatively in between the three anchor outcomes, as illustrated in Figure 4.4 and listed below for each of the tasks.

For the second experiment we analyze participants' affective predictions throughout three different tasks (during the experiment the three studied anchor outcomes for each task were not in any way marked or let known to participants):

- **Task 2.1.** Participants are asked to imagine that they are a musician and are working on composing new songs. The outcomes to consider are:
  - **A.** In the last month they have composed 2 songs.
  - **B.** In the last month they have composed 4 songs (Anchor outcome).
  - **C.** In the last month they have composed 6 songs.
  - **D.** In the last month they have composed 8 songs (Anchor outcome).
  - **E.** In the last month they have composed 10 songs.
  - **F.** In the last month they have composed 12 songs (Anchor outcome).
  - **G.** In the last month they have composed 15 songs.
- **Task 2.2.** Participants are asked to imagine that they have a hairdresser appointment and are planning not to take long. The outcomes to consider are:
  - **A.** The appointment got delayed 10 minutes.
  - **B.** The appointment got delayed 30 minutes (Anchor outcome).
  - **C.** The appointment got delayed 40 minutes.
  - **D.** The appointment got delayed 60 minutes (Anchor outcome).
  - **E.** The appointment got delayed 80 minutes.
  - **F.** The appointment got delayed 90 minutes (Anchor outcome).
  - **G.** The appointment got delayed 120 minutes.
- **Task 2.3.** Participants are asked to imagine that they have created a mobile application and are distributing it on the app markets. The outcomes to consider are:
  - **A.** This month you generated 100 euros.
  - **B.** This month you generated 250 euros (Anchor outcome).
  - **C.** This month you generated 400 euros.
  - **D.** This month you generated 500 euros (Anchor outcome).
  - **E.** This month you generated 600 euros.
  - **F.** This month you generated 800 euros (Anchor outcome).
  - **G.** This month you generated 1000 euros.

4.2 Measures

Taking into consideration the previous literature work analyzed in Chapter 3, we collected data not only regarding the measures that can allow us to evaluate and study Distinction Bias in the InfoVis realm but also regarding participants' personality and their interaction with the visualization, such as hover events and completion times. The variables analyzed during the first experiment are represented in Table 4.1.

Table 4.1: Variables to measure in the first experiment.

Name	Type	Dependency	Description
Affective Forecast	Quantitative Ratio	Dependent	User's prediction of happiness (1,2,3,4,5,6,7,8,9)
Affective Difference	Quantitative	Dependent	Computed difference between two affective forecasts of the same participant
Condition	Categorical	Independent	User's evaluation mode (JE,SE1,SE2,SE3,SE4)
Hover	Quantitative Ratio	Dependent	Quantity of hovers per task
Neuroticism	Quantitative	Independent	Individual neuroticism score
Task	Categorical	Independent	Task identifying number (1,2,3,4)
Time	Quantitative Ratio	Dependent	Time (s) taken to make all predi-ctions referent to a given task

In order to reach the goal of our first experiment, participants are asked to complete four tasks, randomized in order for each participant session, as mentioned in Section 4.1.1.

To study the effects of Distinction Bias for each one of these tasks we must simulate the contrasting evaluation modes, JE mode and SE mode, achieved with the visualizations detailed in Section 4.1 and detailed in Figures 4.1 and 4.2. Moreover, this means that, for each of these tasks, participants are assigned to one of five conditions, one for JE mode, in which participants are faced with all four outcome scenarios, and four for SE, in which participants are faced with only one of the possible outcome scenarios.

Participants of all conditions analyze their visualization and assess their happiness levels, denominated affective forecasts, for the outcome(s) displayed. Similarly to Hsee and Zhang [10]'s studies we have participants forecasting their happiness for each presented scenario in a 9-point scale, from 1 being extremely unhappy to 9 being extremely happy.

Following the sessions, the differences between the reported affected forecasts for consecutive outcomes are computed for each participant, as this derived measure is instrumental for our examination of the Distinction Bias effects and their possible link to some of our measures.

The time each participant takes to complete each given task's set of predictions and the amount of times they hover with their mouse cursor over the visualization bars for a given task is also measured.

As mentioned prior, alongside Distinction Bias, our study also explores the effects of a specific personality trait, neuroticism, within the InfoVis context. For our study the neurotic trait is evaluated in the context of the FFM.

When it comes to our second experiment, some additional variables are required. The variables analyzed during the second experiment are represented in Table 4.2.

Table 4.2: Variables to measure in the second experiment.

Name	**Type**	**Dependency**	**Description**
Affective Forecast	Quantitative	Dependent	User's prediction of happiness (1,2,3,4,5,6,7,8,9)
Affective Difference	Quantitative Ratio	Dependent	Computed difference between two affective forecasts of the same participant
Axis Max	Categorical	Independent	Maximum value of y-axis
Hover	Quantitative Ratio	Dependent	Quantity of hovers per task
Neuroticism	Quantitative	Independent	Individual neuroticism score
Outcome Amount	Categorical	Independent	Quantity of outcomes displayed (3,5,7)
Task	Categorical	Independent	Task identifying number (1,2,3)
Time	Quantitative Ratio	Dependent	Time (s) taken to make all predictions referent to a given task

Firstly, this experiment is set entirely in JE mode, and as such every participant evaluation mode condition is equivalent, and therefore the condition measure is no longer necessary in this context.

Additionally, with this study, we intend to analyse how the number of presented outcomes for a given task can impact the bias effects, by having different quantities of outcomes being presented to the various participants for the same tasks, setting three between-subjects groups for which, in each task, participants can be presented with either 3, 5 or 7 quantitatively different outcomes.

Furthermore, with the second experiment we also intend to explore how visualization design may interfere with Distinction Bias, specifically the repercussions of an identified possible spot of concern when dealing with a bias in which cognition is influenced by comparison with perceived baselines, the chosen visualization's vertical axis range. To achieve this we set three between-subjects groups for which, in each task, participants can be presented with a bar chart visualization with a y-axis maximum value consisting of one of three possible values: the maximum attribute value of the presented task and two progressively larger values picked arbitrarily for each task. For our study we settled on the following levels for the vertical axis maximum value, adjusted to the values encoded in each of the tasks: (15,18,20) for Task 2.1, (120,150,180) for Task 2.2 and (1000, 1500, 2000) for Task 2.3.

4.3 Research Questions and Hypothesis

As described by Dimara et al. [1], Distinction Bias stands as likely relevant in the Information Visualization realm, but not yet discussed. As such, our study took an exploratory approach. By taking inspiration from the available literature on Distinction Bias [10] our experiments start by posing the following research question:

RQ1. Does the Distinction Bias transfer to Information Visualization?

Specifically, with our first experiment we intend to analyze and compare the effects that simulating the two opposing evaluation modes, JE and SE mode, within our bar chart visualization can have on people's happiness forecasts. Thus, considering the possible relevancy of Distinction Bias in the context of visualizations, we derived some hypotheses under the aforementioned research question:

H1. The evaluation mode will have an impact on participants' affective forecasts.

Furthermore:

H1.1. Participants in JE mode will tend to overpredict quantitative differences of affect, when compared to participants in SE mode.

H1.2. Participants in JE mode will tend to not overpredict qualitative differences of affect when compared to participants in SE mode.

As Distinction Bias literature points out, when people are presented with several outcomes for a scenario, what we call JE mode, they tend to compare the presented options in order to predict the utility and level of happiness each outcome could bring them. This tends to result in, for quantitatively different values of a compared attribute, significantly differing affective predictions. In contrast, people in SE mode consider only one outcome, not comparing it to other options, tending to, for quantitatively different attribute values of the same valence, predict small differences of utility and happiness, if any. In circumstances in which outcome values differ only qualitatively, affective predictions reported by both people in JE mode and SE mode are observed to differ significantly [10]. These findings are what characterize what we denominate as Distinction Bias and are illustrated in Figure 2.1.

This essentially means that the differences recalled in JE mode are many times actually inconsequential. As related literature suggests, we believe that this effect may transfer into the realm of visualizations and for this we hypothesize that our results should confirm it, detailed in **H1.1** and **H1.2**, as such:

When comparing differences between affective forecasts from participants in JE mode conditions, these should tend to be larger when compared to the differences between affective forecasts of participants in SE mode conditions, for exclusively quantitatively different outcome values, as described in **H1.1**. Additionally in these cases we expect to find statistically significant differences between outcome predictions in JE mode, in contrast to SE mode, as literature points that the evaluation function for SE mode tends to flatten when distant enough from the baseline, as illustrated in Figure 2.1.

Yet, when comparing exclusively qualitatively different outcome values, differences in affective forecasts from participants in JE mode conditions should not tend to be larger than for participants in SE mode conditions, as described in **H1.2**. Additionally in these cases we expect to find statistically significant differences between affective forecasts from both evaluation modes, as literature points that both modes' evaluation functions tend to be steep when close to the baseline, as illustrated in Figure 2.1.

Furthermore, now taking into account the goals for our second experiment, under the aforementioned research question **RQ1** we derived the following hypotheses:

H2. The amount of quantitatively different outcomes being presented will have an impact on the differences between affective forecasts of anchor outcomes.

H3. The scale in which information is presented in the visualization will have an impact on the differences between affective forecasts of studied options.

Distinction Bias literature denotes that SE mode predictions of happiness level do not follow the steep curve of JE mode predictions for quantitatively different values of an attribute, instead appearing as much more flat in these situations, as evidenced in Figure 2.1. This theory seems to suggest that by variating the number of presented alternatives in JE mode we may expect different results from JE mode predictors regarding the same outcomes, as summarized in **H2**.

Additionally, a possible visualization concern when dealing with a bias in which cognition is influenced by comparison with perceived baselines would be the scale in which different outcomes are displayed in, in the case of our visualization the range of the vertical axis (y-axis), as comparisons between the displayed options and other possible options ideated by predictors may also impact their judgement. This is an hypothesis we also tested, as described in **H3**.

One additional factor was also at play in Task 1.4, the ability to independently evaluate an attribute. Hsee and Zhang [10] discuss that Distinction Bias theory applies mostly to attributes that are not too easy to evaluate independently, yet believe that finding an attribute that is notably easy to evaluate independently is the exception rather than the rule. The authors state school grades as an example of an attribute which people have sufficient knowledge about, resulting in prediction values, when assessed in SE mode, being close to those assessed in JE mode. We took these statements as inspiration to come up with Task 1.4, for which the aforementioned **H1** might not be verified, posing the question:

RQ2. Do the effects of Distinction Bias on predictions persist when information is presented in a bar chart for an attribute that is independently easier to evaluate?

Additionally, our study integrates a component of personality analysis, namely of participants' neuroticism scores. Highly neurotic individuals have a tendency to easily experience negative emotions and to more easily feel stressed, anxious and depressed [15,18], tending to also be more pessimistic [15,63]. According to related literature, participants with higher neuroticism scores tend to feel more pressured to complete tasks with accuracy, being more attentive [65], having been observed to have higher success rates than their lower scoring peers, specially when dealing with tasks revolving around unfamiliar environments [18]. Highly neurotic users were also observed to focus more on negative information than on positive information [66] and have harder times making decisions [67, 68], yet individuals with high neuroticism scores were observed to take less time in search and inference tasks [4]. Furthermore, Brown et al. [6] denoted a correlation between personality traits such as neuroticism, and mouse activity, like hovers.

For the study of participant personality and its repercussions in the context of our study we analyzed the existing literature findings, defining another research question:

RQ3. Do participants' neuroticism scores impact their affective forecasts?

Under the aforementioned research question, we derived the following hypothesis:

H4. Neuroticism scores will have an effect on the differences between affective predictions.

As we request participants for happiness predictions we anticipate the possibility that participants with a higher neuroticism score could generally predict lower values, given literature findings. Yet, with the focus of our research revolving around the effects of Distinction Bias, it is relevant to test if participants with higher neuroticism scores predict their affect for the presented scenario outcomes consistently and proportionally lower in value, which should not in theory significantly impact the affective prediction differences between two given presented outcomes, or if, being primed to incur in Distinction Bias, the neurotic trait displays correlations with the overprediction of these differences, specifically when attribute values differ merely quantitatively.

Additionally, given the diversity of related literature findings regarding user interaction data, we subsequently took the opportunity to study some of the possible effects of the studied trait in a bias-prone context through a more general approach, hypothesizing that:

H5. Neuroticism scores will have an effect on the time participants take to make their predictions in a bias-prone context.

H6. Neuroticism will have an effect on the number of hovers performed by participants in a bias-prone context.

4.4 Data Collection

Throughout the Data Collection Section we examine the participants and their recruitment process, the apparatus used for the sessions, and guide the conducted procedure for each of the sessions.

4.4.1 Participants

Our study was conducted for a total of 82 participants. Participants were recruited through standard convenience sampling procedures including direct contact and through word of mouth. Due to unforeseen circumstances with the handling of our apparatus on two separate occasions, the data for the held sessions was not saved and, since this posed such a small subset and no dependent measures were planned to be collected manually by us, we opted to discard the data for these two participant sessions instead of rebuilding it from video recordings.

The data used for our study was collected from a total of 80 participants (20 females, 57 males and 3 other) between the ages of 18 to 27 years of age (M = 21.71, SD = 2.425). However we were only able to collect personality data referent to the NEO PI-R [69] from 58 of these participants (18 females, 38 males and 2 other). Inspection of boxplots confirmed that no outliers were found for our samples.

4.4.2 Apparatus

In order to collect participants' neuroticism scores, we made use of the European Portuguese version of the NEO PI-R questionnaire [69]. The questionnaire is composed of 240 items, with 48 items per each of the five broad traits, and more specifically eight items per facet of each broad trait.

At the beginning of each experiment, participants were presented with a consent form, being asked to fill out their consent to collect their data both relative to personality as well as the data to be collected during the session. The data collected during the session consists of the full set of dependent variables required as our study's measures, defined in Section 4.2, as well as audio and screen recordings of the session, collected as a safety measure, allowing us to retrieve the session's measures manually in case of possible malfunctions as well as verify unusual retrieved data. The consent form was further utilized to collect the participant's email address as well as some demographic information such as age and gender. A copy of the form was emailed to the participant upon completion. The consent form was developed with Google Forms and can be consulted in Appendix A.

Participants were then also provided with a tutorial animated gif image, being explained the layout they would be presented with for the experiment, how to interpret the data, as well as the course of action regarding what they would be asked to do. In order to prevent the effects of other biases the tutorial featured data unrelated to the prepared tasks.

An additional Google Form was utilized in order to collect participants' affective forecasts for the several tasks of the session, each assessed on a 9-point scale from 1 being extremely unhappy to 9 being extremely happy, as assessed in prior examined literature [10]. Given the several variations of participant conditions for the tasks referent to both the first experiment as well as the second one, several versions of these forms were developed, in order to feature the appropriate number of affective forecast inquiries for each of the tasks in each of the held sessions. An example experiment form can be consulted in Appendix B and seen in the context of our experiment layout in Figure 4.6.

The visualization layout was developed in order to automatically collect the remaining dependent variables not yet assured by the described Google Forms, such as each of the task's completion times, the quantity of bar hovers performed for each of the tasks, the order in which the tasks were being performed as it had been previously randomized, as well as the visualization's independent measures for the specific session, such as experiment one's condition and experiment two's number of presented outcomes and axis range. The data was then exported into a comma-separated value file at the end of each session. Participant sessions were conducted both remotely through Zoom conference sessions as well as in person. The study setting was tested and verified not to have an effect on the study's dependent variables. Participants completed the experiments using a computer.

4.4.3 Procedure

As mentioned prior, our sessions were composed of two experiments, performed sequentially by the same participants. All tasks were translated into Portuguese for our participant sessions. In order to keep away from additional biases the order of the tasks was randomized for each of the experiments and re-labeled with sequential numbers for each session, as mentioned prior.

Each session started as the participant agreed to be part of the study, filling out a consent form (Appendix A) that allowed us to collect and analyze their personality data, the data to be collected during the experiments, as well as the session's audio and screen recordings. Additionally, the same form also required participants to fill out some demographic data.

Participants were then inquired if they were familiar with the bar chart visualization idiom. A tutorial was then provided, and participants were informed about what would be presented and required of them during the experiment and how visualizations can be analyzed and interacted with. Participants were then asked if they had any questions regarding the procedure that would follow.

Once the tutorial was finished and participants confirmed that they understood what they were required to do, we would start the audio and screen recording, letting the participant know of this occurrence.

The first experiment then began taking place. Each participant was randomly assigned a condition, consisting of a simulation of either JE mode or SE mode, consistent throughout the first experiment tasks. For each one of four tasks that compose the first experiment, participants were presented with information regarding a specific scenario and asked to imagine this was occurring to them. Facing this hypothetical situation, participants analyzed the visualization's bar chart, which encodes different outcomes for this situation, several in the case of participants in JE mode conditions and only one for participants in SE mode conditions. Participants in JE mode conditions were verbally encouraged to compare the presented alternative outcomes, as a way to better assure their evaluation mode. SE mode participants were not, for the same reason. Participants were then asked to predict their level of happiness, what we refer to as affective forecast, for each presented outcome, assessed through the experiment form (Appendix B), as shown in Figure 4.6.

Similarly to the prior described experiment, the second experiment is also composed of several tasks, for which participants of all conditions were required to imagine a specific scenario, analyse a bar chart idiom encoding different outcomes for this situation and asked to predict their level of happiness for each, once more assessed through the experiment form. For the second experiment all participant conditions are set on JE mode and, as such, all participants were verbally encouraged to compare the presented alternative outcomes to further assure joint evaluation circumstances. As previously mentioned, throughout the second experiment, for a given task, two different participants can be presented with different quantities of outcomes presented in a bar chart visualization in different scales. These variations were randomly assorted for each participant session. Between-subject groups for each of these factors remained consistent throughout the entirety of the tasks of our second experiment.

As the second experiment was finished and the participant completed the total of seven tasks, we stopped the recordings, thanked them for their time and then finally showed ourselves available to take further questions regarding the purpose of our study, as some participants showed interest and at this stage this would no longer influence their involvement. At last, we stored the session data retrieved in real time by our interface, as well as the data collected through the forms.

4.5 Data Analysis

In order to search for answers to our previously mentioned research questions, we leveraged several statistical analysis methods, which we detail in this section.

In preparation to begin our data analysis, the retrieved data from the sessions was combined, joining the data collected in the forms with the retrieved data from our visualization interface for each of the sessions. For this purpose we formatted our study data into a comma-value separated file, with each line containing each participant's session data. We then proceeded to correct a few formatting anomalies, caused by four participants pressing an incorrect interface button, which consequently introduced clutter into their session's data. Throughout the process of removing it, we made sure to verify the accuracy of the remaining data through the session recordings.

To effectively explore Distinction Bias we then derived the affective forecast difference measures for each of the tasks of each participant session. For each of the tasks, if a session data contained multiple affective forecasts, the differences of the values for every consecutive forecast were calculated and stored as new variables. Depending on the task some of these computed differences would be referent to two exclusively quantitatively different outcomes while others to two exclusively qualitatively different outcomes. As noted previously, this is what allows us to explore the Distinction Bias effects.

For this study we utilized several statistical analysis methods. These included multiple types of t-tests, Two-Way Analysis of Variance (ANOVA) and Spearman Correlations. Table 4.3 explicits the correspondence between our hypothesis and the statistical methods used to evaluate them. Adequate Bonferroni corrections were consistently applied for result analysis.

Table 4.3: Statistical tests used to evaluate the study.

Hypothesis	Independent Variables	Dependent Variables	Statistical Test
H1	Condition	Affective Forecast	Paired-Samples T-test and Independent T-test
H2	Outcome Amount and Axis Max	Affective Difference	Two-Way ANOVA
H3	Outcome Amount and Axis Max	Affective Difference	Two-Way ANOVA
H4	Neuroticism	Affective Difference	Spearman Correlation
H5	Neuroticism	Time	Spearman Correlation
H6	Neuroticism	Hover	Spearman Correlation

A further description of each of the tests utilized, as well as our procedure and predictions regarding them is presented throughout this section.

4.5.1 Paired Samples T-test

The paired-samples t-test is used to determine whether the mean difference between paired observations is statistically significantly different from zero. The participants are either the same individuals tested at two time points or under two different conditions on the same dependent variable. This statistical test requires one dependent variable that is measured at the continuous (i.e., ratio or interval) level, as well as one independent variable consisting of two categorical groups. Additionally (a) there should be no significant outliers in the differences between the two related groups, and (b) the distribution of the differences of the dependent variable between the two related groups should be approximately normally distributed.

We leveraged paired-samples t-tests as one of our methods to examine **H1. The evaluation mode will have an impact on participants' affective forecasts** and furthermore **H1.1. Participants in JE mode will tend to overpredict quantitative differences of affect, when compared to participants in SE mode** and **H1.2. Participants in JE mode will tend to not overpredict qualitative differences of affect when compared to participants in SE mode**.

In order to determine any correlations between the participants' evaluation mode condition in the first experiment and their affective forecasts, in ways that are consistent with Distinction Bias theory, we had to run multiple t-tests: paired-samples t-tests for affective forecast data referent to participants in JE mode conditions, as well as independent t-tests in the cases of affective forecast data referent to participants in SE mode conditions, as for each of the first experiment tasks JE mode participants have been presented with several outcomes to assess, while SE mode participants have only been presented one.

For each of the first experiment tasks, we ran a paired samples t-test on each pair of reported affective forecasts referent to consecutive presented outcomes or, in other words, pair of happiness assessments referent to outcomes encoded by bars situated next to eachother in the presented visualization.

Furthermore we analyzed the calculated statistical significance values as well as the obtained mean differences, comparing them between conditions. To confirm our **H1.1** and **H1.2** assumptions, we expected our paired-samples t-tests to be deemed statistically significant, verifying that JE mode affective forecasts were significantly different from eachother. Yet when considering the mean differences obtained from our JE mode paired samples t-tests ran for the same two consecutive outcomes' affective forecasts they should tend to be: larger than corresponding SE mode ones in the cases in which outcomes differ only quantitatively and, smaller when outcomes differ only qualitatively.

The examination of outliers was achieved through inspection of boxplots. Two categories for outliers were taken into account: outliers and extreme points, consisting of data points more than 1.5 and 3 box-lengths from the end of their box, respectively. For our study we only considered extreme points as outliers, as data points situated in between 1.5 and 3 box-lengths from the end of their box in boxplots are reported to not be too troublesome compared to those considered as extreme points[1]. With this criteria we found the occasional outlier, either posing as a slightly unusual value or being part of an uncommon context, such as a situation in which the majority of data points converged in a single value, resulting in almost any other value being flagged as an outlier. In light of this, we decided to run tests

[1] https://statistics.laerd.com/premium/spss/pstt/paired-samples-t-test-in-spss-9.php

with and without removal of these data points, comparing the results. Ultimately we concluded that the results were not materially affected and inevitably pointed us to the same conclusions. As such, we found no good reason to reject these points as invalid and decided not to remove them. For the majority of performed tests the assumption of normality was violated. However, we utilized Normal Q-Q Plots to confirm that data was fairly normally distributed for all cases and, as paired-samples t-tests are robust to deviations from normality[2], we proceeded with the tests.

4.5.2 Independent T-test

The independent-samples t-test is used to determine if a difference exists between the means of two independent groups on a continuous dependent variable and how statistically significant it is. This test requires one dependent variable measured at the continuous level as well as one independent variable that consists of two categorical, independent groups. Additionally the test requires independence of observations as well as (a) no significant outliers in the two groups of your independent variable in terms of the dependent variable, (b) an approximately normally distributed dependent variable for each group of the independent variable, and (c) homogeneity of variances.

We leveraged independent t-tests as one of our methods to examine **H1. The evaluation mode will have an impact on participants' affective forecasts** and furthermore **H1.1. Participants in JE mode will tend to overpredict quantitative differences of affect, when compared to participants in SE mode** and **H1.2. Participants in JE mode will tend to not overpredict qualitative differences of affect when compared to participants in SE mode**.

As previously detailed, in order to determine correlations between the participants' evaluation mode condition in the first experiment and their affective forecasts, in ways that are consistent with Distinction Bias theory, we had to run multiple t-tests: paired-samples t-tests for affective forecast data referent to participants in JE mode conditions, as well as independent t-tests in the cases of affective forecast data referent to participants in SE mode conditions.

For each of the first experiment tasks, we ran an independent t-test on affective forecasts referent to each pair of consecutive SE mode conditions (SE1, SE2, SE3, SE4), this way allowing us to, similarly to the paired samples t-test procedure, examine the differences between happiness assessments referent to consecutive outcomes, even though each SE mode participant had only been presented with one (as each SE mode condition entailed being presented with a different one).

Furthermore, as detailed for the paired-samples t-test, we analyzed the calculated statistical significance values as well as the obtained mean differences, comparing them between conditions. To confirm our **H1.1** and **H1.2** assumptions, our independent t-tests were not expected to be deemed statistically significant, except in the cases in which we would compare affective forecasts referent to two consecutive

[2] https://statistics.laerd.com/premium/spss/pstt/paired-samples-t-test-in-spss-12.php

outcomes that differ qualitatively. In addition, when considering the mean differences obtained from our SE mode independent t-tests ran for two consecutive outcomes' affective forecasts they should tend to be: smaller than corresponding JE mode ones, in the cases in which outcomes differ only quantitatively and, larger when outcomes differ only qualitatively.

The examination of outliers was achieved through inspection of boxplots in a similar way as the one detailed for paired samples t-tests. For the majority of performed tests the assumption of normality was violated. However, we utilized Normal Q-Q Plots to confirm that data was fairly normally distributed for all cases and, as independent t-tests are robust to deviations from normality[3], we proceeded with the tests. For a few of the tested cases homogeneity of variances was not able to be assured. These cases will be detailed in the following Chapter. A modification, denominated Welch t-test, was made to the standard t-test to accommodate unequal variances[4]. We proceeded with both, analyzing results from the Welch t-test for these cases, however always comparing these results to the ones obtained from the standard t-test.

4.5.3 Two-Way Analysis of Variance (ANOVA)

The Two-Way ANOVA is utilized to determine whether there is an interaction effect between two independent variables on a continuous dependent variable (or in other words if a two-way interaction effect exists). For this, we are required to have one dependent variable that is measured at the continuous level and two independent variables where each independent variable consists of two or more categorical independent groups, as well as independence of observations, meaning that there is no relationship between the observations in each group of the independent variable or between the groups themselves. Additionally, (a) there should be no significant outliers in the cells of the design, (b) the distribution of the dependent variable should be approximately normally distributed in every cell of the design, and (c) homogeneity of variances (i.e., the variance of the dependent variable is equal in each group of the independent variable) should be assured for the dependent variable in the cells of the design.

We utilized two-way ANOVAs as the method to test **H2. The amount of quantitatively different outcomes being presented will have an impact on the differences between affective forecasts of anchor outcomes**, as well as **H3. The scale in which information is presented in the visualization will have an impact on the differences between affective forecasts of studied options**.

When it comes to **H2.**, this test allowed us to assess how the number of quantitatively different outcomes presented to participants (for our study examined as a categorical variable with three between-subject groups: presenting 3, 5 or 7 quantitatively different outcomes for the tasks of our second experiment) could influence the affective forecast differences for each of the second experiment tasks (2.1, 2.2

[3] https://statistics.laerd.com/premium/spss/istt/independent-t-test-in-spss-12.php
[4] https://statistics.laerd.com/premium/spss/istt/independent-t-test-in-spss-17.php

and 2.3). Similarly to what was done for **H2.**, as the scale in which data is presented posed as a possible visualization concern when dealing with a bias in which cognition is influenced by comparisons, for **H3.** this test allowed us to explore the influence of the visualization's vertical axis range, in which the task's data was presented (for our study also examined as a categorical variable with three between-subject groups that featured an increasingly higher y-axis maximum value, adjusted to the values referent to each of the tasks) on our participants' affective forecast differences for each of the second experiment tasks (2.1, 2.2 and 2.3).

The examination of outliers was achieved through inspection of boxplots. Similarly to the aforementioned t-test procedure, for our study we only considered extreme points as outliers, as data points situated in between 1.5 and 3 box-lengths from the end of their box in boxplots are reported to not be too troublesome compared to those considered as extreme points[5]. With this criteria we identified a few outliers, but only in unusual situations in which affective difference data points converged into a single value, resulting in any slightly different value being flagged as an outlier. Due to the context of our study we decided to keep these data points. No other outliers were found.

For a few performed tests the assumption of normality was violated. However, the two-way ANOVA is robust to deviations from normality[6] and therefore we proceeded with the tests. Likewise, the assumption of homogeneity of variances was also violated for some of the performed tests. However, the two-way ANOVA is robust to some heterogeneity of variance[7] and therefore we proceeded with the tests.

4.5.4 Spearman Correlation

The Spearman's rank-order correlation calculates a measure of strength and direction of an association/relationship between two continuous or ordinal variables. It is also required that these two variables represent paired observations. Additionally it is required that the two variables present a monotonic relationship.

We leveraged this test to evaluate **H4. Neuroticism scores will have an effect on the differences between affective predictions**, as well as **H5. Neuroticism scores will have an effect on the time participants take to make their predictions in a bias-prone context** and **H6. Neuroticism will have an effect on the number of hovers performed by participants in a bias-prone context**.

When it comes to **H4.**, we were able to examine possible correlations between our participants' neuroticism scores and the previously computed affective forecast differences in a bias-prone context. This way we were able to analyse if our sessions reported any correlations between this trait and Distinction Bias. For this to be the case, we would expect to see correlations between neuroticism scores and affective forecast differences between two affective forecasts referent to consecutive quantitatively dif-

[5] https://statistics.laerd.com/premium/spss/twa/two-way-anova-in-spss-8.php
[6] https://statistics.laerd.com/premium/spss/twa/two-way-anova-in-spss-9.php
[7] https://statistics.laerd.com/premium/spss/twa/two-way-anova-in-spss-10.php

ferent outcomes. Additionally, in case there would be any correlations between neuroticism and affective forecast differences referent to outcomes that vary qualitatively they would be expected to occur in an opposing direction. This would be the case as Distinction Bias theory notes how its effect derives from the link between: overprediction of affective differences between two outcomes that differ only quantitatively, and; the lack of overprediction of affetive differences between two outcomes that differ only qualitatively.

For **H5.** Spearman correlation tests allowed us to understand if there were any correlations between the time that participants in JE mode conditions took on their tasks, both for the first and second experiment, and their neuroticism scores. For **H6.**, similarly to **H5.**, we were able to examine correlations between the amount of hovers performed by JE mode participants for each of their tasks and their neuroticism scores.

For all the mentioned situations Spearman Correlation tests require paired observations of its two variables. In each of the mentioned cases we always made sure that the two measures compared are referent to the same participants and therefore related to the same sessions. Furthermore, in each of the tests we inspected a scatterplot visualization of the two compared variables in order to account for monotonic relationships. For some of the tests we had some difficulty fully pinpointing this relationship, yet believed there was enough evidence to verify the assumption and therefore proceeded.

5

Results

Contents

Throughout this chapter we discuss the results of our study. Utilizing the measures and statistical analysis tests detailed in the previous chapter we tested our assumptions and hypothesis, with the aim to provide some answers to our research questions for the context of our study.

We start by analyzing the Distinction Bias effects in the context of our visualization, with the aim to provide answers to both **RQ1.** and **RQ2.** We then introduce personality factors onto our analysis, with the aim to provide some answers to **RQ3.** Finally, we discuss our results, reporting the main insights obtained. Data are mean $\pm$ standard error, unless otherwise stated.

5.1 Distinction Bias and Information Visualization

In this section we start by presenting the results that focused on providing answers to **RQ1. Does the Distinction Bias transfer to Information Visualization?** For this we will firstly focus on the results drawn from the measures and procedures intended to evaluate the effects of Distinction Bias. We will then present the results regarding the included task that introduced **RQ2. Do the effects of Distinction Bias on predictions persist when information is presented in a bar chart for an attribute that is independently easier to evaluate?** Finally we describe the results regarding the possible influence that the factors introduced and explored for the second experiment tasks had on the studied bias effects.

5.1.1 Distinction Bias Effects

In order to examine the ways in which Distinction Bias effects transfer to Information Visualization, specifically to our developed visualization detailed in Section 4.1, we took on a similar approach as the one utilized for existing Distinction Bias literature. As such, we made use of both Paired Samples T-Tests as well as Independent T-Tests in order to examine the differences between JE mode and SE mode conditions, respectively, in relation to the assessed affective forecasts. The aforementioned methods were leveraged in order to verify **H1. The evaluation mode will have an impact on participants' affective forecasts** , and furthermore **H1.1** and **H1.2** .

Therefore, in order to explore biased behavior that is congruent with Distinction Bias, for each of the four tasks referent to the first experiment, further detailed in Section 4.1.1, we studied our participants affective forecast data for each of the outcomes presented. Throughout this subsection we focus on the first three tasks, as Task 1.4 refers to a specific context that aims to provide answers for **RQ2.** and will be addressed in the next subsection.

For each of the tasks we ran both a Paired Samples T-Test, for JE mode participant data, and an Independent T-Test, for SE mode participant data, on pairs of reported affective forecasts referent to consecutive presented outcomes or, in other words, pair of happiness assessments referent to outcome bars situated next to eachother in the presented visualization.

5.1.1.A Task 1.1

The following Table 5.1 presents the results obtained from the several Paired Samples, as well as Independent T-Tests, performed on affective forecasts referent to all pairs of consecutive outcomes, for both JE and SE evaluation modes, regarding Task 1.1.

Table 5.1: Task 1.1 - Mean affective forecasts for: Outcome A - sold 5 books, Outcome B - sold 80 books, Outcome C - sold 160 books and Outcome D - sold 240 books; and mean affective differences for consecutive pairs: AB, BC and CD.

	JE		**SE**	
Outcome	**Mean**	**Mean Difference**	**Mean**	**Mean Difference**
A	M = 4.90, SD = 1.535		M = 3.70, SD = 1.263	
		1.872 ± 1.105, t(38) = 10.583, p<.001		3.522 ± 0.846, t(17) = 4.163, p<.001
B	M = 6.77, SD = 1.564		M = 7.22, SD = 1.202	
		1.077 ± 0.174, t(38) = 6.196, p<.001		0.778 ± 0.412, t(21) = 1.886, p = .073
C	M = 7.85, SD = 1.065		M= 8.00, SD = 0.784	
		0.615 ± 0.125, t(38) = 4.915, p<.001		0.143 ± 0.409, t(19) = -0.349, p = 0.731
D	M = 8.46, SD = 1.022		M = 7.86, SD = 1.069	

In Task 1.1, by analyzing the p-values referent to the tested consecutive pairs of outcomes, we can conclude that **JE mode participants reported statistically significant differences in levels of happiness for all consecutive pairs of presented outcomes, in accordance with our initial assumptions** that tests were expected to confirm statistically significant differences for all tested JE mode outcome pairs.

All outcomes displayed for Task 1.1 differ only quantitatively. **SE mode participants did not report statistically significant differences in affective forecasts for only two of our three quantitatively different pairs of outcomes, which only partially supports our initial assumptions**, as quantitative differences between outcomes were predicted not to lead to statistically significant differences for affective forecasts of SE mode participants.

A summary of mean affective forecasts referent to each outcome for the two studied evaluation modes is presented in Figure 5.1, visually depicting mean affective differences between outcomes.

According to Distinction Bias literature concerning qualitative and quantitative differences, we anticipated that JE mode participants would tend to overpredict affective differences between outcomes with only quantitative variation, when compared to SE mode participants.

JE mode participants reported lower affective forecast differences between Task 1.1's Outcome A - sold 5 books - and Outcome B - sold 80 books, when compared to SE mode participants. **This result goes against our assumptions, as Outcomes A and B differ only quantitatively.**

Moreover, **JE mode participants reported higher affective forecast differences** between Task

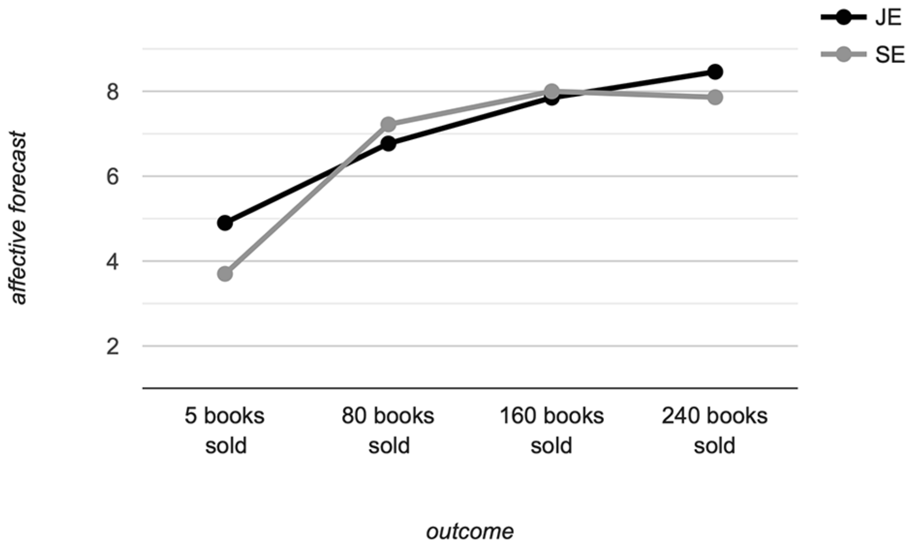


Figure 5.1: Task 1.1 - Mean affective forecast trajectory for each outcome, for both JE and SE evaluation modes.

1.1's Outcome B - sold 80 books - and Outcome C - sold 160 books, as well as between Outcome C - sold 160 books - and Outcome D - sold 240 books, when compared to SE mode participants. **This result supports our assumptions, as Outcomes B, C and D differ only quantitatively.**

A possible explanation for our results for Task 1.1 only being partially in accordance to initial assumptions, given the large and statistically significant difference between SE mode's affective forecasts for Outcome A - sold 5 books - and Outcome B - sold 80 books, could be that Outcome A is situated too close to the baseline for Distinction Bias effects to fully manifest, as Figure 2.1 illustrates, given that SE mode's curve only flattens once distant enough from the baseline, which can get difficult to predict for each situation before experimental observation. Additionally, in fact, Outcome A is the only outcome from Task 1.1 to be added by us, as the Distinction Bias studies from which we took inspiration to craft our tasks considered Outcome A - sold 0 books, which instead differs qualitatively from other presented outcomes. As our study is centered around a visualization, involving actions such as hovers, we found ourselves forced to replace this value with one which could be encoded with bar length, instead opting for 5.

5.1.1.B Task 1.2

The following Table 5.2 presents the results obtained from the several Paired Samples, as well as Independent T-Tests, performed on affective forecasts referent to all pairs of consecutive outcomes, for both JE and SE evaluation modes, regarding Task 1.2.

Table 5.2: Task 1.2 - Mean affective forecasts for: Outcome A - read 25 negative words, Outcome B - read 10 negative words, Outcome C - read 10 positive words and Outcome D - read 25 positive words; and mean affective differences for consecutive pairs: AB, BC and CD.

	JE		SE	
Outcome	**Mean**	**Mean Difference**	**Mean**	**Mean Difference**
A	M = 3.72, SD = 1.486		M = 4.17, SD = 0.937	
		0.538 ± 0.089, t(38) = 6.062, p<.001		0.278 ± 0.494, t(19) = 0.562, p = .581
B	M = 4.26, SD = 1.229		M = 4.44, SD = 1.333	
		1.769 ± 0.319, t(38) = 5.544, p<.001		1.889 ± 0.646, t(13) = 2.922, p = .012
C	M = 6.03, SD = 1.181		M = 6.33, SD = 1.033	
		0.385 ± 0.125, t(38) = 3.072, p = .004		-0.405 ± 0.543, t(19) = -0.746, p = 0.465
D	M = 6.41, SD = 1.464		M = 5.93, SD = 1.141	

In Task 1.2, by analyzing the p-values referent to the tested consecutive pairs of outcomes, we can conclude that **JE mode participants reported statistically significant differences in levels of happiness for all consecutive pairs of presented outcomes, in accordance with our initial assumptions** that tests were expected to confirm statistically significant differences for all tested JE mode outcome pairs.

SE mode participants did not report statistically significant differences in affective forecasts for our only two quantitatively different pairs of outcomes, AB and CD, supporting our initial assumptions, as quantitative differences between outcomes were predicted not to lead to statistically significant differences for affective forecasts of SE mode participants. Additionally, **SE mode participants did report statistically significant differences in affective forecasts for our qualitatively different pair of outcomes, BC, also supporting our initial assumptions**, as qualitative differences between outcomes were predicted to lead to statistically significant differences for affective forecasts of SE mode participants.

A summary of mean affective forecasts referent to each outcome for the two studied evaluation modes is presented in Figure 5.2, visually depicting mean affective differences between outcomes.

According to Distinction Bias literature concerning qualitative and quantitative differences, we anticipated that JE mode participants would tend to overpredict affective differences between outcomes with only quantitative variation, when compared to SE mode participants. On the other hand, we anticipated that JE mode participants would not tend to overpredict affective differences between outcomes with only qualitative variation, when compared to SE mode participants.

JE mode participants reported lower affective forecast differences between Task 1.2's Outcome B - read 10 negative words - and Outcome C - read 10 positive words, when compared to SE mode participants. **This result supports our assumptions, as Outcomes B and C differ only qualitatively.**

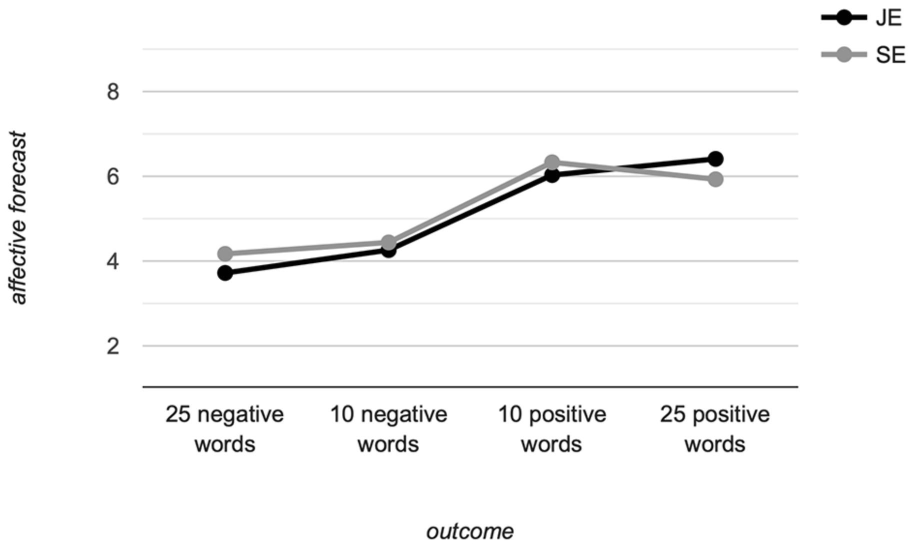


Figure 5.2: Task 1.2 - Mean affective forecast trajectory for each outcome, for both JE and SE evaluation modes.

Moreover, **JE mode participants reported a higher affective forecast difference** between Task 1.2's Outcome A - read 25 negative words and Outcome B - read 10 negative words, as well as between Outcome C - read 10 positive words and Outcome D - read 25 positive words, when compared to SE mode participants, with SE mode participants even reporting a decrease in happiness levels between the last pair. **This result also supports our assumptions, as the outcomes pair AB and CD differ only quantitatively.**

5.1.1.C Task 1.3

Table 5.3 presents the results obtained from the several Paired Samples, as well as Independent T-Tests, performed on affective forecasts referent to all pairs of consecutive outcomes, for both JE and SE evaluation modes, regarding Task 1.3.

In Task 1.3, by analyzing the p-values referent to the tested consecutive pairs of outcomes, we can conclude that **JE mode participants reported statistically significant differences in levels of happiness for all consecutive pairs of presented outcomes, in accordance with our initial assumptions** that tests were expected to confirm statistically significant differences for all tested JE mode outcome pairs.

As with Task 1.2, Task 1.3 is composed by two pairs of consecutive outcomes differing quantitatively, AB and CD, and one pair of consecutive outcomes differing qualitatively, BC. **SE mode participants did not report statistically significant differences in affective forecasts between Outcomes A -**

Table 5.3: Task 1.3 - Mean affective forecasts for: Outcome A - lost 100 euros, Outcome B - lost 50 euros, Outcome C - gained 50 euros and Outcome D - gained 100 euros; and mean affective differences for consecutive pairs: AB, BC and CD.

	JE		SE	
Outcome	**Mean**	**Mean Difference**	**Mean**	**Mean Difference**
A	M = 1.79, SD = 0.923		M = 2.00, SD = 0.926	
		1.026 ± 0.145, $t(38) = 7.094$, $p < .001$		0.273 ± 0.424, $t(17) = 0.643$, $p = .529$
B	M = 2.82, SD = 1.335		M = 2.27, SD = 0.905	
		4.462 ± 0.307, $t(38) = 7.094$, $p < .001$		4.527 ± 0.446, $t(19) = 10.155$, $p < .001$
C	M = 7.28, SD = 0.999		M = 6.80, SD = 1.135	
		1.128 ± 0.117, $t(38) = 9.626$, $p < .001$		1.617 ± 0.389, $t(20) = 4.432$, $p < .001$
D	M = 8.41, SD = 0.818		M = 8.42, SD = 0.515	

lost 100 euros - and B - lost 50 euros, yet did report statistically significant differences in affective forecasts between Outcomes C - gained 50 euros - and D - gained 100 euros, failing to fully support our initial assumptions, as quantitative differences between outcomes were predicted not to lead to statistically significant differences for affective forecasts of SE mode participants. Additionally, **SE mode participants did report statistically significant differences in affective forecasts for our qualitatively different pair of outcomes, BC, in accordance to our initial assumptions**, as qualitative differences between outcomes were predicted to lead to statistically significant differences for affective forecasts of SE mode participants.

A summary of mean affective forecasts referent to each outcome for the two studied evaluation modes is presented in Figure 5.3, visually depicting mean affective differences between outcomes.

According to Distinction Bias literature concerning qualitative and quantitative differences, we anticipated that JE mode participants would tend to overpredict affective differences between outcomes with only quantitative variation, when compared to SE mode participants.

JE mode participants reported lower affective forecast differences between Task 1.3's Outcome B - lost 50 euros - and Outcome C - gained 50 euros, when compared to SE mode participants. **This result supports our assumptions, as Outcomes B and C differ only qualitatively.** Additionally, **JE mode participants also reported lower affective forecast differences** between Task 1.3's Outcome C - gained 50 euros - and Outcome D - gained 100 euros, when compared to SE mode participants. **This result goes against our assumptions, as Outcomes C and D differ only quantitatively.**

JE mode participants reported higher affective forecast differences between Task 1.3's Outcome A - lost 100 euros - and Outcome B - lost 50 euros, when compared to SE mode participants. **This result supports our assumptions, as Outcomes A and B differ only quantitatively.**

A Welch t-test was run in place of the Paired-Samples T-Test to determine the differences in affective

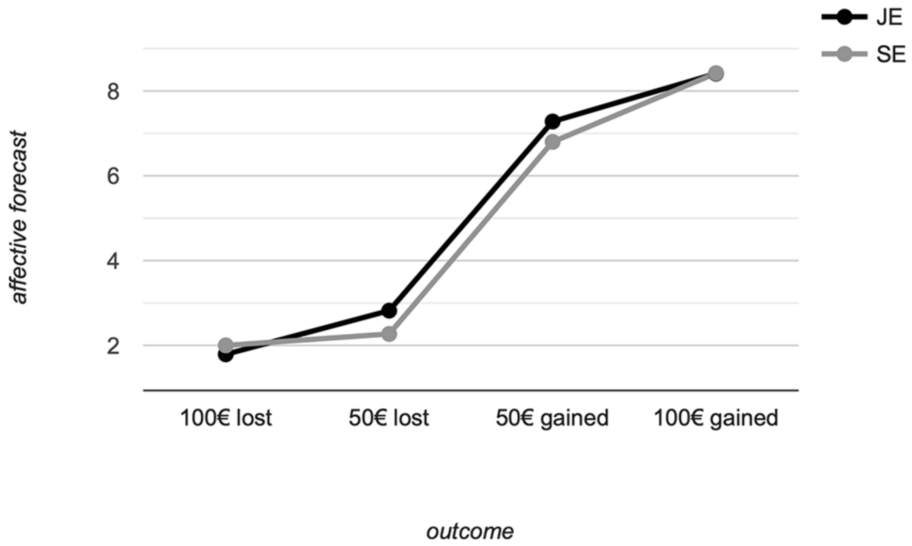


Figure 5.3: Task 1.3 - Mean affective forecast trajectory for each outcome, for both JE and SE evaluation modes.

forecasts between Task 1.3's Outcome C and Outcome D for SE mode data due to the assumption of homogeneity of variances being violated, as assessed by Levene's test for equality of variances (p = .005). However results for both were compared and the same conclusions were drawn.

5.1.2 Independently Easier-to-Evaluate Attributes

5.1.2.A Task 1.4

For Task 1.4, one additional factor was at play, the ability to independently evaluate an attribute. Hsee and Zhang [10] discuss that biased behaviour consistent with Distinction Bias applies mostly to attributes that are not too easy to evaluate independently, yet believe that finding an attribute that is notably easy to evaluate independently is the exception rather than the rule. The authors state school grades as an example of an attribute that people have sufficient knowledge about, resulting in SE mode happiness level predictions that may more closely match the evolution of JE mode ones. We took these statements as inspiration and crafted Task 1.4, regarding a scenario and outcomes centered around this attribute, striving to find answers for **RQ2. Do the effects of Distinction Bias on predictions persist when information is presented in a bar chart for an attribute that is independently easier to evaluate?**

We took on an exploratory approach for this task, aiming to analyze our metrics in the same ways as detailed for the previous tasks, examining resulting affective forecasts that address a reportedly easier-to-evaluate attribute, for both of our studied evaluation modes.

Table 5.4 presents the results obtained from the several Paired Samples, as well as Independent T-Tests, performed on affective forecasts referent to all pairs of consecutive outcomes, for both JE and SE evaluation modes, regarding Task 1.4.

Table 5.4: Task 1.4 - Mean affective forecast for: Outcome A - graded as 4 (out of 20), Outcome B - graded as 8 (out of 20), Outcome C - graded as 12 (out of 20) and Outcome D - graded as 16 (out of 20); and mean affective differences for consecutive pairs: AB, BC and CD.

	JE		SE	
Outcome	**Mean**	**Mean Difference**	**Mean**	**Mean Difference**
A	M = 1.95, SD = 1.025		M = 1.82, SD = 0.874	
		$1.359 \pm 0.0.231$, t(38) = 5.887, p<.001		1.015 ± 0.379, t(21) = 2.679, p = .014
B	M = 3.31, SD = 1.080		M = 2.83, SD = 0.937	
		2.641 ± 0.231, t(38) = 11.441, p<.001		2.367 ± 0.651, t(20) = 3.633, p = .003
C	M = 5.95, SD = 1.376		M = 5.20, SD = 1.874	
		1.769 ± 0.158, t(38) = 11.209, p<.001		2.550 ± 0.643, t(20) = 3.965, p = .002
D	M = 7.72, SD = 0.857		M = 7.75, SD = 0.707	

Given the existing literature, we expected that SE mode's affective forecasts would have a similar trajectory to JE mode ones. Therefore we predicted that, as for JE mode, SE mode participants would tend to report statistically significant differences in levels of happiness for all consecutive pairs of presented outcomes.

Results confirmed our predictions, as every performed T-Test, referent to both JE mode and SE mode forecasts, was deemed statistically significant. This demonstrates that, unlike what was expected for other tasks, **SE mode participants, just like JE mode ones, reported statistically significant divergence of happiness levels for each presented outcome**, even if each SE mode participant was only presented with one, this way not being able to easily compare it with other possible outcomes.

As for what details Distinction Bias literature concerning qualitative and quantitative differences, we did not set any predictions, as Distinction Bias theory details that the presence or lack of overpredictions of affect for JE mode participants is always measured in comparison to SE mode data. Given Task 1.4's attribute choice being expected to cause SE mode happiness level predictions to closely match the evolution of JE mode ones, we concluded that no resourceful predictions could be made in this case.

Nonetheless, the results obtained showed that JE mode participants reported lower affective forecast differences between Outcome C - graded 12 (out of 20) - and Outcome D - graded 16 (out of 20), when compared to SE mode participants. In opposition, JE mode participants reported higher affective forecast differences between the pairs Outcome A - graded 4 (out of 20) - and Outcome B - graded 8 (out of 20), and the pair Outcome B - graded 8 (out of 20) - and Outcome C - graded 12 (out of 20), when compared to SE mode participants.

Due to the assumption of homogeneity of variances being violated, as assessed by Levene's test for equality of variances, for SE mode's affective forecast differences between the Outcome pairs BC (p = .007) and CD (p = .004) a Welch t-test was run in place of the Paired-Samples T-Test in these cases. However results for both were compared and the same conclusions were drawn.

A summary of mean affective forecasts referent to each outcome for the two studied evaluation modes is presented in Figure 5.4, visually depicting mean affective differences between outcomes.

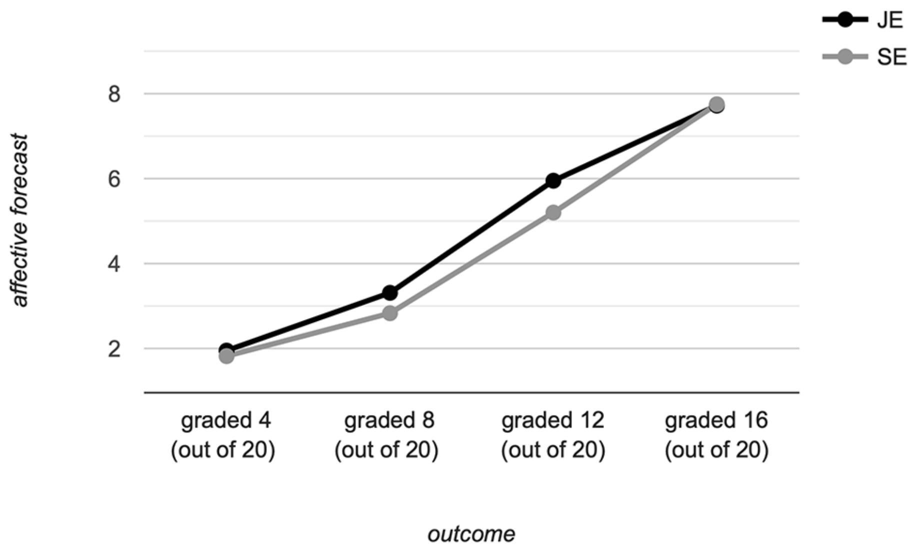


Figure 5.4: Task 1.4 - Mean affective forecast trajectory for each outcome, for both JE and SE evaluation modes.

5.1.3 Added Factors

With our second experiment, referent to Tasks 2.1, 2.2 and 2.3 detailed in Section 4.1.1, we aimed to further explore the ways in which Distinction Bias manifests itself in JE mode for quantitatively different outcomes. For this, we added two new factors, the quantity of presented outcomes for each of the tasks and the range of the vertical axis in which data would be displayed. For each, we set three different between-subjects conditions for the second experiment tasks.

In order to analyze possible correlations between these factors and the effects of Distinction Bias, we had to examine the effects of the quantity of presented outcomes and the vertical axis scale on affective forecast differences, by conducting two-way ANOVAs.

As identified in Section 4.1.1, only the three anchor outcomes presented for participants of every condition were taken into consideration for testing (denominated Outcome B, D, and F). Therefore, for each of the second experiment tasks, we tested the affective differences between the sequential pairs:

the pair Outcome B and Outcome D, as well as the pair Outcome D and Outcome F, included for all conditions. Table 5.5 presents the interaction effects between the quantity of presented outcomes and the vertical axis scale on the affective differences between the two pairs for each of the second experiment tasks.

Table 5.5: Interaction effects between the 'Outcome Amount' measure and the 'Axis max' measure on the affective differences between sequential pairs of Outcomes BD and DF, for each of the second experiment tasks.

Task	Outcome pair BD	Outcome pair DF
2.1	F(4, 71) = 2.669, p = .039, η^2 = .131	F(4, 71) = 1.136, p = .347, η^2 = .060
2.2	F(4, 70) = 0.762, p = .554, η^2 = .042	F(4, 70) = 1.683, p = .069, η^2 = .115
2.3	F(4, 71) = 5.884, p $<$ .001, η^2 = .249	F(4, 71) = 0.485, p = .002, η^2 = .746

Upon analysis, we found two statistically significant interaction effects. In both Task 2.1 and Task 2.3 there was a statistically significant interaction between the quantity of presented outcomes and the vertical axis scale on the affective differences between the pair Outcome B and Outcome D. As such, we decided to further investigate these tests by obtaining simple main effects, following them up with pairwise comparisons, in order to understand both the effect of quantity of presented outcomes at the three different levels of vertical axis range, as well as the effect of vertical axis ranges at the three different levels of the quantity of outcomes. The results are detailed in their respective subsections.

Table 5.6 details the main effects of the quantity of presented outcomes (or the 'Outcome Amount' measure) on affective differences between the two sequential pairs for every second experiment task. Table 5.7 details the main effects of the vertical axis scale (or the 'Axis Max' measure) on affective differences between the two sequential pairs for every second experiment task.

Table 5.6: Main effects of the quantity of presented outcomes on the affective differences between sequential pairs of Outcomes BD and DF, for each of the second experiment tasks.

Task	Outcome pair BD	Outcome pair DF
2.1	F(2, 71) = 1.857, p = .164, η^2 = .050	F(2, 71) = 0.698, p = .501, η^2 = .019
2.2	F(2, 70) = 1.823, p = .169, η^2 = .049	F(2, 70) = 0.188, p = .829, η^2 = .005
2.3	F(2, 71) = 1.764, p = .179, η^2 = .047	F(2, 71) = 2.326, p = .105, η^2 = .061

Table 5.7: Main effects of the vertical axis scale on the affective differences between sequential pairs of Outcomes BD and DF, for each of the second experiment tasks.

Task	Outcome pair BD	Outcome pair DF
2.1	F(2, 71) = 0.582, p = .562, η^2 = .016	F(2, 71) = 0.606, p = .548, η^2 = .017
2.2	F(2, 70) = 0.018, p = .982, η^2 = .001	F(2, 70) = 0.978, p = .381, η^2 = .027
2.3	F(2, 71) = 0.833, p = .439, η^2 = .023	F(2, 71) = 1.701, p = .190, η^2 = .046

Additionally, it is important to note that there were no statistically significant main effects of either the quantity of presented outcomes or the vertical axis scale on affective differences between either outcome pair on any of the second experiment tasks, as seen by examination of Tables 5.6 and 5.7.

5.1.3.A Quantity of Presented Outcomes

As aforementioned, we found two statistically significant interaction effects of our two added factors on the affective differences between the pair Outcome B and Outcome D, for Task 2.1 and 2.3.

As such, we decided to further investigate by assessing simple main effects. In this subsection we detail our study on the simple main effects for the quantity of presented outcomes on affective differences of Outcome pair BD.

Essentially this means that, when considering affective differences between Outcome B and D, we had to examine the effects for the quantity of presented outcomes at each level of the 'Axis Max' measure. We made sure to apply the suitable Bonferroni adjustments. Thus, we only declared a simple main effect as statistically significant if $p < .017$.

Upon analysis we found that, in Task 2.1, there was a statistically significant difference in mean affective difference (BD) between the levels of 'Outcome Amount' for participants interacting with the smaller level of vertical axis range ($F(2, 71) = 6.360$, $p = .003$, $\eta^2 = .152$).

After assessing pairwise comparisons, we verified that, for participants interacting with the smaller level of vertical axis range:

- the mean affective difference (BD) was 1.750 (95% CI, 3.125 to 0.375) lower for participants being presented with seven outcomes than for the ones being presented with three outcomes.
- the mean affective difference (BD) was 1.806 (95% CI, 3.214 to 0.397) lower for participants being presented with seven outcomes than for the ones being presented with five outcomes.

Furthermore, in Task 2.3, there was a statistically significant difference in mean affective difference (BD) between the levels of 'Outcome Amount' for participants being presented with our medium level of vertical axis range ($F(2, 71) = 5.759$, $p = .005$, $\eta^2 = .140$). After assessing pairwise comparisons, we verified that, for participants interacting with the medium level of vertical axis range:

- the mean affective difference (BD) was 1.067 (95% CI, 1.860 to 0.274) lower for participants being presented with five outcomes than for the ones being presented with three outcomes.

While our findings are not consistently substantial to provide a thorough conclusion, results do manifest an interesting trend. Throughout all statistically significant pairwise comparisons we found a decrease in affective differences as we moved towards higher levels of quantities of presented outcomes. **Even though our amount of significant findings was limited we can observe that our statistically**

significant results showed a trajectory in which participants being presented with increasingly higher amounts of outcomes had a tendency to report smaller affective differences regarding the same studied anchor outcomes.

This is especially interesting as, for our few statistically significant results, the trend of decreasing affective differences as participants interacted with increasingly more outcomes stood even though from the first level (3 presented outcomes) to the second level (5 presented outcomes) we added one distractor outcome with a higher value that any other and one distractor outcome with a lower value than any other, yet from the second level (5 presented outcomes) to the third level (5 presented outcomes), instead of doing the same, we added two distractor outcomes, in between our anchor outcomes, as exemplified in Figure 4.4.

As we were dealing with exclusively quantitatively different outcomes, and overprediction of differences of utility congruent with Distinction Bias manifests specifically in these contexts, in practise this means that, generalizing from these cases, overpredictions would be more likely to decrease further as we present participants with higher quantities of outcomes for comparison.

Overall, **while our findings are not consistently substantial to support H2 (the amount of quantitatively different outcomes being presented will have an impact on the differences between affective forecasts of anchor outcomes), results do show a trend that raises some interest on further studying the premise that, in practise, comparing higher quantities of alternatives could be somewhat linked to a decrease in the measured Distinction Bias effects.**

5.1.3.B Scale of Presented Outcomes

In this subsection we detail our study on the simple main effects for the vertical axis scale on affective differences of Outcome pair BD.

As aforementioned, we found two statistically significant interaction effects of our two added factors on the affective differences between the pair Outcome B and Outcome D, for Task 2.1 and 2.3.

As such, similarly to what we did throughout the previous subsection, when considering affective differences between Outcome B and D, we had to examine the effects for the vertical axis scale at each level of the 'Outcome amount' measure. We made sure to apply the suitable Bonferroni adjustments. Thus, we only declared a simple main effect as statistically significant if $p < .017$.

Upon analysis we found that, in Task 2.3, there was a statistically significant difference in mean affective difference (BD) between the levels of 'Axis Max' for participants being presented with 3 outcomes ($F(2, 71) = 6.012$, $p = .004$, $\eta^2 = .145$), as well as for participants being presented with 5 outcomes ($F(2, 71) = 6.279$, $p = .003$, $\eta^2 = .150$).

After assessing pairwise comparisons, we verified that, for participants being presented with 3 outcomes:

- the mean affective difference (BD) was 1.150 (95% CI, .331 to 1.969) higher for participants interacting with the larger level of vertical axis range than for those interacting with the medium level.

And furthermore, for participants being presented with 5 outcomes:

- the mean affective difference (BD) was 1.111 (95% CI, .297 to 1.925) higher for participants interacting with the medium level of vertical axis range than for those with the smaller level.
- the mean affective difference (BD) was 0.889 (95% CI, 1.703 to .075) lower for participants interacting with the larger level of vertical axis range than for those with the medium level.

While investigating **H3.** we expected that, in order to support this hypothesis, results would be consistently substantial and highlight an overall linear trajectory for affective differences throughout the increasing vertical axis maximum values.

Upon analysis, we can observe that our findings point to opposite directions as we move through levels, with some stating an increase of affective differences towards higher 'Axis Max' levels, while others a decrease. It is also important to note that, additionally, no consistent repeating patterns regarding the trajectory of affective differences per groups were identified for the three tasks tested.

As such, **our findings were not consistently substantial and our few statistically significant results do not present a clear direction regarding the importance of the scale of presented data for Distinction Bias theory, failing to considerably support H3 (the scale in which information is presented in the visualization will have an impact on the differences between affective forecasts of studied options)**. Nonetheless, further study on the factor would be incentivized.

5.2 Neuroticism Effects

In this section, we present the results that focused on providing answers to **RQ3. Do participants' neuroticism scores impact their affective forecasts?** We then explore the obtained results regarding any relations between the participant's personality and user interaction data, namely task completion times and quantity of hovers performed per task.

5.2.1 Neuroticism and Distinction Bias

In this subsection, we describe our results regarding the inclusion of personality data in our study. Firstly, we examined possible correlations between neuroticism and the effects of Distinction Bias.

As biased behavior takes place in JE mode, we focused this part of our study on it. In theory, in order to link neuroticism with behavior congruent with Distinction theory we would specifically expect correlations with affective forecast differences referent outcomes that vary quantitatively. Additionally, in case there would be any correlations between neuroticism and affective forecast differences referent to outcomes that vary qualitatively they would be expected to occur in an opposing direction, as Distinction Bias theory notes that JE mode difference overprediction is likely to occur for merely quantitatively different values and unlikely to occur for qualitatively different values.

To test our assumptions we ran Spearman Correlation tests between affective differences reported throughout our first experiment tasks and JE mode participants' neuroticism scores. Table 5.8 presents the results obtained by testing correlations between participants' neuroticism scores and affective forecast differences referent to consecutive Outcome pairs AB, BC and CD, for each of the first experiment tasks.

Table 5.8: Correlations between neuroticism scores and affective differences referent to Outcome pairs AB, BC and CD, for each of the first experiment tasks.

Task	Outcome pair AB	Outcome pair BC	Outcome pair CD
1.1	$r_s(29) = .220$, $p = .251$	$r_s(29) = -.101$, $p = .601$	$r_s(29) = -.073$, $p = .707$
1.2	$r_s(29) = .086$, $p = .656$	$r_s(29) = .026$, $p = .892$	$r_s(29) = .242$, $p = .205$
1.3	$r_s(29) = -.023$, $p = .907$	$r_s(29) = .059$, $p = .761$	$r_s(29) = .274$, $p = .151$
1.4	$r_s(29) = .154$, $p = .424$	$r_s(29) = .183$, $p = .343$	$r_s(29) = -.199$, $p = .301$

Additionally, we also tested for correlations between neuroticism scores and actual affective forecasts reported by JE mode participants, as personality theory denotes that neuroticism is characterized by a tendency to easily experience negative emotions and to be more pessimistic. Therefore we decided to examine this possibility, expecting any possible correlations to be characterized by negative correlation coefficient values. Table 5.9 presents the results obtained by testing correlations between participants' neuroticism scores and actual affective forecasts referent to Outcomes A, B, C and D, for each of the first experiment tasks.

Table 5.9: Correlations between neuroticism scores and affective forecasts referent to Outcomes A, B, C and D, for each of the first experiment tasks.

Task	Outcome A	Outcome B	Outcome C	Outcome D
1.1	$r_s(29) = .080$, $p = .680$	$r_s(29) = .291$, $p = .125$	$r_s(29) = .237$, $p = .216$	$r_s(29) = .360$, $p = .055$
1.2	$r_s(29) = -.092$, $p = .635$	$r_s(29) = -.043$, $p = .826$	$r_s(29) = -.125$, $p = .519$	$r_s(29) = -.011$, $p = .955$
1.3	$r_s(29) = -.044$, $p = .821$	$r_s(29) = -.107$, $p = .581$	$r_s(29) = -.014$, $p = .941$	$r_s(29) = .238$, $p = .215$
1.4	$r_s(29) = .070$, $p = .717$	$r_s(29) = .029$, $p = .882$	$r_s(29) = .204$, $p = .289$	$r_s(29) = .145$, $p = .452$

For our second experiment, as all participants took their assessments in JE mode, we decided to also test for correlations for tasks 2.1, 2.2 and 2.3, similarly to what we did for the first experiment data. In this case, we solely considered Outcomes B, D and F, as these were the outcomes every participant of the second experiment was able to interact with.

As such we ran Spearman Correlation tests between affective differences reported throughout our second experiment tasks and participants' neuroticism scores. Table 5.10 presents the results obtained by testing correlations between participants' neuroticism scores and affective forecast differences referent to consecutive Outcome pairs BD and DF, for each of the second experiment tasks.

Table 5.10: Correlations between neuroticism scores and affective differences referent to Outcome pairs BD and DF, for each of the second experiment tasks.

Task	Outcome pair BD	Outcome pair DF
2.1	$r_s(58) = .107$, $p = .423$	$r_s(58) = -.094$, $p = .484$
2.2	$r_s(57) = -.225$, $p = .093$	$r_s(57) = -.053$, $p = .697$
2.3	$r_s(58) = .089$, $p = .508$	$r_s(58) = .172$, $p = .198$

Similarly to our procedure for the first experiment tasks we also tested for correlations between neuroticism scores and actual affective forecasts throughout Tasks 2.1, 2.2 and 2.3, once more expecting any possible correlations to be characterized by negative correlation coefficient values. Table 5.11 presents the results obtained by testing correlations between participants' neuroticism scores and actual affective forecasts referent to Outcomes B, D and F, for each of the second experiment tasks.

Table 5.11: Correlations between neuroticism scores and affective forecasts referent to Outcomes B, D and F, for each of the second experiment tasks.

Task	Outcome B	Outcome D	Outcome F
2.1	$r_s(58) = -.058$, $p = .668$	$r_s(58) = .122$, $p = .362$	$r_s(58) = .091$, $p = .497$
2.2	$r_s(57) = .069$, $p = .611$	$r_s(57) = -.073$, $p = .588$	$r_s(57) = -.087$, $p = .522$
2.3	$r_s(58) = -.027$, $p = .842$	$r_s(58) = -.028$, $p = .833$	$r_s(58) = .104$, $p = .436$

By applying the suitable Bonferroni corrections we solely considered tests with a p-value $< .007$ as statistically significant correlations, as each dependent variable was simultaneously tested against neuroticism scores and the scores for each of its six facets.

Results showed no statistically significant correlations between neuroticism and JE mode participants' affective forecast differences. These results fail to support H4 (neuroticism scores will have an effect on the differences between affective predictions) and furthermore **fail to exhibit any links between the neurotic trait and the effects that characterize Distinction Bias.**

Additionally, our examination for possible correlations between neuroticism and the actual affective forecast values also showed no statistically significant correlations.

5.2.2 Interaction Data

Additionally, due to the literature findings discussed in Section 3.2, we decided to explore possible correlations between participant's personality and the collected user interaction data, namely task completion times and quantity of hovers performed per task.

5.2.2.A Completion Time

We ran Spearman Correlation tests in order to explore possible correlations between neuroticism scores and task completion time in a bias-prone context. Given the diversity of literature findings, we studied completion times through a more general and exploratory approach.

By applying the suitable Bonferroni corrections we solely considered tests with a p-value $< .007$ as statistically significant correlations, as each dependent variable was simultaneously tested against neuroticism scores and the scores for each of its six facets.

For Task 1.1 results showed there was a statistically significant negative correlation between completion times and participants' neuroticism scores, $r_s(29) = -.543$, $p = .002$, and furthermore there were also statistically significant negative correlations between completion times and participants' scores for two neuroticism facet-level traits, them being N3 - Depression ($r_s(29) = -.542$, $p = .002$), as well as N6 - Vulnerability ($r_s(29) = -.596$, $p < .001$). A comprehensive description of these facets is presented in Section 2.2.1.

For the remaining first experiment tasks results showed no statistically significant correlations between completion times and participants' neuroticism scores: ($r_s(29) = -.059$, $p = .763$) for Task 1.2, ($r_s(29) = -.112$, $p = .563$) for Task 1.3 and ($r_s(29) = -.244$, $p = .202$) for Task 1.4.

We decided to also extend the analysis to the second experiment tasks, but due to the variations in the quantity of presented outcomes and therefore the amount of data being analyzed and assessed on, we determined it would be best to perform Spearman Correlation tests separately for every Outcome Amount condition group, as such:

For Task 2.1 we explored correlations between completion times and participants' neuroticism scores for participants being presented with 3 Outcomes ($r_s(20) = -.007$, $p = .976$), 5 Outcomes ($r_s(20) = -.063$, $p = .791$) and 7 Outcomes ($r_s(18) = -.051$, $p = .842$). Results showed there were no statistically significant correlations.

For Task 2.2 we explored correlations between completion times and participants' neuroticism scores for participants being presented with 3 Outcomes ($r_s(20) = -.026$, $p = .915$), 5 Outcomes ($r_s(20) = -.085$, $p = .721$) and 7 Outcomes ($r_s(18) = -.183$, $p = .467$). Results showed there were no statistically significant correlations.

For Task 2.3 we explored correlations between completion times and participants' neuroticism scores for participants being presented with 3 Outcomes ($r_s(20) = .053$, $p = .825$), 5 Outcomes ($r_s(20) = -.023$, $p = .922$) and 7 Outcomes ($r_s(18) = -.352$, $p = .152$). Results showed there were no statistically significant correlations.

Overall, as detailed, **results did not show consistent statistically significant correlations be-**

tween completion times and participants' neuroticism scores, apart from in the context of Task 1.1, for which neuroticism showed a negative correlation with completion times. This could be due to Task 1.1 perhaps being considered the most straightforward from our first experiment, as well as being the only task presenting exclusively quantitative differences between outcomes. These conditions could have promoted the consequences that some literature studies have already associated with highly neurotic individuals. Additionally, and even though no other statistically significant correlations were identified, **overall correlation coefficients consistently pointed to negative values, which does help contribute to the possibility that neuroticism may exhibit negative correlations with completion times**.

Altogether, **our findings failed to show significant support to H5 (neuroticism scores will have an effect on the time participants take to make their predictions in a bias-prone context) yet did express some evidence in support of the possibility that neuroticism may exhibit negative correlations with completion times**.

5.2.2.B Quantity of Hovers

We ran Spearman Correlation tests in order to explore possible correlations between neuroticism scores and the quantity of hover performed per task in a bias-prone context. Given the diversity of literature findings, we studied this metric through a more general and exploratory approach.

By applying the suitable Bonferroni corrections we solely considered tests with a p-value $< .007$ as statistically significant correlations, as each dependent variable was simultaneously tested against neuroticism scores and the scores for each of its six facets.

For the entirety of first experiment tasks our results showed no statistically significant correlations between the quantity of performed hovers and participants' neuroticism scores: ($r_s(29)$ = -.252, p = .188) for Task 1.1, ($r_s(29)$ = .137, p = .478) for Task 1.2, ($r_s(29)$ = -.187, p = .330) for Task 1.3 and ($r_s(29)$ = -.168, p = .384) for Task 1.4.

We decided to also extend the analysis to the second experiment tasks, and due to the variations in the quantity of presented outcomes and therefore the amount of data being analyzed and assessed on, we determined it would be best to perform Spearman Correlation tests separately for every Outcome Amount condition group, as such:

For Task 2.1 we explored correlations between the quantity of performed hovers and participants' neuroticism scores for participants being presented with 3 Outcomes ($r_s(20)$ = .379, p = .099), 5 Outcomes ($r_s(20)$ = -.094, p = .694) and 7 Outcomes ($r_s(18)$ = -.097, p = .701). Results showed there were no statistically significant correlations.

For Task 2.2 we explored correlations between the quantity of performed hovers and participants' neuroticism scores for participants being presented with 3 Outcomes ($r_s(20)$ = .050, p = .835), 5 Outcomes ($r_s(20)$ = .117, p = .623) and 7 Outcomes ($r_s(18)$ = -.095, p = .708). Results showed there were no statistically significant correlations.

For Task 2.3 we explored correlations between the quantity of performed hovers and participants' neuroticism scores for participants being presented with 3 Outcomes ($r_s(20)$ = -.144, p = .544), 5 Outcomes ($r_s(20)$ = .283, p = .227) and 7 Outcomes ($r_s(18)$ = -.119, p = .638). Results showed there were no statistically significant correlations.

Overall, **results did not show statistically significant correlations between the quantity of hovers performed and participants' neuroticism scores. These findings failed to show support for H6 (neuroticism will have an effect on the number of hovers performed by participants in a bias-prone context)**.

5.3 Discussion

Our study was developed to provide some answers to **RQ1 - Does the Distinction Bias transfer to Information Visualization?**, which we set as the main goal for our first experiment. As previously discussed, Distinction was described as a cognitive bias that, although unexplored in the Information Visualization field, is seen as likely relevant [1]. The gathered results corroborate the importance of further study of the effects of Distinction Bias, particularly in the context of InfoVis. Our results validate the relevance of Distinction Bias in this context while also helping to shed light on other discussions, such as the possible influences of personality on biased-prone environments.

Throughout the four tasks that compose our first experiment, detailed in Section 4.1.1, we compared the results from two evaluation modes, JE mode and SE mode. Distinction Bias is characterized as a tendency for overprediction of happiness levels between two outcomes differing merely quantitatively in JE mode in comparison with SE mode, and a lack of tendency for overprediction in the cases for which two outcomes differ qualitatively [10]. As such, we expected to obtain results that would demonstrate these tendencies, as expressed in **H1**, and furthermore in **H1.1** and **H1.2**. It is important to note in retrospect that assumptions were correctly verified for the entirety of outcomes that were replicated from the literature studies in which we based our study on [10].

For Task 1.1 we found results consistent with Distinction for Outcomes B, C and D, while results involving Outcome A failed to verify our assumptions. As previously noted, a possible explanation for this could be that this outcome could be too close to the baseline to prompt the expected effects, as Distinction Bias effects are only experienced within a certain distance from the baseline, as we can see in Figure 2.1. This explanation is consistent with the statistically significant difference found between affective forecasts for Outcome A and Outcome B reported by SE mode participants, which in theory should not otherwise occur.

Task 1.2 fully exhibited the effects that were deemed consistent with Distinction Bias regarding each of the presented Outcomes, encompassing all the situations we made assumptions on as well as fully supporting each one. This task presented a complete demonstration of how, for each side of a baseline, JE mode affective forecasts present statistically significant differences, developing a steeper evaluation function, when compared to SE mode ones which eventually flatten their evaluation function, no longer presenting statistically significant differences. It is this occurence that produces the main effects we associate with Distinction Bias, as JE mode participants overpredicted affective differences between outcomes varying merely quantitatively but not for outcomes varying merely qualitatively.

For Task 1.3 we found results consistent with Distinction for Outcomes A, B and C, while results involving Outcome D failed to verify our assumptions. Unfortunately, no Outcomes with quantitatively larger values are present for this task, for which we could verify if Distinction Bias effects would be present, and therefore we cannot determine if we are being presented with a similar situation as in Task 1.1. Yet, given the statistically significant difference found between affective forecasts for Outcome C and Outcome D in SE mode data, which in theory should not otherwise occur, there is certainly the possibility. Nonetheless, further studies would be required to precisely determine.

With Task 1.4 we were able to provide some answers for **RQ2**, exploring and confirming our assumptions regarding a type of scenario which, in theory, would not be so likely to prime our participants for the effects of Distinction Bias. Our findings support the idea that some attributes can be easier to evaluate independently than others, resulting in SE mode happiness level predictions that more closely match the evolution of JE mode ones, therefore not being as prone to the effects of Distinction Bias. For this task we incorporated an attribute discussed in the available literature as easier to evaluate independently [10], school grades. Results regarding this task demonstrated how SE mode participants reported differences between affective forecasts for all outcomes which were deemed statistically significant, an occurrence no other task from our study consistently exhibited and which in theory should not otherwise occur. However, neither the available literature nor our study is able to provide many guidelines for identifying these types of attributes, as this effect was not identified for any other of our tasks. Even so, with school grades being our experiments' only attribute operating within a limited scale (school grades always have a minimum and maximum value), we would be keen to highlight this fact as a possible reason to further study these scenarios.

For our second experiment, referent to Tasks 2.1, 2.2 and 2.3, we decided to explore aspects we deemed as possibly impacting, one of them being the quantity of presented outcomes and the other the scale in which data would be presented. As such, we tested three between-subjects condition groups for each of these factors against the affective differences that characterize Distinction Bias.

The results obtained were too limited to exhibit consistent substantial support for **H2**, yet our few statistically significant findings showed a trajectory in which participants being presented with increasingly higher amounts of outcomes had a tendency to report smaller affective differences regarding the same studied anchor outcomes. These results raise some interest on further studying the premise that, in practise, comparing higher quantities of alternatives could be somewhat linked to a decrease in the measured Distinction Bias effects.

Furthermore, results obtained also failed to exhibit support for **H3**, as our few statistically significant results do not present a clear direction regarding the importance of the scale of presented data for Distinction Bias theory. Additionally, no consistent repeating patterns regarding a specific trajectory of affective differences throughout the three between-subjects groups was identified.

Moreover, we also introduced participant personality research onto our study, aiming to provide answers to **RQ3**. Given literature findings, we tested possible correlations between participants' neuroticism scores and the affective differences that characterize Distinction Bias. Results showed no statistically significant correlations between neuroticism and JE mode participants' affective forecast differences. These results fail to support **H4** and furthermore our study showed no links between the neurotic trait and the effects of Distinction Bias.

Additionaly, we also tested for possible correlations between neuroticism scores and the actual happiness levels reported by participants. Likewise, our examination for possible correlations between neuroticism and affective forecast values also showed no statistically significant correlations.

Subsequently, we found relevant to test possible connections between our participants' neuroticism scores and collected user interaction data, namely task completion times and the quantity of performed hovers per task, in a bias-prone context.

We found a statistically significant negative correlation between Task 1.1's completion times and participants' neuroticism scores, essentially meaning that highly neurotic participants took less time to complete affective predictions for Task 1.1 in a bias-prone context. Furthermore, statistically significant negative correlations were also found for neuroticism between Task 1.1's completion times and participants' scores for two neuroticism facet-level traits, them being N3 - Depression and N6 - Vulnerability. Additionally, even though no other statistically significant correlations were identified for any other of our tasks, overall correlation coefficients consistently pointed to negative values. As such, even though we deem our results to be insufficient to definitively answer **H5**, they express some evidence contributing to the premise that neuroticism may exhibit negative correlations with completion times, incentivizing further research on the topic.

Results regarding our exploration for possible correlations between the quantity of hover performed per task and participants' neuroticism scores did not show statistically significant correlations, failing to support **H6**.

6

Conclusion

Contents

In light of the lack of empirical work on the link between cognitive biases and visualizations, we conducted our work focused on understanding the effects of a specific cognitive bias, Distinction Bias, in the context of InfoVis, as literature surrounding this bias is still scarce and possible links with Information Visualization remained unexplored. Subsequently, taking into account prior findings regarding individual differences and their role in human-machine systems, our study considered the possible impacts of a personality trait, neuroticism, in such bias-prone contexts.

Distinction Bias describes the way people's utility and affective predictions for several alternatives can vary depending on the evaluation mode they find themselves in, having these predictions then possibly lead to sub-optimal decisions. As such, we developed a graphical interface simulating the two opposing evaluation modes, JE mode and SE mode, presenting data through a bar chart visualization idiom. Participant personality data was collected through the NEO PI-R questionnaire. Throughout our experiments we asked participants to consider they would find themselves in hypothetical situations, consequently considering either several outcomes for this situation or a single one, in accordance with their evaluation mode condition. Participants were then requested to predict their happiness levels for the outcome(s) being presented to them.

Overall, although a few of the examined outcomes did not elicit the expected consequences, results hinted to the effects of Distinction Bias, exhibiting its relevancy in the information visualization context and the pertinence of its further study in the context of visual analytics systems aiming to support decision-making tasks that entail expected utility estimation for the joint evaluation of alternative attribute values. Our findings highlighted a moderate tendency for overprediction of affective differences when outcomes exhibited merely quantitative differences in JE mode in comparison with SE mode, as well as a lack of the same overprediction when outcomes differed only qualitatively. As aforementioned, it is the junction of these findings that characterizes Distinction Bias.

Results also confirmed our assumptions that some attributes are likely to be easier to evaluate independently, therefore being less likely to lead to overpredictions of utility for contexts in which several alternatives are evaluated simultaneously. Our results regarding a formerly pinpointed attribute evidenced that participants in SE mode consistently reported statistically significantly different values for all quantitatively differing outcomes, a result that corroborates how affective forecasts of participants in SE mode exhibited an evolution function otherwise expected solely of participants in JE mode. Furthermore, our results hinted towards possible relevance of the further study of the evolution of Distinction Bias effects as comparisons entail increasingly larger quantities of alternative outcomes.

Additionally, results regarding possible links between the neurotic trait and Distinction Bias exhibited no statistically significant findings. However, further analysis demonstrated a statistically significant negative correlation between neuroticism and completion times for one of our seven tasks. Moreover, correlation test results for the remaining task completion times overall pointed to negative correlation coefficients, yet their relevance is dubious per lack of any statistical significance. These findings, albeit not consistently substantial in the context of our work, reflect the potential for further research regarding the effects of personality and overall individual differences in interaction metrics.

Altogether, our work sheds new light on the topic of biases and their relevant impact on visual analytics systems. Our study corroborates the relevance and importance of further study surrounding Distinction Bias and remaining cognitive biases, particularly in the context of Information Visualization. Contrarily, our results do not exhibit correlations between neuroticism and the effects of Distinction Bias, yet hint at the relevancy of further research on personality and its impacts on user interaction.

6.1 Limitations and Future Work

The obtained findings allowed us to attribute pertinence to Distinction Bias in the context of Information Visualization. Nonetheless, the current research gap surrounding cognitive biases in the context of visualization tools stands as a relevant limitation of our work. As aforementioned, relevant literature on Distinction Bias itself is still scarce and its possible relevance in InfoVis, while considered likely, currently stood unexplored. Therefore, no consensus on ideal data encodings to study this bias were recognized. As such, our choices were motivated by prior literature and the guidelines regarding the characteristics of our utilized attributes and datasets. Moreover, the scarcity of investigation on problem contexts and recommendations for the creation of datasets that could identify the effects of Distinction Bias led us to base the development of our experiments on the rather small body of research on the bias, additionally developing supplementary scenarios and tasks for experimental observation, which did not always replicate the desired effects.

Future work shall consider the implications that may arise from different problem contexts within the processes of affective forecasting, further experimentally testing existing scenarios while also exploring the consequences that may accompany different scenarios. Moreover, as our visualization approach encompassed solely one data encoding, further work on the exploration of Distinction Bias effects across encodings would be incentivized. On a subsequent note, due to our results hinting a possible relevance of the evolution of Distinction Bias effects as comparisons entail increasingly larger quantities of alternative outcomes, future works encompassing this topic would additionally be encouraged, particularly ones leveraging scenarios already proven by literature to successfully be associated with Distinction Bias.

Despite results regarding the links of personality and Distinction Bias standing as mostly non-significant, our study reveals relevance on the further investigation of the robustness of Distinction Bias, and overall cognitive biases, to other personality traits. On a related note, our study hints on potential for further research regarding the effects of personality and overall individual differences in interaction metrics in the context of visualizations.

At last, while we considered our participant sample size reasonable, taking into account our work's scope, future works with the aim to replicate our study to some extent would be encouraged. Consequently, such works would benefit from a larger sample size, as means to more accurately support findings, and subsequently factor into account a diverse group of individuals' personality data when possible.

Bibliography

[1] 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, pp. 1413–1432, 2 2020.

[2] A. C. Valdez, M. Ziefle, and M. Sedlmair, “Studying biases in visualization research: Framework and methods,” pp. 13–27, 2018.

[3] E. Wall, L. M. Blaha, L. Franklin, and A. Endert, “Warning, bias may occur: A proposed approach to detecting cognitive bias in interactive visual analytics.” IEEE, 10 2017, pp. 104–115.

[4] T. M. Green and B. Fisher, “Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction.” IEEE, 10 2010, pp. 203–210.

[5] 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, pp. 88–94, 11 2012.

[6] E. T. Brown, A. Ottley, H. Zhao, Q. Lin, R. Souvenir, A. Endert, and R. Chang, “Finding waldo: Learning about users from their interactions,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, pp. 1663–1672, 12 2014.

[7] D. A. Keim, F. Mansmann, A. Stoffel, and H. Ziegler, “Visual analytics,” 2008. [Online]. Available: http://infovis.uni-konstanz.de

[8] H. A. Simon, *Models of man; social and rational*, 1957.

[9] D. Kahneman, *Thinking, fast and slow*. New York :Farrar, Straus and Giroux, 2011.

[10] C. K. Hsee and J. Zhang, “Distinction bias: Misprediction and mischoice due to joint evaluation,” *Journal of Personality and Social Psychology*, vol. 86, pp. 680–695, 5 2004.

[11] J. Müller and C. Schwieren, “Big five personality factors in the trust game,” *Journal of Business Economics*, vol. 90, pp. 37–55, 2 2020.

[12] G. W. Allport, *Personality: A psychological interpretation*, 1937.

[13] D. P. Schultz and S. E. Schultz, *Theories of personality*. Cengage Learning, 2016.

[14] 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, pp. 81–90, 1987.

[15] P. T. Costa and R. R. McCrae, "Neo pi-r: Inventário de personalidade neo revisto: Manual profissional," 2000.

[16] R. Tadeusiewicz and A. Horzyk, "Man-machine interaction improvement by means of automatic human personality identification," pp. 278–289, 2014.

[17] Z. Liu, R. J. Crouser, and A. Ottley, "Survey on individual differences in visualization," *Computer Graphics Forum*, vol. 39, pp. 693–712, 6 2020.

[18] 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, pp. 1109–1121, 2013.

[19] J. Baron, "Normative models of judgment and decision making," pp. 19–36, 2004.

[20] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *Science*, vol. 185, pp. 1124–1131, 9 1974.

[21] G. Ellis, *Cognitive Biases in Visualizations*, G. Ellis, Ed. Springer International Publishing, 2018.

[22] R. F. Pohl, *Cognitive illusions: Intriguing Phenomena in thinking*. Psychology Press, 2016.

[23] C. K. Hsee, "Less is better: when low-value options are valued more highly than high-value options," *Journal of Behavioral Decision Making*, vol. 11, pp. 107–121, 6 1998.

[24] J. Scott, "Rational choice theory," pp. 126–138, 2000.

[25] P. J. Corr and G. Matthews, "The cambridge handbook of personality psychology," 2009.

[26] D. C. Funder, *The Personality Puzzle*. WW Norton Co, 2019, vol. 8th Edition.

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

[28] H. J. Eysenck, "The biological basis of personality," *Springfield, IL: Thomas*, pp. 100–117, 1967.

[29] L. R. Goldberg, "An alternative "description of personality": The big-five factor structure." *Journal of Personality and Social Psychology*, vol. 59, pp. 1216–1229, 1990.

[30] M. C. Ashton and K. Lee, "Empirical, theoretical, and practical advantages of the hexaco model of personality structure," *Personality and Social Psychology Review*, vol. 11, pp. 150–166, 5 2007.

[31] M. R. Leary and R. H. Hoyle, *Handbook of individual differences in social behavior*. The Guilford Press, 2009.

[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, pp. 1311–1319, 12 2001.

[33] J. Ormel, B. F. Jeronimus, R. Kotov, H. Riese, E. H. Bos, B. Hankin, J. G. Rosmalen, and A. J. Oldehinkel, "Neuroticism and common mental disorders: Meaning and utility of a complex relationship," *Clinical Psychology Review*, vol. 33, pp. 686–697, 7 2013.

[34] T. Munzner, *Visualization Analysis Design*. CRC Press, 2014.

[35] E. Dimara and J. Stasko, "A critical reflection on visualization research: Where do decision making tasks hide?" *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2021.

[36] T. Verbeiren, R. Sakai, and J. Aerts, "A pragmatic approach to biases in visual data analysis," G. Ellis, Ed., 2014.

[37] G. Ellis and A. Dix, "Decision making under uncertainty in visualisation?" K. Potter and R. Westermann, Eds., 2015. [Online]. Available: http://vda.univie.ac.at/uncertainty2015/submissions/ellis_vdmu.pdf

[38] X. Chen, S. D. Starke, C. Baber, and A. Howes, "A cognitive model of how people make decisions through interaction with visual displays." ACM, 5 2017, pp. 1205–1216.

[39] E. Dimara, A. Bezerianos, and P. Dragicevic, "Conceptual and methodological issues in evaluating multidimensional visualizations for decision support," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, pp. 749–759, 1 2018.

[40] M. Bancilhon and A. Ottley, "Did you get the gist of it? understanding how visualization impacts decision-making," 10 2020.

[41] R. Wesslen, D. Markant, A. Karduni, and W. Dou, "Using resource-rational analysis to understand cognitive biases in interactive data visualizations," 9 2020.

[42] E. Dimara, P. Dragicevic, and A. Bezerianos, "Accounting for availability biases in information visualization," 10 2016.

[43] E. Dimara, A. Bezerianos, and P. Dragicevic, "The attraction effect in information visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, pp. 471–480, 1 2017.

[44] E. Dimara, G. Bailly, A. Bezerianos, and S. Franconeri, "Mitigating the attraction effect with visualizations," vol. 25, pp. 850–860, 2019. [Online]. Available: http://www.ieee.org/publications_standards/publications/rights/index.html

[45] I. Cho, R. Wesslen, A. Karduni, S. Santhanam, S. Shaikh, and W. Dou, "The anchoring effect in decision-making with visual analytics." IEEE, 10 2017, pp. 116–126.

[46] R. R. Bootzin, G. H. Bower, J. Crocker, and N. Hall, *Psychology Today : An Introduction*, 7th ed. McGraw-Hill Publishing Company, 1991.

[47] J. Manoogian and B. Benson, "Cognitive bias codex," 2017.

[48] D. A. Norman, *The Design of Everyday Things*. Basic Books, Inc., 2013.

[49] A. Dillon and C. Watson, "User analysis in hci — the historical lessons from individual differences research," *International Journal of Human-Computer Studies*, vol. 45, pp. 619–637, 12 1996.

[50] L. S. Woolhouse and R. Bayne, "Personality and the use of intuition: individual differences in strategy and performance on an implicit learning task," *European Journal of Personality*, vol. 14, pp. 157–169, 3 2000.

[51] I. Ajzen, *Attitudes, personality, and behavior (2nd ed.)*. Milton-Keynes: Open University Press/McGraw-Hill, 2005.

[52] J. Kim, A. Lee, and H. Ryu, "Personality and its effects on learning performance: Design guidelines for an adaptive e-learning system based on a user model," *International Journal of Industrial Ergonomics*, vol. 43, pp. 450–461, 9 2013.

[53] M. Kosinski, Y. Bachrach, P. Kohli, D. Stillwell, and T. Graepel, "Manifestations of user personality in website choice and behaviour on online social networks," *Machine Learning*, vol. 95, pp. 357–380, 10 2014. [Online]. Available: https://link.springer.com/article/10.1007/s10994-013-5415-y

[54] O. Nov, O. Arazy, C. López, and P. Brusilovsky, "Exploring personality-targeted ui design in online social participation systems," 2013.

[55] A. M. Viveros, E. H. Rubio, and D. E. V. Ceballos, "Equivalence of navigation widgets for mobile platforms," pp. 269–278, 2014.

[56] S. Devaraj, R. F. Easley, and J. M. Crant, "Research note —how does personality matter? relating the five-factor model to technology acceptance and use," *Information Systems Research*, vol. 19, pp. 93–105, 3 2008.

[57] G. B. Svendsen, J.-A. K. Johnsen, L. Almås-Sørensen, and J. Vittersø, “Personality and technology acceptance: the influence of personality factors on the core constructs of the technology acceptance model,” *Behaviour Information Technology*, vol. 32, pp. 323–334, 4 2013.

[58] T. Alves, J. Natálio, J. Henriques-Calado, and S. Gama, “Incorporating personality in user interface design: A review,” 3 2020.

[59] J. S. Yi, “Implications of individual differences on evaluating information visualization techniques,” p. 92, 2010.

[60] C. Ziemkiewicz, R. J. Crouser, A. R. Yauilla, S. L. Su, W. Ribarsky, and R. Chang, “How locus of control influences compatibility with visualization style.” IEEE, 10 2011, pp. 81–90.

[61] A. Ottley, H. Yang, and R. Chang, “Personality as a predictor of user strategy.” ACM, 4 2015, pp. 3251–3254.

[62] A. Kobsa, “User experiments with tree visualization systems,” 2004. [Online]. Available: http://www.ics.uci.edu/~kobsa.

[63] D. Williams, “Dispositional optimism, neuroticism, and extraversion,” *Personality and Individual Differences*, vol. 13, pp. 475–477, 4 1992.

[64] C. M. McColeman, F. Yang, T. F. Brady, and S. Franconeri, “Rethinking the ranks of visual channels,” *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 10 2021.

[65] M. C. Ioannou, K. Mogg, and B. P. Bradley, “Vigilance for threat: effects of anxiety and defensiveness,” *Personality and Individual Differences*, vol. 36, pp. 1879–1891, 6 2004.

[66] K. Noguchi, C. L. Gohm, and D. J. Dalsky, “Cognitive tendencies of focusing on positive and negative information,” *Journal of Research in Personality*, vol. 40, pp. 891–910, 12 2006.

[67] N. Bayram and M. Aydemir, “Decision-making styles and personality traits,” p. 1, 2017. [Online]. Available: https://www.researchgate.net/publication/330832648

[68] O. Radwan El, R. E. Othman, R. Hallit, S. Obeid, and S. Hallit, “Personality traits, emotional intelligence and decision-making styles in lebanese universities medical students,” *BMC Psychology*, vol. 8, p. 46, 12 2020.

[69] M. Lima and A. Simões, *Neo pi-r: Manual Profissional.* Lisboa: CEGOC, 2000.

A

Consent Form

Formulário de Consentimento Informado

*Obrigatório

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!

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

Vasco Pires
vascocfpires@tecnico.ulisboa.pt

Tiago Delgado
tiago.delgado@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

1. Consente participar neste estudo? *

Marcar apenas uma oval.

- Sim
- Não

2. Idade *

____

3. Género *

Marcar apenas uma oval.

- Masculino
- Feminino
- Outra: _______________________

B

Experiment Form

Formulário de Experiência

*Obrigatório

1. Participant Id *

Tarefa 1

2. A *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

3. B *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

4. C *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

5. D *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

Tarefa 2

6. A *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	◯	◯	◯	◯	◯	◯	◯	◯	◯	extremamente feliz

7. B *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	◯	◯	◯	◯	◯	◯	◯	◯	◯	extremamente feliz

8. C *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	◯	◯	◯	◯	◯	◯	◯	◯	◯	extremamente feliz

9. D *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	◯	◯	◯	◯	◯	◯	◯	◯	◯	extremamente feliz

Tarefa 3

10. A *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

11. B *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

12. C *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

13. D *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

Tarefa 4

14. A *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

15. B *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	( )	( )	( )	( )	( )	( )	( )	( )	( )	extremamente feliz

16. C *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	( )	( )	( )	( )	( )	( )	( )	( )	( )	extremamente feliz

17. D *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	( )	( )	( )	( )	( )	( )	( )	( )	( )	extremamente feliz

Tarefa 5

18. A *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	( )	( )	( )	( )	( )	( )	( )	( )	( )	extremamente feliz

19. B *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste	( )	( )	( )	( )	( )	( )	( )	( )	( )	extremamente feliz

20. C *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

21. D *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

22. E *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

Tarefa 6

23. A *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

24. B *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

25. C *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

26. D *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

27. E *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

Tarefa 7

28. A *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

29. B *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

30. C *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

31. D *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

32. E *

Marcar apenas uma oval.

	1	2	3	4	5	6	7	8	9	
extremamente triste										extremamente feliz

Obrigado pela participação