# Studying the Distinction Cognitive Bias in Information Visualization

**Vasco Pires**
Instituto Superior Técnico
Lisboa, Portugal

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

### Author Keywords
Distinction Bias; Cognitive Bias; Human-Computer Interaction; Information Visualization; Visual Analytics; Individual Differences; Personality; Neuroticism

## INTRODUCTION
Nowadays people have to deal with increasingly growing amounts of information, produced and collected at incredible rates, while simultaneously having to make progressively more complex and critical decisions [19, 9]. Information Visualization provides users with graphical representations that allow them to efficiently explore, analyse and communicate patterns of bodies of data with various degrees of complexity [20]. In this sense, in order to design effective visualizations we must consider the process of human reasoning, as well as its limitations [5].

Human cognition deals with finite resources such as working memory capacity and attention [17, 8] and thus, in order to make decisions quickly, with limited information and resources, the human mind makes use of heuristics, rules of thumb and approximations, even when we are not consciously aware of these strategies. The imperfections of these strategies manifest themselves as cognitive biases [5, 19]. Among them and the focus of this work is Distinction Bias, described as the tendency to view two options as more dissimilar when evaluating them simultaneously than when evaluating them separately, influencing our predictions and choices [7].

Moreover, individual differences, such as personality, also have an effect on how humans perceive and process information and therefore may impact the way we live and make our decisions [14]. Throughout history many personality models were proposed, being the Five-Factor Model one of the most widely accepted at present times. Throughout most models, a personality trait that prevails is neuroticism. Highly neurotic individuals are characterized by a tendency to feel worried, nervous, depressed, self-conscious and to overall more easily experience negative emotions [12, 4].

The topic of personality has been considered relevant in several fields, among them the field of Human-Computer-Interaction [18] and research has demonstrated how distinct personality types can make a difference in problem-solving and behavioral patterns [11]. While visualization tools aim to support judgments and decisions in the context of problem-solving, the topic of cognitive biases and its impact on how people use these tools is still relatively unexplored. As for Distinction Bias specifically, while considered potentially relevant, it still has not been discussed in the field [5].

### Objectives
The main focus of this study is to **understand the effects of Distinction Bias in the context of Information Visualization, as well as its possible correlations with a personality trait, neuroticism.** In order to achieve this goal, several intermediate steps were defined to progress through the course of this study, such as the development of the visualizations to be used in the study, user testing, collection of user personality data and further analysis of collected data.

## BACKGROUND ON DISTINCTION BIAS
Distinction Bias describes the way people's utility and affective predictions for several alternatives can vary depending on the evaluation mode they find themselves in, having these predictions then possibly lead to sub-optimal decisions [7]. In essence Distinction Bias consists of a situation in which two options appear to be more dissimilar when we examine them together, what we refer to as Joint evaluation (JE) mode, as opposed to separately, what we refer to as Separate evaluation (SE) mode.

Specifically, Hsee and Zhang [7] corroborated through a set of experiments that when people in JE mode predict the affective state they would get from various outcomes, comparing attribute values relative to one another, they tend to assess significantly different levels of happiness. For example, if someone is a book author and is trying to sell their book, a higher number of book sales would be seen as desirable and when comparing and forecasting the happiness experienced from selling 100 or 200 books in JE mode this person is likely to predict a higher level of happiness for an outcome in which they sell 200 books than for the one in which they sell 100.

In contrast, people in SE mode often tend to not have a precise idea of how quantitatively good or bad an attribute value is by itself, tending to not report significantly different levels of happiness for quantitatively different presented options, though they may be able to evaluate them as qualitatively positive or negative in relation to a baseline or reference point.

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

In this sense, depending on the evaluation mode people find themselves in, the evaluation function of an attribute will tend to differ, as represented in Figure 1. We can observe that JE mode's evaluation function is relatively steep and smooth, as people perceive alternative values relative to one another. In comparison, SE mode's evaluation function is steep when close to the baseline value but somewhat flat elsewhere, as in SE mode people are mostly merely able to distinguish if attribute values are either above or below a baseline. This way people in SE mode tend to predict or experience contrasting levels of utility or happiness for qualitatively different options, yet similar levels when alternatives differ only quantitatively.

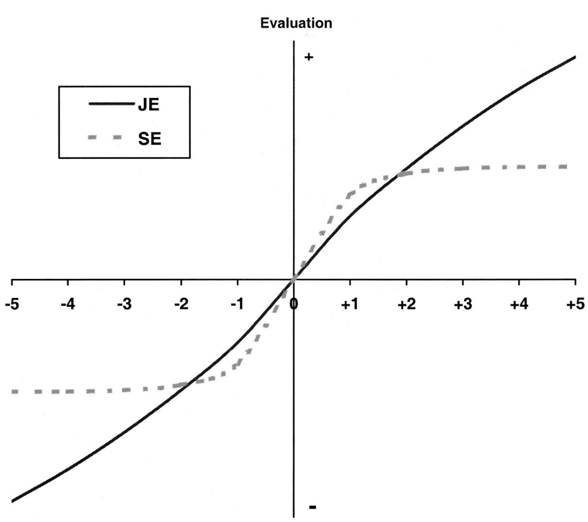


**Figure 1. Joint-evaluation (JE) curve and separate-evaluation (SE) evaluation curve for a hypothetical attribute [7].**

## METHODOLOGY

In order to reach the goal of this study, we took Hsee and Zhang [7] studies and findings as inspiration, crafting visualizations and tasks as models that aim to replicate and test the effects that characterize Distinction Bias, in an InfoVis context.

Likewise, as individual differences have an effect on how we perceive and process information and therefore may impact the way we live and make our decisions [14], we also found pertinent to incorporate the study of personality and its effects on bias-prone scenarios in our work. As such, we will additionally focus on the impacts of a specific personality trait, neuroticism.

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

### Visualization

Taking into consideration the gap of research regarding Distinction Bias within the InfoVis context [5] together with, to the best of our knowledge, the absence of prior work specifically linking visualization idioms to the study of Distinction Bias, for this work we opted to utilize a bar chart idiom to encode the presented data, motivated by the idiom's suitability for the abstract tasks of looking up and comparing individual values, simplicity and the prospects of its familiarity to participants[13]. Additionally, bar charts were observed to have similar results compared to presenting data without the use of visualizations and shown to provide the most relevant findings when compared to other encodings in the InfoVis context [1].

In order to study the Distinction Bias effects, the layouts for our first experiment simulate each of the two evaluation modes: SE mode, presenting one single bar encoding the value referent to a specific outcome of a given hypothetical scenario as in Figure 2, or several for JE mode, as in Figure 3.
We opted to not introduce color to our bar chart layout, as we are dealing with a bias in which cognition is influenced by comparison, and in this regard the monotone layout allows us to work with the largest color contrast levels available.

The crafted visualization reveals a different bar chart for each of the tasks, titled according to the task's specific hypothetical scenario being studied. Each of the bars in the horizontal axis represents a possible outcome to the presented scenario, identified with letters, having their length encode the attribute values referent to each outcome presented, which may differ from each other quantitatively as well as qualitatively.
The valence of each outcome is depicted by representing its attribute value as either positive or negative on the vertical axis. The horizontal axis range is adjusted to the values being encoded at each given moment throughout the first experiment.

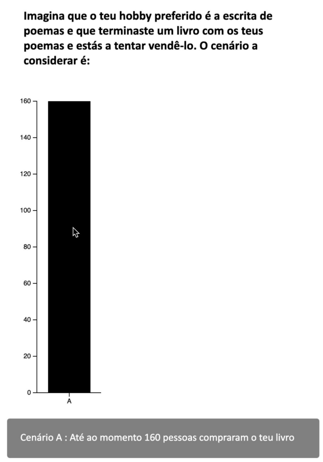


**Figure 2. Visualization layout for an SE mode task, presenting only 1 outcome.**

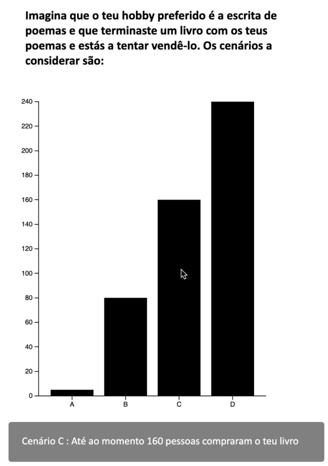


**Figure 3. Visualization layout for a JE mode task, presenting 4 different outcomes.**

Hovering the mouse over a bar reveals a text box containing the value encoded, as well as more context regarding the selected outcome.

With our second experiment we further explore how the Distinction Bias manifests itself in JE mode, in relation to the number of presented outcomes that differ only quantitatively, as these are the conditions that tend to lead to biased predictions. As all participant conditions take place in JE mode for the second experiment, each are presented with several outcome bars for each of the tasks. Yet, between participants, the number of bars displayed will vary, being each participant presented with either 3, 5 or 7 different outcomes for the entirety of tasks in the second experiment.
Additionally, with our second experiment we also explore the possibility that the scale in which the data is displayed may influence biased behavior in our study. We found this factor to be pertinent as we are exploring a bias in which cognition is influenced by comparison not only of several options but also with perceived baselines. Therefore, the bar chart visualization counts with a vertical axis maximum value that will also vary between three possible levels, the maximum attribute value of the presented task and two progressively larger values picked arbitrarily for each task. A single vertical axis condition level was assigned to each participant session, consistent throughout the entirety of the second experiment tasks of that session.

### *Tasks*

For each of our two experiments participants are presented information through the visualizations detailed in the previous section and for each task, as in the Hsee and Zhang [7] studies, are asked to imagine that a hypothetical scenario is occuring to them, having to forecast their happiness levels for different possible outcomes, rating each item from 1 (extremely sad) to 9 (extremely happy), as assessed in prior literature[7].

Our two experiments consist of a total of 7 tasks. As both experiments are held sequentially with the same participants, task order for each of the two experiments was randomized: First experiment tasks were re-labeled from 1 to 4; Second Experiment tasks were re-labeled from 5 to 7. Furthermore for both experiments, participants not being presented with the entirety of the prepared outcomes have their presented outcomes re-labeled with sequential letters. Tasks were translated into Portuguese for participant sessions.

For our first experiment, to better understand the effects of Distinction Bias in the context of visualizations, we made use of data heavily inspired by the outcomes tested in Hsee and Zhang's [7] misprediction studies. Furthermore, throughout the authors' discussion of findings, some additional concepts were briefly brought to light, which we also took as inspiration to design some additional tasks. In the end, we arrived at four different tasks, each requiring participants to make predictions of their affective states for hypothetical outcomes:

**Task 1.1.** Participants are asked to imagine that their favorite hobby is writing poems and that they have compiled a book of their poems and are trying to sell it. The outcomes to consider are: So far you have sold **A.** 5 books; **B.** 80 books; **C.** 160 books; **D.** 240 books.

**Task 1.2.** Participants are asked to imagine that they are requested to read a list of words. The outcomes to consider are: The list contains **A.** 25 negative words; **B.** 10 negative words; **C.** 10 positive words; **D.** 25 positive words.

**Task 1.3.** Participants are asked to imagine that they went to a casino and gambled their own money. The outcomes to consider are: Your bet **A.** lost you 100 euros; **B.** lost you 50 euros; **C.** earned you 50 euros; **D.** earned you 100 euros.

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

It is important to note that the compared attribute in Task 1.4, an assignment grade, is presumed by the original studies authors[7] to be easier to evaluate independently, which in theory may mean that, when presented with this task, participants may be able to predict their affective state in a less biased way, making it a compelling additional situation to test.

Additionally, in order to further explore the Distinction Bias effects in InfoVis, we utilized our second experiment as a way to expand our exploratory analysis and examine how the number of presented alternatives as well as the scale in which they are displayed may affect JE mode predictions.
For the tasks in the second experiment, a participant can be presented with either three, five or seven different outcomes for the same task. In each of these tasks three outcomes serve as an anchor to understand whether more options affect user assessment. This means that, when the condition only presents three options to the participant, only the anchor options are presented in the bar chart. However, for the five or seven options conditions, we leverage the anchors and include distractor quantitatively different outcomes. Participants faced

with five outcomes were presented with the three outcomes that were studied along with an outcome quantitatively larger than any other and an outcome quantitatively smaller than any other. Participants faced with seven outcomes were presented with the five previously described outcomes as well as two additional outcomes quantitatively in between the three studied outcomes.

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

**Task 2.1.** Participants are asked to imagine that they are a musician and have been working on composing new songs. The outcomes to consider are: In the last month you have composed **A.** 2 songs; **B.** 4 songs (Anchor); **C.** 6 songs; **D.** 8 songs (Anchor); **E.** 10 songs; **F.** 12 songs (Anchor); **G.** 15 songs.

**Task 2.2.** Participants are asked to imagine that they have a hairdresser appointment and are planning not to take long. The outcomes to consider are: The appointment got delayed **A.** 10 minutes; **B.** 30 minutes (Anchor); **C.** 40 minutes; **D.** 60 minutes (Anchor); **E.** 80 minutes; **F.** 90 minutes (Anchor); **G.** 120 minutes.

**Task 2.3.** Participants are asked to imagine that they have created a mobile application and are distributing it on the app markets. The outcomes to consider are: This month it generated **A.** 100 euros; **B.** 250 euros (Anchor); **C.** 400 euros; **D.** 500 euros (Anchor); **E.** 600 euros; **F.** 800 euros (Anchor); **G.** 1000 euros.

### Measures

Taking into consideration previous literature, we collected data not only considering the measures that allow us to evaluate and study Distinction Bias in the InfoVis realm but also regarding participants' personality and their interaction with the visualization.

**Independent Variables:** To study the effects of Distinction Bias for the first experiment tasks, participants are assigned to one of five **conditions** (JE, SE1, SE2, SE3, SE4), one for JE mode, being faced with all four outcome scenarios, and four for SE mode, being faced with only one. As mentioned prior, alongside Distinction Bias, our study also focuses on exploring the effects of a specific personality trait, neuroticism, within the InfoVis context. For our study the **neurotic trait score** is evaluated in the context of the Five-Factor Model, assessed through the NEO PI-R questionnaire [10].
For our second experiment some additional independent variables are required. This experiment is set entirely in JE mode, and as such every participant evaluation mode **condition** is equivalent. As we intend to analyse how the **amount of presented outcomes** can impact the bias effects, we set three between-subjects groups for which, throughout the second experiment tasks, each participant can be presented with either 3, 5 or 7 quantitatively different outcomes. Additionally, in order to study possible impacts of the visualization's vertical axis range. To achieve this we set three between-subjects groups for which, in each task, participants can be presented with a bar chart visualization with a **y-axis maximum value** consisting of one of three possible values: the maximum attribute value of the presented task and two progressively larger values picked arbitrarily for each task.

**Dependent Variables:** All participants analyze their visualization and assess their happiness levels for the outcome(s) displayed, denominated as **affective forecasts**, in a 9-point scale, from 1 being extremely unhappy to 9 being extremely happy, as done for Hsee and Zhang [7]'s studies. Following the sessions the **differences between affected forecasts** for consecutive outcomes are computed for each participant, as this derived measure is instrumental for examining correlations between factors and the bias effects. The **time** (s) each participant takes to complete each given task's set of predictions and the amount of times they **hover** with their mouse cursor over the outcome bars for a given task are also collected.

### Research Questions and Hypotheses

By taking inspiration from the available literature on Distinction Bias [7] our experiments start by posing the following research question:

**RQ1. Does the Distinction Bias transfer to Information Visualization?**

Specifically, with our first experiment we intend to analyze and compare the effects that simulating the two opposing evaluation modes, JE and SE mode, within our bar chart visualization can have on people's happiness forecasts. As Distinction Bias literature points out, when people are presented with several outcomes for a scenario, what we call JE mode, they tend to compare the presented options in order to predict the utility and level of happiness each outcome could bring them. This tends to result in, for quantitatively different values of a compared attribute, significantly differing affective predictions. In contrast, people in SE mode consider only one outcome, not comparing it to other options, tending to, for quantitatively different attribute values of the same valence, predict small differences of utility and happiness, if any. In circumstances in which outcome values differ only qualitatively, the differences in utility recalled by both people in JE mode and SE mode are observed to not differ significantly [7]. These findings are what characterize what we denominate as Distinction Bias and are illustrated in Figure 1.

As related literature suggests, we believe that this effect may transfer into the realm of visualizations and for this we hypothesize:

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

Furthermore:

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

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

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

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

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

Distinction Bias literature denotes that SE mode predictions of happiness level do not follow the steep curve of JE mode predictions for quantitatively different values of an attribute, as evidenced in Figure 1. This theory seems to suggest that by variating the number of presented alternatives in JE mode we may expect different results from JE mode predictors regarding the same outcomes, as summarized in **H2**.

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

One additional factor was also at play in Task 1.4, the ability to independently evaluate an attribute. Hsee and Zhang [7] discuss that Distinction Bias theory applies mostly to attributes that are not too easy to evaluate independently, yet believe that finding an attribute that is notably easy to evaluate independently is the exception rather than the rule. The authors state school grades as an example of an attribute which people have sufficient knowledge about, resulting in prediction values, when assessed in SE mode, being close to those assessed in JE mode. We took these statements as inspiration to come up with Task 1.4, for which the aforementioned **H1** might not be verified, posing the question:

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

Additionally, our study integrates a component of personality analysis, namely of participants' neuroticism scores. Highly neurotic individuals have a tendency to easily experience negative emotions and to more easily feel stressed, anxious and depressed [4, 22], tending to also be more pessimistic [21, 4]. Highly neurotic users were observed to focus more on negative information than on positive information [15] and have harder times making decisions [2, 16], yet individuals with high neuroticism scores were observed to take less time in search and inference tasks [6]. Furthermore, Brown et al. [3] denoted a correlation between personality traits such as neuroticism, and mouse activity, like hovers.

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

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

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

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

With our research focus revolving around the effects of Distinction Bias, it is relevant to test if highly neurotic participants predict their affect for the presented scenario outcomes consistently and proportionally lower in value, which should not in theory significantly impact the affective prediction differences between two given presented outcomes, or if, being primed to incur in Distinction Bias, the neurotic trait displays correlations with the overprediction of these differences, specifically when attribute values differ merely quantitatively.

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

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

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

### Data Collection

The data used for our study was collected from a total of 80 participant sessions (20 females, 57 males and 3 other) between the ages of 18 to 27 years of age (M = 21.71, SD = 2.425). However we were only able to collect personality data referent to the NEO PI-R[10] from 58 of these participants (18 females, 38 males and 2 other). Participants were recruited through standard convenience sampling procedures. Prior to the experiments a small questionaire was utilized to inform and assure consent and collect demographic data. Audio and screen were recorded for each of the sessions as a safety measure. Affective forecasts for the several tasks of the session were collected through a questionnaire. Remaining dependent variable collection was assured by our visualization interface.

#### *Procedure*

As aforementioned, sessions were composed of two experiments, performed sequentially by the same participants.
Each session started as the participant read and filled out a consent form. Participants were then inquired if they were familiar with the bar chart idiom. A tutorial was then provided detailing how to interact with the interface and complete tasks. Participants were then asked if they had any questions. As participants confirmed that they understood what they were required to do, we would start the audio and screen recording.

The first experiment then began taking place. Each participant was randomly assigned a condition, consisting of a simulation of either JE mode or SE mode, consistent throughout the first experiment tasks. For each task, participants were presented with information regarding a specific scenario and asked to imagine this was occurring to them. Facing this hypothetical situation, participants analyzed the visualization, encoding the outcomes for this situation, several in the case of participants in JE mode conditions and only one for participants in SE mode conditions. Participants in JE mode conditions were verbally encouraged to compare the presented alternative outcomes, as a way to better assure their evaluation mode. SE mode participants were not, for the same reason. Participants were asked to predict their level of happiness, what we refer to as affective forecast, for each presented outcome.

Similarly to the prior described experiment, the second experiment is also composed of several tasks, for which participants of all conditions were required to imagine a specific scenario, analyse the visualization and asked to predict their level of happiness for each presented outcome. All participants were verbally encouraged to compare the presented alternative outcomes to further assure joint evaluation circumstances.

As the participant completed the total of seven tasks, we stopped the recordings, stored the session data, thanked them for their time and then finally showed ourselves available to take further questions regarding the purpose of our study.

## DISCUSSION OF RESULTS

We started by testing the Distinction Bias effects in the context of our visualization, with the aim to answer **RQ1** and **RQ2**.

### Distinction Bias and Information Visualization

In order to examine the ways in which Distinction Bias effects transfer to our visualization, we took on a similar approach as the one utilized for existing Distinction Bias literature. For each of the tasks we performed both a paired-samples T-Test, for JE mode participant data, and an Independent T-Test, for SE mode participant data, on each pair of reported affective forecasts referent to consecutive presented outcomes or, in other words, pair of happiness assessments referent to outcome bars situated next to eachother in the presented visualization.

*Task 1.1*

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

**Table 1. Task 1.1 - Mean affective forecasts for: Outcome A - sold 5 books, Outcome B - sold 80 books, Outcome C - sold 160 books and Outcome D - sold 240 books; and mean affective differences for consecutive pairs: AB, BC and CD.**

In Task 1.1, by analyzing the p-values referent to the tested consecutive pairs of outcomes, we can conclude that **JE mode participants reported statistically significant differences in levels of happiness for all consecutive pairs of presented outcomes, in accordance with our initial assumptions** that tests were expected to confirm statistically significant differences for all tested JE mode outcome pairs.
All outcomes displayed for Task 1.1 differ only quantitatively.
**SE mode participants did not report statistically significant differences in affective forecasts for only two of our three quantitatively different pairs of outcomes, which only partially supports our initial assumptions**, as quantitative differences between outcomes were predicted not to lead to statistically significant differences for affective forecasts of SE mode participants.

According to Distinction Bias literature concerning qualitative and quantitative differences, we anticipated that JE mode participants would tend to overpredict affective differences between outcomes with only quantitative variation, when compared to SE mode participants.
**JE mode participants reported lower affective forecast differences** between Task 1.1's Outcome A - sold 5 books - and Outcome B - sold 80 books, when compared to SE mode participants. **This result goes against our assumptions, as Outcomes A and B differ only quantitatively.**
Moreover, **JE mode participants reported higher affective forecast differences** between Task 1.1's Outcome B - sold 80 books - and Outcome C - sold 160 books, as well as between Outcome C - sold 160 books - and Outcome D - sold 240 books, when compared to SE mode participants. **This result supports our assumptions, as Outcomes B, C and D differ only quantitatively.**

*Task 1.2*

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

**Table 2. Task 1.2 - Mean affective forecasts for: Outcome A - read 25 negative words, Outcome B - read 10 negative words, Outcome C - read 10 positive words and Outcome D - read 25 positive words; and mean affective differences for consecutive pairs: AB, BC and CD.**

In Task 1.2, by analyzing the p-values referent to the tested consecutive pairs of outcomes, we can conclude that **JE mode participants reported statistically significant differences in levels of happiness for all consecutive pairs of presented outcomes, in accordance with our initial assumptions** that tests were expected to confirm statistically significant differences for all tested JE mode outcome pairs.
**SE mode participants did not report statistically significant differences in affective forecasts for our only two quantitatively different pairs of outcomes, AB and CD, supporting our initial assumptions**, as quantitative differences between outcomes were predicted not to lead to statistically significant differences for affective forecasts of SE mode participants. Additionally, **SE mode participants did report statistically significant differences in affective forecasts for our qualitatively different pair of outcomes, BC, also supporting our initial assumptions**, as qualitative differences between outcomes were predicted to lead to statistically significant differences for affective forecasts of SE mode participants.

According to Distinction Bias literature concerning qualitative and quantitative differences, we anticipated that JE mode participants would tend to overpredict affective differences between outcomes with only quantitative variation, when compared to SE mode participants. On the other hand, we anticipated that JE mode participants would not tend to overpredict affective differences between outcomes with only qualitative variation, when compared to SE mode participants.
**JE mode participants reported lower affective forecast differences** between Task 1.2's Outcome B - read 10 negative words - and Outcome C - read 10 positive words, compared to SE mode participants. **This result supports our assumptions, as Outcomes B and C differ only qualitatively.**
Moreover, **JE mode participants reported a higher affective forecast difference** between Task 1.2's Outcome A - read 25 negative words and Outcome B - read 10 negative words, as well as between Outcome C - read 10 positive words and Outcome D - read 25 positive words, when compared to SE mode participants, with SE mode participants even reporting a decrease in happiness levels between the last pair. **This result also supports our assumptions, as the outcomes pair AB and CD differ only quantitatively.**

### *Task 1.3*

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

**Table 3. Task 1.3 - Mean affective forecasts for: Outcome A - lost 100 euros, Outcome B - lost 50 euros, Outcome C - gained 50 euros and Outcome D - gained 100 euros; and mean affective differences for consecutive pairs: AB, BC and CD.**

In Task 1.3, by analyzing the p-values referent to the tested consecutive pairs of outcomes, we can conclude that **JE mode participants reported statistically significant differences in levels of happiness for all consecutive pairs of presented outcomes, in accordance with our initial assumptions** that tests were expected to confirm statistically significant differences for all tested JE mode outcome pairs.
As with Task 1.2, Task 1.3 is composed by two pairs of consecutive outcomes differing quantitatively, AB and CD, and one pair of consecutive outcomes differing qualitatively, BC. **SE mode participants did not report statistically significant differences in affective forecasts between Outcomes A - lost 100 euros - and B - lost 50 euros, yet did report statistically significant differences in affective forecasts between Outcomes C - gained 50 euros - and D - gained 100 euros, failing to fully support our initial assumptions**, as quantitative differences between outcomes were predicted not to lead to statistically significant differences for affective forecasts of SE mode participants. Additionally, **SE mode participants did report statistically significant differences in affective forecasts for our qualitatively different pair of outcomes, BC, in accordance to our initial assumptions**, as qualitative differences between outcomes were predicted to lead to statistically significant differences for affective forecasts of SE mode participants.

According to Distinction Bias literature concerning qualitative and quantitative differences, we anticipated that JE mode participants would tend to overpredict affective differences between outcomes with only quantitative variation, when compared to SE mode participants.
**JE mode participants reported lower affective forecast differences** between Task 1.3's Outcome B - lost 50 euros - and Outcome C - gained 50 euros, when compared to SE mode participants. **This result supports our assumptions, as Outcomes B and C differ only qualitatively.** Additionally, **JE mode participants also reported lower affective forecast differences** between Task 1.3's Outcome C - gained 50 euros - and Outcome D - gained 100 euros, when compared to SE mode participants. **This result goes against our assumptions, as Outcomes C and D differ only quantitatively.**
**JE mode participants reported higher affective forecast differences** between Task 1.3's Outcome A - lost 100 euros - and Outcome B - lost 50 euros, when compared to SE mode participants. **This result supports our assumptions, as Outcomes A and B differ only quantitatively.**

### *Task 1.4*

For Task 1.4, one additional factor was at play, the ability to independently evaluate an attribute. Hsee and Zhang[7] discuss that biased behaviour consistent with Distinction Bias applies mostly to attributes that are not too easy to evaluate independently, yet believe that finding an attribute that is notably easy to evaluate independently is the exception rather than the rule. The authors state school grades as an example of an attribute that people have sufficient knowledge about, resulting in SE mode happiness level predictions that may more closely match the evolution of JE mode ones. We took these statements as inspiration for Task 1.4, detailing a scenario and outcomes centered around this attribute, striving to answer **RQ2**.

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

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

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

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

### *Added Factors*

With our second experiment tasks, 2.1, 2.2 and 2.3, we aimed to further explore the ways in which Distinction Bias manifests itself in JE mode for quantitatively different outcomes.
As previously detailed, only the three anchor outcomes presented for participants of every condition were taken into consideration for testing (denominated Outcome B, D, and F). In order to examine the effects of the quantity of presented outcomes and the vertical axis scale on affective forecast differences, we conducted two-way ANOVAs for affective differences between consecutive outcome pairs (BD and DF) for each of our second experiment tasks.

**Quantity of Presented Outcomes:** We will firstly focus on our results on the influence of the levels of the between-subjects variable responsible for the amount of outcomes being displayed to participants (3 presented outcomes, 5 presented outcomes or 7 presented outcomes) in our participants' affective differences.
Our statistically significant pairwise comparison findings, while limited, exhibited a linear decrease of mean affective differences as levels progressed. In Task 2.1, for participants interacting with the smaller vertical scale, the mean affective difference between Outcomes B and D was: 1.750 (95% CI, 3.125 to 0.375) lower for participants being presented with seven outcomes than for the ones being presented with three outcomes, and; 1.806 (95% CI, 3.214 to 0.397) lower for participants being presented with seven outcomes than for the ones being presented with five outcomes. Moreover, in Task 2.3, for participants interacting with the medium vertical scale, the mean affective difference between Outcomes B and D was 1.067 (95% CI, 1.860 to 0.274) lower for participants being presented with five outcomes than for the ones being presented with three outcomes.

Overall, while **our findings are not substantial to support H2**, results do manifest an interesting trend. Throughout pairwise comparisons we can observe that our **few statistically significant results show a trajectory in which participants being presented with increasingly higher amounts of outcomes had a tendency to report smaller affective differences regarding the same studied anchor outcomes**. As we were dealing with exclusively quantitatively different outcomes, these findings raise some interest on further studying the premise that, in practise, comparing higher quantities of alternatives could be somewhat linked to a decrease in the measured Distinction Bias effects.

**Scale of Presented Outcomes:** In order to explore the influence of scale on Distinction Bias effects, we examined possible effects of the levels of the between-subjects variable responsible for the variation of the vertical axis range (smaller axis range, medium axis range or larger axis range) on our participants' affective differences.

Likewise, our results were limited and failed to highlight a consistently substantial and overall linear trajectory for affective differences throughout levels. Altogether, statistically significant pairwise comparisons were observed to show correlations in different directions. It is also important to note that, additionally, no consistent repeating patterns regarding the trajectory of affective differences per groups were identified for the three tasks tested. As such, **our findings were neither substantial or consistent, as our few statistically significant results do not present a clear direction regarding the importance of the scale of presented data for Distinction Bias theory, failing to considerably support H3**. Nonetheless, further study on the influence of this factor is incentivized.

## Neuroticism Effects

Given literature findings we introduced participant personality research onto our study, aiming to answer **RQ3**, testing possible correlations between participants' neuroticism scores and the affective differences that characterize Distinction Bias.
As biased behavior takes place in JE mode, we focused this part of our study on it. To test **H4** we ran Spearman Correlation tests between affective differences reported throughout our tasks and JE mode participants' neuroticism scores, applying the suitable Bonferroni corrections for analysis.

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

Additionally, we also tested for correlations between neuroticism scores and actual affective forecasts reported by JE mode participants, as personality theory denotes that neuroticism is characterized by a tendency to easily experience negative emotions and to be more pessimistic. Likewise, no statistically significant correlations were acknowledged.

### *User Interaction Metrics*

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

**Completion Times:** We performed Spearman correlation tests in order to examine possible links of participants' neuroticism scores and task completion times, once more considering the suitable Bonferroni corrections.

**For Task 1.1 results identified a statistically significant negative correlation between completion times and participants' neuroticism scores**, $r_s(29) = -.543$, $p = .002$, and furthermore there were also statistically significant negative correlations between completion times and participants' scores for two neuroticism facet-level traits, them being N3 - Depression ($r_s(29) = -.542$, $p = .002$), as well as N6 - Vulnerability ($r_s(29) = -.596$, $p < .001$). This could be due to Task 1.1 perhaps being considered the most straightforward from our first experiment, as well as being the only task presenting exclusively quantitative differences between outcomes. These conditions could have promoted the consequences that some literature studies have already associated with highly neurotic individuals. However, **for the remaining tasks results showed no statistically significant correlations between completion times and participants' neuroticism scores**. Additionally, even though no other statistically significant correlations were identified for any other of our tasks, **overall correlation coefficients consistently pointed to negative values**. As such, even though we deem our results to be insufficient to definitively answer **H5**, they express some evidence contributing to the premise that neuroticism may exhibit negative correlations with completion times, incentivizing further research on the topic.

**Quantity of Hovers:** Likewise, we performed Spearman correlation tests in order to examine the possible links of participants' neuroticism scores and the overall quantity of hovers performed by participants for each of the tasks, considering the suitable Bonferroni corrections.
Overall, **results did not show statistically significant correlations between the quantity of hovers performed and participants' neuroticism scores. These findings failed to show support for H6**.

## CONCLUSION

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

Overall, although a few of the examined outcomes did not elicit the expected consequences, **results hint to the effects of Distinction Bias**, exhibiting how, for merely quantitatively different outcomes, JE mode participants tended to predict a statistically significantly steep evaluation of values, as opposed to SE mode participants, that predicted values that tended to not show statistically significant differences in these contexts. Moreover, findings highlighted a moderate tendency for overprediction of affective differences when outcomes exhibited merely quantitative differences in JE mode in comparison with SE mode, as well as a lack of the same overprediction when outcomes differed only qualitatively. It is the junction of these findings that characterizes Distinction Bias.

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

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

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

### Limitations and Future Work

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

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

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

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

## REFERENCES


[1] Melanie Bancilhon and Alvitta Ottley. 2020. Did You Get The Gist Of It? Understanding How Visualization Impacts Decision-Making. (10 2020).

[2] Nuran Bayram and Mine Aydemir. 2017. Decision-Making Styles and Personality Traits. (2017). Issue 1. https://www.researchgate.net/publication/330832648

[3] Eli T Brown, Alvitta Ottley, Helen Zhao, Quan Lin, Richard Souvenir, Alex Endert, and Remco Chang. 2014. Finding Waldo: Learning about Users from their Interactions. *IEEE Transactions on Visualization and Computer Graphics* 20 (12 2014), 1663–1672. Issue 12. DOI:http://dx.doi.org/10.1109/TVCG.2014.2346575

[4] P. T. Costa and R. R. McCrae. 2000. NEO PI-R: Inventário de Personalidade NEO revisto: Manual Profissional. (2000).

[5] Evanthia Dimara, Steven Franconeri, Catherine Plaisant, Anastasia Bezerianos, and Pierre Dragicevic. 2020. A Task-Based Taxonomy of Cognitive Biases for Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 26 (2 2020), 1413–1432. Issue 2. DOI: http://dx.doi.org/10.1109/TVCG.2018.2872577

[6] Tear Marie Green and Brian Fisher. 2010. Towards the Personal Equation of Interaction: The impact of personality factors on visual analytics interface interaction. *2010 IEEE Symposium on Visual Analytics Science and Technology* (10 2010), 203–210. DOI: http://dx.doi.org/10.1109/VAST.2010.5653587

[7] Christopher K. Hsee and Jiao Zhang. 2004. Distinction Bias: Misprediction and Mischoice Due to Joint Evaluation. *Journal of Personality and Social Psychology* 86 (5 2004), 680–695. Issue 5. DOI: http://dx.doi.org/10.1037/0022-3514.86.5.680

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

[9] Daniel A Keim, Florian Mansmann, Andreas Stoffel, and Hartmut Ziegler. 2008. Visual Analytics. (2008). http://infovis.uni-konstanz.de

[10] M. Lima and A. Simões. 2000. *Neo pi-r: Manual profissional*. Lisboa: CEGOC.

[11] Zhengliang Liu, R. Jordan Crouser, and Alvitta Ottley. 2020. Survey on Individual Differences in Visualization. *Computer Graphics Forum* 39 (6 2020), 693–712. Issue 3. DOI:http://dx.doi.org/10.1111/cgf.14033

[12] Robert R. McCrae and Paul T. Costa. 1987. Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology* 52 (1987), 81–90. Issue 1. DOI: http://dx.doi.org/10.1037/0022-3514.52.1.81

[13] Tamara Munzner. 2014. *Visualization Analysis Design*. CRC Press.

[14] Julia Müller and Christiane Schwieren. 2020. Big Five personality factors in the Trust Game. *Journal of Business Economics* 90 (2 2020), 37–55. Issue 1. DOI: http://dx.doi.org/10.1007/s11573-019-00928-3

[15] Kenji Noguchi, Carol L. Gohm, and David J. Dalsky. 2006. Cognitive tendencies of focusing on positive and negative information. *Journal of Research in Personality* 40 (12 2006), 891–910. Issue 6. DOI: http://dx.doi.org/10.1016/j.jrp.2005.09.008

[16] Othman Radwan El, Rola El Othman, Rabih Hallit, Sahar Obeid, and Souheil Hallit. 2020. Personality traits, emotional intelligence and decision-making styles in Lebanese universities medical students. *BMC Psychology* 8 (12 2020), 46. Issue 1. DOI: http://dx.doi.org/10.1186/s40359-020-00406-4

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

[18] Ryszard Tadeusiewicz and Adrian Horzyk. 2014. Man-Machine Interaction Improvement by Means of Automatic Human Personality Identification. (2014). DOI:http://dx.doi.org/10.1007/978-3-662-45237-0_27

[19] André Calero Valdez, Martina Ziefle, and Michael Sedlmair. 2018. Studying Biases in Visualization Research: Framework and Methods. (2018). DOI: http://dx.doi.org/10.1007/978-3-319-95831-6_2

[20] Emily Wall, Leslie M. Blaha, Lyndsey Franklin, and Alex Endert. 2017. Warning, Bias May Occur: A Proposed Approach to Detecting Cognitive Bias in Interactive Visual Analytics. *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)* (10 2017), 104–115. DOI: http://dx.doi.org/10.1109/VAST.2017.8585669

[21] D.G. Williams. 1992. Dispositional optimism, neuroticism, and extraversion. *Personality and Individual Differences* 13 (4 1992), 475–477. Issue 4. DOI:http://dx.doi.org/10.1016/0191-8869(92)90076-2

[22] Caroline Ziemkiewicz, Alvitta Ottley, R. Jordan Crouser, Ashley Rye Yauilla, Sara L. Su, William Ribarsky, and Remco Chang. 2013. How visualization layout relates to locus of control and other personality factors. *IEEE Transactions on Visualization and Computer Graphics* 19 (2013), 1109–1121. Issue 7. DOI: http://dx.doi.org/10.1109/TVCG.2012.180