Computers & Graphics 111 (2023) 47–62

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical Section

Exploring the role of conscientiousness on visualization-supported decision-making[☆,☆☆]

Tomás Alves [a,b,*], Tiago Delgado [a], Joana Henriques-Calado [c], Daniel Gonçalves [a,b], Sandra Gama [a,b]

[a] Instituto Superior Técnico, University of Lisbon, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal
[b] INESC-ID, R. Alves Redol 9, 1000-029 Lisboa, Portugal
[c] CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Lisboa, Portugal

ARTICLE INFO

Article history:
Received 10 September 2022
Received in revised form 16 January 2023
Accepted 16 January 2023
Available online 19 January 2023

Keywords:
Visualization
Decision-making
Conscientiousness
Personality

ABSTRACT

Understanding how visualizations can support decision-making continues to be one of the most relevant challenges in current research. However, prior work provides limited knowledge regarding how the decision-maker profile and, in particular, psychological constructs affect decision-making. Weighing how conscientiousness affects one's tendency to follow the rules and prioritize tasks, this work explores if this personality trait plays a role in visualization-supported decision-making. We asked participants to perform a series of multi-attribute choices using visualizations of high-dimensional data. Further, we study if the quality feedback of the past choice affects the decision-making process. Our results suggest that the feedback quality of past decisions affects how much individuals invest in data analysis and decision-making. In addition, user confidence is affected by conscientiousness scores, resulting in conscientious individuals changing their choices more often. Our findings provide new insights into the research challenges of analyzing decision-making in a visualization setting.

1. Introduction

Past studies show how visualization helps users discover, explain, and form decisions based on the information conveyed [1]. Several strategies may aid visualization-supported decision-making, e.g., offering sufficient guidance [2] or emphasizing critical information [3]. However, recently Dimara and Stasko [4] highlighted the substantial need for user studies that contain visualization-supported decision-making tasks. The core objective of this research agenda is to understand how information visualization (InfoVis) can support decision-making. In particular, it is imperative to study how the decision-maker selects the "best" solution between the alternative solutions to a problem [5,6]. These choices can have distinct complexity levels, spanning from a person deciding which car to buy to which energy plan is more friendly to the climate.

Visualization research shows that decision tasks are vulnerable to cognitive biases and uncertainty [7,8]. Moreover, a recent survey highlights that human perception can be affected by several other factors, e.g., cognitive and personality traits[1] (see Liu et al. [10]). The manifestation of individual differences may hinder the cognitive flow of reasoning of the user and, consequently, how the user derives a choice of direction as an outcome. Nevertheless, there is still limited knowledge regarding how the decision-maker profile and, in particular, psychological constructs affect decision-making [4].

Inspired by these findings, we decided to expand the user profile in visualization-supported decision-making and include personality characteristics in our analysis. A user profile is a set of attributes associated with a user such as education level, age range, goals, or psychological factors (e.g., personality and cognitive traits [10]). User profiles can be used as a digital identity of an individual and represent a user model [11,12]. In particular, we focus on conscientiousness since this trait measures the preference for an organized approach to life than a spontaneous one [13]. Weighing this predisposition, we believe that conscientiousness can bias user perception and interaction and,

☆ This article was recommended for publication by K Buehler.
☆☆ This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT), Portugal with references SFRH/BD/144798/2019 and UIDB/50021/2020.
* Corresponding author at: INESC-ID, R. Alves Redol 9, 1000-029 Lisboa, Portugal.
E-mail address: tomas.alves@tecnico.ulisboa.pt (T. Alves).

[1] Allport [9] first defined personality traits as generalized and personalized determining tendencies, consistent and stable modes of an individual's adjustment to his environment. Furthermore, the author built a vast lexical collection of adjectives that describe these traits. In Personality Psychology research, the traits are often factors measured through questionnaires and associated with a score.

https://doi.org/10.1016/j.cag.2023.01.010
0097-8493/

consequently, influence decision-making. This paper presents an explorative user study to *understand whether conscientiousness provides dispositional information that can help us* improve the user profile in visualization-based decision-making.

We designed and implemented a system that leverages high-dimensional data visualization and prompts the user to play a decision-making game. In addition, we analyze how the feedback that the user receives from their choices affects their decision-making. To do so, we ran two experiments. In the first experiment, the game provides fixed feedback to the user regarding the quality of their past choice, i.e., the reward that users receive is independent of their decision. In the second one, the game reports the real output of their choices. Results suggest that the feedback users receive may affect how invested they are in data analysis and decision-making. Further, conscientiousness scores correlate with how confident users were in their choices under the studied conditions. In addition, those with higher scores tend to change their decision across trials more often than their counterparts. Our findings shed new light on incorporating personality traits into the design pipeline of visualization-supported decision-making systems. By designing tools that have a deeper understanding of the target user, we believe it is possible to enhance the development of automatic visualizations that design flexible data visualization tools (e.g., Mackinlay et al. [14], Wills and Wilkinson [15], Cunha et al. [16]) and, consequently, to better support individual users.

2. Related work

We present the underlying related work in this literature review from three connected scopes. First, we discuss state-of-the-art research in decision-making using visualization systems. Next, we cover how past visualization research on human perception has leveraged high-dimensional data. Finally, we explain how past research considers personality constructs to enhance the user profile.

2.1. Decision-making in visualization

We consider that a decision-making task includes an "explicit intent to derive an ultimate choice of direction as an outcome" [4]. Past research identifies decision-making as a task that visualization can or should assist [17,18]. Nevertheless, there is no further elaboration on what decision-making actually is in the visualization context [4]. Additionally, most visual analytic taxonomies do not present decision-making as an explicit task [4]. Therefore, the current research agenda focuses on clarifying decision tasks and validating design guidelines to aid visualization-based decision-making.

This work focuses on *multi-attribute choice task*, a term introduced by Dimara et al. [19] and defined as "finding the best alternative among a fixed set of alternatives, where alternatives are defined across several attributes". This term stems from multi-criteria decision-making, a discipline that analyses procedures to aid decision-making in areas such as business intelligence and finance [20,21]. In particular, multiobjective optimization tackles looking for solutions to optimization problems with multiple conflicting objectives. Recent work by Filipič and Tušar [22] highlights the variety of graphical encodings supporting Pareto front approximations through visualization. For instance, Dimara et al. [19] benchmarked three chart types (parallel coordinates plot (PCP), scatterplot matrix (SPM), and tabular visualizations) in terms of decision support and subjective metrics. Results showed that all visualizations support multi-attribute choice tasks at a similar rate, although the tabular visualization provided a slight advantage in decision time. There are more studies in visualization that use decision-making tasks such as the EZChooser [23] or the Dust & Magnet [24]. However, to the best of our knowledge, no studies consider objective metrics of decision quality or visualizations as a basis of comparison besides Dimara et al. [19].

Given these prospects, we continued our literature review by focusing on how to evaluate decision-making. Stemming from research on uncertainty and cognitive biases in visualization (e.g., Dimara et al. [7], Hullman et al. [25], Padilla et al. [26]), studies often leverage the speed of decision [27], user predictions [8,28], or user assessments such as fairness [29] to understand the decision process. Dimara et al. [19] also considered user accuracy and decision time as objective metrics and technique preference and choice assessment as subjective metrics. Nonetheless, the authors conclude that there is a need for additional work to establish more sensitive metrics of choice quality while considering non-preference-based choices. Moreover, there is still limited knowledge to support decision-making with visualized data after the user builds on the knowledge [30]. This work builds on the mentioned studies, leveraging user interaction and predictions to investigate the decision-making supported by high-dimensional data.

2.2. Visualizing high-dimensional data

There are several techniques to visualize multi-dimensional data [31]. Some approaches use dimensionality reduction (DR) methods to generate an observable overview by collapsing multiple dimensions into a smaller number of subdimensions [32,33]. Research has evaluated DR in terms of class separability [34] or mapping mechanisms and layout performances [35]. Nevertheless, DR leads to the exclusion of raw values during the reduction process, which is critical in multi-attribute choice tasks since the user must be able to read attribute values directly [19].

In contrast, lossless visualizations use simple visualization encodings and traditional multi-dimensional visualization techniques [36]. Several studies have evaluated the quality of lossless charts such as PCP, SPM, and tabular visualizations. Holten and Van Wijk [37] focused on PCP-based cluster identification in terms of time and correctness performances, while Li et al. [38] studied the effectiveness of PCP and SPM for correlation judgment. Dimara et al. [19] also focused on lossless techniques (PCP, SPM, and tabular visualizations), while Zhao et al. [39] benchmarked lossy and lossless methods (PCP, SPM, principal component analysis, and radial coordinate visualizations). Other non-geometric visualization techniques use icons or glyphs [40], pixels [31], or hierarchical- and graph-based structures [31]. Hybrid techniques combining multiple visualizations also exist [41], although the effectiveness of the possible combinations offers limited knowledge to practitioners.

For this work, we leverage only lossless visualizations to support directly multi-attribute choice tasks. In particular, we use the parallel coordinates plot (PCP) and the scatterplot matrix (SPM). Both charts are general-purpose multidimensional visualization tools extensively studied in state-of-the-art research (e.g., Zhao et al. [39], Holten and Van Wijk [37], Li et al. [38], Dimara et al. [19]). The literature shows that scatterplots are better than bivariate PCP for correlation tasks [38]. In cluster-oriented tasks, combining PCP with scatterplots outperforms PCP alone [37], and PCP seems to be better than SPM in this setting [39]. Moreover, Kuang et al. [42] found that a simplification form of SPM could outperform PCP by showing only a subset of the plots. Based on all the mentioned work, we believe that the PCP and SPM are generic and appropriate enough to support a visualization-based decision-making process [41]. However, previous research did not account for the impact that individual differences such as personality may have on the interaction process with a visualization.

2.3. Personality in visualization

Recent research in user modeling [43,44] and adaptive visualization systems [45] uses personality data to improve user interaction. In particular, these studies often try to understand the role of personality through the lens of the Locus of Control (LoC) [46] and the Five-Factor Model (FFM) [47,48]. Regarding the LoC, this personality trait captures what people believe contribute the most to the rewards they receive from the environment. In particular, individuals can believe that the rewards they receive from the environment are more likely explained by their actions (internal LoC) or by external entities such as powerful others or chance (external LoC). Past research found that this personality trait affects search performance across hierarchical [49], time series [50], and item comparison [51] visualization designs, visualization use [52,53], and behavioral patterns [44].

The FFM defines personality through five core traits: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. State-of-the-art visualization research found that neuroticism, extraversion, and openness to experience affect task performance [53–55]. Additionally, neuroticism makes individuals less likely to be deceived by spurious correlations [54]. The mentioned studies collectively suggest that personality traits offer an opportunity to expand the user profile [6], i.e., include more attributes in the user model representation in visualization-based human–computer interaction settings. Nevertheless, there is a set of untapped traits that can offer new insights regarding the role of personality in user interaction. In particular, no studies include reported measurable effects of agreeableness or conscientiousness on visualization settings [10].

We believe that conscientiousness may play a significant role in visualization-supported decision-making. High conscientiousness leads individuals towards efficiency, while low scores on the trait result in people generally being unreliable and prone to going with the flow [56]. Past psychology research shows that conscientiousness plays a role in rational decision-making style [57]. Moreover, conscientious individuals have better perceptual-questioning skills that facilitate effective adaptation [58,59]. In contrast, mixed results are reporting whether conscientiousness affects how competent one is in decision-making, i.e., if one makes "good" decisions. Some studies are not able to find measurable effects regarding this relationship (e.g., Dewberry et al. [60]), while others find a negative (e.g., LePine et al. [61]) or positive (e.g., LePine et al. [62],Ones and Viswesvaran [63]) influence of the trait scores on decision-making competence. However, Svendsen et al. [64] claim that conscientious individuals are more likely to examine ways to use technology that would enable them to enhance their performance in their job [65]. Based on these prospects, we believe that technology and, more specifically, visualization may provide a medium to verify the manifestation of conscientiousness in user decision-making. In particular, we expect conscientiousness to predict decision-making competence since conscientious decision-makers are likely to consider decisions more carefully and thoroughly [13].

3. Methodology

Our objective was to analyze the visualization-supported decision-making process through the scope of personality psychology. In particular, this study aimed to verify how conscientiousness manifests its characteristics in the process based on the output of the decisions. We designed our experimental procedure based on Dimara et al. [19]. In their experiment, the authors asked users to perform decision tasks supported by parallel coordinates (PCP), scatterplot matrix (SPM), and tabular charts. Before the decision tasks phase, participants needed to complete a training composed of analytical tasks to ensure that they had all the necessary technical knowledge to read and interpret the multi-dimensional visualizations. Similarly, we started by asking the participants to complete basic analytical tasks focused on reading and interacting with two multi-dimensional visualizations: the PCP and the SPM. We conducted this training before the decision phase of the experiment to ensure that users understand the visualization and interaction techniques present in the dashboard. In addition, it counters potential confounding factors that may impact the decision-making process. Next, participants had to perform a series of multi-attribute choice tasks. Since the decision process typically involves learning over time, we decided to expand the study of the decision-making process by analyzing the effect of the feedback regarding the quality of the decision.

We decided to run two versions of the same experiment that vary solely in the feedback that users receive from their decisions. In the first experiment, the reward that users received in each trial is predetermined and independent of their decision. In contrast, the second experiment reports realistic feedback to the user. This factor helped us analyze if people with different conscientiousness scores varied their decision-making process by having obtained adequate feedback on their previous choices. Our work extends the work of Dimara et al. [19] by introducing the conscientiousness trait in the analysis of visualization-supported decision-making. According to the American Psychological Association, a personality trait is "a relatively stable, consistent, and enduring internal characteristic that is inferred from a pattern of behaviors, attitudes, feelings, and habits in the individual".[2] Conscientiousness is the tendency to display self-discipline, act dutifully, and strive for achievement against measures or outside expectations [66]. Individuals with high conscientiousness scores prefer planned rather than spontaneous behavior, and low scores are associated with flexibility and spontaneity. Based on these trends, we believe conscientiousness may help explain the decision-making process.

3.1. Research questions

Before data collection, we created two research questions. First, we are interested in observing if user decision behavior depends on the *decision quality* of past choices, i.e., how "good" the past decision was. We addressed the dependency of subsequent decisions on previous decisions to study if a decision-maker may be biased by experiences of a series of former decisions (e.g., Koch et al. [67]). In particular, we believed that the quality of the last choice may affect how individuals decided afterward and explored the data visualizations for alternatives. We expected that the feedback on the former decision may provide further insights regarding how path-dependent the decision-maker is when supported by visualization. For instance, providing negative feedback may lead individuals to spend more time or be less confident when they make the next decision. In contrast, positive feedback may promote inertia in the decision-maker and increase the likelihood of sticking with former decisions despite alternatives being preferable. To our knowledge, there is limited knowledge regarding how feedback on former decisions affects visualization-supported decision-making. As such, our first research question was:

Q1 *Does the visualization-supported decision-making process depend on the quality of the last decision?*

[2] https://dictionary.apa.org/personality-trait

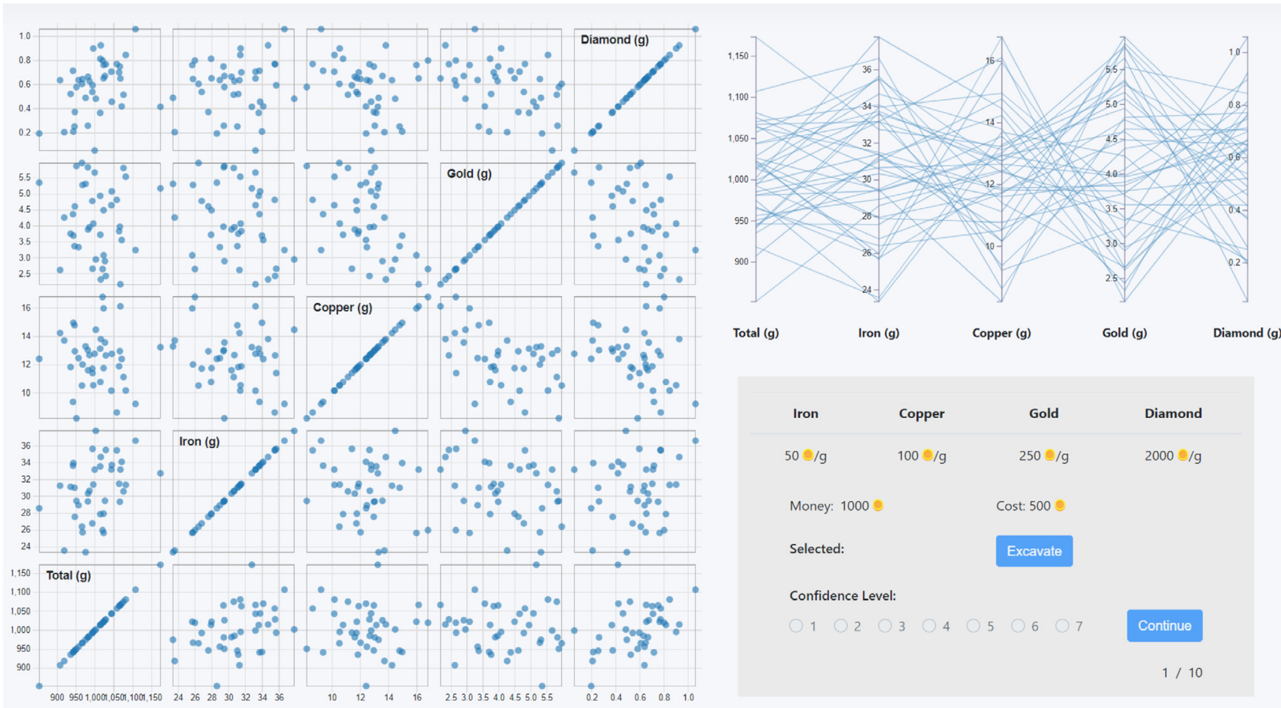


Fig. 1. Dashboard used in the experiment. The left area of the screen has a scatterplot matrix with all correlation pairs. The right part of the screen has a parallel coordinates plot on top, and the bottom half has the controller area.

Moreover, we wanted to understand whether conscientiousness plays a role in the susceptibility to the quality assessment of the last choice and, consequently, affects how users decide later. Since individuals with high scores tend toward efficiency [56], we believed their decisions will likely weigh how good their past choice was. Therefore, we expected conscientious individuals will be more likely to be affected by the quality of their last decision quality on subsequent decisions. In contrast, individuals with lower scores would not pay much attention to their decision quality as they are more easy-going and spontaneous [13]. More specifically, we wanted to study the following question:

Q2 *Does conscientiousness manifest its effects in visualization-supported decision-making?*

3.2. Tasks

We used two task types in the experiment.

Training Tasks For the training, we asked participants to complete a series of analytical task trials individually with the PCP or the SPM. The analytical task was to find the item with either the maximum or the minimum value for the sum of each attribute among all items of the dataset. The objective of the training was to assure that participants understood how to read and interact with either visualization.

Decision Tasks The decision consisted of multi-attribute choices [19], i.e., all alternatives are known in advance and defined across a set of attributes. We framed the decision-making process as a serious game [68] to promote participation engagement, awareness, and understanding of the underlying topic [69,70].

3.3. Visualizations

Researchers develop visualizations to guide users in their decisions through direct (e.g., Willett et al. [71]) and indirect (e.g., Endert et al. [72]) guidance. Further, these visualizations should allow data exploration and help users better understand the information they use in their decisions. Similar to Dimara et al. [19], we leverage standard visualization methods to support decision-making without prior training in decision analysis. We examined two multi-dimensional visualization techniques: the parallel coordinates plot (PCP) and the scatterplot matrix (SPM). Both techniques are lossless [31]. The *parallel coordinates plot* is a polylines diagram where the parallel axes represent dimensions and each item is a polyline that intersects the axes at their corresponding values [73]. Regarding the *scatterplot matrix*, we used the complete grid of scatterplots showing "the bivariate relationships between all pairs of variables in a multivariate data set" [74]. Several visualization textbooks and surveys consider these representations to be the standard versions [36,75]. We kept the design as consistent as possible across the techniques to facilitate item comparison between them. Fig. 1 presents the dashboard containing the visualizations. All visual marks had the same color across all techniques (translucent blue by default or translucent gray when outside a range selection). Axes and fonts were displayed consistently in gray or black. Finally, we developed the visualization dashboard with the d3 library.[3]

3.4. Interaction techniques

Based on previous work [19], we supported two types of interactions that have proven beneficial in decision-making tools. Users could *highlight* individual cases through click and hover events to support value retrieval [76,77]. Both events were supported through *linking* and *detail-on-demand* techniques to help individuals relate items across dimensions and views [75,78]. When users clicked on an item, it changed from blue to orange in all plots. Moreover, hovering a data case changed the item opacity across all plots from the default 50% to 100% and showed a tooltip with the identifier. We also implemented *range selection* to

[3] https://d3js.org/

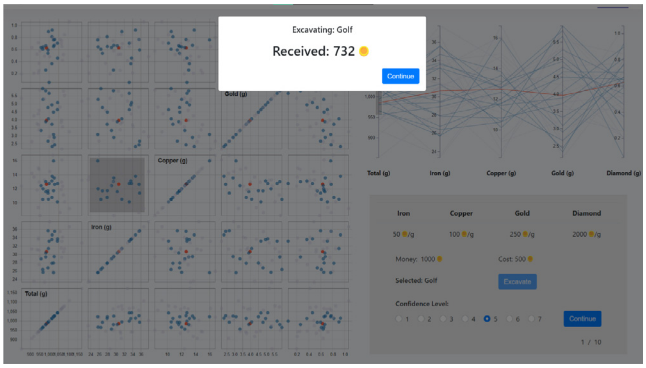


Fig. 2. Instance of the screen after an excavation with the financial return. The screen shows that the excavated site was `Golf` which marks have a red color in the visualizations. It is also possible to observe that the user applied range selection in both visualizations and rated their confidence in choosing `Golf` for that trial.

support dynamic filtering and queries [77,79] (Fig. 2). Users could brush axes which draw selection rectangles on top of the axis. Individuals could select a range across multiple dimensions, and the dashboard only showed the intersection between the conditions. It also grayed out data outside the selection. Notably, brushing on an axis of the PCP only filtered through one dimension, while doing so on the scatterplot selected two ranges simultaneously (one for each scatterplot dimension). Range selections were drag-able and re-sizable through handles that appear on hover.

3.5. Dataset

The synthetic datasets used in the experiment were about soil composition since they had properties amenable to our elicitation techniques. In particular, soil composition was a familiar enough concept that can be decomposed into several types of minerals. Visualizations for multidimensional data could then use this large number of materials as variables and soil samples as data items. Therefore, it allowed participants to understand what each item represents while the distribution is unknown, prompting an exploration. We developed two datasets for our experiment.

Training Tasks For the training, we created a dataset detailing the soil composition of several fictional desert locations. We named each desert according to the international phonetic alphabet. Moreover, the composition of each desert was composed of six continuous variables representing different materials (Clay, Silt, Sand, Gravel, Pebble, and Rock). We generated distributions using a randomly generated covariance matrix with the NumPy Library.[4] We randomly attributed an average value between 80 and 160 to each variable.

Decision Tasks We created the dataset using Python with the Pandas and NumPy libraries. Using NumPy, we generated values using a covariance matrix (see supplemental material) for the variables diamond, gold, copper, iron, and total, with averages of 0.6, 4, 12, 36, and 1000, respectively. We obtained these values empirically to ensure that each site was relatively balanced and, at the same time, the variables did not contain any extreme outliers. In particular, larger sites should have more valuable materials (diamond, gold, copper, iron), and those with a big amount of one material should have lesser amounts of the rest, on average.

4 https://numpy.org/

Table 1
Financial return of the last decision (in coins) of each trial. Color encodes whether the last decision yielded profits (green) or losses (red) compared to the excavation cost (500 coins).

Trial	1	2	3	4	5	6	7	8	9	10
Last decision financial return	NA	732	456	796	245	454	279	87	622	632

3.6. The excavation game

The dashboard domain was an excavation game where the objective was *to maximize the amount of money* the user had across ten consecutive trials. Each user had an initial budget of 1000 coins and had to choose a site (data item) to excavate in each trial. The cost of excavating a site was 500 coins. When the user excavated a site, the dashboard simulated that the user collected 100gr of soil and then calculated the financial return of the precious materials (`iron`: 50 coins/gr, `copper`: 100 coins/gr, `gold`: 250 coins/gr, and `diamonds`: 2000 coins/gr) present in the soil sample. The financial return was added instantly to the current amount of coins.

The process of playing the game consisted of three steps. First, users had to **select** which site they wanted to excavate. The dashboard contained a controller area on the right corner where users could check the financial return per gram of excavated material and the cost of excavating. Each mark in the visualizations (Fig. 1) represented a site. A site could be selected by clicking on the mark that represented that site. When users clicked on a site, its name was displayed on the controller area in the right corner of the screen (Fig. 1, after `Selected:`). Second, users had to **rate** how confident they were about the choice of excavation site ranging from *not confident* (1) to *extremely confident* (7). Third, the user could click on the action button `Excavate`. When the user **excavated** a site, the dashboard instantly added another sample of 100gr to that site following the distribution of the dataset and, consequently, not changing the data displayed. After clicking the button, the dashboard locked the selection and stopped the timer. Then, an alert appeared on the screen indicating the financial return of the excavation (Fig. 2). Finally, users clicked on `Continue` to progress to the following trial. Users could select the same excavation site between trials or change it. Users could also check anytime on the controller area their current amount of money and the number of trials completed so far.

As we mentioned, we ran two experiments to better understand the manifestation of conscientiousness in the decision-making process. In one experiment, we fix the financial return in each trial. In particular, we followed the distribution depicted in Table 1. We empirically designed the distribution to alternate between profits and losses along with the trials. It consecutively alternated between profits and losses in the first half of the trials (Trials 1 to 4) to foster data exploration and give a sense of uncertainty. Then, the returns showed significant losses up to Trial 7 to see how users reacted to constant negative feedback. Finally, the remaining trials (Trials 8 and 9) yielded positive financial returns in contrast to the past negative pattern.

A limitation of fixing the returns path is that it disconnects the user choice and the reward, i.e. users do not see a direct consequence of their choice but rather a fixed value. To minimize this effect, we informed the user before the trials that the precious materials were randomly distributed in the excavation sites and, consequently, the values depicted in the charts estimated the quantities of each material. Further, we explained that, when the user excavated a site, the consequence was that the quantity of each precious material did not directly yield the values displayed on the charts. This explicit dissonance helped the participants accept that there was room for some disconnection between action and reward. We ran pilot tests with a think-aloud protocol to understand whether users understood that we fixed the distribution. Users did not acknowledge it. We also observed in the actual experiment that users did not realize that the returns were fixed. Consequently, we believe that this approach helped tackle the dissonance between the choice and the reward that users received.

In the second experiment, the Excavation Game outputs the real returns from the choice. The returns are calculated based on the sum of the price per gram of each material times the estimated quantity of that material on the excavation site. Contrary to the first experiment, this time there is no dissonance between the choice and the reward that users received. It results in realistic feedback and the return values correspond to the distribution on which the visual information is based. We believe that tackling both approaches allows for a better understanding of the decision-making and, in particular, how conscientiousness plays a role in this process.

3.7. Measures

Objective metrics We measured the quality of the decision process through multiple metrics for each trial. We measured the **task decision time** in seconds from the moment users pressed `Continue`, to when participants pressed `Excavate` to confirm their site choice. To better understand how users explored the visualizations, we collected the **number of events triggered in each chart**. The events we tracked were clicks, hovers, and brushes. They were automatically collected when triggered by the user. We measured clicks and hovers per item when the user clicks or hovers it, respectively. Brush events happened on the axes of the charts or scatterplots when the user attempted to filter some items by clicking on an initial position, dragging the mouse to a final position, and releasing the click. Finally, we analyzed user accuracy. As Dimara et al. [19] stated, accuracy has a core subjective aspect, making it challenging to calculate. We considered using Pareto dominance [80] as a measure. However, we build the dataset with enough alternatives and features to hinder a dominating choice. Further, the *perceived accuracy* is based on the fixed path of financial returns. Consequently, we decided to evaluate the **decision accuracy** based on how well estimated it was. In particular, we used multi-attribute utility theory to determine how adequate the choice is since it is possible to calculate it based on the values presented on the charts, the chosen excavation site, and the price of each precious material. Moreover, the quality of an excavation site was kept the same across a complete experiment since we do not change the distribution or the prices of the materials between trials. Therefore, we normalized the accuracy between 0 and 1, where 0 and 1 corresponded to the expected financial returns with minimal (i.e., the worst excavation site) and maximum values (i.e., the best excavation site), respectively. For instance, if sites A and B yield the best and worst returns, respectively, the score of site C is calculated as follows:

$$C\_score = 1 - \frac{A\_average\_returns - C\_average\_returns}{A\_average\_returns - B\_average\_returns} \tag{1}$$

Subjective metrics After choosing an excavation site, we asked users to rate how **confident** they were in their choice. This evaluation is assessed often in decision support tools [24,77]. It allowed us to complement the objective measures with a personal assessment of the choice quality.

Personality The FFM is the most widespread and generally accepted model of personality [48,81]. In particular, it provides a taxonomy and a conceptual framework that unifies much of the research findings in the psychology of individual differences. Personality data was collected with the Revised NEO Personality Inventory (NEO PI-R) [66,82]. It is the standard questionnaire measure of the FFM by allowing researchers to assess the five personality traits and their 30 facets. The NEO PI-R has a high internal consistency with values ranging from 0.79 to 0.86 [82]. We calculate **conscientiousness** scores by the sum of its Likert Scales based on assertions semantically connected to behaviors and five possible alternatives of agreement: *strongly agree*, *agree*, *undecided*, *disagree*, and *strongly disagree*.

Demographics We recorded the gender, age, and self-reported visual acuity and whether the participant was color-blind through Ishihara tests [83]. We also measured visualization literacy regarding PCP and SPM on a five-point Likert scale ranging from *not familiar* (1) to *very familiar* (7).

3.8. Procedure

As we mentioned, the procedure for both experiments was kept the same. The only difference between the experiments was whether the quality of feedback was controlled. We conducted each user test as a Zoom video meeting with one experimenter at a time due to constraints from COVID-19. Participants consented to have the screen recorded. All participants had normal or corrected-to-normal vision, and there were no color-blind subjects. We conducted a pilot study to ensure the clarity of the instructions and estimate the experiment length. Each experiment lasted on average 30 min and consisted of four phases: (i) demographic questionnaire, (ii) tutorial, (iii) training, and (iv) formal study. We started by asking the participant to fill out the NEO PI-R and the demographic questionnaire. We continued by showing an example of a PCP and SPM to collect visualization literacy. Then, we asked participants to complete the training trials.

Training Tasks Participants completed consecutive trials varying the chart (PCP or SPM), number of dimensions (four, five, or six), the dataset size (13, 20, or 26), and the task (*Which one is the biggest/smallest desert?*) in a total of 36 trials ($2 \times 3 \times 3 \times 2 = 36$). We randomly ordered the trials were a Latin square design to assure that participants interacted with each possible combination and, simultaneously, diminish the learning effects. We assured that participants had no questions regarding how to read and interact with the charts. If the participant was not able to answer correctly to the last ten trials, we asked them to repeat sets of ten trials until they answered correctly ten trials in a row. Otherwise, we continued with the decision tasks.

Decision Tasks First, we explained the context of the excavation game. In particular, we presented the right-bottom corner panel, including the financial returns of each material, the cost of excavating, how to choose and dig an excavation site, and how to assess their choice confidence. Participants were free to interact with the dashboard until they understood how to complete the decision task and progress to the following trial. When the participant reported being comfortable, we started the ten consecutive trials. In each trial, the participant needed to choose an excavation site, assess how confident they were, click on a button to receive the financial return of their excavation, and progress to the next trial. After completing the trials, participants received compensation for their time.

3.9. Research design and data analysis

In our user study, participants performed ten decision tasks in consecutive trials. In each trial, we collect task decision time, decision accuracy, user confidence, and the number of events triggered in each chart. Since there is no previous financial return

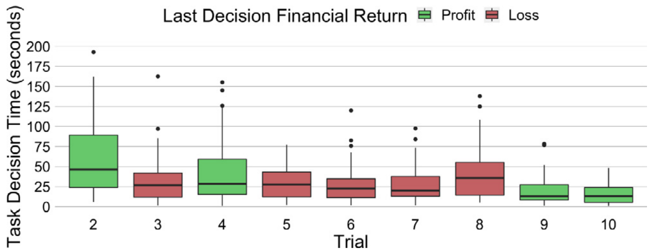


Fig. 3. Task decision time per trial with fixed returns.

on Trial 1, we ignore this trial in our analysis. The independent variable is the conscientiousness score of the user. We analyze, report, and interpret the effect of conscientiousness on task decision time, decision accuracy, and user confidence using Spearman's rank-order correlation tests [84]. We also perform pairwise comparisons between trials using Tukey HSD [85] tests including Bonferroni corrections. For the events triggered in the charts, we analyze the sum of all triggered events per trial, i.e., the sum of the number of hovers, click, and brush events in both charts. We use Spearman's rank-order correlation [84] and Wilcoxon signed-rank tests to study the effect of conscientiousness on the interaction metric. Finally, we study if conscientiousness made individuals change their decision across trials using point-biserial correlation coefficient tests [86] since changing the decision is a dichotomous variable and the conscientiousness score is a continuous one.

4. Results

This section presents the results of our study. We tackle the objective and subjective metrics to understand better how conscientiousness and returns affect the decision-making process. Data are presented as mean ± standard deviation unless otherwise stated.

4.1. Fixed returns experiment

We recruited 38 participants (21 male, 17 female) aged 21 – 53 ($M = 25.24, SD = 5.39$). Each test was 45:39 ± 20:02 (mm:ss) long with minimum and maximum durations of 15:32 and 1:28:32 (hh:mm:ss), respectively. Additionally, participants self-reported being fairly familiar with PCP ($M = 3.37, SD = 2.26$) and SPM ($M = 4.34, SD = 1.76$).

4.1.1. Task decision time

As depicted in Fig. 3, users usually took more time to decide in Trial 2 (57.42 ± 43.37 s) than in the rest of the experiment. As we mentioned, Trial 2 was the first trial after the users received feedback regarding their decision. It appears that the positive feedback led users to spend more time exploring the data and deciding carefully. In contrast, we observed a statistically significant decrease of 23.23 (95% CI, 1.743 to 44.716) s, $p = .023$, from Trial 2 to Trial 3. After receiving the negative financial return, users were considerably faster in Trial 3 (34.19 ± 31.91). Similar to Trial 2, users took more time to decide in Trial 4 (40.84 ± 40.54) after a positive financial return.

Afterward, we designed the distribution to present a negative financial return consecutively between Trials 4 and 7. Users took a similar time to decide in Trials 5 (30.86 ± 21.22), 6 (28.81 ± 25.62), and 7 (29.09 ± 23.57). Nevertheless, after the three consecutive negative return values, results suggest that participants generally took more time to decide in Trial 8 (42.31 ± 34.20). This finding suggests that it takes a short series

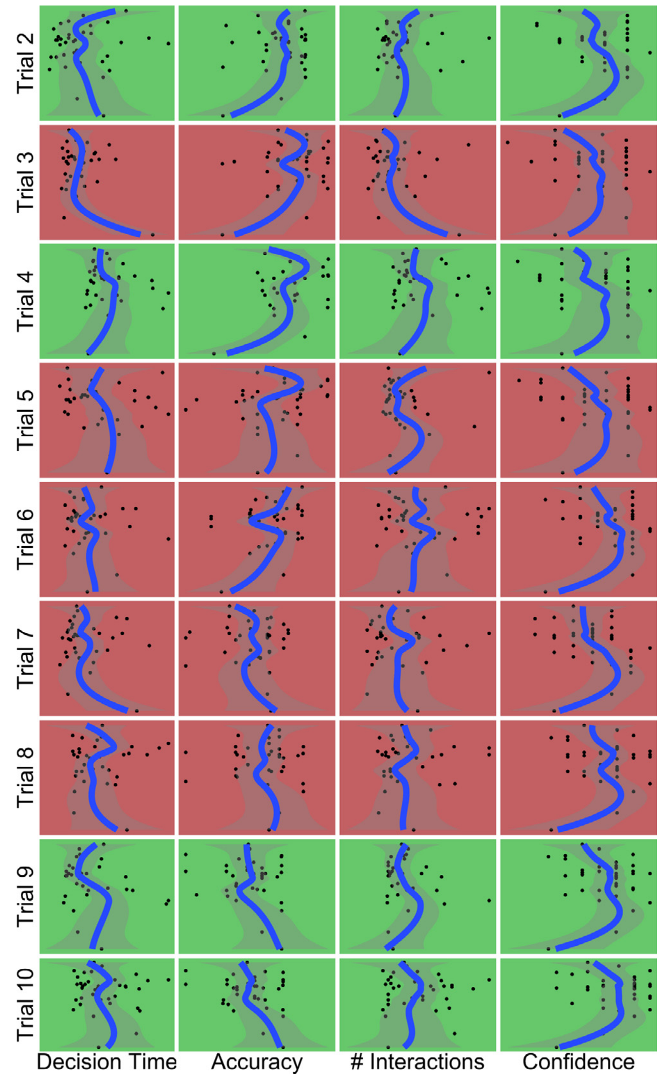


Fig. 4. Scatterplots with LOESS lines for the effect of conscientiousness (vertical axes, scores increase from bottom to top) on the dependent variables per trial (horizontal axes, values increase from left to right) with fixed returns.

of negative feedback before individuals are willing to spend more time deciding on the excavation site.

Another appealing finding emerges from the time individuals took to decide at the end of the experiment. After taking more time to decide in Trial 8, we can observe that the positive feedback from completing that trial led participants to be significantly faster in Trial 9 (19.60 ± 18.94). In particular, we found a statistically significant decrease of 22.71 (95% CI, 1.222 to 44.194) seconds, $p = .029$, from Trial 8 to Trial 9. We found a similar time distribution for Trial 10, where participants spent the least time deciding on the excavation site (15.78 ± 12.21). It suggests that participants believed they converged on a safe excavate choice after receiving the first positive feedback.

Regarding conscientiousness, Fig. 4 shows no clear effect. We found no significant correlation between the time users took to decide and their conscientiousness score. It appears that locally estimated scatterplot smoothing (LOESS) lines do not follow a specific trend. Consequently, the conscientiousness trait does not affect the task decision time.

4.1.2. Accuracy

Next, we examine user accuracy along with the trials. As we mentioned, we normalized decision accuracy between 0 and 1

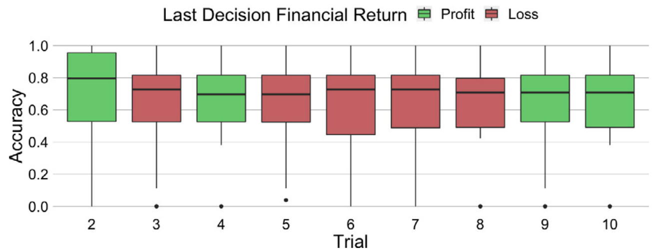


Fig. 5. Accuracy per trial with fixed returns.

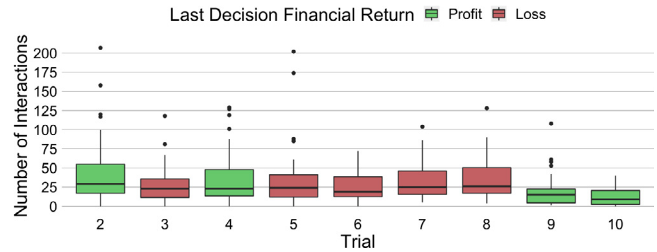


Fig. 6. Number of interactions per trial with fixed returns.

with the excavation site that yields the smallest and the biggest expected financial return, respectively. We can observe in Fig. 5 that the median accuracy was approximately the same across the trials except for Trial 2 (70.85 ± 29.75%). Notably, Trial 2 has the largest interquartile range, which is understandable since the users were at the first trials and there was a large item pool to consider. Interestingly, the valence variation of Trial 3 resulted in a drop in user accuracy to 66.70% ± 26.06%. Trial 4 showed that the decision accuracy slightly increased to 67.55% ± 21.91% after returning to positive decision quality feedback. Next, the trend that we saw in the decision time is not present in decision accuracy. The consecutive negative feedback did not show any effects on user accuracy in Trials 5 (65.49%±22.83%), 6 (64.77%± 28.75%), 7 (65.61% ± 26.01%), and 8 (66.18% ± 23.94%). Finally, returning to positive financial returns also yielded no effect on decision accuracy; Trials 9 and 10 have a decision accuracy of 67.13% ± 26.50% and 66.79% ± 25.59%, respectively.

Nonetheless, Fig. 4 suggests that conscientiousness appears to play a role in user accuracy. In Trial 2, we can observe through the LOESS line that individuals have similar scores independently of the conscientiousness scores. From Trials 3 to 6, there is a slight u-shape effect where individuals with average scores show lower decision accuracy than others. Starting from Trial 7, the LOESS significantly changes from the previous trials. In particular, we can see that individuals with the highest conscientiousness scores start to decrease their decision accuracy. Moreover, the u-shaped effect appears to invert, and individuals with average conscientiousness scores increase their accuracy. Finally, when users start receiving positive feedback (Trials 9 and 10), the shape of the LOESS starts reassembling a linear relationship, which is what one would expect from the manifestation of the conscientiousness trait. We discuss these trends in the Discussion section.

4.1.3. Interaction events

Participants triggered on average 281.24 ± 131.75 interaction events in the experiment, ranging from 57 to 529 events (see Fig. 6). As expected, participants triggered more interaction events in Trial 2 (45.43±45.25). In contrast, there was a clear decrease in events in Trials 3 (30.46±27.08) and 4 (36.35±36.79). Next, participants were exposed to constant negative feedback starting in Trial 5. At first, we observed that participants triggered about the same number of events in Trial 5 (36.27 ± 42.63) compared to Trial 4. Then, participants performed fewer events in Trials 6 (27.16 ± 21.08) and 7 (30.76 ± 23.20). After receiving negative feedback for four consecutive trials, participants increased the number of interaction events to 37.22 ± 32.00 in Trial 8. Finally, the last two trials included positive feedback, and participants triggered fewer interaction events in Trials 9 (18.81 ± 21.90) and 10 (12.24 ± 10.86).

We continued by running Spearman's rank-order correlation test to understand whether conscientiousness affected the number of interactions independently of the trial. We found no significant correlation, $r_s(36) = .047, p = .781$. Afterward, we analyzed

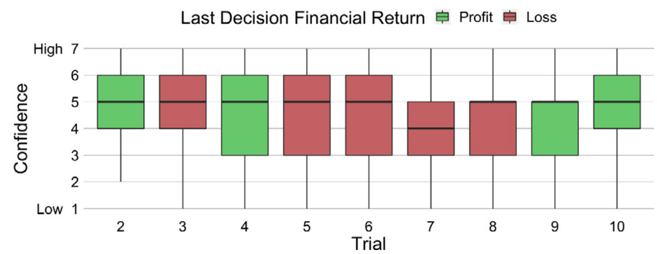


Fig. 7. Confidence per trial.

in each trial whether a significant difference was present to understand if the valence of the financial return had an interaction effect with conscientiousness. Similar to the previous analysis, we found no significant correlation in any trial. The lack of significant effects suggests that user interaction was independent of the financial return and the conscientiousness trait. In particular, Fig. 4 depicts LOESS lines that do not follow any particular pattern.

We decided to study more in-depth if participants interacted more with a particular chart. Of the 38 participants recruited to the study, the PCP elicited an increase in interactions in 34 participants compared to the SPM, whereas four participants interacted more with the latter. A Wilcoxon signed-rank test reported a statistically significant median increase in the number of interactions (203) when subjects used the PCP (Mdn = 208) compared to the SPM (Mdn = 5), $z = -5.141, p < .001$. Interestingly, we found that individuals that interacted more with the PCP had lower accuracy rates on average, $r_s(38) = -.342, p = .035$. In contrast, there was no significant correlation between the number of triggered events on the SPM and the accuracy rates on average, $r_s(38) = .017, p = .920$.

4.1.4. Confidence

For subjective metrics, we analyze how confident users were in their choices along with the trials. As presented in Fig. 7, confidence was highest in Trial 2 (4.73 ± 1.28). In Trial 3, the value was similar to the previous trial (4.59 ± 1.34). The next trials also showed similar distributions (Trial 4: 4.41 ± 1.57; Trial 5: 4.54 ± 1.63; Trial 6: 4.65 ± 1.28). However, we can observe the first major drop in confidence in Trial 7 (4.00 ± 1.51). This trend may be a result of consecutive negative feedback. The following Trials 8 and 9 kept similar confidence rates of 4.30 ± 1.70 and 4.35 ± 1.60, respectively. Finally, the confidence returned to higher values in Trial 10 after the participants received consecutive positive feedback. In particular, we can observe a value of 4.70 ± 1.82, which is close to the level at the beginning of the experiment.

Regarding conscientiousness, we found significant correlations between the trait and the self-assessed confidence in each trial. Although all correlations are weak (Table 2), these results suggest that conscientiousness plays a role in how individuals perceive their choices in a decision-making setting. Moreover, the non-monotonic relationship can be observed in the scatterplots of

Table 2
Spearman's rank-order correlation test results between conscientiousness and user confidence in each trial. Color encodes whether the last decision yielded profits (green) or losses (red) compared to the excavation cost (500 coins).

Trial	2	3	4	5	6	7	8	9	10
Rho	−.143	−.141	−.134	−.135	−.135	−.118	−.143	−.135	−.155
p-value	.014	.015	.021	.020	.020	.042	.014	.021	.007

Table 3
Point-biserial correlation test results between conscientiousness and whether the user changed their choice in each trial with fixed returns. Color encodes whether the last decision yielded profits (green) or losses (red) compared to the excavation cost (500 coins).

Trial	2	3	4	5	6	7	8	9	10
Rho	.170	.152	.143	.167	.156	.155	.159	.171	.151
p-value	.003	.009	.013	.004	.007	.008	.006	.003	.009

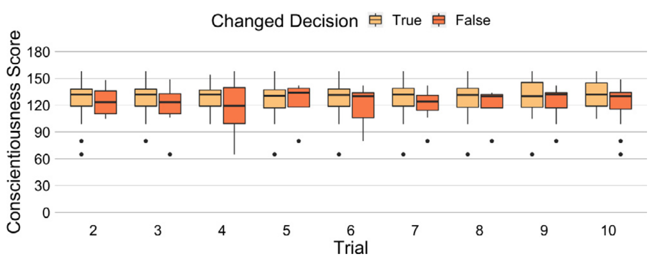


Fig. 8. Conscientiousness scores according to whether individuals changed their excavation site choice from the previous trial with fixed returns.

Fig. 4. Although there are some local, small variations, the general distribution shows a u-shaped effect, hinting that individuals with average scores tend to be more confident in their choices than their counterparts.

4.1.5. Changed decision

Fig. 8 shows that the individuals who kept their choice showed lower conscientiousness scores. In contrast, participants with higher scores appear to pay attention to the result of the financial returns by exploring more in the breadth of the available excavation sites. Point-biserial correlation tests showed these were statistically significant differences in each trial (Table 3). The largest difference was in Trial 4 with the individuals that changed their choice scoring higher (129.31 ± 13.67) than their counterparts (116.13 ± 31.61). In contrast, conscientiousness scores were similar for those who changed their choice (127.07 ± 20.14) and those who did not (124.25 ± 16.53). Afterward, we analyzed whether conscientiousness led individuals to choose a larger variety of excavation sites. Nevertheless, Spearman's rank-order correlation test showed no statistically significant between the conscientiousness score and the number of different chosen excavation sites, $r_s(36) = .04$, $p = .811$.

4.2. Real returns experiment

We recruited 36 participants (20 male, 16 female) aged 22 – 63 ($M = 26.36$, $SD = 8.34$). Out of the 36 individuals, 31 participated in the first experiment. Each test was 24:31 ± 6:54 (mm:ss) long with minimum and maximum durations of 9:54 and 36:22 (mm:ss), respectively. Additionally, participants self-reported being fairly familiar with PCP ($M = 2.53$, $SD = 2.34$) and SPM ($M = 4.11$, $SD = 1.77$).

4.2.1. Task decision time

Fig. 9 depicts the time users took to decide the next excavation point based on how good was their past decision. One can observe

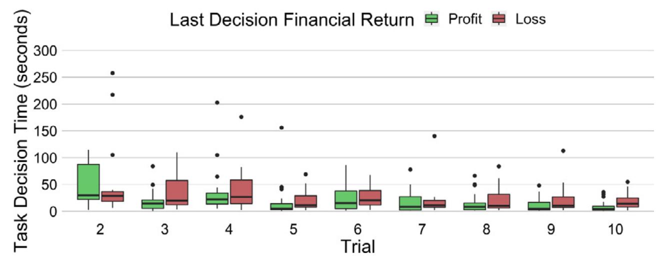


Fig. 9. Task decision time per trial with real returns.

that there was a decreasing amount of time spent on the trials as participants progressed in the experiment. Users took the longest time to decide in Trial 2 (53.04 ± 56.62 s). We also observed a statistically significant decrease of 23.80 (95% CI, 0.268 to 47.335) s, $p = .045$, from Trial 2 to Trial 3. Then, it appears that users continued exploring while making their decisions in Trials 3 (29.24 ± 29.02), 4 (38.57 ± 44.21), 5 (19.75 ± 28.87), and 6 (23.19 ± 21.29). Afterward, we can observe an interesting trend. We observed a continuous decrease of the average decision time in Trials 7 (19.24 ± 26.77), 8 (18.62 ± 21.37), 9 (15.75 ± 21.54), and 10 (12.09 ± 13.42). This distribution points toward users learning over time to play the game and being faster in making their decisions.

We also found that individuals who turned a profit in the previous trial are always faster on average than those who had a choice returning losses. Starting in Trial 2, participants with a previous profit spent 49.55 ± 38.36 s deciding while those with losses took 57.94 ± 76.61 s to decide. A similar trend was present in Trials 3 (Profit: 20.22 ± 21.74, Loss: 36.46 ± 32.48) and 4 (Profit: 36.66 ± 46.57, Loss: 40.70 ± 42.73). Further, we can observe a decrease in the standard deviation after Trial 4. People who had a profit on the past choice were consistently faster than their counterparts in Trials 5 (Profit: 18.85 ± 34.53, Loss: 21.02 ± 19.45), 6 (Profit: 21.60 ± 21.86, Loss: 26.39 ± 20.65), 7 (Profit: 17.10 ± 19.68, Loss: 24.11 ± 39.25), 8 (Profit: 16.49 ± 19.85, Loss: 21.61 ± 23.72), 9 (Profit: 10.71 ± 12.59, Loss: 22.79 ± 29.02), and 10 (Profit: 8.37 ± 10.03, Loss: 19.54 ± 19.50). In particular, between Trials 7 and 10, the difference between individuals with contrasting feedback is evident and consistent based on the smaller standard deviation values.

Regarding conscientiousness, we found no significant correlations after a Bonferroni correction with a significance level of 0.0056 (0.05/9). Nevertheless, Fig. 10 shows moderate, negative correlations between conscientiousness scores and the task decision time in Trials 2, $r_s(36) = -.343$, $p = .041$, and 6, $r_s(36) = -.339$, $p = .043$. The LOESS lines suggest that people with higher conscientiousness scores appear to be faster in the first three trials. Then, between Trials 5 and 8 there are fewer differences between different conscientiousness scores except in Trial 6. We can also observe that individuals with high scores are slightly faster than their counterparts. Consequently, the conscientiousness trait appears to affect the task decision time when the returns are real.

When we consider whether the participant received a profit (Fig. 11) or a loss (Fig. 12) in the previous trial, we found that individuals with higher conscientiousness scores are generally faster when they receive positive feedback except in Trials 5, 8,

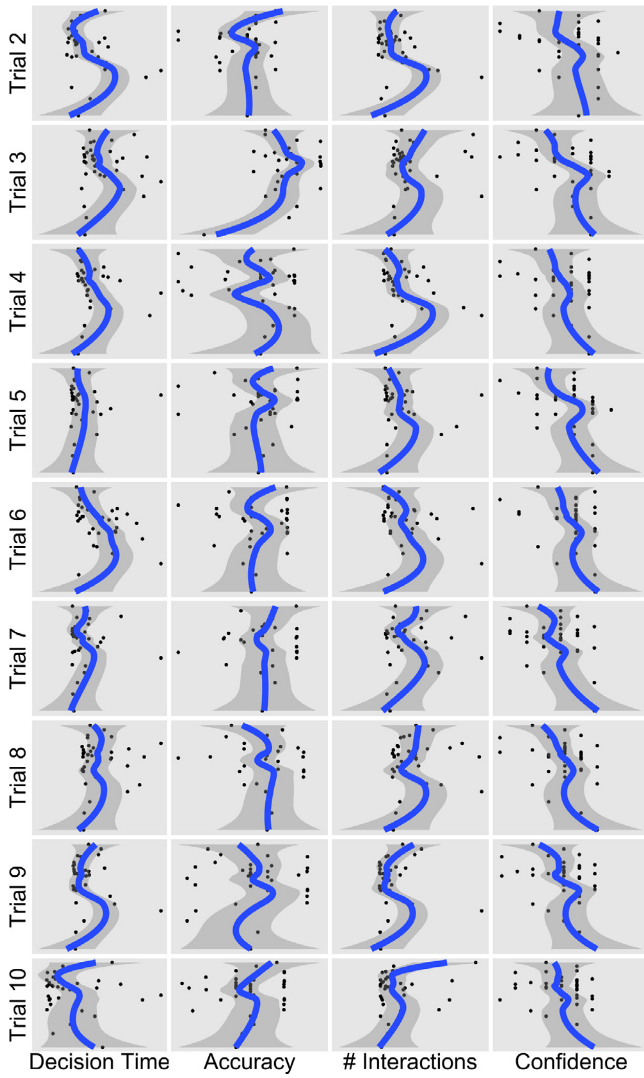


Fig. 10. Scatterplots with LOESS lines for the effect of conscientiousness (vertical axes, scores increase from bottom to top) on the dependent variables per trial (horizontal axes, values increase from left to right) with real returns.

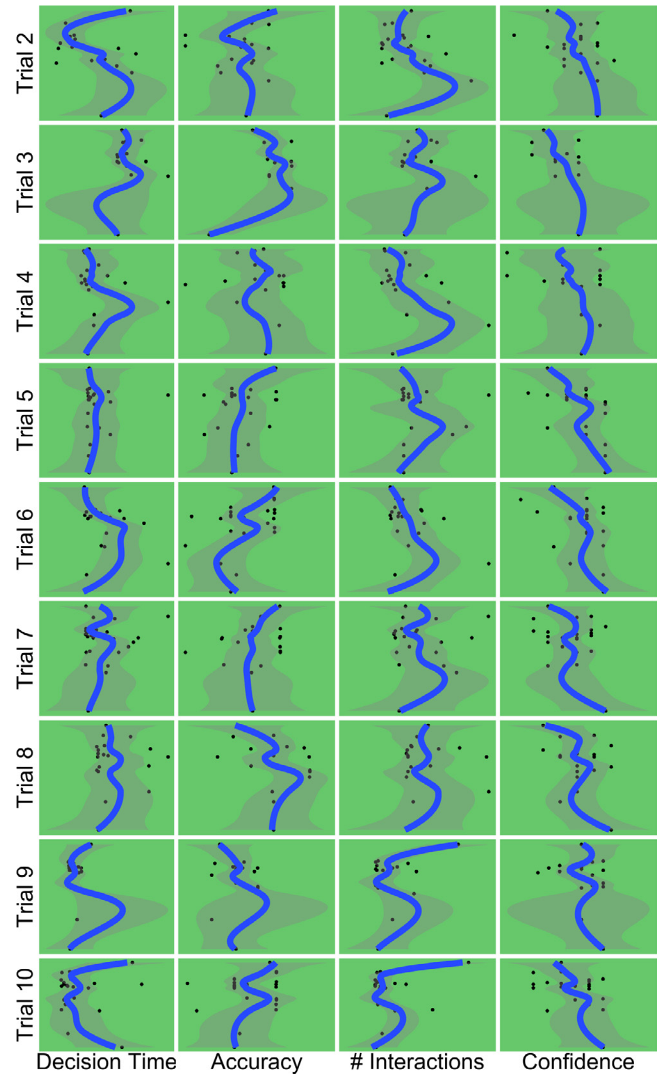


Fig. 11. Scatterplots with LOESS lines for the effect of conscientiousness (vertical axes, scores increase from bottom to top) on the dependent variables per trial (horizontal axes, values increase from left to right) with real profit returns.

and 10. Similar to the previous analysis, the effect is evident in Trial 6 as well as in Trials 2 and 4. When individuals receive losses in the previous trial, we observed that individuals with the higher score also tend to be faster to make decisions with Trial 5 standing out with a contrasting LOESS line. Taking into account our findings, it appears that Trial 5 was the trial when participants started to be more proficient in playing the Excavation Game.

4.2.2. Accuracy

We continued by examining the user accuracy. Fig. 13 that receiving positive feedback on the previous choice led individuals to be more accurate in their next decision. Similar to our previous experiment, it appears that the median accuracy was approximately the same across the trials. As expected, the highest accuracy rate was present in Trial 10 (78.68% ± 17.82%) and the lowest in Trial 4 (67.84% ± 5.16%). These results are in line with the previous findings regarding task decision time where we were able to observe a learning effect from the participants as they progressed in the experiment.

Notably, Trial 2 is the only time that participants who had a loss in the previous trial (71.82% ± 16.20%) showed higher accuracy than those with profits (67.21% ± 29.58%). Afterward, the effect was always the opposite. We observed that individuals who receive positive feedback were more accurate in the next trial than their counterparts and that the difference between them was more evident in the initial trials. In particular, we can see larger differences in Trials 3 (Profit: 76.12% ± 24.94%, Loss: 67.32% ± 15.64%), 4 (Profit: 74.99% ± 25.12%, Loss: 59.83% ± 35.44%), and 5 (Profit: 77.31% ± 14.82%, Loss: 67.34% ± 27.09%). However, the differences between individuals with contrasting feedback become smaller in Trials 6 (Profit: 80.43 ± 18.16, Loss: 67.34% ± 27.09%), 7 (Profit: 72.79% ± 24.52%, Loss: 70.84% ± 16.18%), 8 (Profit: 71.14% ± 20.83%, Loss: 66.04% ± 22.60%), and 9 (Profit: 73.14% ± 19.46%, Loss: 72.99% ± 19.03%). The exception is on the last trial (Profit: 82.34% ± 17.71%, Loss: 71.05% ± 16.32%).

While running Spearman correlation tests independently of the returns, we only found weak nonsignificant correlations. These findings suggest that conscientiousness does not affect user accuracy when we do not consider the valence of the returns. Fig. 10 depicts the LOESS lines. Trials 2 and 10 suggest that the individuals with the highest conscientiousness scores were more accurate. Further, individuals with scores higher than the average appear to struggle more in Trials 3, 5, 6, and 9 than those with average scores. However, we can observe interesting trends when

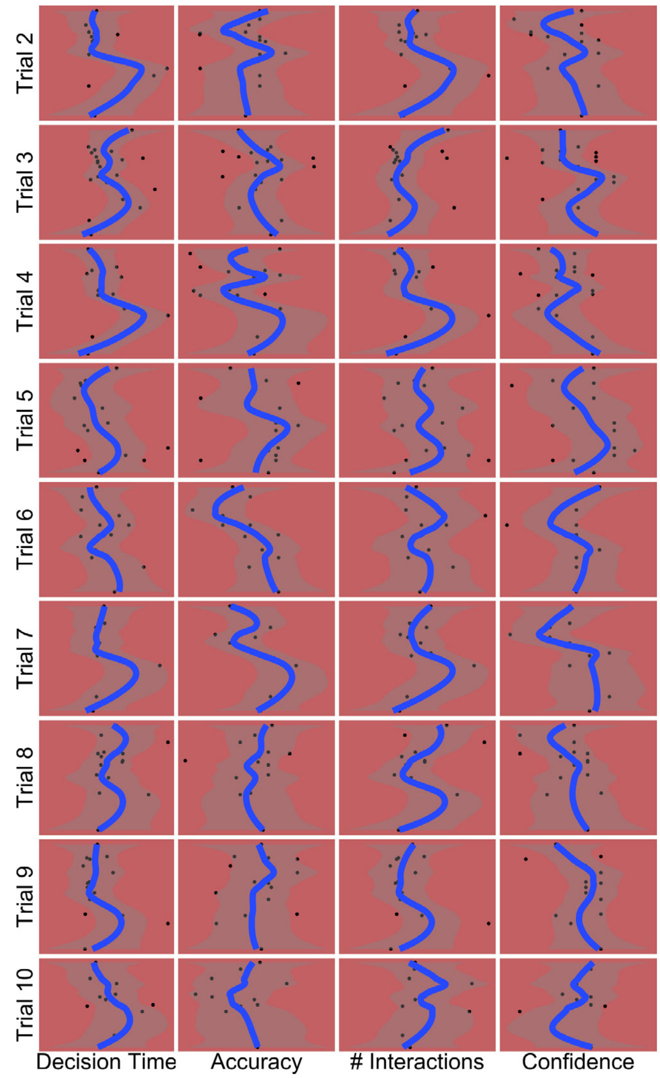


Fig. 12. Scatterplots with LOESS lines for the effect of conscientiousness (vertical axes, scores increase from bottom to top) on the dependent variables per trial (horizontal axes, values increase from left to right) with real loss returns.

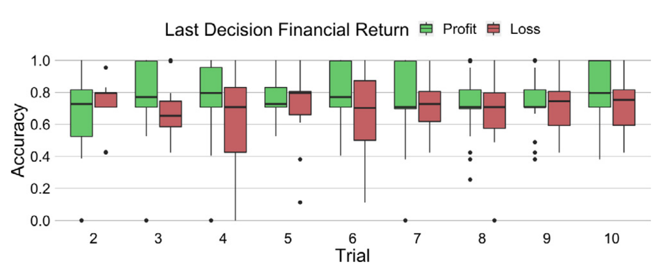


Fig. 13. Accuracy per trial with real returns.

we consider whether the participant received a profit or a loss on their past decision.

Although we did not find any significant correlations, it appears that the distribution of accuracy rates and conscientiousness scores of individuals who had a profit in the previous trial are similar to the ones independently of the return valence (Fig. 11). In contrast, we find dissimilar LOESS lines when we analyze the individuals who had a loss in the previous trial (Fig. 12). The differences are more evident starting in Trial 5. In particular, it appears that participants with higher conscientiousness are more sensitive to negative feedback since they achieve lower accuracy rates compared to the remaining participants in Trials 5, 6, and 7. We also found a nonsignificant, strong, negative correlation between user accuracy in Trial 6 and the conscientiousness scores for individuals who had a loss in the previous trial, $r_s(12) = -.620, p = .031$. These findings hint that the valence of the returns, i.e., whether the participant receives a loss or a profit in the previous trial, interacts with the conscientiousness trait on how accurate individuals are when they decide in the following moment. However, it appears that after Trial 8 this trend diminishes which may hint towards the benefits of learning how to play the Excavation Game.

4.2.3. Interaction events

Due to technical malfunctions, we were not able to collect interaction data from one participant. Consequently, we do not consider it in this analysis. Similar to the previous experiment, participants triggered on average 279.89 ± 230.30 interaction events in the experiment, ranging from 0 to 1241 events. In particular, individuals triggered more interaction events in Trial 2 (62.40 ± 77.09). There was a steep decrease in Trials 3 (34.97 ± 49.20) and 4 (42.51 ± 50.10). However, we can observe that starting in Trial 5 (27.00 ± 38.47) the values started to be very similar (Trial 6: 27.43 ± 29.97; Trial 7: 21.03 ± 21.77; Trial 8: 25.97 ± 29.54; Trial 9: 20.69 ± 31.25; Trial 10: 17.89 ± 26.52). As in our previous statistical analysis, we believe that this finding reinforces how participants finished exploring the dashboard in Trial 4. Further, we inspected through a Wilcoxon signed-rank test if participants interacted more with a specific chart. Results show a statistically significant median increase in the number of interactions (165) when subjects used the PCP (Mdn = 182) compared to the SPM (Mdn = 17), $z = -4.930, p < .001$. Therefore, the PCP appears to be the chart that most people interacted with in both experiments although the number of interactions not leading to better accuracy rates on average, $r_s(35) = -.199, p = .251$. However, it appears that interacting with the SPM led participants to achieve higher accuracy rates on average, $r_s(35) = .393, p = .020$.

Taking into account the real returns from their decisions, we observed that individuals tend to perform fewer interactions with the charts after receiving positive feedback (Fig. 14). The difference between the number of interactions was larger in the initial and final trials between individuals with profits (Trial 2: 51.19 ± 50.71; Trial 3: 20.13 ± 23.70; Trial 9: 12.55 ± 17.15; Trial 10: 14.04 ± 25.78) and those with losses (Trial 2: 79.21 ± 105.26; Trial 3: 46.10 ± 60.12; Trial 9: 31.53 ± 41.85; Trial 10: 25.25 ± 27.46). In contrast, we found that independently of the returns, participants with profits (Trial 4: 41.63 ± 46.60; Trial 5: 31.05 ± 49.33; Trial 6: 24.09 ± 28.42; Trial 7: 19.17 ± 20.40; Trial 8: 24.55 ± 33.94) and those with losses (Trial 4: 43.56 ± 55.51; Trial 5: 21.60 ± 15.37; Trial 6: 33.83 ± 33.05; Trial 7: 25.09 ± 25.04; Trial 8: 27.87 ± 23.44) showed similar numbers of interactions per trial in the remaining trials.

When we consider conscientiousness, we did not find any significant correlation between the trait score and the number of interactions with both charts, $r_s(35) = -.021, p = .904$, with the PCP, $r_s(35) = .012, p = .944$, or with the SPM, $r_s(35) = -.205, p = .237$. The LOESS lines suggest that individuals with higher conscientiousness scores tend to perform fewer interaction events when deciding between the excavation sites (Fig. 10). When we take the returns, we did not find any significant differences between individuals with profits (Fig. 11) or losses (Fig. 12). The LOESS lines also show similar shapes in this case. Consequently, it appears that conscientiousness does not manifest its effects on the number of interaction events triggered by the participants independently of the valence of the returns.

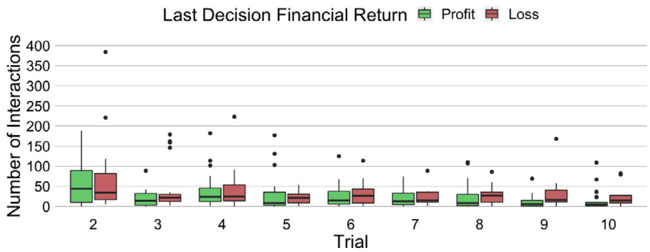
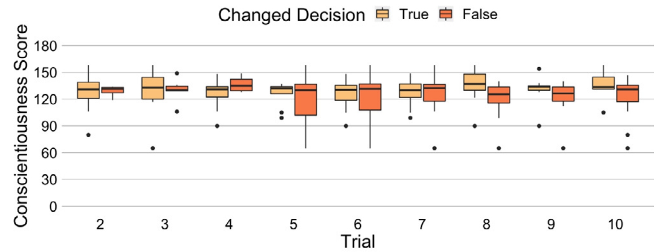
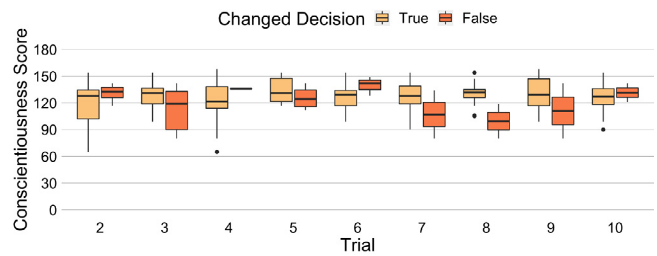


Fig. 14. Number of interactions per trial with real returns.

Fig. 15. Confidence per trial with real returns.



(a) Individuals with profits in the previous decision.



(b) Individuals with losses in the previous decision.

Fig. 16. Conscientiousness scores according to whether individuals changed their excavation site choice from the previous trial with real returns.

4.2.4. Confidence

We continued our analysis by tackling self-reported confidence in the decisions. It appears that the average self-reported confidence was similar across trials. Contrary to the experiment with fixed returns, this time the highest confidence was reported in Trial 10 (5.00 ± 1.60). Trials 8 (4.92 ± 1.46) and 9 (4.94 ± 1.60) showed similar scores thus supporting that users were slightly more confident at the end of the experiment. These values contrast with the self-reported confidence at the beginning of the experiment since the values were lower in Trials 2 (4.50 ± 1.48) and 4 (4.53 ± 1.40). Further, when we consider the real returns of the decisions, we observe that the distributions were similar across all trials except for Trials 2, 6, 7, 9, and 10 (Fig. 15). Trial 2 did indeed led participants with previous losses to report lower confidence in their choice (4.13 ± 1.60) than those with profits (4.76 ± 1.38). Nevertheless, it appears that in Trial 6 there were similar confidence values for the former (4.60 ± 1.6) and the latter (4.67 ± 1.20). Finally, we were able to observe differences similar to the ones in Trial 2 in Trials 7 (Profit: 4.60 ± 1.44, Loss: 5.09 ± 1.45), 9 (Profit: 5.10 ± 1.70, Loss: 4.73 ± 1.49), and 10 (Profit: 4.96 ± 1.65, Loss: 5.08 ± 1.56).

We continued our analysis by tackling the effect of conscientiousness (Fig. 10). Independently of the returns, it appears that there are no significant negative effects in Trials 3 ($r_s(36) = -.395, p = .017$) and 5 ($r_s(36) = -.446, p = .006$) at an adjusted Bonferroni acceptance threshold of 0.0056. Similar to the other experiment, it appears that higher conscientiousness scores lead participants to report lower confidence in their decisions compared to the remaining individuals. Taking the profits into account, we found a relatively strong significant correlation between the trait scores of individuals with previous profits and the self-reported confidence in Trial 5 ($r_s(21) = -.587, p = .005$). Further, there was a strong nonsignificant correlation between the trait scores of individuals with previous losses and the self-reported confidence in Trial 7 ($r_s(11) = -.628, p = .039$). These findings support our hypothesis that conscientiousness does indeed play a role in how individuals interpret their decision-making process.

4.2.5. Changed decision

Fig. 16 shows whether individuals kept their choice based on the returns from their previous decision. For those with profits, we can observe that individuals with higher conscientiousness scores tend to change their choice in the last trials. In contrast, those with lower scores generally keep their choices after receiving positive returns. Then, we analyzed how individuals with losses behaved. Results show a similar trend compared to those with profits. In particular, from Trials 7 to 9 we can observe large differences between those who decided to change their choices when faced with losses. Both findings suggest that individuals with lower conscientiousness tend to keep their choices independent of the results while those with higher scores keep looking for better options after finishing the exploration of the Excavation Game. However, a Spearman's rank-order correlation test reported no significant relationship between the conscientiousness score and the number of different chosen excavation sites, $r_s(36) = .068, p = .692$.

We continued by running point-biserial correlation tests for individuals with profits and losses in their previous decision (Table 4). Although no statistical test reported a significant difference at a significance level of 0.0028 (0.05/18), we found that in Trial 8 there was a relatively strong nonsignificant effect from those with losses. Moreover, the negative coefficients from those with profits in the second half of the experiment suggest that higher scores in conscientiousness tend to change their decision more frequently, as we observed in the past analysis. Finally, there are several coefficients from those with losses ranging from moderate to relatively strong effects. All the mentioned results corroborate the findings in our previous experiment and, more specifically, that conscientiousness appears to manifest its effect in interaction with the valence of the returns on whether individuals choose other alternatives.

5. Discussion

As the first steps toward understanding whether personality plays a role in visualization-based decision-making, we measured how participants make their decisions and gauge their confidence in said decisions while playing a game. Further, we manipulate the decision quality feedback of their decisions by presenting real and fixed returns. We analyzed the effects of the fixed decision quality feedback (**Q1**) and the conscientiousness trait (**Q2**) through objective and subjective metrics.

Table 4
Point-biserial correlation test results between conscientiousness and whether the user changed their choice in each trial with real returns.

	Trial	2	3	4	5	6	7	8	9	10
Profit	Rho	.010	.042	.277	−.152	−.136	−.087	−.322	−.226	−.258
	p-value	.966	.878	.250	.511	.527	.680	.155	.325	.223
Loss	Rho	.235	−.341	.144	−.278	.376	−.354	−.577	−.285	.129
	p-value	.399	.142	.582	.316	.228	.286	.024	.303	.691

5.1. Effect of the quality of the last decision

First, we focused on whether the quality of the last decision affected how much time users took on the next decision. In the first experiment with fixed returns, up until Trial 4, participants received alternate feedback valences to promote early exploration of the dataset. Starting in Trial 5, we provided the participant with two series of constant feedback quality to observe whether there were any effects after the continuous exposition. Results suggested that receiving constant negative feedback led individuals to focus on correcting the quality of the choice by spending more time on it. This trend can be observed from Trials 5 to 8. We assume that the increase in task decision time is a result of participants trying to adjust their decision-making behavior and improve the quality of the excavation site. In the last two trials, we observed a contrasting effect. In particular, participants appeared to be faster at making decisions after receiving positive feedback in a row (Trials 9 and 10). The task decision times in both trials were the fastest in the experiment. However, other factors may have affected the task decision time. For instance, faster decision-making may result from increased expertise by the end of the experiment. By being more accustomed to the dashboard, participants may be faster to find their choice of the excavation site. Nonetheless, the sequential increase in decision time from Trials 6 to 8 hints that participants did indeed spend more time making their decisions.

In the second experiment, we were able to observe through real returns that indeed participants were faster from Trials 8 to 10 when they received positive feedback compared to those with negative feedback. Our results also confirmed that participants playing the Excavation Game more times led participants to be faster deciding in the last trials independently of the feedback quality. Finally, we also observed that the participants appeared to explore the game up until Trial 5 and that, afterward, individuals appear to show a level of proficiency in their decision-making. These results suggest that the quality of the last decision affects how much time they take to decide.

When we assess the user accuracy with fixed returns, we notice a slight increase in the decision accuracy and interaction events from Trials 6 to 8. In contrast, the positive feedback in Trials 9 and 10 led to a slight increase in decision accuracy and few interaction events. We believe that the trends of these variables hint that the consecutive negative feedback led participants to be more invested in the decision-making and when met by positive feedback, they believed they converged on an optimal solution and decided more quickly and with less data analysis. As we mentioned, participants could choose any excavation site in a trial. We normalized the decision accuracy between the excavation sites with the lowest and highest expected financial returns. Nevertheless, participants did not see how accurate their decisions were. We fixed the financial return they received at the end of their trial independently of their selection. Consequently, this trend was not driven by user expertise with the dashboard but by feedback. Results showed that the quality of the last decision did not significantly affect how accurate the next decision was. In contrast, results suggest that when participants received real returns, receiving positive feedback at the initial trials led them to be more accurate. We believe these results suggest that the quality of the feedback plays a role in user accuracy while individuals are exploring the dashboard and learning how to decide but are less relevant afterward.

Furthermore, the second experiment confirmed that the valence of the feedback quality played a role in the number of triggered interaction events. As we mentioned, individuals performed fewer interaction events when they were presented with two fixed consecutive positive feedback instances. We observed that real positive feedback led participants to trigger fewer events than people with losses in their previous decision between Trials 5 and 10. This effect was even more evident in the last trials and, paired with the trends observed in the decision time, we can understand how learning to play the excavation game led participants to be more concise in deciding. We also found that participants who interacted more with the PCP in the first experiment had lower accuracy rates on average. Additionally, interacting more with the SPM led individuals to score higher accuracy rates on average in the second experiment. Although Dimara et al. [19] found that both visualization techniques led to similar accuracy rates when evaluated individually, these findings highlight that the presence of both visualizations at the same time may hinder or boost one's accuracy rate.

Afterward, we studied if the quality of the feedback affected how confident individuals felt about their decisions. In the first experiment, we did not observe a significant effect of the feedback quality in the self-assessed confidence. The null effects may result from the lack of dependency of user confidence on the feedback on their past choice. However, the second experiment showed that real returns did not affect self-confidence as well. We believe these results may derive from the approach we used to assess user confidence. In particular, participants can only assess the present decision based on the current state of the game. We hypothesize that participants only considered their past decision when assessing their self-confidence and, consequently, were anchored by it. By being able to view their progress at any time, we expect it would lead participants to vary how confident they were in the decisions.

5.2. Effect of conscientiousness

Regarding conscientiousness, we were also able to see that the personality trait did not show significant effects on the task decision time, hinting that the time participants take to decide is independent of their tendency to plan carefully their strategies [13]. Nevertheless, results suggested an effect of conscientiousness in the initial trials when the returns were real. Further, receiving positive feedback appears to make individuals with higher conscientiousness scores decide faster than their counterparts. Consequently, the mixed results lead us to conclude that deeper synergies between the conscientiousness trait and the quality of the feedback may affect the task decision time but our study design was not able to measure them.

In the first experiment, we found that conscientiousness appears to manifest its effects in the last trials where we can see an adjustment of the decision accuracy by conscientious individuals when receiving constant negative returns. We found a similar effect in the second experiment from individuals receiving real negative feedback. Both findings hint toward high conscientiousness scores manifesting a higher sensibility to this feedback since they are efficient, goal-oriented, and methodical. Consequently, we expect them to work diligently to achieve the goal of maximizing the financial returns. The constant negative feedback may have led them to adopt other choices that followed less optimized outcomes. We also believe that having explored more data in the final trials directly affected the decision quality of individuals with average scores. The average conscientiousness scores may have led individuals to not be so rigid in their selections and adapt their choices based on the decision quality feedback.

Regarding user interaction, conscientiousness did not affect the number of events triggered by the participants in either experiment. Results showed that participants interacted much more frequently with the PCP than with the SPM independently of whether we controlled the quality feedback. However, it appears that conscientiousness scores do not affect how much individuals interact with either chart. Additionally, we found interesting effects regarding user confidence in the choice taken. There was a clear manifestation of the conscientiousness trait. The non-monotonic relationship shows a u-shaped effect that depicts individuals with average conscientiousness scores as more confident in their choices than their counterparts. In particular, the second experiment highlighted that participants with high conscientiousness scores were less confident in their choices and that the valence of the feedback. We believe this difference may result from the level of choice uncertainty. Individuals with high conscientiousness scores tend to be careful and seek to retain control over a situation [87]. We suggest that the uncertainty factor may make them self-doubt their decisions since they feel more pressured to consider all possible choices. For individuals with lower scores, it may derive from their tendency to be easy-going [13]. Therefore, they would be less interested in making an optimal decision and, consequently, be unsure as to whether it would be a "good" decision.

Finally, we examined whether conscientiousness and the decision quality feedback affected whether users kept their choice across the trials and the number of different excavation sites they selected. Results showed that participants who changed their score had higher conscientiousness scores than those who kept the excavation sites across the trials in both experiments. Additionally, if we consider the valence of the feedback in the second experiment, it appears that losses exacerbate the manifestation of conscientiousness, especially in the last trials. Paired with the findings regarding decision accuracy, we believe that there is a manifestation of conscientiousness. In particular, individuals with higher conscientiousness scores keep trying new excavation sites to achieve higher financial returns, and they converge on higher decision accuracy in the last trials.

5.3. Remarks

We observed that the quality feedback participants receive based on their decisions plays a significant role in the decision-making process when participants interact with the system in a decision-making context. For instance, results indicate that receiving negative feedback leads participants to spend more time deciding than when the feedback is positive. Additionally, we also observed that the conscientiousness trait biases the decision-making process of individuals when they interact with a decision-support system that leverages visualization techniques. In particular, results show that the time individuals took deciding in the initial trials depends on their conscientiousness scores, specifically with conscientious individuals being faster. Similar interactions between the quality of the past decision and conscientiousness scores showed that conscientious individuals focus on improving their decision accuracy and change their choices across the alternatives more frequently after receiving negative feedback.

These types of relationships highlight that when visualization designers are testing their decision-support system, they should not analyze dependent metrics without considering the effect of confounding factors. As we mentioned, our findings show that visualization designers measuring the task decision time may leverage the quality feedback valence and individual differences factors such as conscientiousness to understand the decision-making process better. Specifically, regarding conscientiousness, past research has not identified measurable effects of this trait in visualization contexts though there is already a considerable body of literature [10]. Our findings suggest that visualization researchers should consider conducting more studies on the bias that conscientiousness may introduce in user interaction with visualization systems. Further, visualization designers may need to increase the awareness of personality profiles to break the cycle of one-size-fits-all design approaches by adding the conscientiousness trait to the user profile when designing a decision-support visualization-based system.

5.4. Limitations and future work

There are some limitations to the results of this study. First, we leverage user confidence as a self-assessment metric. Although it provides further insights regarding the decision-making process, different users may calibrate their assessment differently, i.e., although two participants were equally confident in their choice, each reports a different but similar score. As a result, this metric may not have been sensible enough to capture decision quality effectively. Second, our study focuses on the conscientiousness trait. This approach allows us to focus on a trait and study whether it manifests its characteristics in the decision-making process. Nevertheless, the context of decision-making supported by visualization may lead to the manifestation of other trait features and consequently produce synergies between traits that our experiment did not control. Future studies should design the user study to accommodate more personality traits in the analysis. Further, our sample presents two non-extreme outliers in conscientiousness scores. Although the statistical analysis we applied was robust to outliers [88,89], future studies should consider collecting a normal sample.

Third, our study is dependent on the visualization context and user perception. Several studies have shown that the PCP and the SPM allow users to complete analytical tasks accurately. Nonetheless, more chart types can be analyzed in their role in visualization-supported decision-making. Future work should consider other graphs to study how the graphical exposition of elements affects the decision-making process. In particular, it could be interesting to investigate if including uncertainty generated by lossy projection techniques leads conscientiousness to manifest its effects at a stronger degree. Further, our experiment considers a single sequence of trials with fixed returns, and, consequently, the findings may not apply to other feedback sequences. Future work should consider more sequences varying in several factors, e.g., path length or feedback valence order. Additionally, it would be interesting to assess if knowing the number of decision trials participants had to perform introduced noise in the experiment. Future work may also consider not revealing this information to assess if the progress within the set of trials influences user behavior. Finally, the size of the dataset may have also introduced some noise in our results. Future studies should consider how varying the scale of the data affects the decision-making process and how confident users are in their decisions.

6. Conclusions

The state-of-the-art research includes few studies on visualization-supported decision-making. In particular, there is a need to enhance the profile of the decision-maker to understand how designers can empower them [6]. Our work adds to prior research by exploring whether conscientiousness, a personality trait responsible for the tendency to follow goals and prioritize tasks [13], affects the decision-making process. We also study if the conscientiousness trait manifests its effects based on the feedback users receive on their choice with fixed and real returns. Results suggest that the feedback quality may play a role regarding how invested participants were in the decision-making and data analysis during the trials. Additionally, conscientiousness appears to affect how confident individuals are in their choices, leading conscientious users to change their choices more often. We believe that these findings open the research space to consider individual differences in visualization-supported decision-making and support the vital role decision-making has in society [4].

CRediT authorship contribution statement

Tomás Alves: Conceptualization, Methodology, Formal analysis, Writing – original draft, Visualization, Funding acquisition. **Tiago Delgado:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Joana Henriques-Calado:** Methodology, Resources, Writing – review & editing, Supervision. **Daniel Gonçalves:** Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Sandra Gama:** Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cag.2023.01.010. It includes the covariance matrix used in the experiment to create the dataset.

References

[1] Card M. Readings in information visualization: using vision to think. Morgan Kaufmann; 1999.
[2] Zhang Y, Bellamy RK, Kellogg WA. Designing information for remediating cognitive biases in decision-making. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems. 2015, p. 2211–20.
[3] Dimara E, Bailly G, Bezerianos A, Franconeri S. Mitigating the attraction effect with visualizations. IEEE Trans Vis Comput Graphics 2018;25(1):850–60.
[4] Dimara E, Stasko J. A critical reflection on visualization research: Where do decision making tasks hide? IEEE Trans Vis Comput Graphics 2021.
[5] Aseniero BA, Wun T, Ledo D, Ruhe G, Tang A, Carpendale S. Stratos: Using visualization to support decisions in strategic software release planning. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems. 2015, p. 1479–88.
[6] Dimara E, Zhang H, Tory M, Franconeri S. The unmet data visualization needs of decision makers within organizations. IEEE Trans Vis Comput Graphics 2021.
[7] Dimara E, Franconeri S, Plaisant C, Bezerianos A, Dragicevic P. A task-based taxonomy of cognitive biases for information visualization. IEEE Trans Vis Comput Graphics 2018;26(2):1413–32.
[8] Kale A, Kay M, Hullman J. Visual reasoning strategies for effect size judgments and decisions. IEEE Trans Vis Comput Graphics 2020;27(2):272–82.
[9] Allport GW. Concepts of trait and personality. Psychol Bull 1927;24(5):284. http://dx.doi.org/10.1037/h0073629.
[10] Liu Z, Crouser RJ, Ottley A. Survey on individual differences in visualization. In: Computer graphics forum, vol. 39. Wiley Online Library; 2020, p. 693–712.
[11] Massari L. Analysis of myspace user profiles. Inform Syst Front 2010;12(4):361–7.
[12] Piao G, Breslin JG. Inferring user interests in microblogging social networks: A survey. User Model User-Adapt Interact 2018;28(3):277–329.
[13] Halko S, Kientz JA. Personality and persuasive technology: An exploratory study on health-promoting mobile applications. In: International conference on persuasive technology. Springer; 2010, p. 150–61. http://dx.doi.org/10.1007/978-3-642-13226-1_16.
[14] Mackinlay J, Hanrahan P, Stolte C. Show me: Automatic presentation for visual analysis. IEEE Trans Vis Comput Graphics 2007;13(6):1137–44.
[15] Wills G, Wilkinson L. Autovis: Automatic visualization. Inform Visual 2010;9(1):47–69.
[16] Cunha JM, Polisciuc E, Martins P, Machado P. The many-faced plot: Strategy for automatic glyph generation. In: 2018 22nd international conference information visualisation. IEEE; 2018, p. 71–7.
[17] Battle L, Scheidegger C. A structured review of data management technology for interactive visualization and analysis. IEEE Trans Vis Comput Graphics 2020;27(2):1128–38.
[18] Schneider H, Wayrauther J, Hassib M, Butz A. Communicating uncertainty in fertility prognosis. In: Proceedings of the 2019 CHI conference on human factors in computing systems. 2019, p. 1–11.
[19] Dimara E, Bezerianos A, Dragicevic P. Conceptual and methodological issues in evaluating multidimensional visualizations for decision support. IEEE Trans Vis Comput Graphics 2017;24(1):749–59.
[20] Triantaphyllou E. Multi-criteria decision making methods. In: Multi-criteria decision making methods: a comparative study. Springer; 2000, p. 5–21.
[21] Zionts S. Mcdm–If not a Roman numeral, then what? Interfaces 1979;9(4):94–101.
[22] Filipič B, Tušar T. A taxonomy of methods for visualizing pareto front approximations. In: Proceedings of the genetic and evolutionary computation conference. 2018, p. 649–56.
[23] Wittenburg K, Lanning T, Heinrichs M, Stanton M. Parallel bargrams for consumer-based information exploration and choice. In: Proceedings of the 14th annual ACM symposium on user interface software and technology. 2001, p. 51–60.
[24] So. Yi J, Melton R, Stasko J, Jacko JA. Dust & magnet: Multivariate information visualization using a magnet metaphor. Inform Visual 2005;4(4):239–56.
[25] Hullman J, Qiao X, Correll M, Kale A, Kay M. In pursuit of error: A survey of uncertainty visualization evaluation. IEEE Trans Vis Comput Graphics 2018;25(1):903–13.
[26] Padilla LM, Powell M, Kay M, Hullman J. Uncertain about uncertainty: How qualitative expressions of forecaster confidence impact decision-making with uncertainty visualizations. Front Psychol 2021;3747.
[27] Butson CR, Tamm G, Jain S, Fogal T, Krüger J. Evaluation of interactive visualization on mobile computing platforms for selection of deep brain stimulation parameters. IEEE Trans Vis Comput Graphics 2012;19(1):108–17.
[28] Guo S, Du F, Malik S, Koh E, Kim S, Liu Z, et al. Visualizing uncertainty and alternatives in event sequence predictions. In: Proceedings of the 2019 CHI conference on human factors in computing systems. 2019, p. 1–12.
[29] Ahn Y, Lin YR. Fairsight: Visual analytics for fairness in decision making. IEEE Trans Vis Comput Graphics 2019;26(1):1086–95.
[30] Padilla LM, Creem-Regehr SH, Hegarty M, Stefanucci JK. Decision making with visualizations: A cognitive framework across disciplines. Cogn Res: Princ Implic 2018;3(1):1–25.
[31] Keim DA, Kriegel HP. Visualization techniques for mining large databases: A comparison. IEEE Trans Knowl Data Eng 1996;8(6):923–38.
[32] Ingram S, Munzner T, Olano M. Glimmer: Multilevel mds on the GPU. IEEE Trans Vis Comput Graphics 2008;15(2):249–61.
[33] Liii Pearson K. On lines and planes of closest fit to systems of points in space. Lond Edinb Dublin Philos Mag J Sci 1901;2(11):559–72.
[34] Sedlmair M, Munzner T, Tory M. Empirical guidance on scatterplot and dimension reduction technique choices. IEEE Trans Vis Comput Graphics 2013;19(12):2634–43.
[35] Rubio-Sánchez M, Raya L, Diaz F, Sanchez A. A comparative study between RadViz and star coordinates. IEEE Trans Vis Comput Graphics 2015;22(1):619–28.
[36] Munzner T. Visualization analysis and design. CRC Press; 2014.

[37] Holten D, Van Wijk JJ. Evaluation of cluster identification performance for different PCP variants. In: Computer graphics forum, vol. 29. Wiley Online Library; 2010, p. 793–802.
[38] Li J, Martens JB, Va. Wijk JJ. Judging correlation from scatterplots and parallel coordinate plots. Inform Visual 2010;9(1):13–30.
[39] Zhao Y, Luo F, Chen M, Wang Y, Xia J, Zhou F, et al. Evaluating multi-dimensional visualizations for understanding fuzzy clusters. IEEE Trans Vis Comput Graphics 2018;25(1):12–21.
[40] Fuchs J, Isenberg P, Bezerianos A, Keim D. A systematic review of experimental studies on data glyphs. IEEE Trans Vis Comput Graphics 2016;23(7):1863–79.
[41] D. Oliveira MF, Levkowitz H. From visual data exploration to visual data mining: A survey. IEEE Trans Vis Comput Graphics 2003;9(3):378–94.
[42] Kuang X, Zhang H, Zhao S, McGuffin MJ. Tracing tuples across dimensions: A comparison of scatterplots and parallel coordinate plots. In: Computer graphics forum, vol. 31. Wiley Online Library; 2012, p. 1365–74.
[43] Brown ET, Ottley A, Zhao H, Lin Q, Souvenir R, Endert A, et al. Finding waldo: Learning about users from their interactions. IEEE Trans Vis Comput Graphics 2014;20(12):1663–72.
[44] Ottley A, Yang H, Chang R. Personality as a predictor of user strategy: How locus of control affects search strategies on tree visualizations. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems. 2015, p. 3251–4.
[45] Lallé S, Toker D, Conati C. Gaze-driven adaptive interventions for magazine-style narrative visualizations. IEEE Trans Vis Comput Graphics 2019.
[46] Lefcourt HM. Locus of control: current trends in theory & research. Psychology Press; 2014.
[47] Costa P, R. McCrae R. The revised neo personality inventory (NEO-PI-R). In: The SAGE handbook of personality theory and assessment, vol. 2. 2008, p. 179–98. http://dx.doi.org/10.4135/9781849200479.n9.
[48] Goldberg L. The structure of phenotypic personality traits. Am Psychol 1993;48:26–34. http://dx.doi.org/10.1037/0003-066X.48.1.26.
[49] Green TM, Fisher B. Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction. In: 2010 IEEE symposium on visual analytics science and technology. IEEE; 2010, p. 203–10.
[50] Sheidin J, Lanir J, Conati C, Toker D, Kuflik T. The effect of user characteristics in time series visualizations. In: Proceedings of the 25th international conference on intelligent user interfaces. 2020, p. 380–9.
[51] Cashman D, Wu Y, Chang R, Ottley A. Inferential tasks as a data-rich evaluation method for visualization. In: EVIVA-ML: IEEE VIS workshop on evaluation of interactive visual machine learning systems, vol. 7. 2019.
[52] Ziemkiewicz C, Crouser RJ, Yauilla AR, Su SL, Ribarsky W, Chang R. How locus of control influences compatibility with visualization style. In: 2011 IEEE conference on visual analytics science and technology. IEEE; 2011, p. 81–90.
[53] Ziemkiewicz C, Ottley A, Crouser RJ, Yauilla AR, Su SL, Ribarsky W, et al. How visualization layout relates to locus of control and other personality factors. IEEE Trans Vis Comput Graphics 2012;19(7):1109–21.
[54] Oscar N, Mejía S, Metoyer R, Hooker K. Towards personalized visualization: Information granularity, situation, and personality. In: Proceedings of the 2017 conference on designing interactive systems. 2017, p. 811–9.
[55] Ziemkiewicz C, Kosara R. Preconceptions and individual differences in understanding visual metaphors. Comput Graph Forum 2009;28(3):911–8. http://dx.doi.org/10.1111/j.1467-8659.2009.01442.x.
[56] Ozer DJ, Benet-Martinez V. Personality and the prediction of consequential outcomes. Annu Rev Psychol 2006;57:401.
[57] Riaz MN, Riaz MA, Batool N. Personality types as predictors of decision making styles. J Behav Sci 2012;22(2).
[58] Leiba-O'Sullivan S. The distinction between stable and dynamic cross-cultural competencies: Implications for expatriate trainability. J Int Bus Stud 1999;30(4):709–25.
[59] Walczuch R, Lundgren H. Psychological antecedents of institution-based consumer trust in e-retailing. Inform Manag 2004;42(1):159–77.
[60] Dewberry C, Juanchich M, Narendran S. Decision-making competence in everyday life: The roles of general cognitive styles, decision-making styles and personality. Pers Individ Differences 2013;55(7):783–8.
[61] LePine JA, Colquitt JA, Erez A. Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. Personnel Psychol 2000;53(3):563–93.
[62] LePine JA, Hollenbeck JR, Ilgen DR, Hedlund J. Effects of individual differences on the performance of hierarchical decision-making teams: Much more than g. J Appl Psychol 1997;82(5):803.
[63] Ones DS, Viswesvaran C. Relative importance of personality dimensions for expatriate selection: A policy capturing study. Hum Perform 1999;12(3–4):275–94.
[64] Svendsen GB, Johnsen JAK, Almås-Sørensen L, Vittersø J. Personality and technology acceptance: The influence of personality factors on the core constructs of the technology acceptance model. Behav Inform Technol 2013;32(4):323–34.
[65] Devaraj S, Easley RF, Crant JM. Research note-how does personality matter? relating the five-factor model to technology acceptance and use. Inf Syst Res 2008;19(1):93–105.
[66] Cost Jr PT, McCrae RR. The revised NEO personality inventory (NEO-PI-R). Sage Publications, Inc; 2008.
[67] Koch J, Eisend M, Petermann A. Path dependence in decision-making processes: Exploring the impact of complexity under increasing returns. BuR - Bus Res 2009;2(1):2. http://dx.doi.org/10.1007/BF03343529.
[68] Alvarez J, Djaouti D, et al. An introduction to serious game definitions and concepts. Serious Games Simul Risks Manag 2011;11(1):11–5.
[69] Keijser X, Ripken M, Mayer I, Warmelink H, Abspoel L, Fairgrieve R, et al. Stakeholder engagement in maritime spatial planning: The efficacy of a serious game approach. Water 2018;10(6):724.
[70] Lalicic L, Weber-Sabil J. Stakeholder engagement in sustainable tourism planning through serious gaming. Tour Geographies 2021;23(1–2):185–205.
[71] Willett W, Heer J, Agrawala M. Scented widgets: Improving navigation cues with embedded visualizations. IEEE Trans Vis Comput Graphics 2007;13(6):1129–36.
[72] Endert A, Chang R, North C, Zhou M. Semantic interaction: Coupling cognition and computation through usable interactive analytics. IEEE Comput Graph Appl 2015;35(4):94–9.
[73] Johansson J, Forsell C. Evaluation of parallel coordinates: Overview, categorization and guidelines for future research. IEEE Trans Visual Comput Graph 2015;22(1):579–88.
[74] Emerson JW, Green WA, Schloerke B, Crowley J, Cook D, Hofmann H, et al. The generalized pairs plot. J Comput Graph Statist 2013;22(1):79–91.
[75] Wong PC, Bergeron RD. 30 Years of multidimensional multivariate visualization. Sci Vis 1994;2:3–33.
[76] Pajer S, Streit M, Torsney-Weir T, Spechtenhauser F, Möller T, Piringer H. Weightlifter: Visual weight space exploration for multi-criteria decision making. IEEE Trans Vis Comput Graphics 2016;23(1):611–20.
[77] Gratzl S, Lex A, Gehlenborg N, Pfister H, Streit M. Lineup: Visual analysis of multi-attribute rankings. IEEE Trans Vis Comput Graphics 2013;19(12):2277–86.
[78] Siirtola H, Räihä KJ. Interacting with parallel coordinates. Interact Comput 2006;18(6):1278–309.
[79] Riehmann P, Opolka J, Froehlich B. The product explorer: Decision making with ease. In: Proceedings of the international working conference on advanced visual interfaces. 2012, p. 423–32.
[80] Dimara E, Bezerianos A, Dragicevic P. The attraction effect in information visualization. IEEE Trans Vis Comput Graphics 2016;23(1):471–80.
[81] Russell MT, Cattell RB, Cattell A, Cattell HE, Karol DL. 16PF fifth edition administrator's manual. Institute for Personality and Ability Testing, Incorporated; 1994.
[82] Lima M, Simes A. Neo-Pi-R manual profissional. Lisboa: CEGOC; 2000.
[83] de Alwis DV, Kon CH. A new way to use the Ishihara test. J Neurol 1992;239(8):451–4.
[84] Spearman C. The proof and measurement of association between two things. Am J Psychol 1904;15:88–103.
[85] Abdi H, Williams LJ. Tukey's honestly significant difference (HSD) test. In: Encyclopedia of research design, vol. 3, no. 1. 2010, p. 1–5.
[86] Brown JD. Point-biserial correlation coefficients. Statistics 2001;5(3):12–6.
[87] Dinesen PT, Nørgaard AS, Klemmensen R. The civic personality: Personality and democratic citizenship. Political Stud 2014;62:134–52.
[88] Gideon RA, Hollister RA. A rank correlation coefficient resistant to outliers. J Amer Statist Assoc 1987;82(398):656–66.
[89] Mair P, Wilcox R. Robust statistical methods in r using the WRS2 package. Behav Res Methods 2020;52(2):464–88.