Human-centered Information Visualization Adaptation Engine

Panayiotis Andreou
UCLan Cyprus/InSPIRE
Larnaka, Cyprus
pgandreou@uclan.ac.uk

Christos Amyrotos
UCLan Cyprus/InSPIRE
Larnaka, Cyprus
amyrotos@gmail.com

Panagiotis Germanakos
SAP SE/InSPIRE
Waldorf, Germany
panagiotis.germanakos@sap.com

Irene Polycarpou
UCLan Cyprus/InSPIRE
Larnaka, Cyprus
ipolycarpou@uclan.ac.uk

ABSTRACT

Data Analytics is the art of turning data into insights for efficient and effective business decisions. Data visualization is among the most powerful tools in the data analyst's arsenal, enabling the transformation of data into effective visualizations that can be easily comprehended. However, its effectiveness is often affected by the data analysts' experience and their ability to quickly understand and interpret information. Even though business analytics tools have made a significant progress to deliver immersive data visualization environments for improving users' efficiency and effectiveness, they still do not consider the individual differences in the core process that influences the visualization structure, encoding, and readability.

This paper leverages the users' individual differences to deliver a novel human-centered by-design adaptation engine for business users. The adaptation engine aims to improve the comprehension of data visualizations by delivering personalized content (visualization type and adaptation of visual elements), which in turn leads to improved accuracy and time-to-action efficiency. The proposed adaptation mechanism is evaluated using 45 professional business analysts from multiple industry sectors. The results suggest that individual differences can play an important role in the adaptation process of data visualizations enhancing analysts' comprehensibility and decision making.

CCS CONCEPTS

• **Human-centered computing** → *Visualization toolkits*; **Visual analytics**; **Information visualization**; **User models**; **User centered design**; **Visualization techniques**; • **Information systems** → **Personalization**.

KEYWORDS

information visualization, personalization, individual differences, rule-based adaptation

UMAP '23, June 26–29, 2023, Limassol, Cyprus

ACM ISBN 978-1-4503-9932-6/23/06.
https://doi.org/10.1145/3565472.3592961

ACM Reference Format:
Panayiotis Andreou, Christos Amyrotos, Panagiotis Germanakos, and Irene Polycarpou. 2023. Human-centered Information Visualization Adaptation Engine. In *UMAP '23: Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization (UMAP '23), June 26–29, 2023, Limassol, Cyprus.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3565472.3592961

1 INTRODUCTION

The last decade has witnessed a phenomenal growth in the volume of information and data science, revolutionizing many industry domains [35]. With such large volumes of data being generated, companies that want to stay competitive in today's data driven market adopt Business Intelligence and Analytics (BI&A) software [22]. These software support the full data lifecycle, from raw data to visualizations, delivering actionable insights to decision makers [23].

Recently, BI&A platforms adopted techniques, such as self-service analytics, to empower non-expert analysts to seamlessly utilize all the facilities of the BI&A environment [8]. Despite the assistance of such facilities, the non human-centered one-size-fits-all approach adopted for delivering data visualizations [28] may disorient users since they are faced with an abundance of features to select and customize. More specifically, the rendered data visualizations are solely based on hard-coded user preferences (e.g., preferred visualization types, color themes), the selected dataset's metadata (e.g., categorical vs. numerical data) or the current analysis task (e.g., time-series analysis), not considering the user's requirements or individual differences [24].

Research on individual differences in visualization is progressively growing, demonstrating that interaction with data visualizations can be affected by the individual user's cognitive abilities (e.g., visual working memory [31]), cognitive styles (e.g., field dependent [30]), personality factors (e.g., extraversion [14]), and expertise/experience [19, 20]. Recent works [2, 3, 37] have started exploring the impact of individual differences of business users on the performance of understanding and interpreting data visualizations in the business domain for taking prompt and actionable decisions. These works identify the need for the design and development of adaptive data visualization systems that consider the unique characteristics and requirements of business analysts.

This paper presents a novel human-centered adaptation engine, which aims to enhance the comprehensibility of data visualizations to improve user accuracy and time-to-action efficiency. The adaptation engine adopts an ensemble system with a fuzzy rule inference engine to predict the best fit visualization type and style (i.e., set of visual elements) for business data analysts. The studies that supported the design, development, deployment, and integration of the adaptation engine took place between Sep'2021 and Nov'2022. Firstly, the user model of all users was extracted, recording human factors in individual differences, demographics, experience, and expertise, using psychometric tests and questionnaires based on our previous work in [2, 33]. Next, the fuzzy rules were constructed based on several user studies that assessed the impact of cognitive factors on the understandability of data visualizations (see Sections 4.3 and 4.4). These rules were then combined by the inference engine to produce the adapted content (see Section 3).

The adaptation engine was evaluated using 45 professional analysts from multiple industry sectors. Our evaluation results show that the delivered adaptation improves: (i) user's performance (i.e., time taken to address an analysis task) by an average of 8.1s; (ii) task accuracy (i.e., correctness of analysis task response), where 62% of users were more accurate; and (iii) usability, by improving perceived user experience by 9% .

The rest of the paper is structured as follows: Section 2 presents the system model, consisting of the core adaptation engine inputs: user model, analysis tasks, data, and data visualizations. Next, Section 3 introduces the adaptation engine and its two phases: generation of adaptation rules and adaptation process. The rule extraction process is further described in Section 4. Section 5 presents the user study that was conducted to evaluate the adaptation engine's impact on user's efficiency and effectiveness. The results of the study are presented in Section 6. Then, Section 7 presents related work on data visualization adaptation based on individual differences. Finally, Section 8 concludes the paper.

2 SYSTEM MODEL

In this section, we provide an overview of the system model. The system consists of a set of data analyst users $\{u_1, u_2, ..., u_N\} \in U$, and for each user u_i the system maintains a user model $um(u_i)$ to store the user's characteristics (e.g., demographics, psychometric indicators). We assume that each user u_i is assigned a subset of the organization's data analysis tasks $T' \subset T = \{t_1, t_2, ..., t_M\}$ that need to be addressed through appropriate explorations. We also assume that a task t_j is atomic (i.e., not composed of several sub-tasks) and will require the construction of a single query q that will retrieve the required data d to address the task. The data will be visualized using appropriate visualizations within the user interface.

User Model: The user model of a user $um(u_i)$ is a set of triplets of the form (ct, ch, val), where ct represents the category (e.g., d=demographics, p=psychometric characteristics), ch represents the characteristic that belongs to the triplet's category (e.g., age for demographics), and val represents the respective value for that characteristic (e.g., 35 for age). An example of a user model for user u_i can be $um(u_i) = \{(d, age, 35), (p, wm, low), (p, fdi, fd)\}$, denoting that the user u_i has an age of 35, a low Working Memory and is classified as field-dependent.

Tasks: A task ($t_j \in T$) represents a business question, such as "Identify if the glass bottles pack type is growing in terms of sales value in August of 2021 compared to June 2021.", that needs to be addressed through appropriate explorations. A task is a tuple $t_j = (text, ty_i)$ consisting of: (i) the narrative ($text$); and (ii) the task's type ($ty_i \in TY$), as summarised in Table 1 (following the work of Amar et al. [1], which presented a set of low-level analytical tasks that largely capture people's activities while interacting with information visualization tools). An example of a task t_j is ($text$ =, "Identify if all Soft Drinks sales were affected by seasonality in 2019", fan), where the type of the task ty_i is $fan = FindAnomaly$).

Data: The system employs an information retrieval engine that can support the users' data explorations. In the context of this work, we assume that the system maintains a number of high-quality prepared datasets. The datasets can be accessed via a query engine, which can specify queries in a supported language (e.g., q ="Retrieve Month, SalesValue From SalesDataset").

Data Visualizations: The system maintains a Data Visualization Engine that is responsible for rendering all data visualizations. All data visualizations $v_k \in V$ have a set of adaptable visual elements ($ve_i \in VE$) (e.g., the color and width of a bar, enabling/disabling grids), which can be enabled by the adaptation engine for delivering the desired adaptation/personalization. Table 2 lists the set of visual elements VE and their applicability on each of the available data visualization types V. In order to render a data visualization, the Data Visualization Engine requires: (i) the data returned by the query engine; (ii) the type of data visualization v_i to render; and (iii) the set of visual element modifications ve_i as selected by the Adaptation Engine.

Table 1: Available System Task Types (TY)

Task Name	Task Descriptions [1]
Retrieve Value (rvl)	Find and retrieve attribute values.
Compute Derived Value (cdv)	Compute a summary (numeric) of data.
Find Anomalies (fan)	Identify any anomalies within a given set of data cases with respect to a given relationship or expectation, e.g., statistical outliers.
Correlate (cor)	Given a set of data cases and two attributes, determine useful relationships between the values of those attributes.
Simple Comparison (com) (not in [1])	Simple data value comparison e.g., finding the lowest or the highest value in the dataset.

3 ADAPTATION ENGINE

The adaptation engine is responsible for delivering the best fit data visualization to the analyst user by utilizing the user's model $um(u_i)$, an analysis task t_j, and two sets of adaptation rules. The rules are used for selecting: (i) the best fit data visualization type and (ii) a set of visual element modifications that further personalize the data visualization. The construction and utilization of the two rule sets are detailed in the subsequent sections.

3.1 Adaptation Rules

The adaptation rules are fuzzy *if-then* rules that are triggered during the adaptation process. They are stored in two sets: (i) rules for adapting visualization types ($ARVT$) and (ii) rules for adapting

Table 2: Table of Adaptive Elements VE and their applicability to Data Visualization Types

ve_i	Description	Visualization Type*					
		B	C	L	R	P	T
hgl	Enables horizontal and vertical grid lines	×	×	×	×		
$cpt1$	Switches to color palette 1 (duller colors compared to $cpt2$)	×	×	×	×	×	
$cpt2$	Switches to color palette 2 (brighter colors compared to $cpt1$)	×	×	×	×	×	
dt	Enables dark background and white text (dark theme)	×	×	×	×	×	×
$esiz$	Changes the default size of primary elements (bars, columns, lines)	×	×	×	×		
$prox$	Changes the default proximity between primary elements (bars and columns)	×	×				
dl	Displays data values on top of elements (e.g., bars)	×	×	×	×	×	

*Visualization types: B:Bar, C:Column, L:Line, R:Radar, P:Pie, T:Table

visualization elements ($ARVE$). Example rules from the $ARVT$ and $ARVE$ rule sets can be seen in Table 3 and Table 4, respectively.

The rules consist of several metadata attributes that allow them to be triggered according to the user, task, and data characteristics. For example, the *Factor* and *Level* columns in both rule sets refer to user's classification for a human factor (e.g., Working Memory - High). The rule sets are also augmented with several other metadata attributes (e.g., TaskType) to cater for filtering each of the rule sets prior to adaptation. This structure is flexible to cater for future metadata enhancements. The filtered rules are then combined to recommend the best fit visualization type.

Similarly, the $ARVE$ rule set contains the *ChartType* and *Element* columns, which refer to the chart type and visual element modification that a specific rule applies. The adaptation engine combines the filtered $ARVE$ rule set for determining the adaptation of a specific visual element (e.g., dark theme) on the selected data visualization.

Table 3: Example Adaptation Rules for Visualization Type (ARVT)

Factor	Level	TaskType	Vis. Type* Weight (%))					
			B	R	C	L	P	T
WM	High	fan	4	0	96	0	0	0
WM	High	fan	0	92	0	4	0	4

*Visualization types: B:Bar, C:Column, L:Line, R:Radar, P:Pie, T:Table

The selected representation of the rules is twofold: (i) to facilitate *efficient execution* of the rules and (ii) to improve the *explainability* of the inference process carried out by the system. A simplified example of the first rule in Table 3 is presented below: IF Factor="WM" and Level="High" and TaskType="Find Anomaly" THEN VisualizationType=Column=95%, Bar=4%.

Table 4: Example Adaptation Rules for Visual Elements (ARVE) for Dark Theme (dt)

Factor	Level	ChartType	Element	ENABLE	DISABLE
WM	High	Bar	dt	60 %	40 %
WM	High	Line	dt	32 %	68 %

3.2 Adaptation Procedure

The adaptation procedure can be conceptually illustrated as the function $ap(u_i, t_j, q) = (d, v_k, VE')$. The adaptation procedure function receives a query q for a task t_j by user u_i, and returns the desired data and visualization instructions in the form of a triplet that contains the data d, best fit data visualization type (v_k), and any visual element modifications ($VE' \subset VE$). The output of this function is used as input to the Data Visualization Engine. The detailed steps of the process are presented in Algorithm 1.

The procedure starts by retrieving the data d for query q (line #1) and the task metadata tm (line #2). If the task has a predefined data visualization (line #3) the algorithm stops execution and returns the pre-selected data visualization type for the task ($tm.predefVisType$), the retrieved data, and a default set of visual element adaptations $\{default\ elements\}$ (line #4). Otherwise, adaptation is initiated by first retrieving the user's model ($um(u_i)$) (line #6) and subsequently identifying the best fit data visualization type (v_k) by combining adaptation rules. This is done by first filtering down the $ARVT$ rule set (line #7) according to the user's model (um) and task metadata (tm). The resulting $ARVT'$ rules set is forwarded to the fuzzy classifier that aggregates the votes for each of the available visualization types and returns the *winner* data visualization (line #8).

Algorithm 1 Adaptation Algorithm $ap(u_i, t_j, q)$

1: $d \leftarrow queryEngine(q)$
2: $tm \leftarrow getTaskMetadata(t_j)$
3: **if** Not $tm.predefVisTypeisnull$ **then**
4:     **return** $(d, tm.predefVisType, \{default\ elements\})$
5: **end if**
6: $um \leftarrow um(u_i)$
7: $ARVT' \leftarrow filterARVT(um, tm)$
8: $v_k \leftarrow visTypeClassifier(ARVT')$
9: $VE' \leftarrow \emptyset$
10: **for each** $el \in elements$ **do**
11:     $ARVE' \leftarrow filterARVE(um, v_k, el)$
12:     $enableElement \leftarrow visElementVoter(ARVE')$
13:     **if** $enableElement$ **then**
14:         $VE' \leftarrow VE' \cup \{el\}$
15:     **end if**
16: **end for**
17: **return** (d, v_k, VE')

The algorithm then selects a number of visual element modifications ($VE' \subset VE$) that are specific to the winner data visualization. Initially an empty set (VE') is initialized (line #9) for maintaining visual element modifications that will be applied to the selected data visualization type. Line #10 begins examining all possible visual element modifications (VE) supported by the visualization engine iteratively, and filters rules from the $ARVE$ rule set using the user's model (um), the winner data visualization type (v_k), and the visual element (el) that is to be voted. The filtered set of rules ($ARVE'$) is forwarded to the voter function (line #12) that combines the votes and decides if the current visual element modification el will be enabled or disabled. The result of the vote is added in the VE' set (line #14). The procedure finishes once all elements are evaluated (line #16) and returns the data (d), winner data visualization type (v_k), and the visual element modifications set (VE') (line #17). The result will be forwarded to the Visualization Engine for rendering.

4 USER STUDY – PHASE A: ADAPTATION RULES EXTRACTION

This section describes how the adaptation rule sets and user models were extracted through a user study that analyzed the interactions (performance and accuracy) of 60 business data analyst participants with different data visualizations types and visual elements.

4.1 Study Setup

Participants: The study recruited 60 business data analyst participants from two industry organizations (RAI Consultants LTD[1] and KPMG Cyprus[2]) working in diverse industry fields such as Retail, and Consumer Marketing, Advisory Services, Audit, and Risk Assessment, who had on average at least 2 years of experience in the field of data analytics and were using data visualizations on a daily basis. Moreover, the sample consisted of 30 male and 30 female participants, with age ranging from 24 to 57 (mean age 33.9 ± 7.8). All participants had varying expertise levels, including decision makers and managers, and executive-, senior- and junior-analysts.

User Modeling: The user model included human factors in individual differences, demographics, experience, and expertise, of all users. The parameters were extracted using psychometric tests and questionnaires based on our previous work in [2, 33]. The resulting model consisted of the Field Dependent-Independent (FDI) cognitive style construct, cognitive processing abilities (control of attention, speed of processing, visual working memory) and demographic information (gender, age, educational status, and experience/expertise characteristics.

Dataset: To cater for the diverse user analysts' expertise, a synthetic sales dataset was constructed, consisting of comic book sales with typical dimensions, such as time, product characteristics, location characteristics, and distributive and algebraic measures, such as quantity, price, amount, average price, and weighted price.

Data Visualization Types: According to our previous findings [2] we opted for the most frequently used data visualizations of our user group, which included the Bar Chart, Line Chart, Pie Chart, Column Chart, Data Table, and Radar Chart.

[1]https://www.rai.com.cy/
[2]https://home.kpmg/cy/

Data Visualization Tasks: Using the synthetic comic book sales dataset, we produced 160 visual exploration tasks, each requiring participants to interact with a data visualization and answer a question. In the background, the system recorded the participant's response time (ms) and assessed whether the response was correct. The tasks were organized into four experiments: (i) *Chart Type* experiment; (ii) *Task Complexity* experiment; (iii) *Dimensionality* experiment; and (iv) *Visual Elements* experiment, to capture user interactions with visualizations of varying types (see Table 1), complexity, dimensionality, and appearance, respectively. Tasks were delivered using default visual settings (i.e., no modifications on visual elements) and acted as control tasks to the Visual Elements experiment. The tasks for the Visual Elements Experiment were divided into seven sets, each one introducing a visual element modification ($ve_i \in VE$) as seen in Table 2.

4.2 Study Procedure

Due to the implications of the national restrictions in response to the COVID-19 pandemic, it was decided that the study had to be conducted in a remote manner. The study received ethics approval from Cyprus National Bioethics Committee and followed all the relevant protocols. The study was conducted for seven days and participants had to complete all four experiments, consisting of 160 visual exploration tasks. Each participant was given access to the experiment platform. Once an experiment began, the platform loaded the analysis tasks and presented them in a random order. A task was completed once the participant provided a response. One important constraint enforced was that once an experiment started it could not be stopped until all tasks were addressed. At the beginning, users were given instructions on how to set their screen environment, such as setting the minimum screen size and screen resolution, to ensure that the study experience was as identical as possible across different participants. Moreover, prior to being able to engage with the experiments, instructions were provided to the participants describing the experiment process and the approximate amount of time required to complete each experiment. Finally, all participants were given a set of training analysis tasks similar to those of the four experiments.

4.3 Adaptation Rules for Visualization Types

It must be noted that for generating the $ARVT$ rule set, only responses from the three control experiments were used (i.e., responses for the Visual Elements experiment were excluded). The following five steps describe the procedure performed to extract rules for every human factor group. To facilitate our description, we use as an example a single group of participants, the ones with High Working Memory, as the process is identical for all other groups. Step 1 started by filtering the responses to the selected group of participants. In Step 2, the response time values were aggregated, according to the task's metadata and chart type, to produce the average response time of each participant on every data visualization type, task type, and data characteristics. Step 3, ranked the data visualization types for each participant at different task configuration levels, such as task type and data characteristics. Ranking was based on the average response time the participant achieved when using a specific data visualization under a set of specific analysis task configuration. Step 4 the results were further filtered to the data visualization with the highest rank in combination with the analysis task configuration. The results were then summarized for all participants leading to six records, each representing the score of each of the six available data visualizations used for a specific analysis task configuration. Essentially, the score in this context represents the number of times a specific data visualization had the best performance in terms of time response (in milliseconds) for a specific analysis task configuration across all participants. In the final step, the data visualization scores were normalized to the range [0..1], assuming that all rules represent homogeneous fuzzy rules (i.e., all rules are of equal weight). The resulting $ARVT$ was used by the adaptation engine to select the best fit data visualization as described in Section 3.

4.4 Adaptation Rules for Visual Elements

In this section, we define the rule extraction process of the $ARVE$ set, which utilized the responses captured from both, the control experiments and the visual elements experiment. For instance, rule extraction for the "dark theme" visual element/setting required the utilization of responses from the visual elements experiment, which had "dark theme" enabled. All corresponding analysis tasks from the rest of the experiments acted as control analysis tasks. Similarly to the extraction of adaptation rules for Data Visualization Types, the $ARVE$ rule set was generated in five steps, aggregating the response times of participants when a visual element was enabled and disabled, across all data visualization types. The resulting rules were used by the adaptation engine to configure the best fit data visualization elements as described in Section 3.

5 USER STUDY – PHASE B: ADAPTATION ENGINE EVALUATION

This section presents a user study that evaluates the use of personalized data visualizations to improve the efficiency and effectiveness of business data analysts. The adaptation engine utilized the adaptation rules generated in the previous section. The study collected the following evaluation metrics: (i) performance and accuracy of participants when addressing visual analysis tasks; and (ii) perceived user experience. For capturing the adaptation engine's impact we followed a within-subjects study design, enabling us to record all evaluation metrics when the participant navigated over: (i) the original non-adapted/personalized content, which included analysis tasks with predefined data visualizations without any alterations or enhancements and (ii) data analysis tasks that included dynamically adapted/personalized data visualizations. Our null hypotheses are the following:

H0: The performance of the participants in terms of milliseconds taken to address an analysis task, between the two conditions (i.e., adaptation enabled and disabled) will not be significantly different.

H1: The accuracy of the participants in terms of the total number of tasks addressed correctly, between the two conditions (i.e., adaptation enabled and disabled) will not be significantly different.

H2: The system's user experience score between the two conditions (i.e., adaptation enabled and disabled) will not be significantly different.

5.1 Study Setup

Participants: For this study, 45 business data analyst participants were recruited (24 male and 21 female) with their age ranging between 25 and 60 years (mean age 35.4 ± 8.8). On average, the participants had at least 2 years of experience in the field of data analytics. 21 out of the 45 participants volunteered from our previous user study (Phase A), described in Section 4. The new participants were recruited from the same organizations and adhered to the same participation preconditions as in the above mentioned user study. It must be noted that the new 24 participants were given enough time to complete the necessary questionnaires/test to construct their user models (i.e., perform all psychometric tests and answer all questionnaires).

Dataset: Provided that some study participants were also part of the previous study, a new dataset was constructed in collaboration with the partner organizations to prevent the learning effect of repeated experiments. A real dataset consisting of soft drinks sales was selected to be used for the study. The dataset comprised of 19 attributes[3] and three dimensions (time, product, and outlet), and consisted of 731,446 real sale transactions recorded over a period of three years (2019-2021). The outlet names, brands, and product names of the dataset were anonymized upon request of the providing organization.

Analysis Tasks: Using the realistic sales dataset, 38 analysis tasks were created in pairs of equal complexity, resulting in 19 pairs of tasks. Moreover, each pair of tasks had a specific analysis task type assigned using the taxonomy of tasks presented in Table 1. Each pair included: (i) the control non-adapted/non-personalized task generated using a specific/predefined visualization type applicable to the task; and (ii) the personalized data visualization generated by the adaptation engine for each participant and task. The data visualization types used for constructing the set of tasks for this experiment are the data visualization types used during the previous user study in Section 4, which were identified as the most commonly used by the user group. In the background, the system recorded the time required for the participant to interact and respond to the task (in milliseconds) and assessed whether the response of the participant was correct.

Study Material: Besides recording performance and accuracy for each task, the study aimed at capturing the users' experience factor. To this end, we utilized two accredited system evaluation questionnaires, which we combined into a web-based questionnaire. Specifically, for measuring the participants' user experience we used the User Experience Questionnaire Short Version (UEQ-S) [34]. According to the questionnaire's authors, this questionnaire's scales "cover a comprehensive impression of user experience. Both classical usability aspects (efficiency, perspicuity, dependability) and user experience aspects (originality, stimulation) are measured".

5.2 Study Procedure

Due to the implications of the national restrictions in response to the COVID-19 pandemic, we decided that the study was conducted online. Following this decision, we also ensured that we maintained

[3]Date, Year, Month, Day, Quarter, Brand, Product Name, Promotion, Pack Type, Diet, Outlet Name, Outlet Type, Urban Or Rural, Area Name, Size, Price, Quantity, Sales Value and Sales Volume

(a) the *ecological validity* of the study, since the experimental design, procedure, and setting approximated the real-life context that is under investigation [4], and (b) its *internal validity* as the accuracy of the data and the respective conclusions are drawn from users who were a good fit to the subject of the investigation (given their business roles and daily business routine) [10].

At the beginning, all participants were invited to an online meeting where the study and its goals were introduced. It was decided that during the study participants would be assigned the role of a Brand Manager employed at a fictitious company that sells soft drinks of the brand "IdealCola". A training session was conducted, allowing participants to experience the system by addressing a few demo analysis tasks through the user interface.

At the end of the training session, a recorded version of the presentation, system demonstration videos, guidelines and access credentials to access the platform were shared with the participants. This served as a second training step, so that each participant could login during their free time and practice an additional set of demonstration analysis tasks. This ensured that all participants were well familiarized with the platform's interface and analysis tools prior to engaging with the study. Moreover, it must be noted that when participants logged in for the first time, they were greeted with a welcome presentation that demonstrated all the features of the data exploration system and also GDPR-related features (e.g., enabling a participant to request the deletion of their information). Finally, since the study was conducted in a remote fashion, appropriate controls were developed to prevent the system from tracking the user's view time/performance in the cases where the participant was distracted by external factors. In particular, if the participant was found to be non-responsive for 30 seconds the platform prompted the user to validate that they are still present, otherwise all recorded metrics were reset and the user had to restart the task. The 30-second interval was decided with the industry partners after performing some tests with tasks of varying complexity.

The rest of the study was divided into two parts. The first part utilized the 19 non-adapted set of analysis tasks (i.e., the system returned the same predefined data visualization for the specific analysis task to all participants). The second part utilized the 19 adapted analysis tasks (i.e., the data visualizations were adapted by the adaptation engine, according to the user model and task characteristics). For each part, a participant could navigate in the list of available tasks (presented in random order), study their question/narrative, and select one that will serve as the current task. The the participant could then navigate to the "Analysis Wizard", a tool which enabled participants to perform data explorations in a step-by-step manner. The functionality of the Analysis Wizard was limited to providing the capabilities required to address the study tasks. In particular, during the exploration process, the participant was required to perform three steps of the Analysis Wizard: (i) select analysis; (ii) select attributes; and (iii) view result (i.e., adapted or non-adapted visualization). Furthermore, the current analysis task's narrative was available to users on the top of the Analysis Wizard to remind them of what was required. Finally, once the participant had an answer for the analysis task, they could navigate back to the list of analysis tasks and provide their answer. During the exploration process, the platform was monitoring the interaction time of participants for all steps and assessed the correctness of their responses.

After completing all analysis tasks the participants were required to complete online questionnaires for assessing the system's user experience.

6 EVALUATION RESULTS

This section presents the results of our evaluation by reporting the impact of data visualization adaptation on the participant's performance, accuracy, and perceived user experience as stated by the Hypotheses in Section 5.

6.1 Impact on Analysis Task Performance

The analysis of performance consisted of comparing the time required for participants to address paired (non-personalized vs. personalized) analysis tasks. At the beginning, the results were filtered to include only pairs where the participant responded accurately to both analysis tasks. An outlier analysis was performed for assessing each participant's and each task's response times to reveal abnormal observations. Our analysis revealed that for Task 13, which featured computation of a derived value, all participants took an extreme amount of time to complete the personalized task. After talking to participants, it was understood that the majority of the participants were not aware of how to approach addressing Task 13. Additionally, some participants reported having to use a calculator to find the correct answer for this task. Consequently, it was decided to remove the task results from the analysis.

The analysis across the two study conditions revealed that adaptation had a positive effect on participants' performance enabling them to achieve an average decrease of 8.1 ± 6.9s with regards to task completion time. Moreover, with adaptation enabled, performance improved for an average of 9 ± 2 tasks per participant (max:15, min:5).

Analysis on the impact of adaptation with regards to performance across different task types suggests that adaptation had a positive effect on participants' performance enabling them to achieve: (i) a statistically significant average decrease of 7.8s for Retrieve Value tasks ($p < .01$); (ii) a statistically significant average decrease of 25.9s for Correlation tasks ($p = .01$); (iii) a statistically significant average decrease of 8.2s on Simple Comparison tasks ($p < .01$); and (iv) a non-statistically significant average decrease of 10.6s on Compute Derived Value tasks ($p = 0.24$).

Since Simple Comparison tasks was the larger group of analysis tasks (10 task pairs), it was decided to further explore this group of tasks by independently analyzing Simple Comparison tasks that used time series data. Results suggest that with adaptation enabled, participants achieved: (i) a statistically significant average decrease of 9.9s on Simple Comparison tasks that used time series data ($p < .01$) and (ii) a statistically significant average decrease of 4.5s on the remaining Simple Comparison tasks ($p < .01$). Moving on, the adaptation enabled performance improvements for an average of 84 ± 82 task responses across all analysis task types (max:199 for Simple Comparison, min:5 for Compute Derived Value tasks). Unfortunately, the sample of our tasks was limited to a single Find Anomaly pair of tasks, for which the majority of participants only responded correctly to the personalized variant of the task, leaving a very small sample of only four responses that could not be incorporated in this analysis.

In conclusion, the above results reveal that when the adaptation engine was enabled, participants' performance was positively affected, and thus, we reject the null hypothesis **H0**.

6.2 Impact on Analysis Task Accuracy

The analysis of accuracy considered the participants' ability to address (non-personalized vs. personalized) analysis tasks correctly. For each of the study conditions a participant was able to achieve a maximum score of 19 (i.e., the total number of tasks). Analyzing the accuracy scores of each participant reveals that 62% of participants were more accurate when addressing analysis tasks with adapted/personalized data visualizations. Moreover, 18% of participants were not affected in terms of accuracy across the two study conditions, while the remaining 20% of participants were negatively impacted by adaptation in terms of accuracy. In contrast to analysis tasks with no data visualization adaptation, participants were able to address on average an additional 8% of analysis tasks correctly when working with personalized visual elements tasks. Analysis of accuracy scores across task types for both conditions revealed that participants were generally much more accurate in addressing tasks when adaptation was enabled for Simple Comparison, Compute Derived Value, and Find Anomaly tasks. Specifically, participants were more accurate by 6.6% for Simple Comparison tasks, 34.2% for Computer Derived Value tasks, and 90% for Find Anomaly tasks. In contrast, for Correlation and Retrieve Value task types we were not able to see a significant impact in terms of accuracy when participants were using adapted/personalized data visualizations for addressing the analysis tasks.

In conclusion, the overall analysis of task accuracy revealed that when the adaptation engine was enabled, participants' accuracy was positively affected, and thus, we reject the null hypothesis **H1**.

6.3 Impact on User Experience

During the evaluation study 35 out of 45 participants provided voluntary responses to the User Experience Questionnaire (UEQ-S) [34], right after they addressed all control analysis tasks (i.e., those tasks with a predefined/non-adapted data visualizations). Additionally, the same 35 participants responded to the two questionnaires right after they had addressed all analysis tasks for which the system produced an adapted data visualization. The collected data was analyzed by an automated process offered by the questionnaire's authors in order to investigate **H2**.

The User Experience Questionnaire has in total 8 scales, 4 scales measuring Pragmatic Quality (a metric that focuses on the task-oriented nature of an experience, e.g., considers the task's efficiency and ease of use) and 4 scales measuring Hedonic Quality (a metric that focuses more on the fun, appeal, and more generally on the originality aspects of the experience offered by a system). Using the responses of all participants we calculated Cronbach's alpha (or coefficient alpha) for each set of scales belonging to each metric (i.e., pragmatic quality and hedonic quality) for data collected in both conditions (i.e., adaptation disabled/enabled). Alpha values for both metrics across the two conditions were higher than 0.7, which is considered acceptable. Generally, scales that belong to the same group should show a high correlation. Therefore, using the Cronbach's alpha, which is a measure for the consistency of a scale [11], helped us ensure that the different scales of the questionnaire were interpreted as intended by the participants. The baseline scores (i.e., adaptation disabled condition) for (i) pragmatic quality was 1.35, (ii) hedonic quality was 0.86, and (iii) the overall user experience was 1.11. Moreover, with adaptation enabled the score for (i) pragmatic quality was increased to 1.45, (ii) hedonic quality was increased to 0.97, and (iii) the overall user experience was increased to 1.21. The user experience scores achieved by the system across the two conditions are above the value of 0.8, and thus, are considered a positive evaluation [34]. The evaluation revealed that enabling data visualization adaptation when participants interact with the given analysis tasks facilitated an increase of their perceived user experience, and thus, we reject the null hypothesis **H2**.

6.4 Discussion

Our user study evaluation shows that the current adaptation engine with the adopted rule generation procedure improves the participants' performance and accuracy across a variety of data analysis tasks. The platform positively affects the participants' perceived user experience score. Furthermore, the fuzzy rule-based classification framework enables the quick integration of new rules based on new data visualization interaction data. Additionally, the fuzzy adaptation logic and the ensemble processing approach used by the adaptation engine makes it easier to combine and utilize in parallel multiple adaptation driving factors (e.g., human factors), which might interact with each other. The characteristics that we used in the user profiles reflect human factors that complement each other and have a specific impact on the type and presentation of data visualization in relation always to the type of task and intent (e.g., working memory monitors the complexity, whereby cognitive style the type). The computational model makes sure that aggregated views (clusters) of specific user profiles are created dynamically, matching specific adaptive conditions. This "cooperative" learning process is optimized in time, improving the quality of the results considering the current dataset. Moreover, the extraction of data visualization interaction data for adaptation rule generation can take place during a user study using appropriate data collection tools, similar to the ones used in this work, or can be be extracted from other sources, such as data analysis tools offering user interaction logging. Overall, the flexibility of the current framework allows for quickly collecting interaction data and efficiently transform it in adaptation rules. The framework is also flexible to be used for experimentation in other domains.

While our work focused on the improvement of the overall efficiency and effectiveness of the business data analyst when addressing data analysis tasks, there are some limitations that we would like to address in the future. The sample of analysis tasks used during evaluation was not balanced in terms of task type, since more focus was given on simpler comparison tasks. Moreover, this work assumes homogeneous weak learners adaptation rules. An alternative approach would be to use a boosting framework by identifying which adaptations/interventions had the most influence on the improvement with regards to accuracy and performance.

Some questions arising from this work that we plan in addressing as part of future endeavors include: (i) How could our approach offer a transparent explanation to the business analyst with regards to why the best fit data visualization was selected? (ii) How can we more effectively process the resulting user's interaction with the adapted output and further gain insight on which adaptation/intervention was the most helpful for that type of user? and (iii) How does our adaptation perform with unexplored data visualizations and analysis task types? Our goal is to attempt to address these questions by first extending our sample of users and gathering more data visualization interaction data that can yield more diverse adaptation rules; thus, facilitating further exploration of the interaction of human factors on data visualizations, but also the exploration of this interaction as a driving force to the current adaptation engine.

7 RELATED WORK

Designing a user adaptive system involves the consideration of three questions: *what to adapt to, when to adapt, and how to adapt* [5]. Our work turns the focus on *what to adapt to* and further explores *how to adapt* aspects.

With regards to *what to adapt to*, research on information visualization reveals that individual differences influence how a user interacts, understands, and utilizes data visualizations for performing analytical tasks. In fact, the growing interest on the effect of individual differences in information visualization resulted in comprehensive survey publications on the subject [21]. A subset of works on individual differences and their effect on information visualization includes, but is not limited to, the exploration of human factors such as cognitive abilities [6, 7, 19, 28, 31, 32, 36, 39], cognitive styles [17, 25, 29, 30, 33], personality traits [9, 14, 38, 39], and expertise/experience [19, 20, 31].

The question of *how to adapt* data visualizations is usually addressed (i) at the visualization type level, i.e., using recommendations for a best fit data visualization or (ii) at the individual visual element level, i.e., applying modifications or additions of visual elements on a data visualization. The works of Gotz et al. [12] and Grawemeyer [13] focused on adaptation with regards to delivering data visualization type recommendations based on user's interaction behavior or task features. On the other hand, visual element modifications are equally important to note. The work of Carenini et al. [5] investigated how the effectiveness of a data visualization (specifically a bar chart) can be increased with four different adaptive interventions. Additionally, in the context of Security Information and Event Management systems, Yelizarov et al. [37] proposed a graph of computer hosts that highlights the most significant hosts (i.e., graph nodes) and dims (using opacity) the rest according to the current cognitive load of the user for increasing efficiency when dealing with system threats. While Yelizarov et al. [37] leveraged the user's cognitive load for adaptation, others have utilized the underlying data for adapting the color of visual elements according to their mapped data category in order to reduce the user's cognitive load [27].

By contrast to the above works, our goal is to build a flexible human-centered by-design adaptation engine that leverages the power of a multidimensional human-centered user model for delivering the best fit data visualization (both in terms of data visualization type and visual element modifications). Specifically, our work targets data analyst users who perform visual data exploration in the context of a business environment, aiming to increase their comprehensibility of information leading to improved accuracy and time-to-action efficiency. Our adaptation engine utilizes a fuzzy rules-based recommendation system based on established techniques for multi classifier fusion [18, 26]. Similar to our approach, fuzzy rules are obtained by computing the grade of certainty of a rule [15, 26] and then a subset is selected based on specific criteria to improve prediction accuracy [16]. However, our approach does not include a step for fuzzy rules refinement as currently there is no assessment of which rules contribute to the best recommendation.

8 CONCLUSIONS AND FUTURE WORK

The paper presented a novel adaptation engine that adopts a fuzzy rule-based classification system, consisting of a fuzzy rule generation procedure and a classification procedure that selects the best-fit data visualization for a user. The engine includes two steps: (i) selection of the best visualization type and (ii) selection of the visualization element modifications to be applied. Within this paper, we presented the architecture of the adaption engine and also the rule extraction process as this was performed using data visualization interaction data that was captured during a user study.

The evaluation of the adaptation engine, using realistic data and 45 business data analysts, revealed that the majority of participants were positively affected by the delivered data visualization adaptation, in terms of their ability to correctly address analysis tasks. Additionally, we found that data visualization adaptation enabled our participants to execute their tasks more efficiently. This latter effect of performance improvement was more evident for simpler types of analysis tasks (e.g., Simple Comparison and Retrieve Value tasks). Moreover, the results show that the perceived user experience factor before and after adaptation was enabled was improved.

The designed adaptation rule extraction process is generic and can be easily expanded to introduce additional factors, more fine-grained factor levels, and new chart types and visual elements. In addition, the user modeling platform can accommodate, with minimum effort, additional parameters for existing (i.e., demographics, experience, and expertise) or new dimensions (e.g., cultural background). This enables the research team to test the proposed framework to other industry domains, generating a larger sample of more diverse adaptation rules. This, in turn, will enable the team to further validate the adaptation approach with more complex data visualizations and adaptive interventions. We also aim to investigate additional factors that contribute to the understandability and comprehension of data visualizations, such as transparency and explainability. Finally, we will investigate the benefits that can be gained if the system adapts to the task and not to the user.

ACKNOWLEDGMENTS

This research was partially funded by the Cyprus RIF under the project IDEALVis (EXCELLENCE/ 0918/ 0366) and the European Union under the project SLICES-PP (No. 101008468).

REFERENCES

[1] Robert Amar, James Eagan, and John Stasko. 2005. Low-Level Components of Analytic Activity in Information Visualization. In *Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization (INFOVIS '05)*. IEEE Computer Society, USA, 15. https://doi.org/10.1109/INFOVIS.2005.24

[2] Christos Amyrotos, Panayiotis Andreou, and Panagiotis Germanakos. 2021. Human-Centred Persona Driven Personalization in Business Data Analytics. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization* (Utrecht, Netherlands) *(UMAP '21)*. Association for Computing Machinery, New York, NY, USA, 175–180. https://doi.org/10.1145/3450614.3462241

[3] Panayiotis Andreou, Christos Amyrotos, Andreas Pamboris, and Panagiotis Germanakos. 2021. RABIT: Reflective Analytics for Business InTelligence. In *CHI Greece 2021: 1st International Conference of the ACM Greek SIGCHI Chapter* (Online (Athens, Greece), Greece) *(CHI Greece 2021)*. Association for Computing Machinery, New York, NY, USA, Article 13, 8 pages. https://doi.org/10.1145/3489410.3489423

[4] Marilynn B. Brewer and William D. Crano. 2014. *Research Design and Issues of Validity* (2 ed.). Cambridge University Press, 11–26. https://doi.org/10.1017/CBO9780511996481.005

[5] Giuseppe Carenini, Cristina Conati, Enamul Hoque, Ben Steichen, Dereck Toker, and James Enns. 2014. Highlighting Interventions and User Differences: Informing Adaptive Information Visualization Support. In *Proc. ACM SIGCHI* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1835–1844. https://doi.org/10.1145/2556288.2557141

[6] Chaomei Chen. 2000. Individual differences in a spatial-semantic virtual environment. *Journal of the American Society for Information Science* 51, 6 (March 2000), 529–542. https://doi.org/10.1002/(SICI)1097-4571(2000)51:6<529::AID-ASI5>3.0.CO;2-F

[7] Chaomei Chen and Mary Czerwinski. 1997. Spatial ability and visual navigation: an empirical study. *New Review of Hypermedia and Multimedia* 3, 1 (July 1997), 67–89. https://doi.org/10.1080/13614569708914684

[8] Claudia Imhoff and Colin White. 2011. Self-Service Business Intelligence Empowering Users to Generate Insights. http://docs.media.bitpipe.com/io_10x/io_106625/item_583281/TDWI_Best_Practices_Report_Self-Service_BI_Q311%5B1%5D.pdf. Online; accessed 26 January 2021.

[9] Cristina Conati and Heather Maclaren. 2008. Exploring the Role of Individual Differences in Information Visualization. In *Proc. AVI* (Napoli, Italy) *(AVI '08)*. Association for Computing Machinery, New York, NY, USA, 199–206. https://doi.org/10.1145/1385569.1385602

[10] Thomas D Cook and D T Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Houghton Mifflin.

[11] Lee J. Cronbach. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 3 (1951), 297–334. https://doi.org/10.1007/BF02310555

[12] David Gotz and Zhen Wen. 2009. Behavior-Driven Visualization Recommendation. In *Proceedings of the 14th International Conference on Intelligent User Interfaces* (Sanibel Island, Florida, USA) *(IUI '09)*. Association for Computing Machinery, New York, NY, USA, 315–324. https://doi.org/10.1145/1502650.1502695

[13] Beate Grawemeyer. 2006. Evaluation of ERST – An External Representation Selection Tutor. In *Diagrammatic Representation and Inference*, Dave Barker-Plummer, Richard Cox, and Nik Swoboda (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 154–167.

[14] Tear M. Green and Brian Fisher. 2010. Towards the Personal Equation of Interaction: The impact of personality factors on visual analytics interface interaction. In *Proc. VAST* (Salt Lake City, UT, USA). IEEE, 203–210. https://doi.org/10.1109/VAST.2010.5653587

[15] Hisao Ishibuchi, Tomoharu Nakashima, and Takehiko Morisawa. 1999. Voting in fuzzy rule-based systems for pattern classification problems. *Fuzzy Sets and Systems* 103, 2 (1999), 223–238. https://doi.org/10.1016/S0165-0114(98)00223-1

[16] Hisao Ishibuchi, Tomoharu Nakashima, and Tadahiko Murata. 2001. Three-objective genetics-based machine learning for linguistic rule extraction. *Information Sciences* 136, 1 (2001), 109–133. https://doi.org/10.1016/S0020-0255(01)00144-X

[17] Marta Koć-Januchta, Tim Höffler, Gun-Brit Thoma, Helmut Prechtl, and Detlev Leutner. 2017. Visualizers versus verbalizers: Effects of cognitive style on learning with texts and pictures – An eye-tracking study. *Computers in Human Behavior* 68 (March 2017), 170–179. https://doi.org/10.1016/j.chb.2016.11.028

[18] Ludmila I. Kuncheva, James C. Bezdek, and Robert P.W. Duin. 2001. Decision templates for multiple classifier fusion: an experimental comparison. *Pattern Recognition* 34, 2 (2001), 299–314. https://doi.org/10.1016/S0031-3203(99)00223-X

[19] Sébastien Lallé, Cristina Conati, and Giuseppe Carenini. 2017. Impact of Individual Differences on User Experience with a Real-World Visualization Interface for Public Engagement. In *Proc. UMAP* (Bratislava, Slovakia) *(UMAP '17)*. Association for Computing Machinery, New York, NY, USA, 369–370. https://doi.org/10.1145/3079628.3079634

[20] Sukwon Lee, Sung-Hee Kim, Ya-Hsin Hung, Heidi Lam, Youn-Ah Kang, and Ji S. Yi. 2016. How do People Make Sense of Unfamiliar Visualizations?: A Grounded Model of Novice's Information Visualization Sensemaking. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 499–508. https://doi.org/10.1109/TVCG.2015.2467195

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

[22] Michael Schrage. 2016. How the Big Data Explosion Has Changed Decision Making. https://hbr.org/2016/08/how-the-big-data-explosion-has-changed-decision-making. Online; accessed 20 January 2021.

[23] Solomon Negash. 2004. Business Intelligence. *Communications of the Association for Information Systems* 13 (Jan. 2004), 177–195. https://doi.org/10.17705/1CAIS.01315

[24] James Richardson, Rita Sallam, Kurt Schlegel, Austin Kronz, and Julian Sun. 2020. *Gartner Magic Quadrant for Analytics and Business Intelligence Platforms*. Retrieved Jan 20, 2021 from https://www.gartner.com/en/documents/3980852/magic-quadrant-for-analytics-and-business-intelligence-p

[25] Richard Riding and Graeme Douglas. 1993. The effect of cognitive style and mode of presentation on learning performance. *British Journal of Educational Psychology* 63, 2 (June 1993), 297–307. https://doi.org/10.1111/j.2044-8279.1993.tb01059.x

[26] Johannes A. Roubos, Magne Setnes, and Janos Abonyi. 2003. Learning fuzzy classification rules from labeled data. *Information Sciences* 150, 1 (2003), 77–93. https://doi.org/10.1016/S0020-0255(02)00369-9

[27] Vidya Setlur and Maureen C. Stone. 2016. A Linguistic Approach to Categorical Color Assignment for Data Visualization. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 698–707. https://doi.org/10.1109/TVCG.2015.2467471

[28] Ben Steichen, Giuseppe Carenini, and Cristina Conati. 2013. User-Adaptive Information Visualization: Using Eye Gaze Data to Infer Visualization Tasks and User Cognitive Abilities. In *Proc. IUI* (Santa Monica, California, USA) *(IUI '13)*. Association for Computing Machinery, New York, NY, USA, 317–328. https://doi.org/10.1145/2449396.2449439

[29] Ben Steichen and Bo Fu. 2019. Towards Adaptive Information Visualization - A Study of Information Visualization Aids and the Role of User Cognitive Style. *Frontiers in Artificial Intelligence* 2 (2019). https://doi.org/10.3389/frai.2019.00022

[30] Ben Steichen, Bo Fu, and Tho Nguyen. 2020. Inferring Cognitive Style from Eye Gaze Behavior During Information Visualization Usage. In *Proc. UMAP* (Genoa, Italy) *(UMAP '20)*. Association for Computing Machinery, New York, NY, USA, 348–352. https://doi.org/10.1145/3340631.3394881

[31] Dereck Toker, Cristina Conati, Giuseppe Carenini, and Mona Haraty. 2012. Towards Adaptive Information Visualization: On the Influence of User Characteristics. In *Proc. UMAP*, Judith Masthoff, Bamshad Mobasher, Michel C. Desmarais, and Roger Nkambou (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 274–285.

[32] Dereck Toker, Cristina Conati, Ben Steichen, and Giuseppe Carenini. 2013. Individual User Characteristics and Information Visualization: Connecting the Dots through Eye Tracking. In *Proc. ACM SIGCHI* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 295–304. https://doi.org/10.1145/2470654.2470696

[33] Nikos Tsianos, Panagiotis Germanakos, Zacharias Lekkas, Costas Mourlas, and George Samaras. 2009. Eye-Tracking Users' Behavior in Relation to Cognitive Style within an E-learning Environment. In *2009 Ninth IEEE International Conference on Advanced Learning Technologies*. 329–333. https://doi.org/10.1109/ICALT.2009.110

[34] UEQ-S. 2022. User Experience Questionnaire Short Version (UEQ-S). https://www.ueq-online.org. Online; accessed 25 January 2023.

[35] Konstantinos Vassakis, Emmanuel Petrakis, and Ioannis Kopanakis. 2018. *Big Data Analytics: Applications, Prospects and Challenges*. Springer International Publishing, Cham, 3–20. https://doi.org/10.1007/978-3-319-67925-9_1

[36] Maria. C. Velez, Deborah Silver, and Marilyn Tremaine. 2005. Understanding visualization through spatial ability differences. In *Proc. IEEE VIS* (Minneapolis, MN, USA). IEEE, 511–518. https://doi.org/10.1109/VISUAL.2005.1532836

[37] Anatoly Yelizarov and Dennis Gamayunov. 2014. Adaptive Visualization Interface That Manages User's Cognitive Load Based on Interaction Characteristics. In *Proc. VINCI* (Sydney NSW, Australia) *(VINCI '14)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/2636240.2636844

[38] Caroline Ziemkiewicz, Jordan R. Crouser, Ashley R. Yauilla, Sara L. Su, William Ribarsky, and Remco Chang. 2011. How locus of control influences compatibility with visualization style. In *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 81–90. https://doi.org/10.1109/VAST.2011.6102445

[39] Caroline Ziemkiewicz and Robert Kosara. 2009. Preconceptions and Individual Differences in Understanding Visual Metaphors. *Computer Graphics Forum* 28, 3 (July 2009), 911–918. https://doi.org/10.1111/j.1467-8659.2009.01442.x