The Phantom Effect in Information Visualization

Carolina Micaela Pinto Duarte Pereira

Thesis to obtain the Master of Science Degree in

Computer Science and Engineering

Supervisors: Prof. Sandra Pereira Gama
Prof. Tomás Almeida e Silva Martins Alves

Examination Committee

Chairperson: Prof. Luís Manuel Antunes Veiga
Supervisor: Prof. Sandra Pereira Gama
Member of the Committee: Prof. Catarina Pinto Moreira

November 2023

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

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgments

Dou graças por terem sido tantas as pessoas que me acompanharam neste percurso académico, cada uma apoiando-me à sua maneira. Por isso, nestes agradecimentos não há ordem de importância.

Começo por agradecer à Professora Sandra Gama por ter orientado esta dissertação, colocando-me nas mãos a responsabilidade de trazer à comunidade científica mais uma peça de conhecimento sobre a expressão de vieses cognitivos em InfoVis. Agradeço também ao Tomás Alves, meu co-orientador, cujos céleres conselhos me ajudaram a ter um percurso suave no desenvolvimento deste trabalho. Aos dois agradeço por terem ido além do simples orientar. Motivaram-me, fizeram-me sentir acolhida e deram-me confiança para avançar sem medo. Foi uma experiência que superou em muito as minhas expectativas, deixando boas memórias.

Agradeço a todos os professores que de certo modo me foram moldando ao longo de todos estes anos. Entre todos eles, há um nome que não poderia deixar de ser aqui mencionado. Obrigada ao meu professor e amigo Manuel Nunes por me ter ensinado sobre estátuas interiores e sombras em cavernas. Sempre me incentivou a tratar a minha escrita com a devida dignidade.

Se os professores tiveram no meu percurso escolar e académico um papel importante, então o papel da família é elevado a fulcral. Por isso, deixo aqui um profundo agradecimento aos meus pais por todo o esforço e sacrifícios que fizeram para que eu chegasse aqui. Obrigada por me darem todas as condições para que eu estudasse e tivesse gosto em aprender. Obrigada pela educação que me deram desde pequenina e por estarem sempre presentes. Todas as vossas ações contaram para eu me tornar quem sou hoje. Obrigada, pai, por me teres colocado um computador no meu quarto apesar da minha birra. O meu gosto pela informática e esta escolha de percurso académico deve-se muito a ti. Obrigada, mãe, por puxares pelo lado mais emocional de mim e me relembrares que é importante partilhar emoções e cuscar quanto-basta. Agradeço também aos meus avós, especialmente à avó São por ter fé em mim e por tanto me ter ajudado nos tempos mais difíceis. Obrigada à tia Eliana por nunca falhar em pôr toda a gente a rir.

Direciono agora o holofote à parte da família cujos laços não se ditam por sangue ou genética. Foi no meio da pandemia, por ordem do mais frágil encadeamento de acontecimentos, que conheci o Nuno Saavedra, o rapaz que me ensinou o significado de romance. É contigo que tenho aprendido a amar, algo que me tem mostrado todo o tipo de emoções na sua forma mais crua e intensa. Adiciona este parágrafo à lista de intermináveis momentos em que agradeci estares ao meu lado e fazeres parte da minha vida. Obrigada por teres banido as nuvens de alguma vez entrarem no meu céu e por me teres ajudado a pousar o peso psicológico que carregava. Obrigada, também por teres sido um exemplo académico a seguir. Espero que continues a brilhar na ciência e mais ainda na minha vida.

Lembro-me de nas férias de verão de 2018, antes de entrar para o Técnico, me questionar sobre os amigos que iria fazer nesta nova ”escola”. Apesar de continuar a achar que tal pessoa não pode ser real, há documentos que provam que a Daniela Castanho existe e por isso posso agradecer aqui a nossa amizade sem me diagnosticarem com esquizofrenia. Obrigada, por todas as vezes em que rimos para não chorar. Obrigada por todos os anos aturares a minha mensagem a perguntar se depois de todas as desgraças ainda querias fazer grupo comigo. Espero que as nossas multidões continuem a gerar dramas fictícios para que tenhamos motivos para fazer pressão ao Pedro. Aproveito a deixa para agradecer ao Pedro Matias, um amigo que já vem dos tempos do Secundário. Obrigada por seres o cromo mais engraçado, confuso e simpático que conheço. Estás sempre aqui para me apoiar, oferecendo sempre um abraço e dois gigabytes de memes. Depois de dois anos de desafios académicos ultrapassados em conjunto, o Pedro Matono também merece ter o nome aqui. Por último, obrigada Eduardo Espadeiro por me chamares weeb e me teres pedido para ir ao Discord para defender que o DM é fofinho. Mudaste a minha vida.

Há uma pessoa especial a destacar na transição entre amigos e relações terapêuticas. Um grande abraço para a Rita Wahl que com tanta compaixão me acolheu no NDA. Obrigada, Rita, por toda a ajuda, aconselhamento e encorajamento. As suas palavras e energia são suficientes para me alegrar o dia, tornando a vida mais fácil. Obrigada pela sua preocupação com os estudantes e por trabalhar tanto para que ganhemos asas. Agradeço também ao Dr. Mark Serrano pelas longas conversas e divagações. Cavámos fundo e encontrámos uma saída. Obrigada à Dra. Ana Botas por me ter ajudado a organizar as ideias numa altura de redescoberta de mim mesma.

Finalmente, agradeço a todos aqueles que participaram neste estudo. Sem a vossa participação, esta dissertação não teria dado frutos. Obrigada, Frederico Rosado, pelas muitas horas que passámos a recolher dados.

Abstract

With the increased usage of visualizations to draw conclusions from data, the need to guarantee that cognitive biases do not interfere with InfoVis also rises. The Phantom Effect is a cognitive bias in the category of decoy effects where the decoy is an optimal yet unavailable alternative. Decoy effects affect decision tasks by leading the decision toward the decoy. To study the Phantom Effect in InfoVis, we conducted a mixed design experiment where 76 people performed decision tasks with the aid of information represented in five different visualization idioms. The close relationship between cognitive biases and individual differences motivated us to check for correlations between those and the expression of the Phantom Effect. Therefore, we measured participants' cognitive abilities, namely Perceptual Speed and Visual Working Memory, as well as their visualization literacy. The results of this study present proof that the Phantom Effect, similar to other decoy effects, does have an expression in InfoVis. Some visualization idioms, such as two-sided bar charts, are more effective at mitigating this cognitive bias. We also tested three different presentation delay conditions for disclosing the phantom alternative's unavailability and found that one of the conditions can reduce the Phantom Effect. On the other hand, we found no statistically significant correlation between the measured individual differences and the expression of the Phantom Effect. We motivate future research and replication of our study in order to consolidate our results and expand the knowledge on this cognitive bias that occurs in one of the most common visualization tasks.

Keywords

Cognitive Bias; InfoVis; Phantom Effect; Individual Differences; Perceptual Speed; Visual Working Memory

Resumo

Com o crescente recurso a visualizações para extrair conclusões sobre dados, surge a necessidade de garantir que os vieses cognitivos não interferem com InfoVis. O *Phantom Effect* é um viés cognitivo da categoria dos *Decoy Effects*, onde o isco é uma alternativa ótima indisponível. Os *Decoy Effects* são conhecidos por afetar tarefas de decisão ao direcionar a escolha para o isco. Para estudar o *Phantom Effect* em InfoVis, realizámos um estudo híbrido no qual 76 pessoas completaram tarefas de decisão com o auxílio de informação representada em cinco idiomas de visualização diferentes. A relação entre os vieses cognitivos e as diferenças individuais motivou-nos a procurar correlações entre estas e a expressão do *Phantom Effect*. Para tal, medimos capacidades cognitivas dos participantes, nomeadamente a Velocidade Percetiva e a Memória Visual Operacional, bem como a sua literacia em visualização. Os resultados provam que, à semelhança de outros efeitos de isco, o *Phantom Effect* está presente em InfoVis. Alguns idiomas de visualização, como os gráficos de barras bilaterais, são eficazes a colmatar este efeito. Testámos também três formas diferentes de revelar a indisponibilidade da alternativa fantasma e descobrimos que uma destas reduz o *Phantom Effect*. Por outro lado, não encontrámos correlações estatisticamente significativas entre as diferenças individuais medidas e a expressão do *Phantom Effect*. Realçamos a necessidade de replicar e estender este estudo para consolidar o nossos resultados e expandir o conhecimento sobre este viés cognitivo que surge numa das mais comuns tarefas de visualização.

Palavras Chave

Viés Cognitivo; Visualização de Informação; Phantom Effect; Diferenças Individuais; Velocidade Percetiva; Memória Visual Operacional

Contents

1 Introduction **1**
- 1.1 Objectives . . . 4
- 1.2 Research Contributions . . . 5
- 1.3 Organization of the Document . . . 5

2 Background **7**
- 2.1 Cognitive Biases . . . 9
- 2.2 Asymmetrically Dominated Alternatives . . . 10
- 2.3 The Phantom Effect . . . 11

3 Related Work **15**
- 3.1 Effect of Individual Differences on InfoVis . . . 17
  - 3.1.1 Personality Traits . . . 19
  - 3.1.2 Cognitive Abilities . . . 20
- 3.2 Measuring Cognitive Biases on InfoVis . . . 22
- 3.3 Mitigation of Cognitive Biases on InfoVis . . . 23
- 3.4 Discussion . . . 25

4 Methodology **29**
- 4.1 Research Questions and Hypotheses . . . 31
- 4.2 Apparatus . . . 35
  - 4.2.1 The Phantom Effect App . . . 35
  - 4.2.2 Pre-study Questionnaire . . . 35
  - 4.2.3 Perceptual Speed Test . . . 36
  - 4.2.4 Visual Working Memory Test . . . 37
  - 4.2.5 The Phantom Effect Test . . . 37
    - 4.2.5.A Scenarios . . . 38
    - 4.2.5.B Visualization Design . . . 39
    - 4.2.5.C Presentation Delay . . . 42
  - 4.2.6 Familiarity Test . . . 42

4.3 Experimental Design . . . 42
4.3.1 Variables . . . 43
4.3.1.A Independent Variables . . . 43
4.3.1.B Dependent Variables . . . 44
4.4 User Tests . . . 44
4.4.1 Procedure . . . 44
4.4.2 Participants . . . 47
4.4.2.A Demographics . . . 48
4.4.2.B Familiarity . . . 48
4.5 Data Analysis . . . 49
4.5.1 Data Collection . . . 49
4.5.2 Data Cleaning . . . 50
4.5.3 Data Processing . . . 50
4.5.4 Validation Methods . . . 51
5 Results **53**
5.1 Phantom Effect . . . 55
5.1.1 Idioms . . . 56
5.1.2 Metrics for Formative Analysis . . . 56
5.1.3 Interactions . . . 58
5.2 Presentation Delay . . . 60
5.3 Individual Differences . . . 63
5.3.1 Perceptual Speed . . . 63
5.3.2 Visual Working Memory . . . 63
5.3.3 Demographics . . . 64
5.3.4 Visualization Literacy . . . 65
5.4 Discussion . . . 66
5.4.1 Limitations and Future Work . . . 69
6 Conclusion **71**
Bibliography **75**
A VLAT Questions **83**
A.1 One-Sided Bar Chart . . . 85
A.2 Two-Sided Bar Chart . . . 87
A.3 Parallel Coordinates Plot . . . 89
A.4 Scatterplot . . . 91

B Experiment Guide 95

C Informed Consent 97

List of Figures

2.1 Overview of 154 cognitive biases organized by experimental task. 9
2.2 Visual representation of each of the approaches for presentation delays. 13

3.1 Types of visualization tasks observed in literature. 19
3.2 Visualization layouts used by Jeong et al. [1]. 24
3.3 The 29 key articles reviewed by Liu et al. [2] 26

4.1 Locations of the two alternatives and decoys [1]. 32
4.2 Sequence of images in the change detection test. 37
4.3 Visualization idioms used as experimental stimuli in our experiment. 40
4.4 Main steps of the experiment in chronological order. 43
4.5 Number of participants for each of the presentation delay conditions. 48
4.6 Distribution of the sample in terms of education. 48
4.7 Box-plots of the familiarity of participants with each scenario. 49

5.1 Ratio of participants who chose the target option with and without a phantom alternative. 55
5.2 Distribution of interactions across the different visualization idioms. 59
5.3 Table interactions network. 60
5.4 One-sided bar chart interactions network. 60
5.5 Two-sided bar chart interactions network. 61
5.6 Scatterplot interactions network. 61
5.7 Parallel coordinates plot interactions network. 62
5.8 Boxplots of Priming Score across the different presentation delay conditions. 62
5.9 Boxplots of Visualization Literacy Assessment Test (VLAT) scores across the different idioms. 65

List of Tables

3.1 Definitions of the cognitive traits that are common in the visualization literature [2]. 18
3.2 Wall et al.'s formative metrics for measuring cognitive biases [3]. 22

4.1 Independent variables. 45
4.2 Dependent variables. 46
4.3 Classification of attribute distribution based on the user's interactions. 51
4.4 Statistical methods and respective variables. 52

5.1 Summary of chi-square test results and choice ratios among different visualization idioms. 56
5.2 Summary of Kendall's tau-b test results for the formative analysis. 57
5.3 Summary of Point-biserial correlation test results for the impact of Perceptual Speed (PS) in the expression of the Phantom Effect for each visualization idiom. 64
5.4 Summary of Kendall's tau-b correlation test results for the impact of Visual Working Memory (VWM) in the expression of the Phantom Effect for each visualization idiom. 64
5.5 Summary of Kendall's tau-b correlation test results for the impact of VLAT scores in the expression of the Phantom Effect for each visualization idiom. 66

Acronyms

FFM Five Factor Model

InfoVis Information Visualization

PS Perceptual Speed

VLAT Visualization Literacy Assessment Test

VWM Visual Working Memory

1

Introduction

Contents

1.1 Objectives 4

1.2 Research Contributions 5

1.3 Organization of the Document 5

When exposed to big amounts of data, visualizing the information becomes almost inevitable, especially in decision tasks [4]. Information Visualization (InfoVis) is the subject that studies the best ways to visualize and interact with information through different encodings and interaction techniques. When the goal is to draw conclusions from the data, InfoVis highlights trends and points of interest. InfoVis is also a valuable tool for decision-making allowing for a visual comparison of the attributes' values of the different options and representation of alternatives in space [4–6]. Decision tasks, which are the tasks this proposal focuses on, consist of choosing the optimal alternative out of a set of options, by finding the alternative that better corresponds to the needs of each attribute. Using InfoVis, this decision is made with the support of an interactive visualization.

The close relationship of human interaction with InfoVis during task performance makes visualizations prone to cognitive biases. Cognitive biases are involuntary systematic violations of judgment and decision-making logic followed by people, deviating the interpretation of data from reality. In order to guarantee valid and unbiased insights taken from visualizations, there must be an effort to mitigate these cognitive biases. While the causes and effects of these biases have been long studied in the field of psychology, research on how much these biases have an impact in InfoVis is still scarce [2].

The Phantom Effect is a cognitive bias in the category of decoy effects. Decoy effects are biases where the presence of a decoy alternative in a decision task impacts the final decision by shifting the preference towards the decoy. In the case of the Phantom Effect, this decoy is a superior but unavailable option [7, 8]. The information about an option's unavailability can be displayed at different times, a condition known as presentation delay [9]. That information can be disclosed once the visualization is presented (onset), a fixed time interval after (known), or upon making a decision (unknown).

Despite the existence of literature in the field of decoy effects in InfoVis, including work on mitigating decoy effects [1], there are no studies about the Phantom Effect in InfoVis in particular, to our knowledge. Our study replicates Jeong et al.'s experiment with a focus on the Phantom Effect and extends it by testing different presentation delay conditions.

As much as InfoVis brings advantages in data visualization, it can also lead to misinterpretations of said data. This may happen due to design and/or human factors. In terms of design, there are techniques and good practices that have been studied to optimize interfaces and interaction techniques in order to mitigate unwanted effects [10, 11]. On the other hand, human factors are related to cognitive traits such as personality and cognitive abilities, known as individual differences. Human factors are more challenging to mitigate since these depend on each group of individuals with identical cognitive traits, such as more neurotic users or users with External Locus of Control. These groups define types of users that can be used to generalize characteristics when personalizing visualizations to meet most needs of these users in a group adaption strategy. In this study, we measure individual differences and test for a statistically significant variation in the results which may suggest individual differences affect the expression of the Phantom Effect.

Visualizations which are used for tasks that require comparing different attributes with a strong visual component are prone to be affected by cognitive abilities related to visual reasoning, such as Perceptual Speed (PS) and Visual Working Memory (VWM). PS is the cognitive ability that describes the ability of an individual to make accurate comparisons between objects in a given time period [12]. This cognitive ability affects information visualization tasks in which it is needed to compare data points, such as in decision tasks. The ability to store visual information between eye fixations for immediate use [13], known as VWM, equally affects performance on visual tasks, namely decision tasks where people retain the visual encoding of attributes' values for comparison.

Visualizations can be adapted to meet the needs of a type of users grouped by similar individual differences. These adaptions consist of changes in layout, interaction techniques, and/or encoding of attributes resulting in a different visualization for each group. One example is applying elimination by aspects (successive rejection of alternatives within a threshold) or highlighting optimal options in decision problems in order to mitigate the Attraction Effect [14]. This adaptation can be also implemented at the individual user level by gathering information on the user's individual differences such as cognitive traits and demographics and having the visualization adapt itself according to that score. Adapting visualizations to each user group according to their individual differences not only allows for a better user experience but also improves task performance [2]. Besides tweaking the layout and type of visualization for each group, there is a need to explore ways to present the information that mitigate cognitive biases. In the case of our study, we focus on identifying the general groups of users with different needs and preferences which require distinct approaches of customization [14].

1.1 Objectives

The main goal of this study is to **investigate whether the Phantom Effect persists on data presented visually**. In the case of the Phantom Effect having an expression in InfoVis, we extend our goal to study its occurrence by analyzing the **impact of visual encoding and the moment of presentation of the unavailable alternative**. Moreover, we seek to explore **the influence of PS and VWM on the expression of the Phantom Effect**. To achieve such objectives, the following intermediate steps must be followed.

1. Implement the cognitive abilities tests, since there are no questionnaires already available for this effect which fulfill our needs;

2. Implement the presentation delay mechanisms, i.e. when the information about the unavailability of an option is shown;

3. Create the visualizations to be used in this study;
4. Design the experiment, including scripts, equipment needed, and variables to be measured;
5. Conduct a mixed design experiment with 80 participants;
6. Analyse the data and discuss if there are statistically significant changes between the different conditions.

Once the procedure is complete, the goals can be met.

1.2 Research Contributions

We consider this study to be an important addition to the research community in the field of cognitive biases in InfoVis, contributing to answering the need to further explore this subject. Firstly, the **data we gathered on the participants' individual differences and the interactions they performed on the different visualizations** poses a significant contribution to the study of the Phantom Effect in the context of InfoVis, allowing for future extension of the analysis presented in this document. The dataset resulting from our user study is available on Figshare under the DOI: 10.6084/m9.figshare.24224785.v1. We also developed two free-to-use open-source tools for measuring PS and VWM which are available on GitHub at github.com/CarolinaMPereira/Perceptual-Speed-Test-Tool and github.com/CarolinaMPereira/Visual-Working-Memory-Test-Tool.

Moreover, we **consolidated existing literature** on how to measure cognitive biases in InfoVis and confirmed previous results for decoy effects. The results from our analysis contribute to the knowledge about cognitive biases as we verified the expression of the Phantom Effect in InfoVis, which has not been explicitly studied before. Simultaneously, we expanded the investigation of the impact of cognitive abilities on InfoVis, offering new views on their relationship with cognitive biases.

Finally, we **prepared and submitted two scientific articles** to the Transactions on Visualization and Computer Graphics journal (Q1, Scimago). One of the articles describes the two free open-source tools we have developed to measure cognitive abilities while the other article focuses on the results we obtained regarding the expression of the Phantom Effect on InfoVis.

1.3 Organization of the Document

The theoretical concepts needed in order to capture the context and relevance of the Phantom Effect problem are presented in Chapter 2. Once these terms are explained, it gives an overview of the literature around cognitive biases and the effects of individual differences in InfoVis. Before the literature on the Phantom Effect is discussed, we introduce Asymmetrically Dominated Options, a key concept for the understanding of decoy effects.

Chapter 3 discusses studies about cognitive biases in the context of InfoVis. It focuses on ways of measuring cognitive biases through the interaction with a visualization and attempts at mitigating cognitive biases such as the Attraction Effect and decoy effects. It includes a discussion of the literature and identification of research gaps.

Chapter 4 describes the core of our practical study. First, the Research Questions and Hypotheses are introduced. Then we detail the apparatus needed, going through implementation details. Then, the design of the experiment is described in detail in terms of phases and variables. Next, we provide information about the user study we conducted. Finally, we explain which data we collected and how we processed and analyzed it. Follows Chapter 5, describing the results obtained from our data analysis, presenting the needed information for answering our research questions. This chapter finishes with a discussion of the obtained results and an enumeration of our study's limitations, as well as an exploration of future work opportunities. Finally, we close the document with Chapter 6, consisting of a summary of the document and an emphasis on the key ideas.

2

Background

Contents

2.1 Cognitive Biases 9
2.2 Asymmetrically Dominated Alternatives 10
2.3 The Phantom Effect 11

In this chapter, we introduce the key concepts for this study. We start with the definition of cognitive biases in order to contextualize the problem. Then we introduce asymmetrically dominated alternatives, the base concept for the cognitive bias under the spotlight - the Phantom Effect. Finally, the literature on this cognitive bias is reviewed.

2.1 Cognitive Biases

Cognitive biases are systematic violations of judgment and decision-making axioms followed by people. Dimara et al. defined cognitive biases as previously proposed by Pohl [15]. This definition characterizes cognitive biases as "a cognitive phenomenon which: 1) reliably deviates from reality, 2) occurs systematically, 3) occurs involuntarily, 4) is difficult or impossible to avoid, and 5) appears rather distinct from the normal course of information processing" [16].

Until now, cognitive biases have been described and categorized under an explanatory view, i.e. why they occur. There was a need for categorizing the biases in a task-based taxonomy in order for researchers and designers to easily detect which cognitive biases are more relevant to their work, depending on the type of task being asked from the users. Dimara et al. answered this need for a new taxonomy classifying cognitive biases. The task-based taxonomy consists of categorizing by the experimental situation in which the cognitive biases have occurred. The proposed taxonomy goes beyond the previously proposed taxonomy by Pohl, having more low-level tasks for better detail in classification. This makes it the largest cognitive biases taxonomy available. The task-categories were established by first identifying the tasks performed when a cognitive bias was measured on relevant literature. Then, the authors grouped tasks with similar labels and proposed the final taxonomy presented in Figure 2.1.

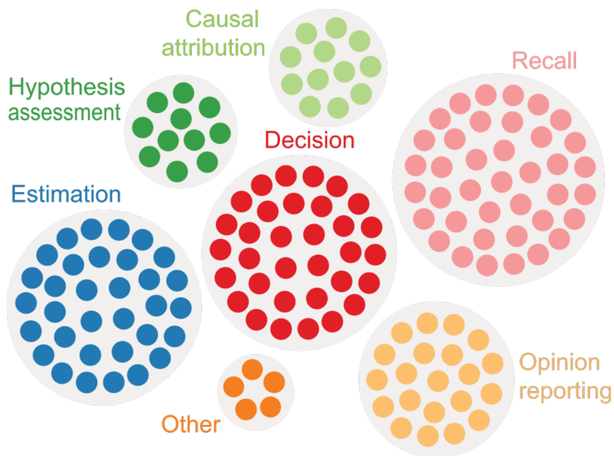


Figure 2.1: Overview of 154 cognitive biases organized by experimental task. Dots represent the different cognitive biases [16].

Below, we give a few examples of cognitive biases from each task-category:

- **Estimation** (assessing a value or quantity), e.g. the *anchoring effect* - the estimation is influenced by specific information (the anchor, or reference point) gathered before drawing conclusions [17];
- **Decision** (selecting an option), e.g. the Phantom Effect - the selected option is influenced by a superior but unavailable alternative [18];
- **Recall** (recognizing of previously seen content), e.g. the *bizarreness effect* - bizarre items are easier to recall [19];
- **Opinion reporting** (sharing opinions and beliefs), e.g. the *bandwagon effect* - the opinion of others influence one's beliefs [20];
- **Causal attribution** (explaining the cause of events or behaviors), e.g. the *actor-observer asymmetry* - when others fail the cause is attributed to their behavior or personality, while one's own failures are due to the current circumstances [21];
- **Hypothesis assessment** (assessing the validity of hypothesis), e.g. the *clustering illusion* - seeing patterns in random distributions of data, for example, in a dot field [22];
- **Other** (tasks that do not fit in any of the above task-categories) - *attentional bias* (weighting the processing of information [23]), *risk compensation* (when constantly exposed to a risk, it is perceived as less dangerous [24]), *ostrich effect* (being selective against negative information [25]).

In this taxonomy, both the Decoy Effect (refered to as Attraction Effect) and the Phantom Effect belong to the category *Decision*, more specifically to the subcategory *Baseline*. Dimara classifies each bias in terms of relevance to InfoVis, giving the Phantom Effect a 4 and the Attraction Effect a 8, where 4 means "Not discussed in visualization but likely relevant" and 8 means "Evidence for the alleviation of the cognitive bias in visualization".

Although Dimara et al. produced the taxonomy to ground the research field, researchers must conduct more experimental studies on most documented cognitive biases. Up to today, cognitive biases in the categories of estimation and decision tasks have been target of several studies, resulting in existing proof of their existence and mitigation in visualization. There is a lack of experimental studies on most documented cognitive biases. Researchers must keep in mind that they too are prone to biases in their experiments, especially in InfoVis. Even if not directly, every bias can have an impact on the conclusion drawing from InfoVis.

2.2 Asymmetrically Dominated Alternatives

In decision tasks, a person must choose one of the given options. Each of these options is called an alternative. An alternative can be worse than other, i.e. dominated by others depending on the relevant attributes and goal of the decision. When an alternative is dominated by at least one item in the set but not by at least one other, we say it is asymmetrically dominated [26].

Asymmetrically dominated alternatives are the core of decoy effect cognitive biases such as the attraction effect and the Phantom Effect. The attraction effect is a cognitive bias where a decision is influenced by the presence of a similar, but slightly inferior alternative. Although several studies conclude that the attraction effect only occurs on tabular numeric data, others contradict such affirmation. Dimara et al. [27] produced tabular data and scatterplots with sets that purposefully include a decoy to manifest the attraction effect. For comparison, the same visualizations were replicated without decoys. This study justifies that if the attraction effect occurs in information visualization, then it is not limited to occur on tabular numeric data. One of the experiments conducted by the authors consisted of having enough alternatives for a tabular representation to be not practical. Adding more alternatives can be achieved by:

1. Adding more alternatives on the Pareto Front, which is a "set of formally uncomparable or non-dominated alternatives";
2. Adding more decoys;
3. Adding "distractors", which are irrelevant alternatives that are not part of the set of targets, competitors, or decoys.

This experiment confirmed that more participants choose the target option when decoys are present than those who do not. This result is visible both on tables and scatterplots. Therefore, not only the attraction effect is present in information visualization, but also it persists when more alternatives are added. The authors proposed a few hypotheses on possible ways to mitigate the attraction effect, which were tested in a later study [14]. Debiasing techniques for the attraction effect include showing only the Pareto Front or highlighting it, de-emphasizing dominated options, and offering debiasing customization options activated on command. The results of this mitigation attempt are discussed in Section 3.3.

2.3 The Phantom Effect

The Phantom Effect is a type of cognitive bias in the category of decoy effects. The decoy effects work by adding a third alternative to a two-choice set composed of a target and a competitor. When the preference between two alternatives depends on a third existing superior but unavailable one, we are facing the Phantom Effect. The unavailability of that option is related to internal or external constraints, highlighting the importance of the attribute on which it excelled [7]. This cognitive bias was first coined as the "Phantom Effect" by Pratkanis and Farquhar in 1986, despite already being studied by social psychologists. Trueblood et al. [8] give an intuitive example of this cognitive bias: "your dinner options are tilapia, vegetable lasagna, and spaghetti with meatballs. You immediately prefer the spaghetti with meatballs, but find out later that it is unavailable. In this situation, you choose the vegetable lasagna over the tilapia". This example shows how a person's choice is influenced by which of the available options is more similar to the unavailable initial choice.

According to Pratkanis and Farquhar [9] from the point of view of the decision maker, a phantom alternative is defined by three dimensions: "(a) knowledge of availability at the time a decision is made, (b) perceptions of the nature of the constraint or reason for unavailability, and (c) sense of entitlement to possess the phantom". The "perception of the nature of the constraint or reason for unavailability" refers to how difficult it is to remove the unavailability constraint. The unavailability can be absolute if it is impossible to change the restrictions, or conditional if additional efforts can be made to remove it. The "sense of entitlement to possess the phantom" comes from the decision-maker's beliefs and expectations of what they deserve. Several studies have been conducted to test different scenarios for the first dimension ("knowledge of availability at the time a decision is made"). Most studies available hide the unavailability of the decoy so that the user needs to consider it later. This is accomplished either by using a known delay presentation or an unknown presentation or marking the phantom as unavailable from the beginning. Figure 2.2 schematizes these three different approaches. Known delay presentation consists of displaying the information of unavailability with a delay before the choice is made. Studies by Pettibone and Wedell [18, 28] conclude this delay should be of three seconds. In the unknown presentation method, we wait until the participant makes the decision before the decoy is presented as unavailable. The participants are then required to make a different choice. Alternatively, participants may know from the beginning which option is unavailable. A study by Scarpi and Pizzi [29] concluded that the unknown presentation is more efficient than the known delay presentation since the latter can have a negative impact on choosing the target. This happens due to the unexpected loss of freedom of choice, meaning a decrease in the target's attractiveness.

In their review of research on the Phantom Effect [9], Pratkanis and Farquhar point out seven empirical generalizations about human behavior in the presence of unavailable options:

1. **The scarcity-attractiveness effect -** unavailable or rare items are more attractive.
2. **Phantom as a reference point -** available options will be weighted using the unavailable one as a reference.
3. **The contrast effect -** the evaluation of available alternatives becomes more negative.
4. **Decision criterion shift -** the strength of the phantom's attributes dictates the criteria used to weight each of the other alternatives.
5. **Phantom fixation -** the unavailable option becomes the focus point.

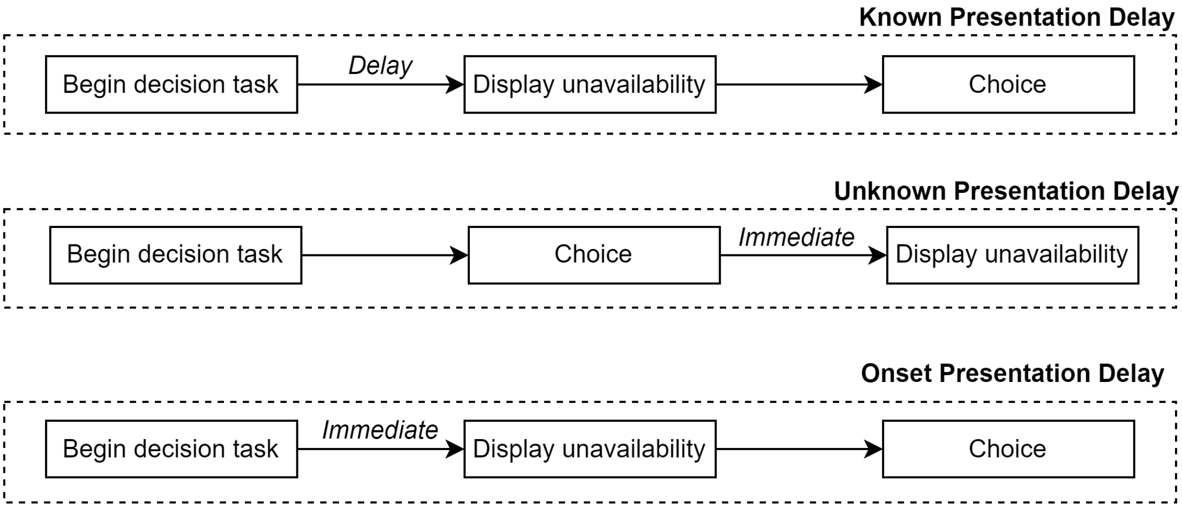


Figure 2.2: Visual representation of each of the approaches for presentation delays. On onset presentation delay, the unavailability is presented at the same time as the task begins.

6. **The frustration-deprived effect -** people feel frustrated for being unable to get the unavailable option, resulting in dissatisfaction with the available alternatives.

7. **Phantom as a self-threat -** being unable to choose the phantom decreases self-esteem.

There are four models of the Phantom Effect that explain users' choice pattern in the presence of a phantom decoy [8]. Trueblood et al. summarize that both value-shift models (context-dependent variation in the subjective assessment of attributes) and dynamic preference accumulation models (accumulation for different attributes during a decision process) "predict that the phantom decoy will increase preference for the competitor over the target." The relative-advantage model (selecting the alternative that represents a smaller loss compared to the reference alternative, i.e. loss-aversion) and comparison-induced distortion theory (the type of comparison made depends on the goal of the decision maker) predicts the opposite. A study conducted by Highhouse et al. [7] concludes that loss-aversion drives the decision-making process in the presence of a phantom.

Farquhar and Pratkanis [30] point out the need to counteract the Phantom Effect in decision-making. One of the approaches is adding complementary phantoms, i.e. alternatives that stand out in an attribute different from the one that makes the original phantom preferable. This approach mitigates the contrast effect and the decision criterion shift. Another approach is rewriting the problem in order to present an alternative to an unavailable option as "the closest substitute" instead of "the only choice". This wording will reinforce self-consistency as well as the motivation for a choice, decreasing the effect of the phantom as a self-threat. The authors defend that promoting these contingency measures towards phantom options reduces the impact of discovering an option is unavailable.

Phantom alternatives offer valuable information on a decision problem's limits, helping the restructuring of a decision problem in order to generate new options. Note that the elimination of such phantoms will exclude the opportunity to use that information and to be critical as to why they are unavailable. However, phantoms can also have adverse effects such as biases, deception, and sub-optimal decisions. Thus, it is important to become aware that some options may not be available so that possibility has the less possible negative impact.

3

Related Work

Contents

3.1 Effect of Individual Differences on InfoVis 17
3.2 Measuring Cognitive Biases on InfoVis 22
3.3 Mitigation of Cognitive Biases on InfoVis 23
3.4 Discussion 25

Research on the consequences of cognitive biases on InfoVis related tasks is still taking its first steps, although it has been studied both in psychology and marketing. A foundation of knowledge about the effect of each cognitive bias is being constructed. Once there is consolidated and consistent knowledge and bias metrics, there is a need to study how those consequences can be mitigated in order to lower the effects of the biases. In this chapter, we summarize and discuss relevant studies within the fields of interest for this study in order to compile the existing experimental knowledge and developed methods. First, we present how individual differences, namely cognitive traits, impact InfoVis. Then, we explore proposed metrics for measuring cognitive biases in interaction. Finally, we give an overview of experimental attempts to mitigate already studied cognitive biases and discuss the present state of the art.

3.1 Effect of Individual Differences on InfoVis

InfoVis is the visual representation of abstract data, often resorting to interactive interfaces. By visualizing the data, humans can easily extract information from large data sets and become aware of relationships between data points, as well as detect tendencies and other relevant insights about the data.

The way humans interpret and interact with data from visualizations is affected by individual differences in personality traits and cognitive abilities that are common in visualization literature. Personality traits are behavioral, feeling, and thinking patterns, characterizing consistent and unique individual ways of interacting with the world [31]. On the other hand, Cognitive abilities are mental skills referring to how individuals reason about information, including how information is retained and the capability to solve problems using such information [32].

Although individual differences have been long studied in the field of psychology, **the investigation of the effects of individual differences on InfoVis is still at its dawn**. Liu et al. [2] compiled the available literature on a state-of-the-art review about the effect of personality traits and cognitive abilities on InfoVis. The authors condensed the results from several studies and discussed them grouped by type of cognitive abilities and personality traits, naming them according to the dimensions of the Five Factor Model (FFM). The FFM is the theory that identifies the big five personality traits: extraversion, neuroticism, openness to experience, agreeableness, and conscientiousness [33]. Besides this categorization, the literature was also grouped by type of visualization design, giving an overview of the most studied ones. There is a balance between simple visualizations, graphs, and high-dimensional designs. However, statistical visualizations have the lowest number of studies associated. The type of tasks performed in the experiments of the 27 key articles reviewed by the authors are mostly of the types *search* (looking for an element that fulfills the requirements in a set), *find extremum* (finding extreme values in

Table 3.1: Definitions of the cognitive traits that are common in the visualization literature [2].

<table>
<tr><th colspan="4">Cognitive Traits</th></tr>
<tr><td rowspan="7">Personality Traits</td><td rowspan="5">Five Factor Model</td><td>Extraversion</td><td>The tendency to engage with the external world.</td></tr>
<tr><td>Neuroticism</td><td>The tendency to experience negative emotions such as stress, depression, or anger.</td></tr>
<tr><td>Openness to Experience</td><td>The propensity to seek, appreciate, understand, and use information.</td></tr>
<tr><td>Agreeableness</td><td>The tendency to consider the harmony among a group of individuals.</td></tr>
<tr><td>Conscientiousness</td><td>The propensity to control one’s impulse and display self-discipline.</td></tr>
<tr><td colspan="2">Locus of Control</td><td>The extent to which a person believes the external world is influenced by their own actions, and/or whether they have control over the outcome of events occurring around them. Individuals are considered Internal or External depending on their level of Locus of Control.</td></tr>
<tr><td colspan="2">Need for Cognition</td><td>The tendency to engage in and enjoy activities that involve thinking.</td></tr>
<tr><td rowspan="5">Cognitive Abilities</td><td colspan="2">Spacial Ability</td><td>The ability to generate, understand, reason, and memorize spatial relations among objects.</td></tr>
<tr><td colspan="2">Perceptual Speed</td><td>The rate at which an individual is able to make accurate visual comparisons between objects.</td></tr>
<tr><td colspan="2">Visual / Spatial Memory</td><td>The capacity to remember the appearance, configuration, location, and/or orientation of an object.</td></tr>
<tr><td colspan="2">Working Memory</td><td>The capacity to store information for immediate use.</td></tr>
<tr><td colspan="2">Associative Memory</td><td>The ability to recall relationships between two unrelated items.</td></tr>
</table>

the dataset) or *compute derived value* (computing a numeric representation for a group of data) [34]. Another five task types are documented and the distribution can be seen in Figure 3.1. The results of

Figure 3.1: The types and distribution of tasks observed in the literature on individual differences in visualization use. Adapted from Liu et al. [2]

the reviewed articles are measured using different recorded data. The most common measures are speed and accuracy, both used to describe task performance. Speed measures the time a user takes to complete a task. Accuracy measures the correspondence between the user's interaction results and the reference values. For example, how much the users choose the target data point in a scatterplot instead of choosing another given point. Some experiments include eye tracking, allowing to follow the user's gaze while interacting with an interface. Finally, users can be asked to report their findings and conclusions so that researchers can evaluate the amount and quality of insights gathered. A similar process is followed for subjective data on the experience, through written or verbal feedback.

The authors divided the article into categories of cognitive traits mentioned in Table 3.1. Each individual has distinct strengths and flaws according to their cognitive traits. Understanding the bigger differences between user groups allows visualization and interface designers to better ponder their options in order to mitigate the consequences of such user group differences. The generalizations on the effect of individual differences in visualization were discussed by Liu et al. in this same article [2].

3.1.1 Personality Traits

In the literature, we found a few studies the personality traits described by the FFM in the context of InfoVis. These studies focus on categorizing how people with different personalities interact with the visualizations, describing problem-solving strategies different people employ, and also how they perform in the presented tasks.

<u>Extraverts</u> exhibited different problem-solving approaches from introverts. Extraverts need less time to understand the problem and complete the task, namely search tasks [35]. More <u>neurotic</u> individuals

have faster completion times as well as higher accuracy on hierarchical search tasks. Possible explanations [35–37] for these results include more neurotic people: paying more attention to tasks; being more in control of interfaces, manipulating them better; pressuring themselves to have a good performance; taking advantage of the out-of-control feeling when facing visualizations they are unfamiliar with. Delgado et al. [38] confirmed that more neurotic people have better task performance. Moreover, neurotic individuals perform better in scatterplot matrices.

Openness to Experience, Conscientiousness, and Agreeableness are under-explored traits in the field of visualization. The latter two are mentioned and measured in the context of the FFM but have no significant impact documented. On the other hand, Ziemkiewicz et al. [39] relate openness capability to overcome conflicting visual and verbal metaphors when dealing with hierarchical visualizations. Conscientiousness affects user confidence on decision tasks, as highly conscientious individuals change their choices more [40].

Locus of Control can be evaluated in a continuous spectrum from "Internals" (individuals who tend to attribute causality to internal factors such as their own actions and the environment around them) and "Externals" (individuals who tend to attribute causality to external factors, consequently being better at "adapting their thinking to external representations") [2]. The layout is a key aspect of the relationship between Locus of Control and interaction with the visualization. Internals find it hard to interact with more contained layouts, performing better in parallel-coordinates plots [38] and Externals can easily adjust to various visual layouts [41]. While Internals complete procedural tasks faster, Externals are faster and more accurate with more contained layouts [36]. The strategy used in hierarchical search tasks changes according to external or internal Locus of Control. Since Locus of Control can be temporarily varied even if it stays stable throughout adulthood [41,42], it can be taken advantage of when designing an interface, since it was proven that the design is a factor for this cognitive ability.

More studies need to be carried out in order to consolidate the effects of Need for Cognition on visualization use. However, there are already a few conclusions drawn from existing research. Need for Cognition can significantly increase the user's accuracy in recall tasks [43]. There is no significant relationship between need for cognition and speed as studies often contradict each other's results.

3.1.2 Cognitive Abilities

Similarly to what happens in the context of personality traits, researchers have been studying the impact of cognitive abilities on InfoVis. Most studies focus on measuring differences on task performance and preference for certain visual encodings.

Spatial Ability highly impacts task performance. The higher the Spatial Ability, the better the recall, and the lower the Spatial Ability the higher the precision [44]. Higher Spatial Ability means better performance when dealing with real or computer-generated 3D objects [45].

PS is related to comparison tasks as higher PS means greater accuracy in identifying different objects or patterns. PS positively affects task performance, benefiting the accuracy of computing derived values [46]. Findings also suggest that higher PS leads to a higher learning rate [47], meaning completion times and accuracy quickly get better with each interaction with the visualization. PS is also related to an individual's ease of learning and search performance [48]. The interaction with a system impacts learning rate and search performance differently depending on the user's PS and mechanisms available on the system. In that same study, Allen et al. point out that features that allow for an improvement and expansion of the search vocabulary such as query expansion have a positive effect on search quality. Studies on the usability of Enterprise Resource Planning (ERP) systems claim that individuals with high cognitive abilities, especially high PS can easily overcome poor usability than those with lower abilities [49, 50]. Those individuals with higher PS also presented better task performance, results also achieved by Ziefle et al. [50]. PS can be measured through tasks that require the individual to decide if two images are the same or different or in which the individual needs to identify certain objects or patterns [12].

Visual/Spatial Memory has no strong connection with search performance or strategy [51], although it can slightly improve accuracy in projection tasks [45,46]. The biggest difference between various levels of Visual and Spatial Memory can be detected while analyzing eye-tracking data, meaning participants use visualizations differently [52].

Working memory is the short-term memory associated with cognitive tasks. There are two types of working memory: **VWM (retention of visual information between eye fixations)** and **Verbal Working Memory (temporary storage and manipulation of language-related information)**. Luck and Vogel [53] concluded that VWM stores information about objects as a whole and not singular attributes, meaning that a person may memorize sixteen different attributes when these are presented through four different objects. A study by Hayhoe et al. [54] indicates that each eye fixation retains a particular bit of information that depends on the task demands at the moment of the fixation. The effect of working memory depends on the type of visualization used. Radar graphs are preferable when users have higher VWM while bar graphs are easier for lower Verbal Working Memory users [55]. Even if layout impacts performance, users differing in working memory show no change in the willingness to customize an interface when given the choice [56]. Verbal Working Memory negatively impacts task performance. However, it has no statistically significant correlation with accuracy, understanding, or interest [43]. Additionally, eye-tracking data allowed researchers [57] to conclude thatindividuals with higher verbal working memory spend less time reading textual information in visualizations.

Finally, there is a scarce investigation on Associative Memory. The only publication [51] on this matter suggests that higher associative memory participants can easily perform well in information retrieval tasks. Subjective feedback on this study refers that good associative memory users think spatial

Table 3.2: Formative metrics for measuring cognitive biases. Each metric computes a specific behavior that can be analyzed to detect bias. Adapted from Wall et al. [3].

Metric	Description	Example Behaviour
Data Point Coverage (DPC)	Measures *how much of the dataset the user has interacted with.*	User interacted with only 3 of 100 players.
Data Point Distribution (DPD)	Measures *how evenly the user is focusing their interactions across the dataset.*	User interacted with some data points dozens of times while ignoring others.
Attribute Coverage (AC)	Measures the *range of an attribute's values explored by the user's interactions.*	User interacted with only players over 84 inches tall, when height ranges from 67 to 88 inches.
Attribute Distribution (AD)	Measures the *difference in the distribution of the user's interactions to the distribution of a particular attribute in the dataset.*	User interacted with a uniform sample of data while the attribute follows a normal curve.
Attribute Weight Coverage (AWC)	Measures the *range of weights for a particular attribute explored by the user's interactions.*	User sets weight values between 0–0.2, ignoring weight values less than 0 and greater than 0.2.
Attribute Weight Distribution (AWD)	Measures the *difference in the distribution of the user-defined weights for an attribute to a baseline of unbiased information weighting.*	User weights follow an exponential distribution, with higher probability for low weight values than high attribute weights.

interfaces are useful.

Analyzing the distribution of each of these categories allows the production of an overview that **highlights future research opportunities by detecting missing or scarce studies on each group**. This review [2] shows evidence that most cognitive traits can impact visualization use. However, more studies in the field are needed to consolidate knowledge about the relationships between individual differences and visualization use and later use that knowledge to implement ways of working around cognitive traits.

3.2 Measuring Cognitive Biases on InfoVis

Efforts have been made to study a way of measuring how biased a user is while analyzing data or making a decision in a visualization. Wall et al. [3] designed a data analysis and decision-making prompt propitious to the anchoring bias with the goal of applying and formalizing the interactive bias metrics described in previous work by Wall et al. [58]. The application of such metrics enables the measurement of the expression of biases through interaction with a visualization. The metrics used in this study are described in Table 3.2.

As the designations suggest, these metrics focus on which data points and attributes the user interacts with. The positive results of this study allowed the authors to prove the usability and accuracy of those metrics, as well as determine the best way to analyze the data gathered during the user's interactions. Wall et al. point out the importance of designing for analyzing versus designing for usability. In the future, a study with more participants, the weighting of time in interactions, and an unbiased interaction interface can provide an even better analysis of these metrics.

Dimara et al. [14, 27] have been measuring cognitive biases, namely the Attraction Effect, using decision-making. When the cognitive bias leads to the existence of target alternatives, it is possible to check if an individual was affected by the bias by analyzing the type of option they chose. Cho et al. achieved similar results by applying an exploratory data analysis method for analyzing users' interaction logs and path of interactions [59].

3.3 Mitigation of Cognitive Biases on InfoVis

The study of cognitive biases allows researchers to come up with ways of reducing the impact or taking advantage of those. This is considered to be one of the greatest challenges in InfoVis. Jeong et al. [1] studied how using visual interfaces to display information might mitigate decoy effects. The experiments consisted of representing a decision problem consisting of three options (in the case of decoy effects) or two options (in the particular case of the Attraction Effect) using different visualization techniques. The decision context was chosen to minimize individual differences in preference for attributes and to portray daily life decisions. In these experiments, the types of visualizations used were the ones that allowed the representation of several attributes as well as made it possible to compare options. The visualizations used can be seen in Figure 3.2. The one-sided bar charts and the two-sided bar charts were the most effective in decreasing the decoy effect. The authors suggest this happens due to the ability these have to support decision-making strategies used by the participants. On the other hand, the scatterplot did not help in making the decision. People are not familiar with parallel-coordinate plots, consequently, it becomes harder for them to interpret this type of visualization which justifies the negative results.

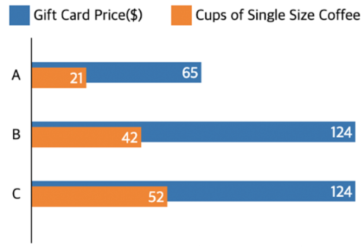
The authors also pointed out that the expression of the decoy effect depends on the type of tasks and scenarios. There are scenarios where people usually have the same preference and ignore the decoy. For example, when choosing a data plan, people tend to choose the one with the largest data size independently of the other attributes. Therefore, for more reliable results, **it is needed that the scenarios of the experiment are as abstract as possible**, while still having attributes the users are familiar with to allow reasoning.

Jeong et al. hypothesized that the person's decision-making strategy influences how much one can be affected by the decoy effect. However, the results of the experiments show no statistically significant relationship. The authors suggest a later study on which visualization techniques better support each decision-making style.

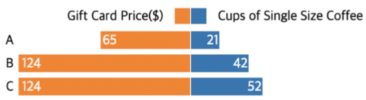
Despite the results being valid for decoy effects in general, and the Phantom Effect being included in the category of decoy effects, **we cannot be sure if the results remain true for the Phantom Effect**. Moreover, the authors refer to the decoy option without discriminating what type of decoy it is (symmetrically dominated, asymmetrically dominated, or phantom decoys), which means **there is no evidence this study contemplates the Phantom Effect**.

Option	Gift Card Price ($)	Cups of Single Size Coffee
A	65	21
B	124	42
C	124	52

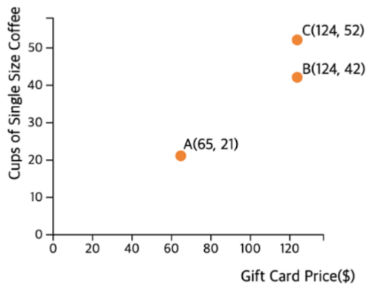
(a) Table



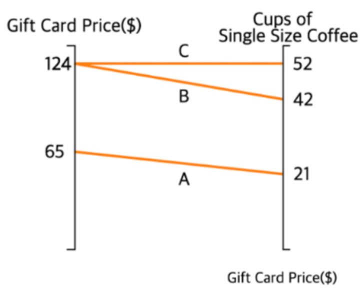
(b) One-Sided Bar Chart



(c) Two-Sided bar Chart



(d) Scatterplot



(e) Parallel Coordinate Plot

Figure 3.2: Examples of experimental stimuli. Four visualizations and a table for a baseline: (a) Table, (b) One-Sided Bar Chart, (c) Two-Sided Bar Chart, (d) Scatterplots, and (e) Parallel-Coordinate Plot. [1].

Dimara et al. [14] enumerated two different perspectives for bias mitigation: decision-maker- and environment-centered. Most existing studies focus on the first strategy, implementing ways of educating the user about cognitive biases. Although this technique is proven successful, the effect is temporary.

The study consisted of testing the participants at every given time interval. The results show that the more time passed since the first test session in which the training was made, the more the participants were primed by the bias.

Dimara et al. followed an environment-centered approach in an attempt to mitigate the Attraction Effect. In this study, the authors put to test the mitigation hypotheses suggested in their previous study [27]. Two experiments were conducted separately, both using scatterplots, each one focusing on a different approach. The first one, the “Pareto Experiment”, consisted of changing the visual representation of data in order to highlight the Pareto front, i.e. the optimal options. The second one referred to as the “Deletion Rationale”, allowed the users to reject all alternatives outside a given threshold testing the non-compensatory strategy “elimination by aspects”. Elimination by aspects consists of repeating the process of eliminating alternatives within the threshold until only one is left, avoiding this way the explicit examination of trade-offs. The two experiments revealed positive results in attraction bias mitigation. Furthermore, the elimination by aspects strategy shows a stronger decrease in the bias expression than highlighting the Pareto front.

In a recent study, Wall et al. [60] developed a technique known as *interaction traces* with the goal of showing a user’s interaction history. This research appears in the context of testing new debiasing strategies. Although interaction traces can increase the awareness of unconscious biases, they may also amplify conscious ones. Thus, the results gathered from the three experiments conducted are inconclusive in terms of bias mitigation.

3.4 Discussion

The studies mentioned in the previous sections identify a correlation between individual differences and InfoVis. **The way an individual perceives a visualization changes depending both on their personality traits and cognitive abilities**.

Some interesting insights with interest for our study can be taken from the table presented in Liu et al.’s state-of-the-art review [2] (Figure 3.3). This overview of traits, visualizations, tasks, and measures as well as which traits have no effect reported in the reviewed literature, allow us to gather insights from the state of the art. For example, it is visible that a big fraction of the literature that studied the relationship between personality traits (FFM, Locus of Control, and Need for Cognition) and visualizations concluded there is no measurable effect of those in InfoVis. While 45.4% of the studies that evaluated personality traits found no relationship between those and visualizations, the fraction of works evaluating cognitive abilities reporting no relationship with visualizations corresponds to 32.7%, significantly less. Therefore, **the existing literature suggests that cognitive traits are more prone to have some kind of impact in InfoVis than personality traits**. Another interesting insight with relevance to our study is the fact that eleven out of 15 studies that evaluated PS, also evaluated VWM, and ten out of twelve studies that evaluated VWM also evaluated PS. This might indicate **there is a relationship between PS and VWM.**

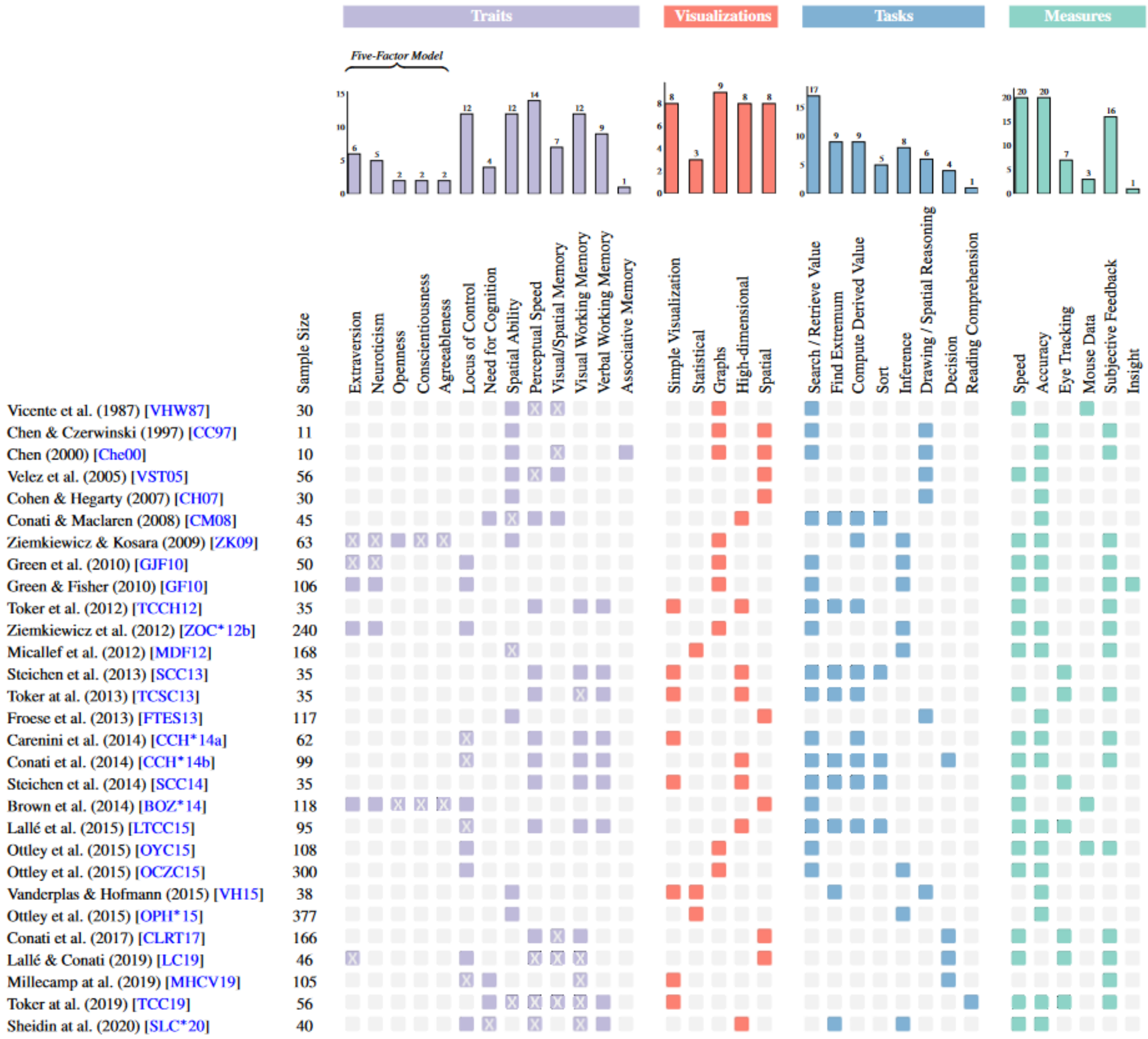


Figure 3.3: The 29 key articles reviewed by Liu et al. The filled boxes indicate the traits (lilac), visualizations (red), tasks (blue), and measures (green) studied in the literature. The "X" indicates "traits that were evaluated, but no measurable effect was reported under the studied conditions". Adapted from Liu et al. [2].

Cognitive biases have been studied mostly in the fields of psychology and marketing. Studies such as Dimara et al.'s [27] are pioneers in exploring the expression of cognitive biases in visualizations. Moreover, this study shows cognitive biases can have an impact in InfoVis by deviating the user's perception of data from reality. Wall et al. proposed a way of measuring cognitive biases in InfoVis through a set of measures [3]. Another study by Dimara et al. shows significant evidence that the Attraction Effect can be mitigated by implementing an interaction technique known as elimination by aspects, or even by highlighting the Pareto Front. This means **not only cognitive biases can be measured in**

visualizations, but also they can be mitigated. With the widespread of visualizations and the more and more frequent exposure of people to large amounts of information, biases should be one of the biggest concerns. However, **there are still few studies in this field**. Currently, besides the Attraction Effect, there are also studies on decoy effects, namely the one by Jeong et at. [1]. There is still a large list of cognitive biases yet to be studied in the context of InfoVis. Most cognitive biases are related to tasks that are often performed with the aid of visualizations such as estimation, decision, recall, and hypothesis assessment tasks.

One example of a cognitive bias with expression in decision tasks is the Phantom Effect, the bias addressed in this study. Since the Phantom Effect is a cognitive bias in the category of decoy effects, Jeong et al.'s study [1] takes the first step in analyzing the expression of these effects in InfoVis. However, this study lacks the focus on the Phantom Effect since this cognitive bias is not mentioned nor the experiment included unavailable options. Moreover, Dimara et al. [16] classified the Phantom Effect as "Not discussed in visualization but likely relevant", hinting for the need to investigate its expression in InfoVis. Therefore, **there is a research gap around the Phantom Effect** which we propose to explore.

One of the aspects frequently mentioned and studied alongside the Phantom Effect is the **presentation delay** [30]. The literature on this subject is inconsistent in conclusions taken on which of the three presentation delay methods is more effective in mitigating this cognitive bias. Moreover, none of the studies specifically mentions InfoVis. Therefore, including this condition in our study not only contributes to solidifying research on which presentation delay better mitigates the Phantom Effect, but also contributes to the implementation of these methods in visualizations.

4

Methodology

Contents

4.1 Research Questions and Hypotheses 31
4.2 Apparatus 35
4.3 Experimental Design 42
4.4 User Tests 44
4.5 Data Analysis 49

In this section, we describe the method to detect the expression of the Phantom Effect on visualizations. We introduce the research questions we aim to answer with this experiment as well as the relevant hypotheses to be tested. Then, we describe how the experiment was designed and conducted in terms of apparatus, procedure, and participants. We also give an overview of how we collected and processed data. Finally, we summarize the validation methods we used to answer our research questions using the obtained data.

4.1 Research Questions and Hypotheses

The fields of InfoVis and cognitive biases lack experiments on whether or not presenting information visually influences the expression of the Phantom Effect. In order to investigate the expression of the Phantom Effect on information visualization, it is mandatory to develop an InfoVis environment in which the presented information can bias the user choice through the existence of an unavailable option.

The expression of the Phantom Effect depends on the presence of a phantom alternative that dominates the others. Besides that unavailable option, there is a target alternative and a competitor. While both are optimal, the target excels in one dimension while the competitor is better in the other, in a two-dimensional decision problem. As shown in Figure 4.1, the phantom alternative should be placed near the target since the consequence of the Phantom Effect is choosing the most similar alternative to the unavailable one. Thus, **by verifying if the target is chosen more often than the competitor, we are able to tell if the individual has been primed by the Phantom Effect or not**.

Despite the studies on decoy effects such as the Phantom Effect mentioned both in Chapters 2 and 3, to our knowledge no study has been conducted on the Phantom Effect in InfoVis with more than three choice alternatives. So far, studies have been made using simple decision problems with only three alternatives. **Expanding the set of alternatives to more than three allows the creation of a scenario identical to a real-life visualization, as it allows testing if the Phantom Effect persists in the presence of more than three alternatives**.

Jeong et al. [1] used four different InfoVis idioms (one-sided bar chart, two-sided bar chart, scatterplot, and parallel-coordinate plot) as stimuli for their experiment on mitigating decoy effects using visual interfaces, as previously shown in Figure 3.2. The results show that since people are less familiar with parallel-coordinate plots, this visualization idiom is not effective in mitigating decoy effects. On the other hand, both one-sided bar charts and two-sided bar charts successfully prevent decoy effects. Scatterplots turn out not to provide help in decision-making tasks where the individual is primed by a decoy effect. The results were obtained by comparing with a baseline tabular representation of a decision problem data. Motivated by Jeong et al.'s study's [1] results and lack of focus on the Phantom Effect itself, we aim to answer the following research question:

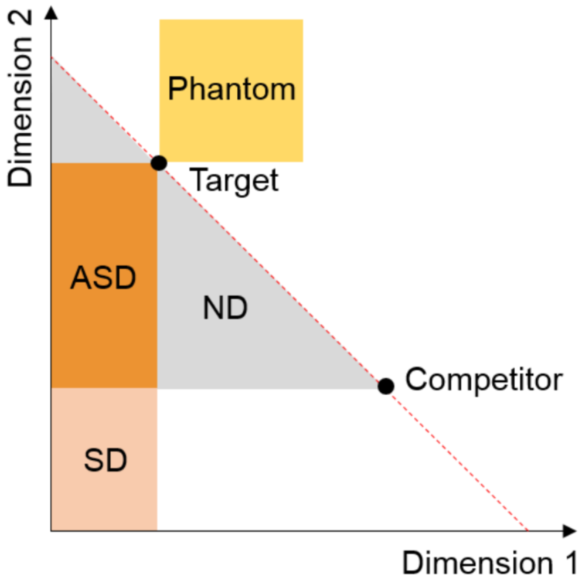


Figure 4.1: Locations of the two alternatives and decoys [1]. The yellow area represents where the phantoms should be located in order to result in the Phantom Effect. Similarly, ASD (asymmetrically dominated alternatives), ND (non-dominated alternatives), and SD (symmetrical dominated alternatives) are placed on the designated regions in the plot to be effective. The target and the competitor can be located anywhere along the red dashed line which represents incomparable, and possibly optimal, options.

RQ1. *How does the visualization idiom affect the expression of the Phantom Effect?*

Considering these results and having in mind the familiarity of people with visualizations, we will test if the visualization idiom impacts the expression of the Phantom Effect. We will replicate the experience conducted by Jeong et al. [1], implementing **one-sided bar charts, two-sided bar charts, scatterplots, and parallel-coordinate plots, as well as a table working as baseline**.

Our motivation for choosing those idioms in particular follows the reasons given by Jeong et al. All five idioms can easily represent two (or more) attributes. Tables have been used in a vast multitude of cases, from communication purposes to data analysis. Tables are known to be efficient in terms of space when a large amount of data needs to be presented in an organized manner. The title of each column and/or row represents the entity being referred to by them. Each cell is the corresponding value of the key (row, column). Despite existing ways of highlighting trends and important information, tables are also known for not being the most effective way of deducting such insights.

Bar charts are one of the most common visual representations of data, especially when the task involves the comparison of values from different categories [61]. The length of the bars is used to represent the value of an attribute and color is often used to differentiate bars from different categories. There are various types of bar charts: one-sided bar charts, two-sided bar charts, stacked bar charts, grouped bar charts, etc. These graphs can be used both vertically and horizontally. When there is more than one attribute to be represented, one-sided bar charts, two-sided bar charts, and stacked bar charts are the most used ones. Stacked bar charts can lead to difficulties in interpreting the data when there are many attributes as it gets harder to compare the bars' lengths.

Scatterplots are another common visualization [61], and also one of the most flexible ones. Scatterplots can represent multiple different attributes since the position, color, size of the data points, and shape of the data points can be used to encode properties of the data. Using scatterplots makes each data point comparable to the others. This idiom has been used in visualization literature, namely to study decoy effects [14, 26, 62].

Parallel-coordinate plots are the less familiar and less used idioms out of these five, according to Jeong et al. [1]. However, they too are capable of handling two attributes. The interpretation of parallel-coordinate plots is made mainly through the slope of the line connecting both axes. Each vertical axis represents a variable, having its values along the line as a regular axis. The degree of the slope describes how different two values are, facilitating the comparison of the two variables and the comparison between sets of pairs of variables. The encodings and sketches for each idiom will be further discussed in Chapter 4.

Although we aimed to replicate Jeong et al.'s work [1], we expanded its domain by addressing presentation delays [9] and individual differences [2]. According to Pratkanis and Farquhar [9], **one of the dimensions of the Phantom Effect is the moment the phantom's unavailability is presented**. Trueblood et al. [8] identified three presentation delay approaches: **known presentation delay** (displaying the information of unavailability with a delay before the choice is made), **unknown presentation delay** (displaying the information of unavailability once the choice is made), and **onset presentation delay** (displaying the information of unavailability from the beginning of the task). Figure 2.2 represents each approach. Pettibone and Wedell [28] concluded that the most effective approach is the know delay presentation with a delay of three seconds. Another study by Scarpi and Pizzi [29] claims that the unknown presentation delay is more efficient as it prevents the frustration caused by the known presentation delay.

Given the lack of consensus about this subject, we seek to contribute by implementing the three different forms of presentation delay in our experiment. Therefore, we introduce the second research question:

RQ2. *How does the presentation delay impact the expression of the Phantom Effect?*

In order to mitigate the learning effect, we will conduct a **3x4 mixed design experiment** in which each participant will only interact with an approach for the presentation delay (between subjects), but they will interact with all the visualization idiom conditions within subjects). Additionally, there will be a control group that will not be submitted to this condition, i.e. no unavailable option will be shown during the experiment. We will guarantee a similar number of participants for each approach by choosing the manual selection of the approach versus a random selection.

Research has proved that the experience while interacting with information displayed visually varies depending on the individual differences among users [2]. Namely, there is evidence summarised in Liu et al.'s state-of-the-art review proving that cognitive abilities impact task performance and the strategies used to interact with a visualization. **Due to the visual-spatial nature of InfoVis, it is intuitive to assume that cognitive abilities that describe an individual's performance on visual tasks will have an impact on how one perceives a visualization**. This also means that different experiences with a visualization may influence how easily an individual can be primed by a cognitive bias such as the Phantom Effect. Moreover, memory and perception-related cognitive abilities interfere with how an individual apprehends and retains information, therefore impacting how susceptible they are to being biased. The third proposed research question addresses this topic:

RQ3. *How do individual differences impact the expression of the Phantom Effect?*

Both personality traits and cognitive abilities have been studied in their relationship with visualization. However, **we choose to focus on cognitive abilities**. Note that with this choice, we are not discarding the hypothesis that personality traits may have an impact on the expression of the Phantom Effect. The focus on cognitive abilities is justified by the greater probability of cognitive abilities having an impact on the Phantom Effect. This results from the observation of the overview presented by Liu et al. [2] (Figure 3.3). As previously discussed in Section 3.1, there is a higher fraction of literature with conclusive results on the relationship between InfoVis and cognitive abilities, than concerning personality traits.

One of the cognitive abilities we approach in this study is PS, defined by **the rate at which an individual can accurately compare two distinguished objects**. Given its definition, higher PS is responsible for an individual's better performance in comparison tasks [46, 55, 63]. Solving a decision problem requires accurately comparing the different alternatives until finding the optimal option. Therefore, we present the following hypothesis:

H3.1. *Individuals with higher PS are primed more frequently by the Phantom Effect.*

We also study VWM, which is the retention of visual information between eye fixations, i.e. **how easily one can recall a previous image after being stimulated with a different one**. While studying a visualization, an individual moves their gaze throughout various areas of the display, changing the focus between data points. Carenini et al. [64] measured the performance of users with different VWM in comparison tasks using bar charts. The authors concluded that individuals with lower VWM had worse performance on those tasks. **VWM is responsible for how much the person remembers from the previous data points**. Therefore, we hypothesize that this cognitive trait can impact the expression of the Phantom Effect due to the visual recall of the phantom's attributes.

H3.2. *Individuals with higher VWM are primed more frequently by the Phantom Effect.*

The results of the research made by Toker et al. [55] relate VWM with the preference for different visualization idioms, specifically between bar charts and radar plots. These results reinforce the relevance of Hypothesis H1.2. and motivate the study of how VWM may impact the expression of the Phantom

Effect in InfoVis.

Besides the logical and scientific reasons for choosing to focus on PS and VWM, the possible relationship between PS and VWM we concluded from Figure 3.3 is one of the other motivations for focusing on these two cognitive abilities in particular. Another reason is the particularity in terms of time and deadlines required for a Master Thesis, alongside the aim for a single session per participant testing phase. This forces us not to go for a broad approach including every cognitive ability, in order to keep the sessions within a reasonable duration and workload while making it possible to design and conduct a well-structured study.

4.2 Apparatus

In this section, we describe the materials we developed in order to gather the needed data for answering our research questions. All these materials were used during the test sessions with the participants and the resulting data was later analyzed.

4.2.1 The Phantom Effect App

We developed a web application for this study that guides the participant through the different phases and allows us to collect the data we need. **All questionnaires, cognitive abilities tests, and visualizations were implemented in this application.** The front-end was developed in ReactJS[1], a front-end Javascript library. The demographics questionnaire, the PS tests, and the familiarity test were implemented using SurveyJS[2] a JavaScript library to build forms. The visualizations were implemented using D3.js[3], a JavaScript library for information visualization. The back-end was developed in a Node.js[4] environment, using the web framework Express[5] in order to handle HTML requests and use the database. We used a PostgreSQL[6] database to save all the needed data recorded during the user tests. We ran the web application locally and the participants accessed it through the web browser installed in the laptop used during the sessions. The interface includes all buttons the user uses to submit their answers.

4.2.2 Pre-study Questionnaire

Before starting the study, an **informed consent** (Appendix C) is presented to the participant where we explain the possible risks of the experiment and let them know what data will be collected. The

[1] https://reactjs.org/
[2] https://surveyjs.io/
[3] https://d3js.org/
[4] https://nodejs.org/
[5] https://expressjs.com/
[6] https://www.postgresql.org/

participant receives a copy of this form. If the participant agrees to participate in the study, they are asked to fill in a **questionnaire on demographic data**. This questionnaire includes five questions regarding the participant's **familiarity with the scenarios** which will be later used in the decision-making tasks.

4.2.3 Perceptual Speed Test

The PS test consists of a Discrimination and Comparison test adapted from the Kit of Factor-Referenced Cognitive Tests [12], namely the tests P-1, P-2, and P-3. This battery of tests has been widely used for measuring PS, including some studies mentioned before in this document [46, 55, 63]. Therefore, the results from these tests are trusted by the community. We adapted the three tests of this Kit on PS to a digital questionnaire consisting of three types of questions:

1. **"Find A's":** Find and select five words with the letter "A" (in the beginning, middle, or end of the word) in a set of 41 words.
    - Length: 25 sets of 41 words
    - Time limit: 120 seconds
    - Scoring: number of correct words selected
    - Metrics: score and time taken to complete the task
2. **"Number Comparison":** Compare two multi-digit numbers and indicate if they are the same or not.
    - Length: 48 pairs of numbers
    - Time limit: 90 seconds
    - Scoring: number of correct answers
    - Metrics: score and time taken to complete the task
3. **"Identical Figures":** Find a figure with the same shape as the reference.
    - Length: 48 sets of shapes
    - Time limit: 90 seconds
    - Scoring: number of correct answers
    - Metrics: score and time taken to complete the task

The final PS task score of each individual is the sum of the task scores (number between 0 and 121) and the total time taken is the sum of the times in seconds (number between 0 and 300). Finally, **the task score and the inverse of the time are multiplied to obtain the PS Score**. Individuals are classified and compared according to this score.

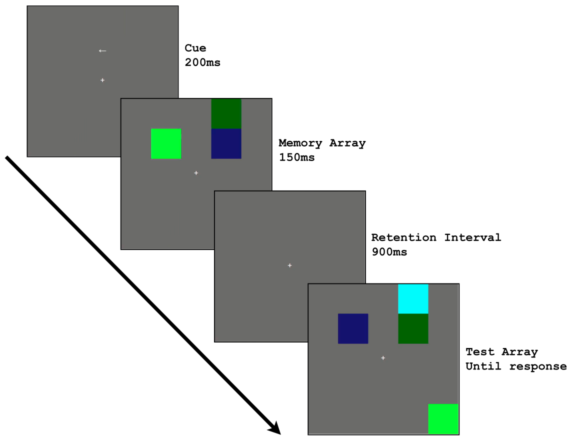


Figure 4.2: Sequence of images in the change detection test.

4.2.4 Visual Working Memory Test

VWM is measured through an Image Change Detection task adapted from Luck and Vogel's [53]. This study has been widely used to measure VWM, as mentioned by Liu et al. in their survey [2, 55, 56, 64]. The test consists of **100 trials of change-detection tasks**. Each trial consists of a series of colored squares which may be the same or different between the Memory Array and the Test Array. A cue is shown for 200ms. The cue is represented by an arrow pointing left, indicating the side the participants must shift their attention to, along with a central fixation point. Then the Memory Array, i.e. the image the participant needs to memorize, is presented for 150ms, followed by 900ms of the same screen with only the fixation point in it. Finally, the Test Array, i.e. the set of squares for comparison with the Memory Array, is shown and stays on the screen until the participant answers. Figure 4.2 represents the sequence of images as described. Participants must focus on the fixation point in order to guarantee that the perception of the image is not disturbed by focusing on other images and to keep the participant focused on the task. To indicate whether the squares on that side changed or not, the participant needs to **press the corresponding key on the keyboard**. The color scale for the squares is composed of eight highly distinguishable colors, even if the participant is colorblind. The background of the trial area is colored a neutral grey with RGB values (108,107,105). Each trial has either four or eight colored squares. There are approximately 50 trials of each size and they are presented in a random order. **The VWM score is the number of correct answers given.**

4.2.5 The Phantom Effect Test

The Phantom Effect phase consists of five decision problems in which the participants need to select the best alternative. Each decision problem was represented using a different InfoVis idiom (table, one-sided bar chart, two-sided bar chart, scatterplot, and parallel-coordinate plot). In this section, we describe the different visualizations used in this phase. We detail the presentation delay condition's implementation, the scenarios of the tasks, the design for each one of the five visualizations, and the interaction techniques they allow.

4.2.5.A Scenarios

We defined **five different hypothetical scenarios** for the decision-making task. Jeong et al. [1] referred to a problem with the familiarity of the scenario, where one of the scenarios resulted in a different pattern from the other scenarios since people always preferred that the options would excel at one attribute independently of the value of the other. To prevent this, we created scenarios less prone to common knowledge consensus to avoid the choice being biased by the general preference for an option. While the decision problem question itself is unfamiliar, the two dimensions of each scenario were chosen to be attributes people recognize and are able to compare. The attributes are quantitative values that can be represented in the chosen idioms. Below, we state the scenarios that were used in this experiment:

- *"You won a giveaway and now you need to choose one of the hampers. Which hamper is the most valuable and balanced one?" ("Ganhou um sorteio e agora tem de escolher um dos cabazes. Qual dos cabazes tem mais valor e melhor equilíbrio de frutas e legumes?")*
  - Attribute X: Amount of Fruit (kg) (Quantidade de Fruta (kg))
  - Attribute Y: Amount of Vegetables (kg) (Quantidade de Legumes (kg))
- *"You want to use social robots is your nursing home to keep elderly people active. Robots are scored in a scale from 1 to 10 in terms of friendliness and trustworthiness, being 10 the highest score. Considering these scores, which model of robots is the best?" ("Quer usar robôs sociais na sua casa de repouso para manter os idosos ativos. Os robôs são pontuados numa escala de 1 a 10 em termos de amabilidade e confiança, sendo 10 a pontuação mais alta. Considerando esta pontuação, qual é o melhor robô?")*
  - Attribute X: Friendliness (Amabilidade)
  - Attribute Y: Trustworthiness (Confiança)
- *"Which of the stock distributions attracts more customers?" ("Qual destas distribuições de stock atrai mais clientes?")*
  - Attribute X: Items on Sale (%) (Produtos em Promoção (%))
  - Attribute Y: Popular Items (%) (Produtos Populares (%))
- *"Our restaurant is short in staff, so we need to minimize the amount of time spent washing tableware while still serving the greater number of tables possible. Which of these options is a better*

scenario for us?" ("Temos poucos empregados no nosso restaurante, pelo que precisamos de minimizar o tempo gasto a lavar a loiça e servir o máximo de mesas possível. Qual destas opções é o melhor cenário?")

  - Attribute X: Minutes Saved Washing Tableware (Minutos Poupados a Lavar a Louça)
  - Attribute Y: Number of Tables Served (Número de Mesas Servidas)

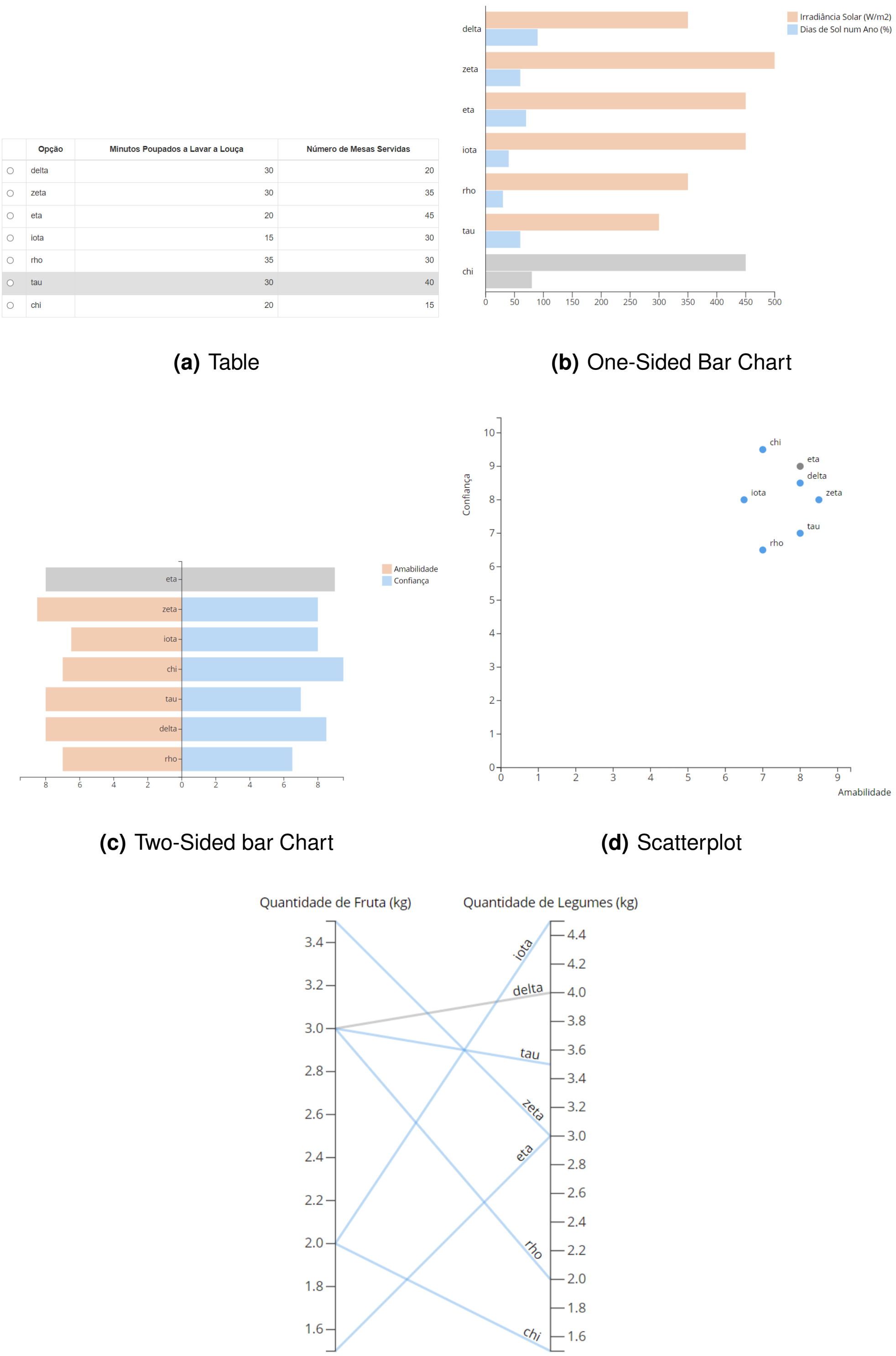
- *"Which is the best place for installing solar panels?" ("Qual o melhor local para montar um painel solar?")*
  - Attribute X: Solar Irradiance (W/m^2) (Irradiância Solar (W/m^2))
  - Attribute Y: Sunny Days in a Year (%) (Dias de Sol num Ano (%))

The scenarios are randomly assigned to each of the five idioms mitigating biases related to the scenario itself. The same scenario was used only once per session. The values of three key alternatives (phantom, target, and competitor) guarantee the point distribution seen in Figure 4.1, and the remaining ones have the attributes distributed in a way not to interfere with the phantom or the target. This guarantees that the Phantom Effect primes the participants as expected and that the target and competitor are the designated ones. **All scenarios have a similar data point distribution** to avoid variations in other biases such as the clustering effect. The values of the attributes for each scenario were equal for every participant.

4.2.5.B Visualization Design

To minimize the impact of visual encodings other than the visualization's idiom, the **colors, spacing, and typography are as homogeneous as possible throughout the visualizations**. Every decision problem has **seven alternatives**, including the phantom, the competitor, and the target. This number of alternatives allows for an experiment with more than three options which is the typical scenario for an InfoVis application. Furthermore, seven is known to be the number of things a human can hold in short-term memory [65]. The idioms were presented in a random order and the scenario was also randomly selected for each idiom. The different visualization idioms used in this experiment can be seen in Figure 4.3.

The phantom alternative, i.e. the unavailable option, is encoded through a change in color. This option is **greyed out** accordingly with the presentation delay condition. Each attribute is encoded by color in order to make it easier to distinguish between the two. The chosen colors were **orange and blue**. Besides being complementary colors, they are able to be told apart even in the presence of different types of color blindness.



	Opção	Minutos Poupados a Lavar a Louça	Número de Mesas Servidas
○	delta	30	20
○	zeta	30	35
○	eta	20	45
○	iota	15	30
○	rho	35	30
○	tau	30	40
○	chi	20	15

(a) Table

(b) One-Sided Bar Chart

(c) Two-Sided bar Chart

(d) Scatterplot

(e) Parallel Coordinate Plot

Figure 4.3: Visualization idioms used as experimental stimuli in our experiment. Four visualization idioms and a table for baseline. The greyed-out option represents the phantom alternative.

The tabular layout (Figure 4.3(a)) was used as a **baseline for the idiom condition**. The first row of the table has the labels of each attribute. Each row of the table corresponds to an option. The first column identifies the option. The second and third columns represent attributes one and two respectively. There is a radio button for each row that allows the corresponding option to be selected, although clicking anywhere inside that row has the same result as clicking the button. The selected row acquires a blue color. When hovered, the rows are highlighted with a light orange color. The data collected from the interactions with the table provides a baseline control group used to evaluate the idiom condition.

The one-sided bar chart (Figure 4.3(b)) consists of **horizontal bars grouped by option**. Each option is described by two bars, one for each attribute. The upper bar and lower bar represent attributes one and two respectively. Different options are separated from the others to help distinguish between groups. A label for each option is presented on the left end of the bars. The length of the bar encodes the value of each attribute, according to the axis. Aside from the grey color for the phantom alternative, each attribute is encoded by color as mentioned before. When selected or hovered, the bars are highlighted through an increase in opacity.

The two-sided bar chart (Figure 4.3(c)) shares many of the design options with the one-sided bar chart. However, the positioning of the bars is different. Instead of having the bars one on top of the other, the **two bars are placed side by side** originating from a symmetrical axis in the middle of the chart. The color encoding and the interactivity are the same as in the one-sided bar chart.

The scatterplot (Figure 4.3(d)) consists of a **two-dimensional space with two axes originating on zero**. Each point in the space, visually encoded using a small dot, represents an option. The name of the option is presented next to the dot. The position of the dot encodes both attributes. The value on the X-axis represents attribute one and the value on the Y-axis represents attribute two. Each axis has a label identifying the attribute. The unavailable option is grey colored while the others are blue. When hovered or selected, the option is highlighted by turning orange. Since the most common cases of colorblindness cause blue tones to be easier than orange to be distinguished from grey, we chose that color for both the scatterplot and the parallel-coordinate plot.

Finally, the parallel-coordinate plot (Figure 4.3(e)) is a **type of line chart that represents each attribute through a vertical axis**. There are as many vertical axes as attributes. In this case, since we have two attributes, there are two axes, making this idiom similar to a slope graph. The line connects the value of the first attribute to the value of the second one. The slope of the line is used for comparing attributes and options. However, we must pay attention to the units and range of each axis when making the comparison in order to avoid making wrong conclusions from the slopes. The color encoding follows the same pattern as in the scatterplot, using grey for the unavailable option and orange for the remaining ones. When an option is hovered or selected, not only does the color change to orange but also the thickness of the line increases.

4.2.5.C Presentation Delay

At the beginning of the experiment, **the participant is randomly assigned to one of the presentation delay conditions (known, unknown, or onset)**. The phantom alternative, i.e. the unavailable option, is encoded through a change in color. This option is greyed out accordingly with the presentation delay. In the case of the known presentation delay, the phantom alternative is greyed out **after three seconds**, following the findings by Pettibone and Weddel [28]. This is achieved by the implementation of a timer that triggers the change of color three seconds after it finishes rendering the visualization. For the unknown presentation delay, the phantom alternative is greyed out **once the participant clicks on it** for the first time. If the participant does not interact with the said option, it is not greyed out. Finally, the onset presentation delay condition presents the unavailability of the option **the moment the visualization is presented to the user**.

4.2.6 Familiarity Test

The familiarity test's goal is to gather information on the participants' knowledge and experience with the different visualization idioms we tested. We adapted and extended the **Visualization Literacy Assessment Test (VLAT)**, a test developed by Lee et al. [66] with the purpose of measuring people's familiarity with different visualization types.

The original VLAT covers twelve different visualizations and eight task types. The score is calculated based on the correctness of the answers given for each of the 53 multiple-choice questions. These multiple-choice questions focus on completing tasks that require interpreting the presented visualization.

We used the questions on scatterplots from the original VLAT and developed questions for tables, one-sided bar charts, two-sided bar charts, and parallel coordinates plots. These new questions cover the most common tasks these visualizations support. The questionnaire can be seen in Appendix A. We scored the test as suggested by Lee et al. [66], by **counting the number of correct answers separately for each type of visualization**. Our version of the VLAT has **32 multiple-choice questions**, covering **five different types of visualizations**. Each question was shown separately to the participant who had to choose the option that best answered each question within **25 seconds**.

4.3 Experimental Design

Given the conditions at study and in order to mitigate the learning effect, our experiment follows a **3x4 mixed design**. The two conditions were the presentation delay and the visualization layout. There were three presentation delay conditions (unknown, known, and onset). Each participant only interacted with one of them. Additionally, there was a control group that was not submitted to this condition, i.e. no unavailable option was shown during the experiment. There were four layout conditions (one-sided bar chart, two-sided bar chart, scatterplot, and parallel coordinates plot), plus a control one, which was the table. Each participant interacted with all of the layouts in random order.

Our study was divided into three parts: the **Individual Differences phase, the Phantom Effect phase, and the Familiarity phase**. The Individual Differences phase took place before the Phantom Effect phase which took place immediately before the Familiarity phase. The three experiments were conducted sequentially in the same test session with the participant. Figure 4.4 represents the sequence of steps taken for each experiment.

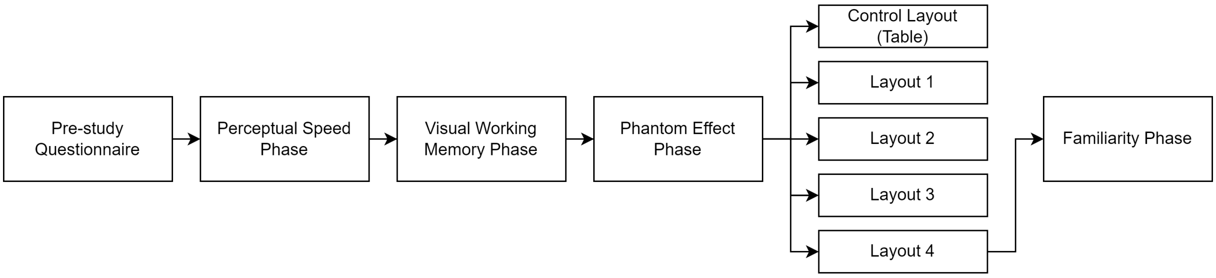


Figure 4.4: Main steps of the experiment in chronological order. The Phantom Effect phase consists of five layouts (table, one-sided bar chart, two-sided bar chart, scatterplot, and parallel-coordinate plot) which were presented in random order.

The Individual Differences phase had the goal of measuring the participants' PS and VWM, gathering data needed to answer **RQ3**. The Phantom Effect phase focused on a decision-making task with visually represented data in which the superior option is unavailable (except in the control condition). This experiment's goal is to study the expression of the Phantom Effect in InfoVis, focusing on the first two research questions (**RQ1 and RQ2**). Finally, The Familiarity phase measured the participants' knowledge and experience with each of the layouts.

4.3.1 Variables

During our study, we collected data for a set of variables that will be further discussed in the following sections. These variables allow us to study whether the participants were primed by the Phantom Effect, discover possible patterns, and correlations between variations in the expression of the bias and motivated by different conditions. The evaluation methods and plan for the analysis of these variables will be further discussed in Section 4.4.1.

4.3.1.A Independent Variables

The independent variables described in Table 4.1 allow us to group results by condition to further study how the expression of the Phantom Effect changes according to each condition. Variables such as the layout and the presentation delay are deeply connected to the design of our experiment. We will also gather information on multiple individual differences such as familiarity with each scenario, visualization literacy and cognitive abilities. PS Score and VWM Score are cognitive variables which depend only on the participant's own characteristics and will be used to test for correlations with the expression of the Phantom Effect.

4.3.1.B Dependent Variables

Data from the participants' interactions with each visualization was saved in the dependent variables in Table 4.2. These variables were later used to study if participants were primed by the Phantom Effect. By comparing the rate of participants who chose the target option with the rate of participants who chose the competitor option, we can determine if they were primed by the Phantom Effect.

During the user tests, we sought to gather as much information about the performance during the tests and the interactions with the visualizations as possible. We measured the time taken to make a decision for each visualization during the Phantom Effect phase, as well as the time needed to finish each of the PS tests as a mean to better score each participant. All these variables allow us to scrutinize the cues for the priming effect. The metrics suggested by Wall et al. [3] for measuring cognitive biases were used to evaluate how the participant interacts with the visualization.

4.4 User Tests

In this Section, we describe the procedure followed during a typical session with a participant. The user test guide we followed can be consulted in Appendix B. We also detail and discuss the characteristics of our sample of participants. Instituto Superior Técnico's Ethics Committee approved the procedure, questionnaires, and data collection and processing methods described.

4.4.1 Procedure

Sessions had an average duration of **45 minutes** and were conducted **in person**. 72 of the trials were conducted at INESC-ID and 4 at Taguspark. We always used the same laptop and mouse for the user tests and kept the distractions minimal during the session. The luminosity and laptop position conditions were kept as constant as possible.

Each session started with a brief explanation of the structure of the session and of our study. Then, we presented the informed consent to the participant who needed to agree with it in order to proceed with the experiment. The participant would sit comfortably at the desk in front of the laptop. Upon opening the Phantom Effect App, the participant was randomly assigned to one of the presentation delay conditions

Table 4.1: Independent variables used for analyzing data gathered during the sessions with the user.

Variable Name	Type	Description and Goal
Age	Discrete	Age of the participant.
Gender	Nominal	Gender of the participant.
Education	Nominal	Educational level of the participant.
Fruit and Vegetables Hampers Familiarity	Ordinal	Used for evaluating familiarity with the fruit and vegetables hampers scenario.
Restaurant Management Familiarity	Ordinal	Used for evaluating familiarity with the restaurant management scenario.
Social Robots Familiarity	Ordinal	Used for evaluating familiarity with the social robots scenario.
Solar Panels Familiarity	Ordinal	Used for evaluating familiarity with the solar panels scenario.
Stock and Stores Familiarity	Ordinal	Used for evaluating familiarity with the stock and stores scenario.
PS Score	Ordinal	Score achieved on the PS experiment. Used for measuring this cognitive ability.
VWM Score	Ordinal	Score achieved on the VWM experiment. Used for measuring this cognitive ability.
Presentation Delay	Nominal	Condition for disclosing the unavailability of the phantom alternative.
Idiom	Nominal	Visualization idiom being interacted with.
Table Literacy	Ordinal	Used for evaluating familiarity with tables.
One-Sided Bar Chart Literacy	Ordinal	Used for evaluating familiarity with one-sided bar charts.
Two-Sided Bar Chart Literacy	Ordinal	Used for evaluating familiarity with two-sided bar charts.
Parallel Coordinates Plot Literacy	Ordinal	Used for evaluating familiarity with parallel coordinates plots.
Scatterplot Literacy	Ordinal	Used for evaluating familiarity with scatterplots.

Table 4.2: Dependent variables used for analyzing data gathered during the sessions with the user.

Variable Name	Type	Description and Goal
Phantom Rate	Ratio	Number of times the Phantom alternative was selected. Used for studying the control condition in which there is no unavailable alternative.
Target Rate	Ratio	Number of times the Target alternative was selected. Used for measuring the expression of the Phantom Effect.
Competitor Rate	Ratio	Number of times the Competitor alternative was selected. Used for measuring the expression of the Phantom Effect.
Other Alternative Rate	Ratio	Number of times an alternative other than the Target or the Competitor was selected. Used for measuring the expression of the Phantom Effect.
Primed in Idiom	Binary	Used to categorize participants. Value is True if the participant was affected by the Phantom Effect in the respective idiom, False otherwise.
Priming Score	Ordinal	Score between 0 and 5. Number of idioms in which the participant was affected by the Phantom Effect.
Duration of Interaction	Continuous	Time, in seconds, the participant spent interacting with each of the visualizations.
Data Point Coverage	Discrete	Number of options the participant interacted with. [3]
Data Point Distribution	Nominal	Description of data point coverage in terms of homogeneity. [3]
Attribute Coverage	Interval	Interval of values for an attribute that are investigated by the participant. [3]
Attribute Distribution	Nominal	Description of attribute coverage in terms of distribution compared to the distribution of the attribute in the dataset. [3]

as well as a unique identifier to guarantee the anonymization of data. Then, the participant would be ready to start answering the demographics and scenario familiarity questionnaire.

Once the participant submits the pre-study questionnaire, the PS test begins. Before each of the three types of questions introduced in Section 4.2.3, the participant would read the instructions to make sure they understood the task. The test started with the “Find A’s” task, followed by the “Number Comparison” task, and ended with the “Identical Figures” task. All these tasks had a time limit to be completed, after which the answers were automatically submitted. The participant could submit the answers once they finished each task. The VWM test begins with a short tutorial. The participant is able to start a training set of trials in which the data is not recorded. Once the participant is ready, the actual VWM test begins.

Right after the VWM test, the Phantom Effect phase begins. The experiment consists of five idioms with randomized scenarios assigned to each one. The idioms are presented in random order too. Every participant interacted with every idiom but only interacted with one presentation delay condition. In this phase, participants had to perform a decision task. The prompt was presented in text form at the top of the page. During the experiment, the participant needed to interact with the visualizations and select the alternative corresponding to the final decision for each decision problem presented.

Then, the Familiarity questionnaire (VLAT) took place. The questionnaire consisted of 32 multiple-choice questions accompanied by a visualization. Participants had 25 seconds to answer each question. Once the test finishes we thank the participant once again for participating in our study. The equipment is sanitized and the system is prepared to receive the next participant.

4.4.2 Participants

A total of **76 people** participated in our experiment. We recruited the participants through convenience sampling, both online and at campi. None of the participants were discarded since we found no anomalies in the gathered data nor there were outliers in the dependent variables. In order to enroll in the study, participants had to schedule a date and time for the test. As a reward, we hosted a giveaway of three FNAC Gift Cards worth 20€ among the participants of our study. The results of the giveaway were then shared with the participants via email.

Each participant was randomly assigned to a presentation delay condition at the beginning of the experiment. Despite the random distribution of this condition ensuring a balance in the number of participants for each type of presentation delay, we made an effort to assign relatively more participants to the Unknown presentation delay condition since we knew this condition would be prone to missing values. This way, we still ensured we had a sample size that allows accurate data analysis for each condition. Figure 4.5 represents the number of participants in each presentation delay condition.

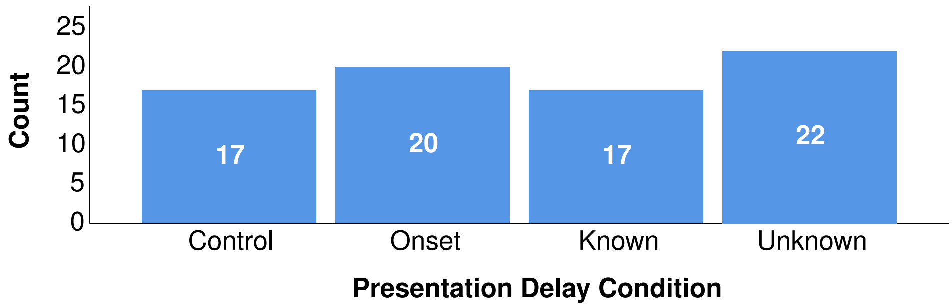


Figure 4.5: Number of participants for each of the presentation delay conditions.

4.4.2.A Demographics

Participants ranged from 18 to 64 years old ($M = 24.96, SD = 8.81$), with a large portion of participants belonging to the age bracket ranging from 21 to 24. The concentration of participants in this age bracket is a consequence of recruiting college students and friends. In terms of gender, our sample shows a satisfactory ratio of 30 females (39.5%) to 46 males (60.5%). The education of the sample shows a normal distribution. 1 participant (1.3%) reported less than High School education, 24 participants (31.6%) completed High School or equivalent, 35 participants (46.1%) completed a Bachelor's Degree or equivalent, 14 participants (18.4%) completed a Graduate Degree or equivalent and 2 participants (2.6%) have higher education or equivalent. Figure 4.6 represents the counts for the educational level of the sample and the gender distribution within each category.

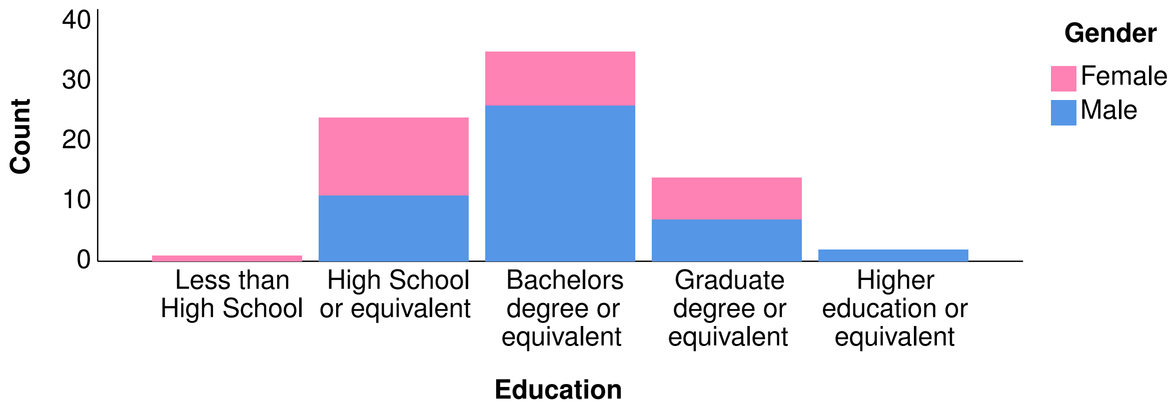


Figure 4.6: Distribution of the sample in terms of education.

4.4.2.B Familiarity

In order to test for possible bias in the sample due to a skew in familiarity with the subjects of the visualization scenarios used in this experiment, we questioned the participants about their familiarity with each of the themes. Participants had to use a Likert scale to describe their knowledge of the subjects, where 1 means “Not Familiar” and 7 means “Very Familiar”. Figure 4.7 represents the box plots of familiarity for each of the 5 scenarios. The box plots show that the scenarios have a variety of levels of familiarity, meaning we have a diverse sample. However, the Hamper Giveaway scenario shows a bigger concentration of participants very familiar with the theme. A similar situation happens with the Managing Restaurants scenario, this time on the low familiarity side. Nevertheless, since the scenarios are randomly attributed to each visualization and overall, we have participants with different levels of familiarity with each scenario, we can discard the hypothesis of having skewed results due to familiarity biases.

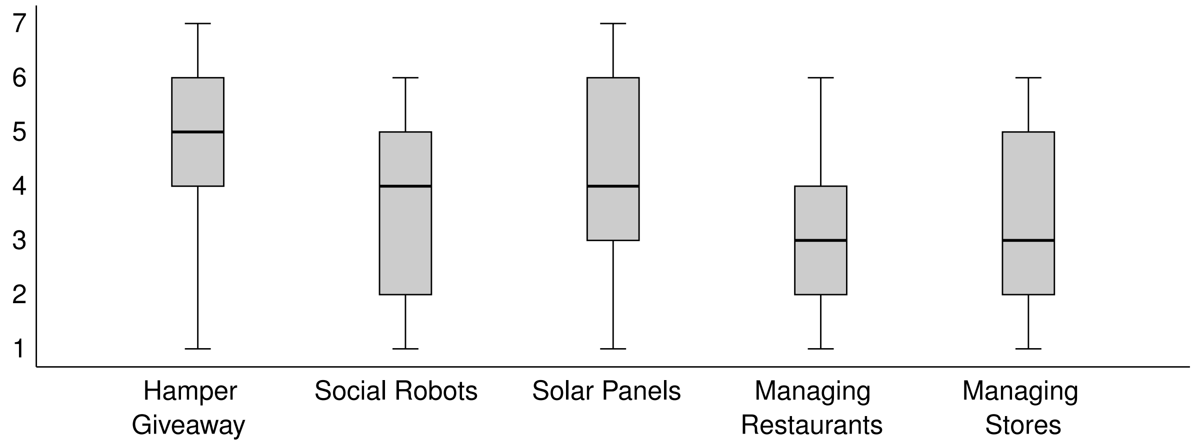


Figure 4.7: Box-plots of the familiarity of participants with the scenarios presented in the visualizations. Familiarity scores range from 1 to 7, where 1 means “Not Familiar” and 7 means “Very Familiar”.

4.5 Data Analysis

In this section, we describe the method we used to prepare and analyze the data resulting from the user tests. We present the following subsections in chronological order, starting with the gathering of data and preliminary data processing and culminating in the validation methods that allowed us to achieve the results described in Chapter 5.

4.5.1 Data Collection

We gathered data on demographics, interaction with the visualizations, and familiarity with each type of visualization idiom through the Phantom Effect App. All the data was stored in a database using a unique identifier for each participant in such a way it would **not be possible to identify the participants**.

The pre-study questionnaire was used to collect demographic data and information about the familiarity of the scenarios presented in the visualizations. During the interaction with the visualizations, we recorded the number of clicks and hovers for each option, as well as the history of interactions. This was useful for analyzing if each participant was primed by the Phantom Effect during the interaction with each plot. We also recorded the score obtained in the visualization literacy assessment test. The score for each visualization is kept separately along with the answers given to each question in particular. While the scores allowed us to study the visualization literacy of our sample, the answers were useful in identifying the type of tasks where participants showed a bigger difficulty.

4.5.2 Data Cleaning

None of our dependent variables presented extreme or out-of-order values justifiable for being considered outliers, hence **no records were removed from our sample**. Regarding the data on the participants who were in the group of the Unknown presentation delay condition, we had to **filter missing values in the priming rate variables**. In this condition, the information on the unavailability of an option was only disclosed at the moment of the decision, that is, when the participant clicked the said option. If the participant never clicks the phantom option, they will not be informed of the unavailability, therefore, the bias could not be expressed. In these cases, instead of filling the column with a value zero for "non-primed" or one for "primed", we left it with a "none", or missing value. Before analyzing the data, we removed these missing values, ignoring them for each idiom independently.

4.5.3 Data Processing

We developed a series of Python scripts in order to process the data. These scripts read the CSV file with the raw data from the database and output a CSV file with the processed data. The output of those scripts was then imported into an SPSS data sheet for analysis.

We started by creating a script that would unfold the JSON data saved in the demographics column of the database. We also created a script for computing the PS score, by multiplying each test's score by the inverse of the time needed to complete the said test and then summing these values to achieve the final score. In this script, participants are also labeled according to their PS and VWM scores.

For analyzing whether a participant was primed or not by the Phantom Effect, we developed a script that analyses the type of option chosen by the participant for each plot. For each idiom, we store in the output a Boolean variable that states if the participant was primed or not. For the Control presentation delay condition, this variable is set to True if the participant chose the Phantom option and to False otherwise. For the Unknown presentation delay condition, we set the variable to True if the participant chose the Target option and clicked the Phantom option. If the participant did not click the Phantom

Table 4.3: Classification of attribute distribution based on the user's interactions with the 7 data points.

Number of options interacted with	Interacted with min value	Interacted with max value	Classification
<7	Yes	No	Min
<7	No	No	Mid
<7	No	Yes	Max
7	Yes	Yes	Full

option, we set the variable to None. If they choose another type of option and clicked the phantom option, we set the variable to False. For the Known and Onset presentation delay conditions, we set the variable to True if the participant chose the Target option and to False otherwise.

In order to perform the formative analysis of cognitive biases in InfoVis proposed by Wall et al. [3], we created a script that counts the number of data points the participant interacted with and computes the interval of values of those data points for each of the attributes. Based on the distribution of that interval, we classify the attribute distribution into one of the categories as described in Table 4.3.

Finally, we implemented a script for processing the Visualization Literacy Assessment Test data. For each tested idiom, we compute the participant's score on a scale from 0 to 100, since the score saved in the database is saved as the number of correct answers. Then, we count the total number of correct answers for each question of the test for later analysis.

4.5.4 Validation Methods

In order to choose the most suitable statistical method for validating our hypotheses and answering our research questions, we grouped the relevant variables by sub-problem. This grouping allowed for an exploration of data in terms of the type of variables, normality, and homogeneity of variance, which are decisive when choosing a statistical method. Table 4.4 summarizes the result from this data arrangement and the validation method chosen for each research question or hypothesis.

Based on Jeong et al.'s study [1], the effect of the visualization idiom on the expression of the Phantom Effect will be analyzed using the Chi-Square test which allows to study if two categorical variables are related. PageRank will allow us to analyze the network of interactions, highlighting the most important nodes. Since the PS Score has a normal distribution, we are able to use the Bi-Serial Correlation Test which checks if quantitative variables (Priming Score and PS Score) are related. Since the remaining tests feature data that does not have a normal distribution, we will apply Kendall's Tau-b, which is a non-parametric correlation test suitable for qualitative variables that can be assigned to numeric values ordered by weight or importance. Chapter 5 focuses on describing the analysis and results of each of these problems, as well as discussing the findings.

Table 4.4: Statistical method and respective variables chosen to study each of the proposed research questions and hypotheses.

Research Question / Hypothesis	Description	Dependent Variables	Independent Variables	Validation Method
RQ1	Effect of the visualization idiom in the expression of the Phantom Effect	Phantom Rate, Target Rate, Competitor Rate, Other Alternative Rate	Idiom, Presentation Delay Condition	Chi-Square
	Effect of the Phantom Effect in attribute distribution	Primed in Idiom, Attribute Distribution	Idiom	Kendall's Tau-b
	Effect of the Phantom Effect in data point coverage	Primed in Idiom, Data Point Coverage	Idiom	Kendall's Tau-b
	Effect of the Phantom Effect in data point distribution	Primed in Idiom, Data Point Distribution	Idiom	Kendall's Tau-b
	Effect of the Phantom Effect in interaction history	Primed in Idiom, Interaction History	Idiom	PageRank
RQ2	Effect of the presentation delay in the expression of the Phantom Effect	Priming Score, Primed in Idiom	Presentation Delay Condition, Idiom	Kendall's Tau-b
H3.1	Effect of PS in the expression of the Phantom Effect	Priming Score	PS Score, Presentation Delay Condition, Idiom	Bi-Serial Correlation Test
H3.2	Effect of visual working memory in the expression of the Phantom Effect	Priming Score	VWM Score, Presentation Delay Condition, Idiom	Kendall's Tau-b
RQ3	Effect of visualization literacy in the expression of the Phantom Effect	Primed in Idiom	Visualization Literacy, Idiom	Kendall's Tau-b
	Effect of education level in the expression of the Phantom Effect	Priming Score	Education	Kendall's Tau-b
	Effect of gender in the expression of the Phantom Effect	Priming Score	Gender	Kendall's Tau-b

5

Results

Contents

5.1 Phantom Effect 55
5.2 Presentation Delay 60
5.3 Individual Differences 63
5.4 Discussion 66

In this chapter, we describe how we analyzed and validated the variables concerning each of our research questions and hypotheses regarding the expression of the Phantom Effect. We start by analyzing the presence of the Phantom Effect in InfoVis, observing which visualization idioms are more prone to this cognitive bias. Then, we compare the results from the expression of the Phantom Effect with different presentation delays in order to conclude which one better mitigates it. We finalize the exposition of results with an analysis of the possible correlations between individual differences, namely cognitive abilities, and the presence of the Phantom Effect. We close this chapter with a discussion of our findings in terms of answering our research questions and contributing to the literature. We also discuss the limitations of our study and explore future work opportunities.

5.1 Phantom Effect

The selection of the target option, when a phantom alternative is introduced, provides evidence that the Phantom Effect is present in InfoVis. Participants in the control conditions, which did not have a phantom alternative, chose the target option with less frequency than those who had a phantom alternative as shown in Figure 5.1.

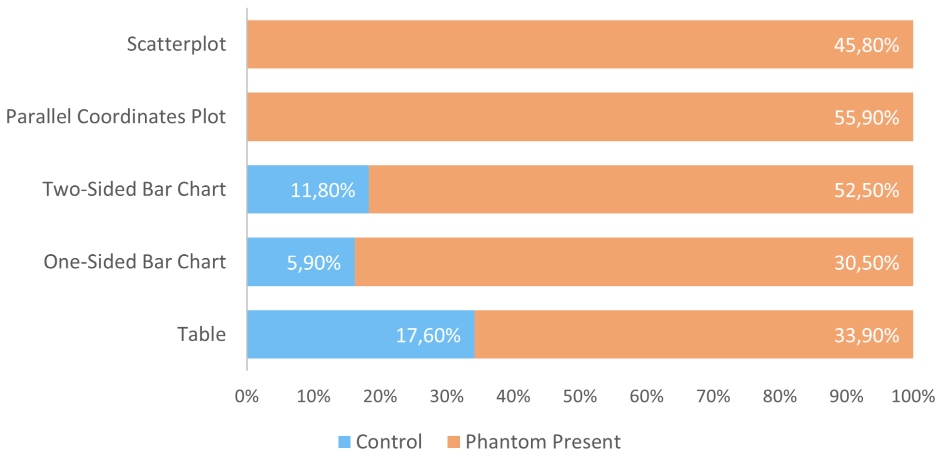


Figure 5.1: Ratio of participants who chose the target option in the control condition versus when a phantom alternative was present.

There is a difference among visualization idioms in the ratio of participants who chose the target option. In the Scatterplot × Control and Parallel Coordinates Plot × Control conditions, no participants chose the target option. The impact of the visualization idiom in the expression of the Phantom Effect will be further discussed in Section 5.1.1.

Table 5.1: Summary of chi-square test results and choice ratios among different visualization idioms.

Idiom	Phantom Alternative	Other	Competitor	Target	Phantom	χ^2	*p*-value
Table	With	2.3	64.4	33.9	-	2.579	.108
	Without	11.8	17.6	17.6	52.9		
One-Sided Bar Chart	With	0.0	69.5	30.5	-	4.309	.038
	Without	0.0	47.1	5.9	47.1		
Two-Sided Bar Chart	With	1.7	45.8	52.5	-	.048	.827
	Without	0.0	41.2	11.8	47.1		
Parallel Coordinates	With	5.1	39.0	55.9	-	6.099	.014
	Without	5.9	11.8	0.0	82.4		
Scatterplot	With	0.0	54.2	45.8	-	3.566	.059
	Without	5.9	29.4	0.0	64.7		

5.1.1 Idioms

For each idiom, a chi-square test for association was conducted between the presence of a phantom alternative and participants being affected by the Phantom Effect, i.e. choosing the target alternative. All expected cell frequencies were greater than five. There was not a statistically significant association between the presence of a phantom alternative and participants being affected by the Phantom Effect for information represented in a table, $\chi^2(1) = 2.579$, $p = .108$. The same result applies to information represented visually using two-sided bar charts, $\chi^2(1) = .048$, $p = .827$, with an even less significant association between the idiom and the participants choosing the target alternative. The remaining idioms show a statistically significant association between these two factors, as shown in Table 5.1. The lack of association between the visualization idiom and the rate of option choice type both in tabular layouts and two-sided bar charts implies that participants' choices were not biased by the presence of a phantom alternative, suggesting these two idioms are able to mitigate the Phantom Effect, answering **RQ1**. The scatterplot presents an almost statistically significant value ($p = .059$), however, it is not high enough for us to consider there was no change in participants' choice patterns.

5.1.2 Metrics for Formative Analysis

We used the metrics for formative analysis proposed by Wall et al. [3] to detect possible patterns in interaction that differ between participants who were affected by the Phantom Effect and those who were not. We started by analyzing the data coverage, i.e. how many options the participants interacted with. Given that the population did not follow a normal distribution for these metrics, we used a non-parametric correlation test.

Table 5.2: Summary of Kendall's tau-b test results for the formative analysis.

	Data Coverage		Attribute X Distribution		Attribute Y Distribution		Data Distribution	
	τb	***p*-value**	**τb**	***p*-value**	**τb**	***p*-value**	**τb**	***p*-value**
Table	.042	.688	.020	.856	-.060	.583	-.022	.835
One-Sided Bar Chart	.130	.231	-.048	.677	.011	.928	-.012	.911
Two-Sided Bar Chart	.119	.277	-.174	.128	-.073	.529	-.169	.114
Parallel Coordinates Plot	.031	.782	.053	.658	-.538	$<$.001	-.135	.220
Scatterplot	-.052	.643	-.164	.158	-.320	.006	-.404	$<$.001

A Kendall's tau-b correlation was run to determine the relationship between data coverage and being primed by the Phantom Effect. There was a weak, positive association between data coverage and being primed by the Phantom Effect in one-sided bar charts and two-sided bar charts, which was not statistically significant, $\tau b = .130$, $p = .231$ and $\tau b = .119$, $p = .277$, respectively. For the remaining three idioms, there was no association between data coverage and being primed by the Phantom Effect ($\tau b < .07$), also not statistically significant (Table 5.2).

In terms of attribute distribution, Kendall's tau-b correlation gave us mixed results. We highlight the weak, negative association between attribute X's distribution and being primed by the Phantom Effect in two-sided bar charts and scatterplots, despite not being statistically significant, $\tau b = -.174$, $p = .128$ and $\tau b = -.164$, $p = .158$, respectively. There was a strong, negative association between the distribution of attribute Y and being primed by the Phantom Effect in parallel coordinates plots, which is statistically significant, $\tau b = -.538$, $p < .001$. Scatterplots have a medium, positive association between these two variables, despite not being statistically significant, $\tau b = .320$, $p = .006$. Two-sided bar charts have a weak, negative association between the distribution of attribute Y and being primed by the Phantom Effect, despite not being statistically significant, $\tau b = -.073$, $p = .529$.

To determine the relationship between data distribution and being primed by the Phantom Effect we ran a Kendall's tau-b correlation. Similarly to the previously mentioned metrics, this test gave us mixed results. There was a strong, negative association between data distribution and being primed by the Phantom Effect in scatterplots, which was statistically significant, $\tau b = -.404$, $p < .001$. There was a weak, negative association between data distribution and being primed by the Phantom Effect in two-sided bar charts and parallel coordinates plots, despite not being statistically significant, $\tau b = -.169$, $p = .114$ and $\tau b = -.135$, $p = .220$, respectively.

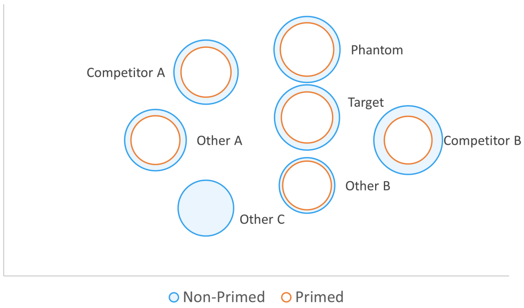
In order to get an overview of the options participants interacted more with, we conducted a frequency analysis. We labeled each option according to its type and relative position and counted the number of participants who clicked or hovered each option, dividing the participants into two groups: Primed and Non-Primed, depending if they chose the target option or not. Figure 5.2 represents the results from this analysis. The position of the bubbles corresponds to the relative position of the option, based on one of the datasets we used in the experiment. The size of the bubble is the number of participants who interacted with the option. Overall, the group of participants who were not affected by the Phantom Effect interacted with a broader set of options and showed a more homogeneous pattern of interactions. Those who were affected by the Phantom Effect tended to focus on the phantom, target, and competitor options, which were optimal options. In the table and two-sided bar chart, which were the two idioms in which the Phantom Effect did not have an expression as mentioned in the previous section, the distribution of the interactions is visibly different from the other idioms, suggesting that a vast majority of participants interacted with all options.

5.1.3 Interactions

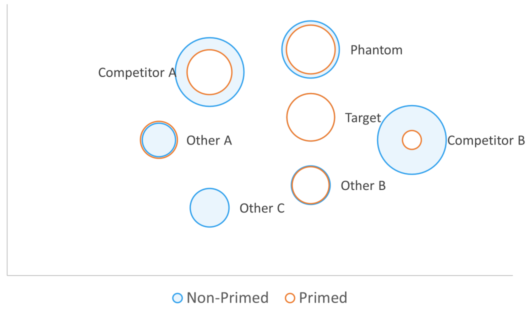
We built a set of graphs based on the history of interactions for each of the idioms. In order to compare the importance of each option, we used a subgroup of participants that did not include the control condition and analyzed interactions from participants who were primed, i.e. chose the target option separately from the interactions from participants who did not choose the target option. The interactions are hovers and clicks, both with the same weight. The width of the edges represents the number of occurrences of the corresponding interaction. The size of the nodes encodes the PageRank score, hence bigger nodes are more important. The relative positions of the nodes result from an expansion of a ForceAtlas2 layout, where nodes are treated as charged particles that repulse each other.

When analyzing the PageRank scores, there is a visible difference in the importance of each option between participants who were affected by the Phantom Effect (primed) and those who were not (non-primed). For primed participants, the target alternative was always the most important option, followed by one of the competitors. In the particular case of the table (Figure 5.3(a)), the target alternative is significantly more important than any other options, which have similar PageRank scores among them. We also highlight a difference in the order of importance of the options in the scatterplot (Figure 5.6(a)), where the phantom alternative closely follows the target alternative, instead of the competitors as in the remaining cases.

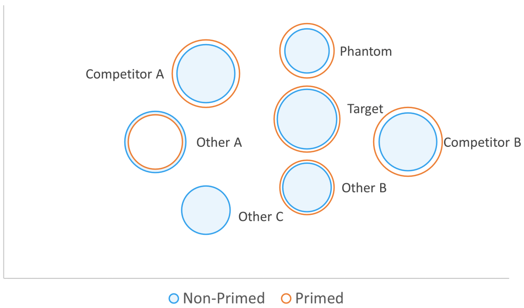
For the group of participants who were non-primed, the target alternative significantly drops in importance and the first place is occupied by one of the competitors, followed by the phantom alternative. There is also a greater number of thicker edges, meaning participants tended to diversify their interactions. In the two-sided bar chart (Figure 5.5(b)), not only the target alternative loses importance, as it is more repelled by the competitors and the phantom alternatives. In the case of the scatterplot (Fig-



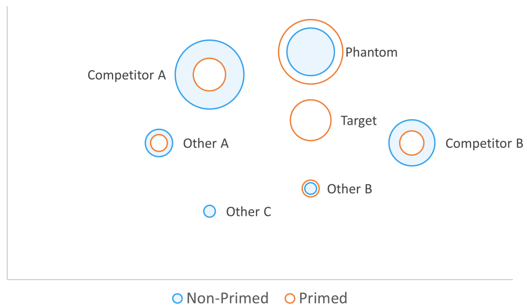
(a) Table



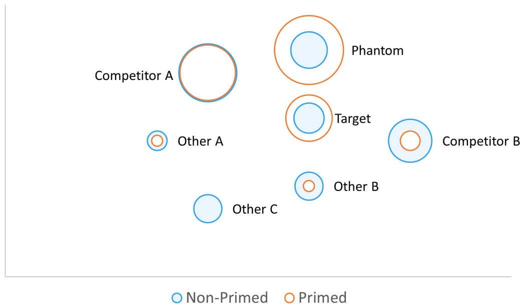
(b) One-Sided Bar Chart



(c) Two-Sided bar Chart



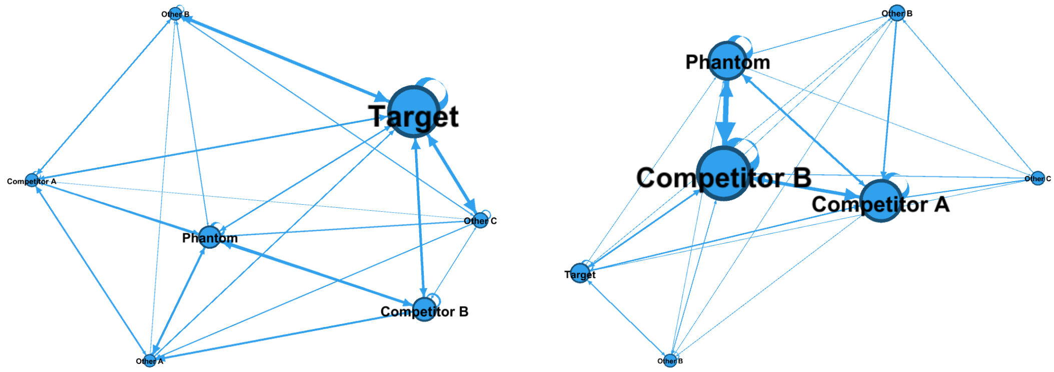
(d) Scatterplot



(e) Parallel Coordinate Plot

Figure 5.2: Distribution of interactions across the different visualization idioms.

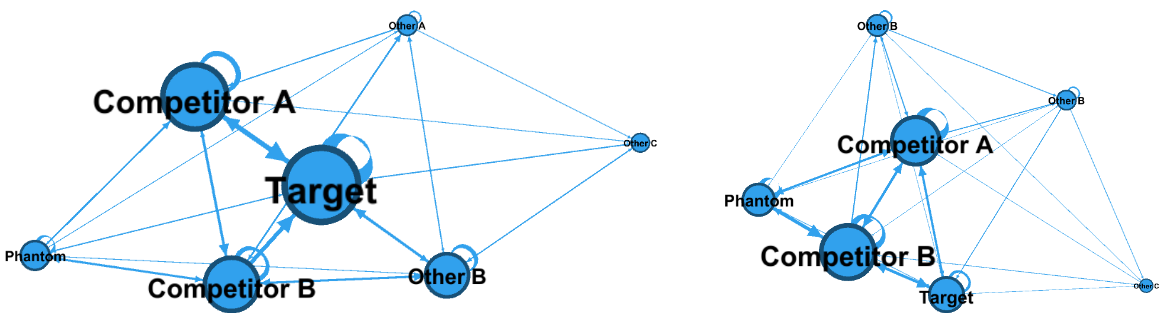
ure 5.6(b)), the competitor option which excelled in the attribute represented in the Y axis (Competitor A) has a PageRank score much greater than the other competitor, compared to the other idioms where they usually have similar scores. Additionally, no interactions were made with one of the suboptimal options, Other B, so it is not represented in the network. The same happens with Other C, which is the worst option, in the Parallel Coordinates Plot (Figure 5.7(b)).



(a) Primed

(b) Non-Primed

Figure 5.3: Table interactions network.



(a) One-Sided Bar Chart - Primed

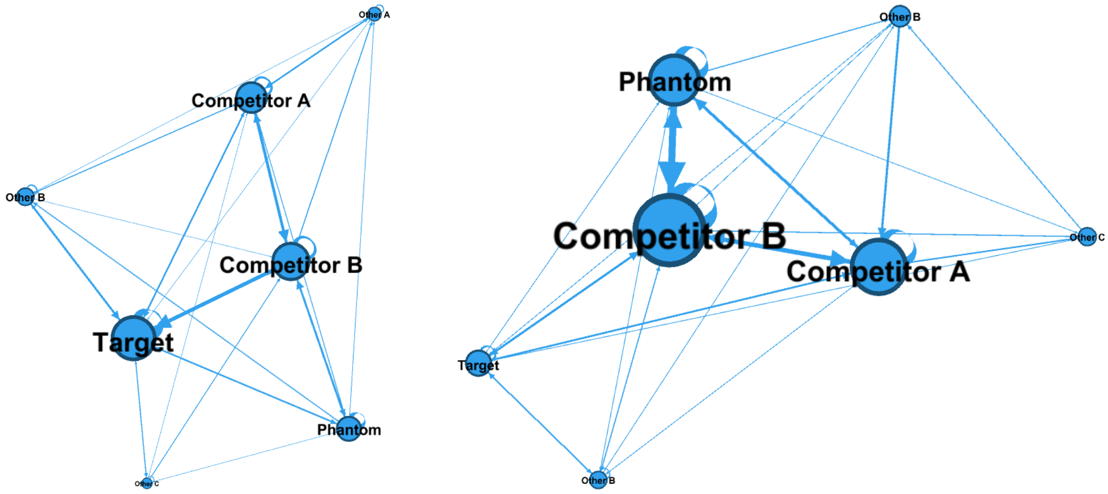
(b) One-Sided Bar Chart - Non-Primed

Figure 5.4: One-sided bar chart interactions network.

5.2 Presentation Delay

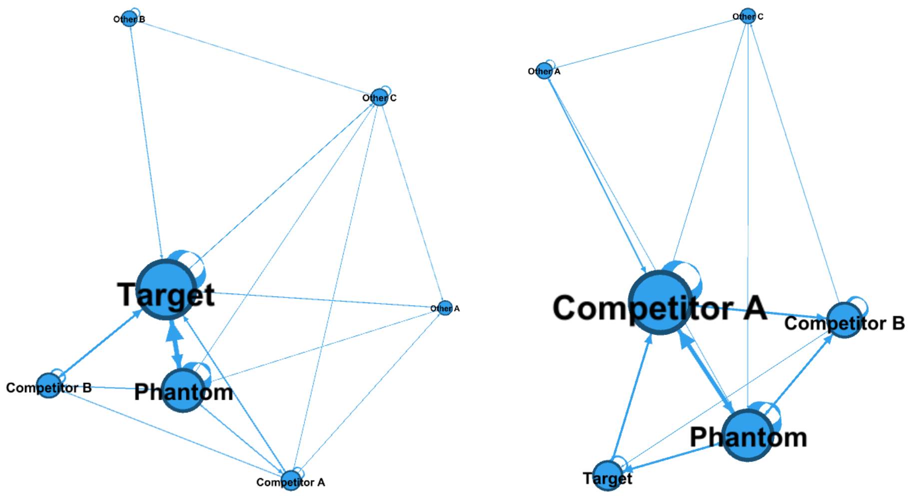
After analyzing the presence of the Phantom Effect and the impact of the visualization idiom used to represent information, we focused on studying the different ways to disclose that an option is unavailable, i.e. the presentation delay, answering **RQ2**.

A Kendall's tau-b correlation was run to determine the relationship between presentation delay and the number of idioms in which participants were affected by the Phantom Effect (Priming Score). There



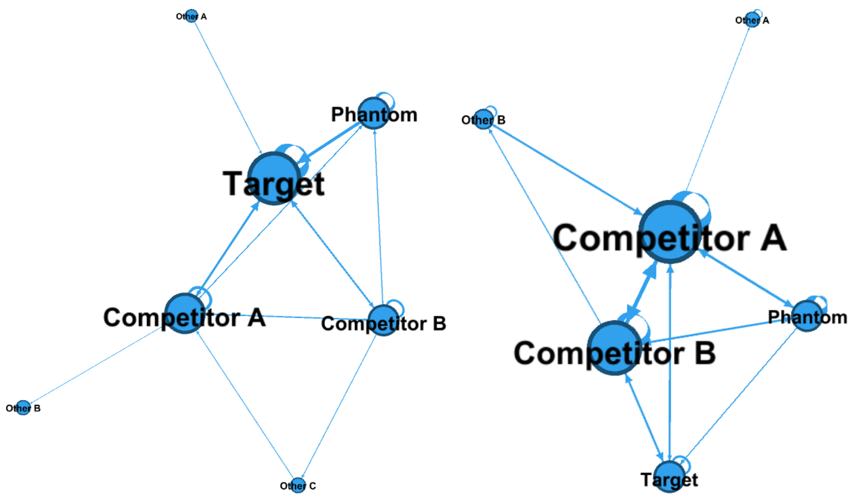
(a) Primed **(b)** Non-Primed

Figure 5.5: Two-sided bar chart interactions network.



(a) Primed **(b)** Non-Primed

Figure 5.6: Scatterplot interactions network.



(a) Primed **(b)** Non-Primed

Figure 5.7: Parallel coordinates plot interactions network.

was a strong, negative association between the presentation delay and the Priming Score, which was statistically significant, $\tau b = -.493$, $p < .001$.

The inspection of the boxplots represented in Figure 5.8 allowed us to notice that the distribution of the population across the different Priming Scores is very broad both for the control and onset conditions. For the known condition, there was a clustering around scores between 2 and 3. The unknown condition showed the highest success rate in mitigating the Phantom Effect, as its median lies on a Priming Score of 0, and a Priming Score of 2 is the upper bound.

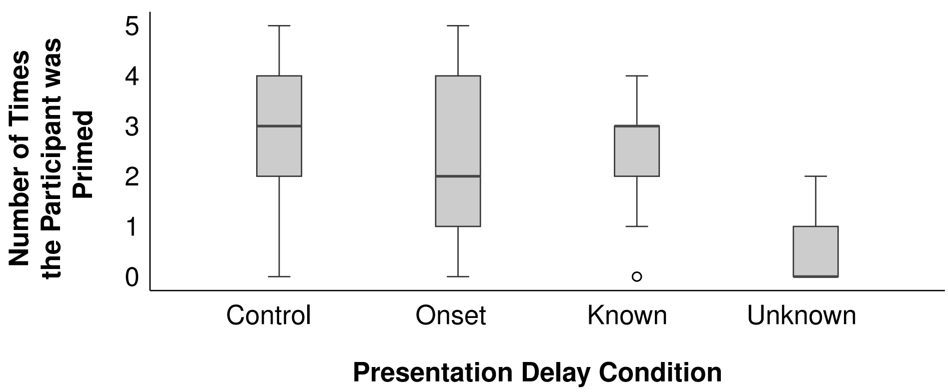


Figure 5.8: Boxplots of Priming Score across the different presentation delay conditions.

5.3 Individual Differences

Our last research question focused on the impact of individual differences in the expression of the Phantom Effect. Therefore, we conducted a series of correlation tests in order to validate **RQ3** in terms of cognitive abilities, demographics, and visualization literacy.

To analyze the impact of individual differences on the expression of the Phantom Effect, we only used the group of participants who had a phantom alternative present in the visualizations, i.e. those who were not in the control group ($N = 59$). We will refer to this group as the "phantom present group". Distributions of the various individual differences values for the control group and the phantom present group were similar, as assessed by visual inspection. Due to the existence of null values resulting from the processing of data for the Unknown presentation delay condition, sample size varies from idiom to idiom as follows - table: $N = 56$; one-sided bar chart: $N = 52$; two-sided bar chart: $N = 50$; parallel coordinates plot: $N = 50$; scatterplot: $N = 52$.

5.3.1 Perceptual Speed

We started by analyzing the data we gathered on cognitive abilities, namely PS. The average PS of the total population was $.8503$ ($N = 76, SD = .1380$), while the average PS of the phantom present group was $.8480$ ($N = 59, SD = .1294$).

A point-biserial correlation was run between being affected by the Phantom Effect and PS score. Data are mean $\pm$ standard deviation unless otherwise stated. Preliminary analyses showed there were (a) no outliers, as assessed by boxplot; (b) PS score distribution showed no evidence against normality, as assessed by Shapiro-Wilk's test ($p > .05$); and (c) there was homogeneity of variances, as assessed by Levene's test for equality of variances. Table 5.3 represents the individual results for each idiom. Overall, there was no statistically significant correlation between being affected by the Phantom Effect and PS score. However, there is a small association between the two variables in one-sided bar charts, despite not being statistically significant, $p > .05$.

5.3.2 Visual Working Memory

The average VWM of the total population was 82.28 ($N = 76, SD = 15.207$), while the average VWM of the phantom present group was 82.61 ($N = 59, SD = 15.395$).

Since VWM scores failed the normality and homogeneity of variance assumptions for a point-biserial correlation test, we ran a non-parametric correlation test. A Kendall's tau-b correlation was run to determine the relationship between being affected by the Phantom Effect and VWM score amongst 59 participants. There was a weak, positive association between being affected by the Phantom Effect and

Table 5.3: Summary of Point-biserial correlation test results for the impact of PS in the expression of the Phantom Effect for each visualization idiom.

Idiom	rpb	Association	*p*-value	Non-Primed		Primed		Variability (%)(rpb^2)
				M	SD	M	SD	
Table	.003	No	.984	.8433	.1404	.8441	.1037	.0009
One-Sided Bar Chart	-.197	Small	.162	.8529	.1286	.7904	.1340	3.8809
Two-Sided Bar Chart	.048	No	.743	.8397	.1149	.8281	.1335	.2304
Parallel Coordinates Plot	.021	No	.883	.8411	.1519	.8467	.1110	.0441
Scatterplot	-.095	No	.503	.8585	.1418	.8327	.1194	.9025

Table 5.4: Summary of Kendall's tau-b correlation test results for the impact of VWM in the expression of the Phantom Effect for each visualization idiom.

Idiom	τb	Association	*p*-value	Non-Primed		Primed	
				M	SD	M	SD
Table	.145	Weak	.192	82.92	12.865	87.80	14.630
One-Sided Bar Chart	-.017	No	.884	82.97	14.740	85.73	11.217
Two-Sided Bar Chart	.164	Weak	.168	82.95	12.867	83.95	14.979
Parallel Coordinates Plot	-.049	No	.683	82.89	14.800	83.87	13.230
Scatterplot	.077	Weak	.510	81.95	15.139	85.16	12.171

VWM score in the case of tables, two-sided bar charts, and scatterplots, but in none of the cases, it was statistically significative, $p < .05$. Table 5.4 summarizes the result of this analysis.

5.3.3 Demographics

In order to study the impact of demographic individual differences, we focused on the participants' educational level and gender. Despite the control group having no females, we still analyzed the relationship between gender and being affected by the Phantom Effect since it had no impact on this test.

A Kendall's tau-b correlation was run to determine the relationship between educational level and the priming score amongst 59 participants. There was a weak, negative association between educational level and the priming score, which was not statistically significant, $\tau b = -.123$, $p = .272$. The more advanced the educational level, the lower the priming score.

The same correlation test was run to determine the relationship between gender and the priming score amongst the same 59 participants. There was a weak, negative association between gender and the priming score, which was not statistically significant, $\tau b = -.091$, $p = .443$. Females ($M =$

$1.76, SD = 1.508$) were slightly more affected by the Phantom Effect than males ($M = 1.50, SD = 1.462$).

5.3.4 Visualization Literacy

Finally, we analyzed the scores obtained on the VLAT. To get an overview of the participants' familiarity with each of the idioms, we inspected the boxplots (Figure 5.9) of VLAT scores amongst the total population. The vast majority of the population displays good visualization literacy since most scores are positive (> 50). One-sided bar charts, however, show the greatest variety of scores and also the lowest median with an average score of 63.596 ($SD = 20.676$). On the other hand, participants have the highest literacy on tables ($M = 86.053, SD = 20.138$).

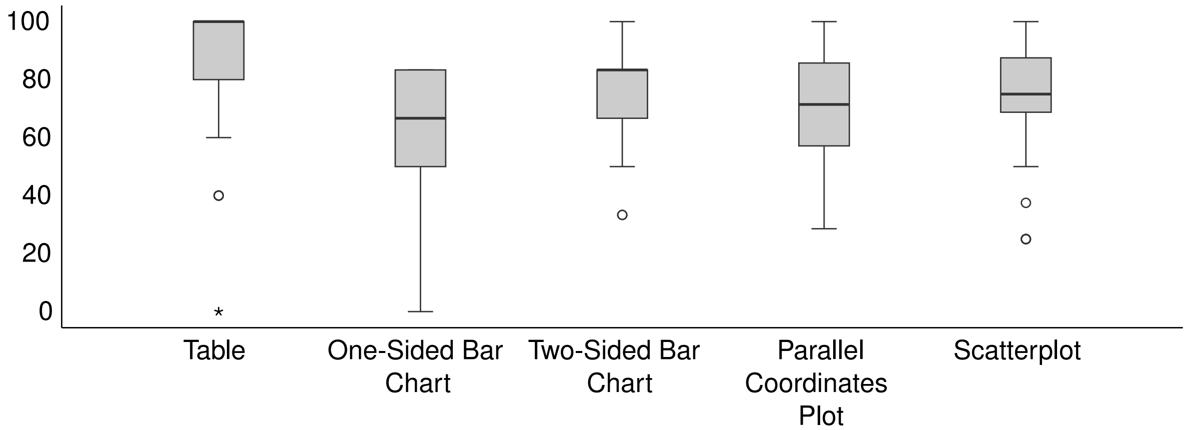


Figure 5.9: Boxplots of VLAT scores across the different idioms. Scores range from 0 to 100, where 100 is the top score.

In order to study the impact of visualization literacy in the expression of the Phantom Effect, we focused on the phantom present group and checked for extreme outliers through boxplot inspection. This resulted in the removal of one extreme outlier from the table VLAT scores. Then, we check for normality and homogeneity of variance. Since these assumptions were not met, we ran a non-parametric correlation test.

A Kendall's tau-b correlation was run to determine the relationship between being affected by the Phantom Effect and VLAT score amongst 59 participants. There was a medium, positive association between being affected by the Phantom Effect in one-sided bar charts and the participant's familiarity with that idiom, which was statistically significant, $\tau b = -288$, $p = .025$. There was also a small, positive, and negative association between being affected by the Phantom Effect and VLAT score for two-sided bar charts and scatterplots, respectively, despite none of them being statistically significant, $p < .05$. Table 5.5 summarizes the result of this analysis.

Table 5.5: Summary of Kendall's tau-b correlation test results for the impact of VLAT scores in the expression of the Phantom Effect for each visualization idiom.

Idiom	**τb**	**Association**	***p*-value**	**Non-Primed**		**Primed**	
				M	**SD**	**M**	**SD**
Table	.013	No	.919	88.421	17.786	90.000	15.718
One-Sided Bar Chart	.288	Medium	.025	60.975	20.279	75.758	11.459
Two-Sided Bar Chart	.151	Small	.249	77.381	17.102	82.576	16.649
Parallel Coordinates Plot	.064	No	.624	69.231	18.819	73.214	12.146
Scatterplot	-.075	Small	.553	74.609	21.647	73.750	18.979

5.4 Discussion

From the analysis of the data we gathered, we concluded that **the Phantom Effect can be present in InfoVis**. Moreover, when observing Figure 5.1 we notice a higher rate of participants being affected by this cognitive bias in the cases where information is represented visually, than when the information is presented in a table. This suggests **InfoVis might even accentuate the Phantom Effect**.

The visualization idiom chosen to represent the information affects the expression of the Phantom Effect. The results presented in Section 5.1.1 **confirm Jeong et al.'s [1] results**, since in their study the two-sided bar charts also presented no difference in the participant's choice with or without a decoy. A possible reason for the Phantom Effect having no expression in two-sided bar charts is the fact that **attributes are compared one by one rather than comparing the values of both attributes at the same time**. This means users can more efficiently identify equivalent options and evaluate each attribute without being biased by the relative position of the phantom alternative. While one-sided bar charts are identical, both attributes are presented on the same side of the chart, making it easier to evaluate both values simultaneously without losing focus on the phantom. As summarized on Table 5.1, **tables can mitigate the Phantom Effect too**, despite having a higher significance compared to two-sided bar charts. **The required effort to compare options without visual aid leads to losing track of the phantom alternative**. Attributes with higher values will stand out, independently of the conjugation of both attributes, leading to a scattering of the choices across all types of options. The tabular layout presented the greatest proportion of participants who chose non-optimal alternatives. **Scatterplots presented a minimal efficiency at mitigating the Phantom Effect** ($p = .059$), but this result is much closer to the cases where idioms do not mitigate the Phantom Effect than those that do. The little statistical significance requires more studies are needed to confirm the results for this visualization idiom. The fact that **scatterplots highlight both optimal options and the value of each option**

combining both attributes explains the results for this idiom. **Depending on how the scatterplot is read**, i.e. focusing on the identification of an imaginary line where balanced options will be located versus focusing on the alternatives that are spatially close to the phantom, might dictate if a person has a biased view of the data or not. The significative differences in the distribution of the points biased and not biased participants interacted with (Table 5.2) demonstrates this statement.

Concerning the metrics proposed by Wall et al. [3], besides the previously mentioned results for the scatterplot, there was only one more statistically significant correlation. In the parallel coordinates plot idiom, the distribution of attribute Y's values for the options participants interacted with was different between participants who were affected by the Phantom Effect and those who were not. Participants who were affected by the Phantom Effect interacted more with attributes with lower Y values or with the whole range of values than those who were not affected, which focused on medium or high Y values. **This means the value chosen on the rightmost axis of the parallel coordinates plot was critical for the expression of the Phantom Effect** since people would use that axis to compare options with identical X values. **The remaining results might not present any statistical significance or strong correlations**, but when manually inspecting the distribution of interactions across the different idioms (Figure 5.2), we can make interesting conclusions. Overall, participants who were not affected by the Phantom Effect interacted more broadly with all options, while **primed participants focused on the phantom, competitors, and target options**. This suggests that participants who inspected all options were able to get a better overview of the data and understand that the target and both competitors were equivalent options. Primed participants would often ignore suboptimal options ("Other"), especially the one that was inferior in both attributes. Another interesting takeout is the fact that both the table and two-sided bar charts, which were the idioms presented as efficient at mitigating the Phantom Effect, **share the same distribution of interactions**, which is different from the remaining idioms. All options were interacted with approximately the same number of times, with no preference for optimal options. Moreover, There is little difference in the distribution between the primed and non-primed groups.

Inspecting the interaction graphs represented in Section 5.1.3 leads to conclusions similar to the ones already described. In most idioms, **nodes corresponding to superior options (phantom, target, or competitors) are more important**, i.e. have a higher PageRank score. In the case of interactions made by **participants who were affected by the Phantom Effect, the target option is always the most important**, from which derives edges with greater weight. The most frequent interactions occurred between the phantom alternative and the competitors, or between the competitors and the target options. This pattern of interactions illustrates the comparisons made by the participants in order to compare equally optimal options. The table, however, presents heavy edges from the target to suboptimal options in the case of primed participants (Figure 5.3(a)), even if the weight differences of the edges are not minimal. This suggests **participants would compare other options to the one that was more similar**

to the phantom alternative.

In order to answer **RQ2**, we analyzed the correlation of the Presentation Delay condition with the number of times participants chose the target option, i.e. were primed by the Phantom Effect. The existence of a statistically significant correlation suggests that **the moment the information about an option's unavailability is presented impacts the expression of the Phantom Effect**. When inspecting the distribution of the priming scores for each presentation delay (Figure 5.8, we conclude that the **Unknown Presentation Delay condition**, which is the condition in which participants would only know an option is unavailable after clicking it, **mitigates the Phantom Effect**. Such results are aligned with the conclusions Scarpi and Pizzi took from their study [29]. The fact that participants are oblivious to the unavailability of an option until they click it to make a choice, allows them to have an **unbiased view of the data with nothing to bring special focus to the phantom alternative**. The sudden loss of freedom of choice makes the phantom alternative less attractive. After analyzing all options and making a decision, choosing an alternative to the phantom will only be dependent on the already established order of preference.

Answering **RQ3** required studying the correlation of multiple individual differences and the Priming Score. We started by analyzing cognitive abilities, namely PS. The distribution of PS scores follows a normal curve as assessed by the Shapiro-Wilk normality test ($p = .295$). The only idiom in which there was a correlation was the one-sided bar chart, but it was not statistically significant and the correlation was small (Table 5.3). Therefore, **we reject hypothesis H3.1.** which stated that individuals with higher PS are more primed by the Phantom Effect. Higher or lower PS scores have no significant impact on how many times individuals are affected by the Phantom Effect. We also got **no statistically significant results for VWM**, as depicted in Table 5.4. There is slight proof that higher VWM scores lead to individuals being more primed by the Phantom Effect. However, our results do not allow us to confidently accept hypothesis **H3.2.**.

We also tested for the impact of individual differences outside cognitive abilities. **Both for education and gender there is a weak negative correlation which is not statistically significant**, meaning the higher the educational level, the less people are affected by the Phantom Effect. Experiencing higher levels of education leads to a greater understanding of the data presented, making people less prone to being attached to the phantom alternative's values. In terms of gender, females were slightly more affected by the Phantom Effect than males, despite the correlation not being statistically significant.

Finally, we analyzed the results from the VLAT used to measure visualization literacy. **Participants demonstrated better knowledge of tables**. Unlike the results from Jeong et al.'s study [1], the parallel coordinates plot was not the visualization idiom participants were less familiar with, at least when assessed through the VLAT. In our case, the **one-sided bar chart provided the lowest literacy scores**. We also tested for correlations with how much people were affected by the Phantom Effect depending on their familiarity with each idiom. **Individuals who scored higher in the one-sided bar chart and two-sided bar chart questions, were more affected by the Phantom Effect**, despite only the first result being statistically significant. The opposite happens in scatterplots since the correlation is negative, but not statistically significant, i.e. people are less affected by this cognitive bias if they are more literate with scatterplots. As seen in Table 5.5, all idioms besides scatterplots present positive tau-b values, suggesting that, **overall, being more familiar with a visualization idiom results in a greater probability of being affected by the Phantom Effect**. When individuals are not familiar with a visualization idiom, they present difficulties reading data points and using the visualization's properties to extract information about the data. Therefore, it is more likely that these people will not realize that there is an option that is similar to the phantom alternative, so they will not feel compelled to choose that option. In the case of scatterplots, literacy brings interpretation techniques such as identifying the Pareto Front or the line over which balanced options are situated. These techniques help to see beyond the spatial location of the phantom alternative, lowering the impact of its existence.

Individual differences generate slight variations in the expression of the Phantom Effect, despite our study revealing few statistically significant results. From all individual differences, visualization literacy is the one with greater impact on the expression of the Phantom Effect.

In summary, the analysis of the data gathered during the user studies allows us to answer the three research questions and reject both our hypotheses. The Phantom Effect does have an expression in InfoVis, but it is possible to mitigate it by using two-sided bar charts complemented by the unknown presentation delay condition. Individual differences influence the expression of the Phantom Effect, namely visualization literacy. However, none of the cognitive abilities we measured revealed an impact on the number of times a person is affected by the Phantom Effect. More studies are needed in order to confirm these results since most of them are not statistically significant.

5.4.1 Limitations and Future Work

Our experiment enabled us to break through the exploration of a cognitive bias that to date has not been studied under the light of InfoVis and allowed us to get relevant results. However, there are still opportunities for further research. It is also important to expand our study by investigating the correlation between the expression of the Phantom Effect and other cognitive abilities not explored in this study, personality traits, and other relevant individual differences we have not explored.

First, we used a preliminary version of the PS and VWM measurement tools that we developed. These tools have been later improved by us, listening to the feedback we got from the users and our experience using the tools. In the latest version of the tools, usability was improved, allowing users to waste less time inputting their answers and interacting with the interface. Measured cognitive abilities' scores might be lower than reality due to usability issues. Additionally, we did not use the memory capacity formula presented by Fukuda and Vogel [**?**], which was corrected on the most version of the tool. Therefore, it would be beneficial to repeat this study using the enhanced version of the tools.

Secondly, there was an issue with the implementation of the table layout in React that sometimes caused mouse events not to correctly trigger actions. Therefore, occasionally participants had to perform extra and undesired interactions with the options presented in the table so that the phantom alternative would be correctly marked as grey or the submit button would stop throwing error messages because the system did not record the participant's choice. Correcting this bug would allow for more accurate results for interaction patterns.

Still concerning the Phantom Effect Phase of the study, we noticed a majority of the participants would not interact with the options while thinking about which one to choose. Some would only click their final answer, resulting in a single interaction despite our advice to use the mouse to click and hover the options. This leads to a big loss of potentially important data on users' interaction patterns. Using eye-tracking would potentially solve this issue.

Following our steps to understand which visualization idioms are less prone to the Phantom Effect, other layouts should be tested, as well as different visual encodings and highlighting techniques proposed by Dimara et al. [14] in order to study if techniques used to mitigate different cognitive biases are also efficient at mitigating the Phantom Effect. The data used in this study was fictitious and the scenarios were designed to avoid personal preferences or familiarity issues. However, this study would benefit from using real data from a meaningful decision task where participants were engaged and motivated with the responsibility to make the best decision possible.

In this experiment, we tested our extension of the VLAT for the first time, excluding pilot tests. Due to the layout of the application, some of the visualization's text appeared too small making it difficult for participants to understand what was being represented. Moreover, the one-sided bar chart test data could be misleading since the attribute shown on the top bar was named “Company B” and the one on the bottom bar was “Company A”. More tests needed to be conducted to verify the validity of this extension.

We share a limitation with an unfortunate amount of user studies which is both the size of our sample and the lack of variety in the population. While 76 people is a number very close to our goal of 80 participants, the more people participate in our experiment, we more valid our data will be. Since we recruited at campus and at academic-related online communities, most of our participants were students with approximately the same age interval and educational background. This translates into a clustering of cognitive ability scores and visualization literacy. A repetition of this study with a more diverse sample, even in cultural terms, would result in less clustered and demographically biased results.

6

Conclusion

Reasoning over big amounts of data requires tools that allow getting an overview of the data while highlighting features. InfoVis presents interactive and varied solutions for data-related tasks, namely decision-making tasks for comparing several different options. The presence of a human in the loop makes InfoVis prone to cognitive biases that affect reasoning over data presented visually, namely in decision-making tasks when we are in the presence of a decoy effect such as the Phantom Effect. The study of these cognitive biases is important for achieving an adjustment of the visual encodings and interfaces in order to mitigate unwanted effects that may irrationally affect the conclusions taken from the data. To do so, it is necessary to check for the expression of cognitive biases in InfoVis, and then study how visual encodings and individual differences impact that expression.

Our main contribution to the field of cognitive biases in InfoVis consisted of an in-person user study in which 76 people participated. These participants had to perform two cognitive abilities tests to measure their PS and VWM. The expression of the Phantom Effect was measured through a set of decision tasks with the aid of visualizations in multiple idioms. Since this was a mixed study, there was a between-subjects condition corresponding to how the information about an option's availability was disclosed. Additionally, participants answered a visualization literacy test (VLAT) that covered the five visualization idioms corresponding to the within-subjects condition of the study.

The data gathered during the user study was methodically analyzed, enabling us to answer our three research questions. We were able to confirm the expression of the Phantom Effect in InfoVis since there was a difference in participants' choices and interaction patterns depending on the presence of a phantom alternative. Moreover, we concluded that the Phantom Effect had a stronger expression on data represented visually than it had on tabular data, since tables, alongside two-sided bar charts showed no difference in the type of option chosen by participants between the control group and the group with the phantom alternative. Therefore, when a visualization's goal is to support decision tasks where some options, namely optimal ones, might be unavailable, it is preferable to use two-sided bar charts, aided by the unknown presentation delay condition as our results suggest. The final step of our data analysis was studying the impact of individual differences in the expression of the Phantom Effect. Only familiarity with one-sided bar charts presented a statistically significant correlation with the expression of the Phantom Effect. All other individual differences studied either had a small impact with no statistical significance or no impact at all. We concluded that VWM and PS have a small impact with no statistical significance, which justifies the rejection of both our hypotheses.

This study poses an important contribution to the investigation of cognitive biases in InfoVis since we were pioneers in the analysis of the Phantom Effect in this context. We were able to confirm results from studies on the decoy effect, a category of cognitive biases to which the Phantom Effect belongs, consolidating the knowledge on which visualization idioms are more efficient at mitigating such bias. We also investigated how individual differences impact the expression of the Phantom Effect, which is a crucial step in order to achieve the customization of interfaces and visualizations to fit the user behind the screen. Furthermore, we contributed to the research community even outside the field of InfoVis by making available two free open-source tools for measuring PS and VWM which consist of the enhanced version of the tools we used in our experiment.

Bibliography

[1] Y. Jeong, S. Oh, Y. Kang, and S.-H. Kim, “Impacts of visualizations on decoy effects,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 23, p. 12674, 2021.

[2] Z. Liu, R. J. Crouser, and A. Ottley, “Survey on Individual Differences in Visualization,” *Computer Graphics Forum*, vol. 39, no. 3, pp. 693–712, Jun. 2020. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1111/cgf.14033

[3] E. Wall, L. Blaha, C. Paul, and A. Endert, “A Formative Study of Interactive Bias Metrics in Visual Analytics Using Anchoring Bias,” in *Human-Computer Interaction – INTERACT 2019*, D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, and P. Zaphiris, Eds. Cham: Springer International Publishing, 2019, vol. 11747, pp. 555–575, series Title: Lecture Notes in Computer Science. [Online]. Available: http://link.springer.com/10.1007/978-3-030-29384-0_34

[4] L. Falschlunger, O. Lehner, and H. Treiblmaier, “Infovis: The impact of information overload on decision making outcome in high complexity settings,” *Proceedings of the Special Interest Group on Human-Computer Interaction, Dublin, Ireland*, p. 6, 2016.

[5] K. Eberhard, “The effects of visualization on judgment and decision-making: a systematic literature review,” *Management Review Quarterly*, 08 2021.

[6] K. Miettinen, “Survey of methods to visualize alternatives in multiple criteria decision making problems,” *OR spectrum*, vol. 36, no. 1, pp. 3–37, 2014.

[7] S. Highhouse, “Context-dependent selection: The effects of decoy and phantom job candidates,” *Organizational Behavior and Human Decision Processes*, vol. 65, no. 1, pp. 68–76, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0749597896900060

[8] J. S. Trueblood and J. C. Pettibone, “The phantom decoy effect in perceptual decision making,” *Journal of Behavioral Decision Making*, 2017.

[9] A. R. Pratkanis and P. H. Farquhar, “A brief history of research on phantom alternatives: Evidence for seven empirical generalizations about phantoms,” *Basic and Applied Social Psychology*, 1992.

[10] M. Dörk, P. Feng, C. Collins, and S. Carpendale, "Critical infovis: Exploring the politics of visualization," in *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 2189–2198. [Online]. Available: https://doi.org/10.1145/2468356.2468739

[11] C. Kostelnick, "The visual rhetoric of data displays: The conundrum of clarity," *IEEE Transactions on Professional Communication*, vol. 50, no. 4, pp. 280–294, 2007.

[12] R. Ekstrom and H. Harman, *Manual for Kit of Factor-referenced Cognitive Tests, 1976*. Educational Testing Service, 1976. [Online]. Available: https://books.google.pt/books?id=vgkIAQAAIAAJ

[13] A. Baddeley, "Working memory," *Science*, vol. 255, no. 5044, pp. 556–559, 1992. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.1736359

[14] E. Dimara, G. Bailly, A. Bezerianos, and S. Franconeri, "Mitigating the Attraction Effect with Visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 850–860, Jan. 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8440809/

[15] R. Pohl, *Cognitive illusions: Intriguing phenomena in thinking, judgment and memory*. Routledge, 07 2016.

[16] E. Dimara, S. Franconeri, C. Plaisant, A. Bezerianos, and P. Dragicevic, "A Task-Based Taxonomy of Cognitive Biases for Information Visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 2, pp. 1413–1432, Feb. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8476234/

[17] A. Furnham and H. C. Boo, "A literature review of the anchoring effect," *The Journal of Socio-Economics*, vol. 40, no. 1, pp. 35–42, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053535710001411

[18] J. C. Pettibone and D. H. Wedell, "Examining Models of Nondominated Decoy Effects across Judgment and Choice," *Organizational Behavior and Human Decision Processes*, vol. 81, no. 2, pp. 300–328, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0749597899928807

[19] M. Mcdaniel and G. Einstein, "Bizarre imagery as an effective memory aid. the importance of distinctiveness," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 12, pp. 54–65, 01 1986.

[20] R. Nadeau, E. Cloutier, and J.-H. Guay, "New evidence about the existence of a bandwagon effect in the opinion formation process," *International Political Science Review*, vol. 14, no. 2, pp. 203–213, 1993. [Online]. Available: https://doi.org/10.1177/019251219301400204

[21] M. R. Wolfson and G. R. Salancik, "Observer orientation and actor-observer differences in attributions for failure," *Journal of Experimental Social Psychology*, vol. 13, no. 5, pp. 441–451, 1977. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022103177900294

[22] D. Kahneman and A. Tversky, "Subjective probability: A judgment of representativeness," *Cognitive Psychology*, vol. 3, no. 3, pp. 430–454, 1972. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0010028572900163

[23] A. Janes, D. Pizzagalli, S. Richardt, B. Frederick, A. Holmes, J. Sousa, M. Fava, A. Evins, and M. Kaufman, "Neural substrates of attentional bias for smoking-related cues: An fmri study," *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, vol. 35, pp. 2339–45, 11 2010.

[24] J. Hedlund, "Risky business: Safety regulations, risk compensation, and individual behavior," *Injury Prevention - INJ PREV*, vol. 6, pp. 82–89, 06 2000.

[25] N. Karlsson, G. Loewenstein, and D. Seppi, "The ostrich effect: Selective attention to information," *Journal of Risk and Uncertainty*, vol. 38, pp. 95–115, 04 2009.

[26] J. Huber, J. W. Payne, and C. P. Puto, "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis." *Journal of Consumer Research*, 1982.

[27] E. Dimara, A. Bezerianos, and P. Dragicevic, "The Attraction Effect in Information Visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 471–480, Jan. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7539348/

[28] J. C. Pettibone and D. H. Wedell, "Testing alternative explanations of phantom decoy effects," *Journal of Behavioral Decision Making*, 2007.

[29] D. Scarpi and G. Pizzi, "The impact of phantom decoys on choices and perceptions," *Journal of Behavioral Decision Making*, vol. 26, 12 2013.

[30] P. H. Farquhar and A. R. Pratkanis, "Decision structuring with phantom alternatives," *Management Science*, vol. 39, no. 10, p. 1214–1226, 1993.

[31] G. W. Allport, *Personality: A psychological interpretation*. Henry Holt And Company, 1971.

[32] D. Ispas and W. Borman, "Personnel selection, psychology of," *International Encyclopedia of the Social & Behavioral Sciences*, 12 2015.

[33] L. R. Goldberg, "The structure of phenotypic personality traits." *American Psychologist*, vol. 48, no. 1, p. 26–34, 1993.

[34] R. A. Amar, J. R. Eagan, and J. T. Stasko, “Low-level components of analytic activity in information visualization,” *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.*, pp. 111–117, 2005.

[35] C. Ziemkiewicz, A. Ottley, R. Crouser, A. Yauilla, S. 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. 1109–21, 07 2013.

[36] T. Greensmith and B. Fisher, “Impact of personality factors on interface interaction and the development of user profiles: Next steps in the personal equation of interaction,” *Information Visualization - IVS*, vol. 11, pp. 205–221, 06 2012.

[37] M. C. Ioannou, K. Mogg, and B. P. Bradley, “Vigilance for threat: effects of anxiety and defensiveness,” *Personality and Individual Differences*, vol. 36, no. 8, pp. 1879–1891, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191886903003088

[38] T. Delgado, T. Alves, and S. Gama, “How neuroticism and locus of control affect user performance in high-dimensional data visualization,” *Computers & Graphics*, vol. 109, pp. 88–99, 2022.

[39] C. Ziemkiewicz and R. Kosara, “Preconceptions and individual differences in understanding visual metaphors,” *Computer Graphics Forum*, vol. 28, no. 3, pp. 911–918, 2009. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2009.01442.x

[40] T. Alves, T. Delgado, J. Henriques-Calado, D. Gonçalves, and S. Gama, “Exploring the role of conscientiousness on visualization-supported decision-making,” *Computers & Graphics*, 2023.

[41] 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,” in *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2011, pp. 81–90.

[42] M. Johnston, P. Gilbert, C. Partridge, and J. Collins, “Changing perceived control in patients with physical disabilities: An intervention study with patients receiving rehabilitation,” *British Journal of Clinical Psychology*, vol. 31, no. 1, p. 89–94, 1992.

[43] D. Toker, C. Conati, and G. Carenini, “Gaze analysis of user characteristics in magazine style narrative visualizations,” *User Modeling and User-Adapted Interaction*, vol. 29, pp. 1–35, 11 2019.

[44] C. Chen and M. Czerwinski, “Spatial ability and visual navigation: an empirical study,” *New Review of Hypermedia and Multimedia*, vol. 3, no. 1, pp. 67–89, 1997. [Online]. Available: https://doi.org/10.1080/13614569708914684

[45] M. Velez, D. Silver, and M. Tremaine, “Understanding visualization through spatial ability differences,” in *VIS 05. IEEE Visualization, 2005.*, 2005, pp. 511–518.

[46] C. Conati and H. Maclaren, “Exploring the role of individual differences in information visualization,” in *Proceedings of the Working Conference on Advanced Visual Interfaces*, ser. AVI ’08. New York, NY, USA: Association for Computing Machinery, 2008, p. 199–206. [Online]. Available: https://doi.org/10.1145/1385569.1385602

[47] S. Lallé, D. Toker, C. Conati, and G. Carenini, “Prediction of users’ learning curves for adaptation while using an information visualization,” *International Conference on Intelligent User Interfaces, Proceedings IUI*, vol. 2015, pp. 357–368, 03 2015.

[48] B. Allen, “Perceptual speed, learning and information retrieval performance,” in *SIGIR ’94*, B. W. Croft and C. J. van Rijsbergen, Eds. London: Springer London, 1994, pp. 71–80.

[49] V. Mittelstädt, P. Brauner, M. Blum, and M. Ziefle, “On the visual design of erp systems the – role of information complexity, presentation and human factors,” *Procedia Manufacturing*, vol. 3, pp. 448–455, 2015, 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2351978915002085

[50] M. Ziefle, P. Brauner, and F. Speicher, “Effects of data presentation and perceptual speed on speed and accuracy in table reading for inventory control,” *Occupational Ergonomics*, vol. 12, pp. 119–129, 09 2015.

[51] C. Chen, “Individual differences in a spatial-semantic virtual environment,” *Journal of The American Society for Information Science - J AMER SOC INFORM SCI*, vol. 51, pp. 529–542, 01 2000.

[52] C. Conati, S. Lallé, M. A. Rahman, and D. Toker, “Further results on predicting cognitive abilities for adaptive visualizations,” in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, 2017, pp. 1568–1574. [Online]. Available: https://doi.org/10.24963/ijcai.2017/217

[53] S. Luck and E. Vogel, “The capacity of visual working memory for features and conjunctions,” *Nature*, vol. 390, no. 6657, p. 279—281, November 1997. [Online]. Available: https://doi.org/10.1038/36846

[54] M. M. Hayhoe, D. G. Bensinger, and D. H. Ballard, “Task constraints in visual working memory,” *Vision Research*, vol. 38, no. 1, pp. 125–137, 1998. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0042698997001168

[55] D. Toker, C. Conati, G. Carenini, and M. Haraty, "Towards adaptive information visualization: On the influence of user characteristics," in *User Modeling, Adaptation, and Personalization*, J. Masthoff, B. Mobasher, M. C. Desmarais, and R. Nkambou, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 274–285.

[56] S. Lallé and C. Conati, "The role of user differences in customization: A case study in personalization for infovis-based content," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, ser. IUI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 329–339. [Online]. Available: https://doi.org/10.1145/3301275.3302283

[57] B. Steichen, G. Carenini, and C. Conati, "User-adaptive information visualization: Using eye gaze data to infer visualization tasks and user cognitive abilities," in *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, ser. IUI '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 317–328. [Online]. Available: https://doi.org/10.1145/2449396.2449439

[58] 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," in *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2017, pp. 104–115.

[59] I. Cho, R. Wesslen, A. Karduni, S. Santhanam, S. Shaikh, and W. Dou, "The Anchoring Effect in Decision-Making with Visual Analytics," in *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*. Phoenix, AZ, USA: IEEE, Oct. 2017, pp. 116–126. [Online]. Available: https://ieeexplore.ieee.org/document/8585665/

[60] E. Wall, A. Narechania, A. Coscia, J. Paden, and A. Endert, "Left, Right, and Gender: Exploring Interaction Traces to Mitigate Human Biases," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 966–975, Jan. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9555923/

[61] A. Shadare, S. Musa, C. Akujuobi, and R. Perry, "Data visualization," *International Journal of Engineering Research and Advanced Technology (IJERAT)*, vol. 12, pp. 2454–6135, 12 2016.

[62] S. Ratneshwar, A. D. Shocker, and D. W. Stewart, "Toward Understanding the Attraction Effect: The Implications of Product Stimulus Meaningfulness and Familiarity," *Journal of Consumer Research*, vol. 13, no. 4, pp. 520–533, 03 1987. [Online]. Available: https://doi.org/10.1086/209085

[63] K. J. Vicente, B. C. Hayes, and R. C. Williges, "Assaying and isolating individual differences in searching a hierarchical file system," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 29, no. 3, p. 349–359, 1987.

[64] G. Carenini, C. Conati, E. Hoque, B. Steichen, D. Toker, and J. Enns, "Highlighting interventions and user differences," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014.

[65] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *The Journal of Psychology*, 1956.

[66] S. Lee, S.-H. Kim, and B. C. Kwon, "Vlat: Development of a visualization literacy assessment test," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 551–560, 2017.

A

VLAT Questions

Table

Budget and Box Office of Movies			
Budget (million $)	**Real Worldwide Box Office (million $)**	**Genre**	**Title**
3,9	402	Drama	Gone with the Wind
237	2790	Action	Avatar
200	2194	Drama	Titanic
11	776	Action	Star Wars: Episode IV - A New Hope
356	2798	Action	Avengers: Endgame
8,2	286	Biography	The Sound of Music
10,5	793	Family	E.T. the Extra-Terrestrial
13	123	Adventure	The Ten Commandments
11	245	Drama	Doctor Zhivago

Item 1	
Task Name	Retrieve Value
Stem	What is the budget of the movie ”The Sound of Music”?
Options (highlight the correct option)	A) 8.20 million \$ B) 10.5 million \$ C) 286 million \$ D) 356 million \$

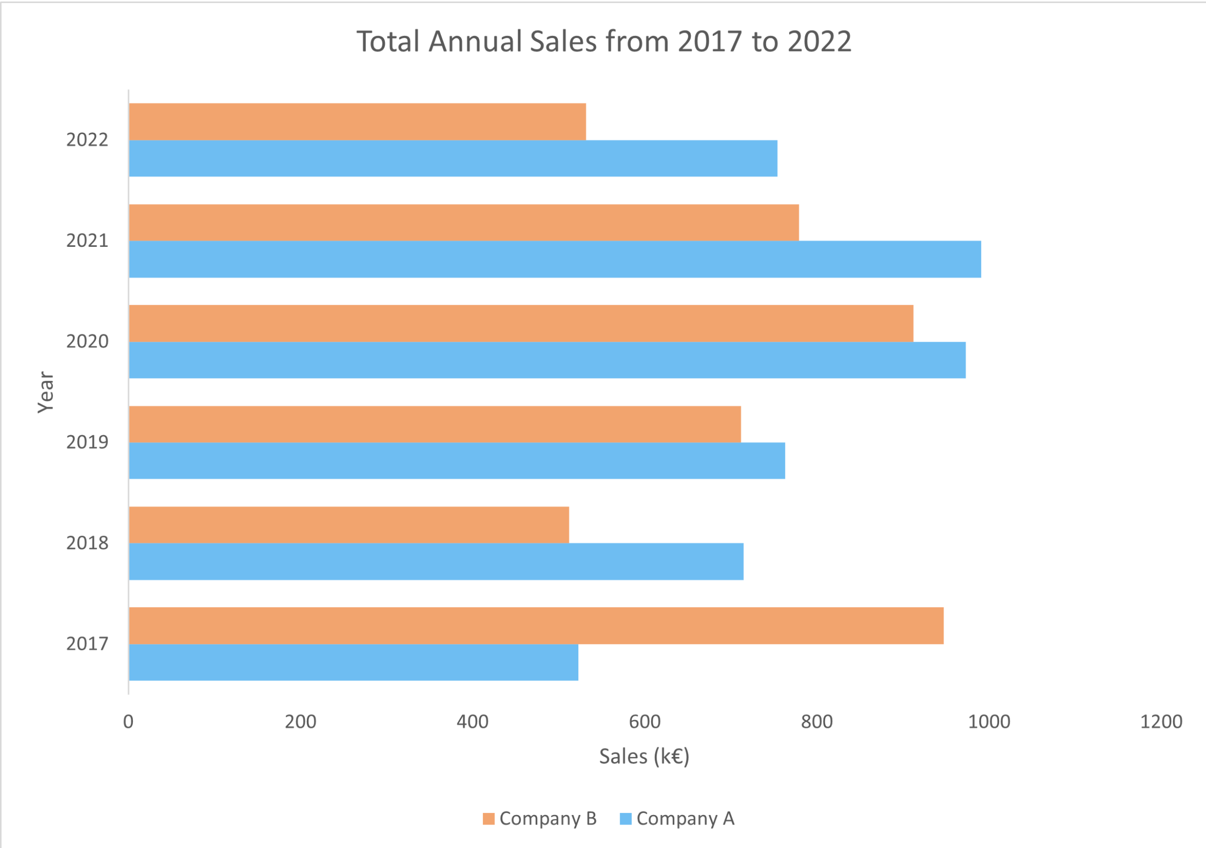
Item 2	
Task Name	Retrieve Value
Stem	What is the real worldwide box office of the only adventure movie in the table?
Options (highlight the correct option)	A) 13 million \$ B) 123 million \$ C) 245 million \$ D) 793 million \$

Item 3	
Task Name	Find Extremum
Stem	What movie had the highest budget?
Options (highlight the correct option)	A) Avatar B) Avengers: Endgame C) Gone with the Wind D) Titanic

Item 4	
Task Name	Find Extremum
Stem	What movie had the lowest real worldwide box office?
Options (highlight the correct option)	A) Avengers: Endgame B) Doctor Zhivago C) Gone with the Wind D) The Ten Commandments

Item 5	
Task Name	Make Comparisons
Stem	Titanic is the drama movie with the highest budget.
Options (highlight the correct option)	A) True B) False

A.1 One-Sided Bar Chart



Item 1	
Task Name	Retrieve Value
Stem	What were the annual total sales of Company A in 2019?
Options (highlight the correct option)	A) 512k € B) 712k € C) 763k € D) 973k €

Item 2	
Task Name	Find Extremum
Stem	What was the year with lowest annual total sales for Company B?
Options (highlight the correct option)	A) 2017 B) 2018 C) 2021 D) 2022

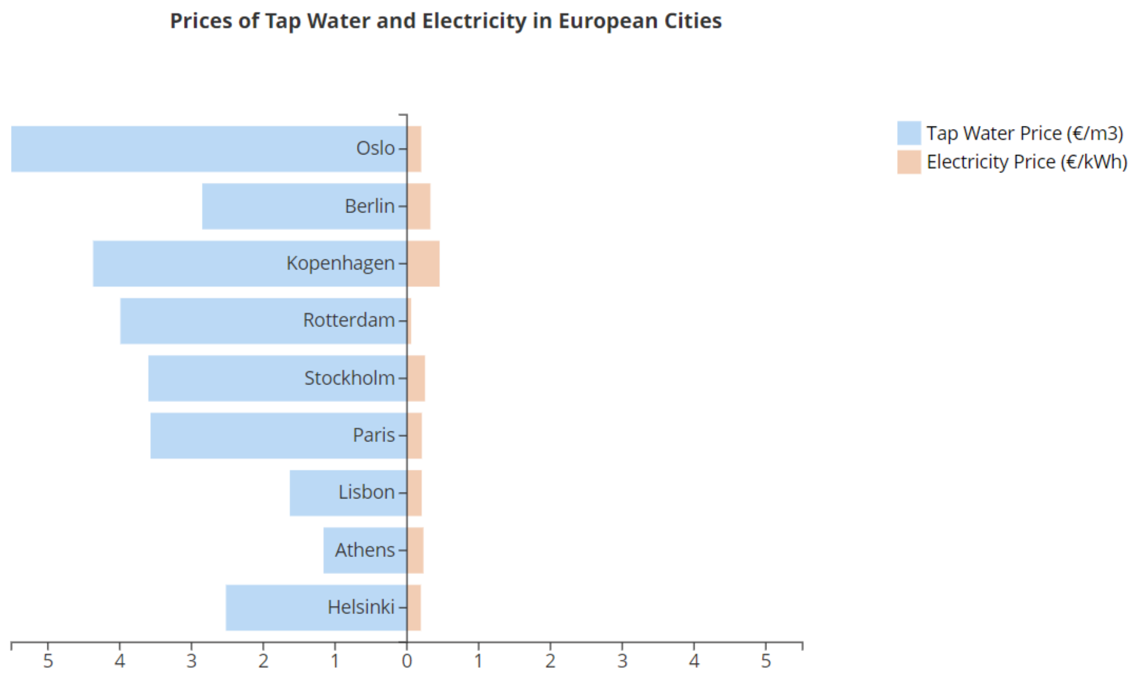
Item 3	
Task Name	Find Extremum
Stem	What was the year with higher annual total sales for Company B?
Options (highlight the correct option)	A) 2017 B) 2018 C) 2020 D) 2021

Item 4	
Task Name	Determine Range
Stem	What is the range of annual total sales for Company A?
Options (highlight the correct option)	A) 0 - 991k € B) 512k - 947k € C) 523k - 991k € D) 947k - 991k €

Item 5	
Task Name	Make Comparisons
Stem	How many years the annual total sales of Company A were greater than the annual total sales of Company B?
Options (highlight the correct option)	A) 0 B) 1 C) 3 D) 5

Item 6	
Task Name	Make Comparisons
Stem	What was the year the two companies had the most similar total sales?
Options (highlight the correct option)	A) 2017 B) 2019 C) 2020 D) 2021

A.2 Two-Sided Bar Chart



Item 1	
Task Name	Retrieve Value
Stem	What is the price of tap water in Berlin?
Options (highlight the correct option)	A) 1.63 €/m3 B) 2.85 €/m3 C) 3.279 €/kWh D) 4.559 €/kwh

Item 2	
Task Name	Find Extremum
Stem	What city has the highest tap water price?
Options (highlight the correct option)	A) Athens B) Helsinki C) Copenhagen D) Oslo

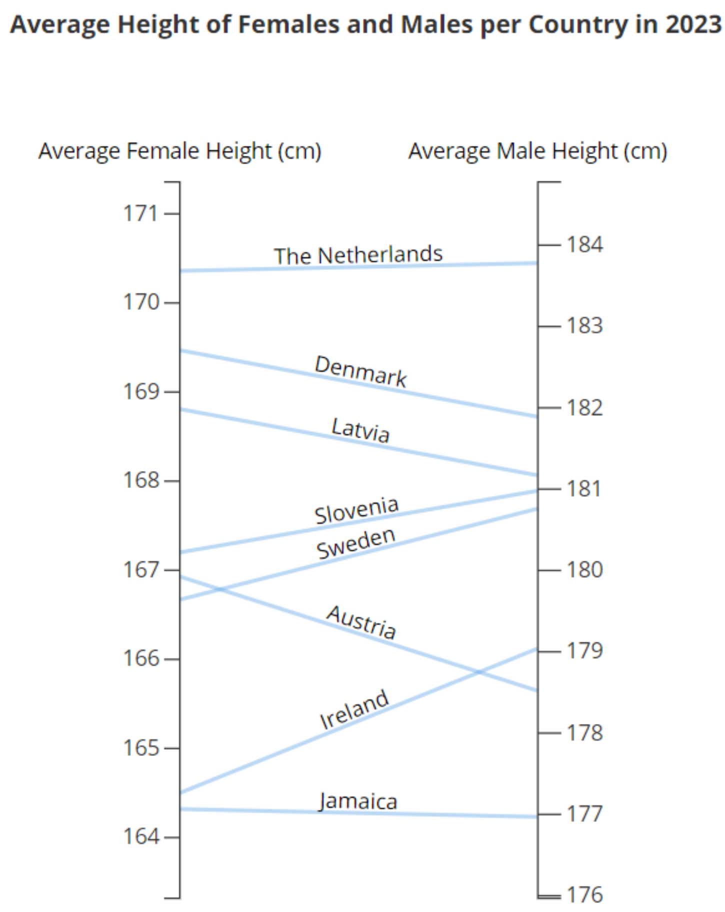
Item 3	
Task Name	Find Extremum
Stem	What city has the lowest electricity price?
Options (highlight the correct option)	A) Athens B) Berlin C) Copenhagen D) Rotterdam

Item 4	
Task Name	Determine Range
Stem	What is the range of tap water prices in these cities?
Options (highlight the correct option)	A) 0 - 5.51 €/m3 B) 1.16 - 5.51 €/m3 C) 1.16 - 4.37 €/m3 D) 1.63 - 5.51 €/m3

Item 5	
Task Name	Make Comparisons
Stem	Stockholm has a higher electricity price than Helsinki.
Options (highlight the correct option)	A) True B) False

Item 6	
Task Name	Retrieve Value (relative value)
Stem	The ratio of the price of electricity to the price of tap water in Paris is lower than in Rotterdam.
Options (highlight the correct option)	A) True B) False

A.3 Parallel Coordinates Plot



Item 1	
Task Name	Retrieve Value
Stem	What is the average female height in Austria?
Options (highlight the correct option)	A) 164.32 cm B) 166.93 cm C) 167.20 cm D) 178.52 cm

Item 2	
Task Name	Retrieve Value
Stem	What is the average male height in Denmark?
Options (highlight the correct option)	A) 168.81 cm B) 169.47 cm C) 181.17 cm D) 181.89 cm

Item 3	
Task Name	Find Extremum
Stem	What is the country with greater average male height?
Options (highlight the correct option)	A) Denmark B) Ireland C) Jamaica D) The Netherlands

Item 4	
Task Name	Make Comparisons
Stem	What is the country with biggest difference between average female height and average male height?
Options (highlight the correct option)	A) Austria B) Ireland C) Jamaica D) Sweden

Item 5	
Task Name	Make Comparisons
Stem	Males and females have the same average height in Jamaica.
Options (highlight the correct option)	A) True B) False

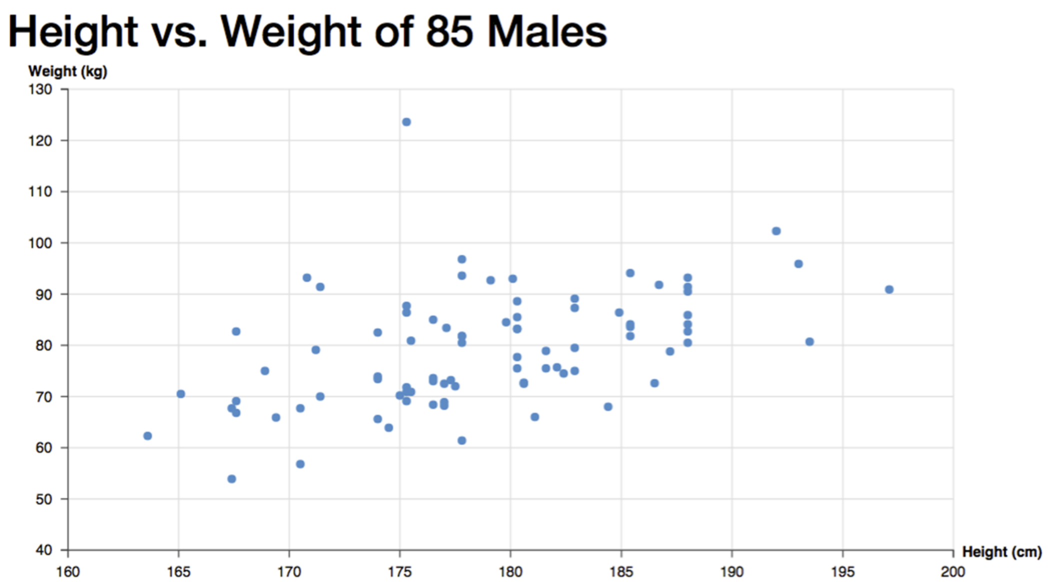
Item 6	
Task Name	Make Comparisons
Stem	Males have a greater average height than females in Denmark.
Options (highlight the correct option)	A) True B) False

Item 1	
Task Name	Retrieve Value
Stem	What is the weight for the person who is 165.1 cm tall?
Options (highlight the correct option)	A) 53.9 kg B) 67.7 kg C) 70.5 kg D) 82.7 kg

Item 7	
Task Name	Make Comparisons
Stem	What country has the most similar difference of average heights to Latvia?
Options (highlight the correct option)	A) Austria B) Denmark C) Slovenia D) Sweden

A.4 Scatterplot

Questions and visualization adapted from Lee et al. [66].



Item 2	
Task Name	Find Extremum
Stem	What is the height for the tallest person among the 85 males?
Options (highlight the correct option)	A) 175.3 cm B) 192 cm C) 197.1 cm D) 200 cm

Item 3	
Task Name	Determine Range
Stem	What is the range in weight for the 85 males?
Options (highlight the correct option)	A) 40 - 130 kg B) 62.3 - 90.9 kg C) 53.9 - 102.3 kg D) 53.9 - 123.6 kg

Item 4	
Task Name	Find Anomalies
Stem	What is the height for a person who lies outside the others the most?
Options (highlight the correct option)	A) 167.4 cm B) 175.3 cm C) 193 cm D) 197.1 cm

Item 5	
Task Name	Find Clusters
Stem	A group of males are gathered around the height of 176 cm and the weight of 70 kg.
Options (highlight the correct option)	A) True B) False

Item 6	
Task Name	Find Correlations/Trends
Stem	There is a negative linear relationship between the height and the weight of the 85 males.
Options (highlight the correct option)	A) True B) False

Item 7	
Task Name	Make Comparisons
Stem	The weights for males with the height of 188 cm are all the same.
Options (highlight the correct option)	A) True B) False

Item 8	
Task Name	Characteristic Distribution
Stem	About the height for the 85 males is normally distributed.
Options (highlight the correct option)	A) True B) False

B

Experiment Guide

- **Planning and Preparation**
  - **Recruitment of Participants**
    * Open slots in a scheduling app within the schedule compatible with both researchers in charge of conducting the user tests.
    * Invite participants through the sharing the scheduling app link among members of the community along with a description and duration of the task.
  - **Experiment Apparatus Setup**
    * Setup the laptop and mouse. Guarantee the laptop has enough battery for the trial and that the screen luminosity is the same for all trials.
    * Sanitize BITalino sensors, laptop keyboard and mouse.
    * Pair BITalino with the computer.
    * Run the Phantom Effect app (frontend and backend, one terminal for each).
    * Before a new participant comes, refresh the app to generate a new UID.

- **Introduction**
  - **Onboarding**
    - Introduce Carolina and Frederico to the participant.
    - Thank the participant for agreeing to take part in the study.
    - Tell them why we are making this test and briefly describe the experiment, including how the app works.
  - **Consent Form**
    - Display the consent form to the participant and ask them to read it carefully.
    - Display the consent form to the participant and ask them to read it carefully.
    - Hand in a copy of the consent.
  - **Prepare Participant**
    - Once the participant gives their consent, proceed to equip them with BITalino sensors.
    - Have the participant seating comfortably in front of the laptop.
- **Practical Segment**
  - Take note of UID for data merging.
  - Ask the participant to take note of the UID (will be necessary to answer any requests to delete data).
  - Open the demographics form and request the participant to begin answering.
  - Make sure the participant reads the instructions for each step.
  - Once the participant enters the Visual Working Memory instructions page, make sure they train a few trials before starting the actual test.
  - The participant finishes the remaining phases.
- **Session Closing**
  - Once the participant finishes the tasks, make sure all needed data has been saved.
  - Remove sensors from participant.
  - Thank the participant one more time.
  - Detail our study, explaining the measured individual differences and biologic signals. Explain the phantom effect. Briefly explain the objectives of our study.
  - Farewell.

C

Informed Consent

Objetivo

Estamos a conduzir um estudo sobre como indivíduos são afetados por vieses cognitivos durante a interação com visualizações, nomeadamente pelo Phantom Effect. Pretendemos também estudar o efeito das capacidades cognitivas e a alteração dos sinais biométricos perante a expressão desses mesmos vieses cognitivos. O objetivo desta sessão é responder a um conjunto de testes e executar uma tarefa de decisão com diferentes visualizações e recolher as suas observações sobre as mesmas.

Características da sessão

A duração desta sessão não deve ser maior que 45 minutos e terá lugar no INESC-ID. Alertamos para a presença de sequências rápidas de imagens que podem levar a ataques epiléticos. Se sofre desse tipo de ataques ou de fotossensibilidade, por favor informe o assistente para parar a sessão. Caso não o faça e assine em como não sofre destes problemas, assumimos que consente com os riscos, tomando como verdadeira a sua afirmação ao assinar o consentimento. Não identificámos outros 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 (ex. seleção de elementos, cliques, tempo despendido), (ii) respostas a questionários, (iii) atividade eletrodérmica, e (iv) eletroencefalograma. Todos os dados recolhidos serão mantidos em sigilo. Os dados mencionados poderão também ser utilizados para apresentação ou exibição de resultados, devidamente pseudonimizados, 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 dados pseudonimizados da experiência (não incluem identificador, ou seja, sem características pessoais identificáveis) serão analisados, exclusivamente, pelos membros da equipa de investigação.

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. Para tal, deverá guardar o identificador único que lhe foi atribuído. Por favor, não partilhe esse identificador com ninguém, exceto os membros deste estudo. 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 capacidades cognitivas recolhidos anteriormente.

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

Obrigado pela sua colaboração!

Eu, ______________________________________, declaro sob a minha honra que não sofro de epilepsia ou fotossensibilidade.

____

(participante)

____

(investigador responsável)

____

(data)

Ao participante será entregue uma cópia assinada deste formulário.

Equipa

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

Tomás Alves
tomas.alves@tecnico.ulisboa.pt

Carolina Pereira (Responsável pelo tratamento de dados)
carolina.m.pereira@tecnico.ulisboa.pt

Frederico Rosado (Responsável pelo tratamento de dados)
frederico.caiado.rosado@tecnico.ulisboa.pt