Discovering Cognitive Biases in a Visual Analytics Environment

Michael A. Bedek*, Alexander Nussbaumer, Luca Huszar, Dietrich Albert

Graz University of Technology, Austria

ABSTRACT

Cognitive biases as systematic reasoning errors may have severe consequences in law enforcement agencies. The European project VALCRI aims to create a visual analytic environment that supports human reasoning and sense-making processes. VALCRI´s goal is to avoid that cognitive biases occur in the first place or at least to minimalize their potential negative effects. To empirically prove this goal, cognitive biases need to be operationalized and measured to compare VALCRI with other existing software solutions. Three approaches, a theory-driven, a behavioral observation and a data-driven approach, have been applied in parallel to measure and discover a selected set of cognitive biases.

Keywords: Cognitive Biases, Visual Analytics, Decision Making, Operationalization.

Index Terms: [Human-centered computing]: Visual analytics, Visualization systems and tools; [Computing methodologies]: Cognitive science.

1 INTRODUCTION

We are continuously exposed to a great amount of information: by newspapers, by advertisement, by the internet. At the same time, we have to make hundreds of more or less important decisions every single day: *Should I buy this or that ice tea*? *Should I leave the country or not*? When people have to make decisions in a situation of uncertainty and when they are overwhelmed by too much information, they often apply heuristics. Heuristics are often referred to as rules of thumb to make a decision. In many cases, such mental shortcuts are extremely useful (see for example [6]). However, they often don´t follow the rules of logic and are not based on exhaustive information search and evaluation. If the heuristics fail in our search for efficiency, they can lead to systematic errors. In such instances, we call them cognitive biases.

The European project VALCRI (www.valcri.org), which stands for Visual Analytics for Sense-making in CRiminal Intelligence analysis, has the goal to address the challenges of today´s law enforcement agencies by creating a system that supports human reasoning and sense-making and either avoids cognitive biases at all or reduces their effects. This goal is pursued by developing appropriate data analytic tools following the principles of visual analytics.

*michael.bedek@tugraz.at

We are in the final phase of the project, and thus, cognitive scientists are focusing on summative evaluations. The main question is, therefore: "does VALCRI induce or hinder the occurrence of cognitive biases, compared to other software solutions?" However, during the whole project an iterative series of formative evaluations have been carried out, aiming to identify suggestions for improvement of the tools, features, and functionalities. For example, in the initial phase, an encompassing set of design guidelines for the software developers has been elaborated. These design principles aimed to facilitate human reasoning and sense-making processes, and in consequence, to either avoid cognitive biases at all or to reduce their effects. In the second, intermediate phase of the VALCRI project, prototypes and tools have been evaluated if and to which extend the design principles have been fulfilled by the software developers. In addition to that, cognitive scientists evaluated if interactions with certain tools may lead to some cognitive biases, resulting in a "tool x bias matrix" as it is described in more detail in section 4.1.2.

In the following sections, we focus on the summative evaluation activities or to be more precise, on the prerequisites of carrying out summative evaluations: making non-directly observable constructs measurable, to enable the measurement of cognitive biases while interacting with a visual analytics environment.

2 THE VALCRI PLATFORM AND ITS TOOLS

In this section we shortly describe the most important tools of the VALCRI platform.

The *Search tool* allows for searching and filtering crime incidents and according documents from the whole data set. Start date and end date limit the data set to a time period and a key word search limits the data to relevant pieces of information.

The *Time tool* (see Figure 1) shows a line chart indicating the number of crime incidents as a function of the time. The time period can be changed interactively, in order to get more details or better overview and to synchronize the other tools to the selected range. Similar to the time tool, a *Statistical process control tool* (SPC-tool) shows standard deviations of the number of occurred crimes in a time period. This allows to quickly grasp if something unusual happened.

The *Location tool* (see Figure 1) depicts crime incidents on a map. On an interactive map, crimes are represented as single dots or as rectangles if a larger set of crimes are available in that area (> 200). In such cases, the size of filled-out rectangles within a particular area indicates the number of crime incidents. The area with the most crime incidents is completely filled out and the sizes of the rectangles in other areas are relative to this maximum. The map can be interactively zoomed in and out, which changes automatically the visual representation and synchronizes the other tools with the updated data set selection.

Figure 1: The *Time tool* (left), the *Location tool* (middle) and the *Bar Chart tool* (right)

The *Bar Chart tool* (see Figure 1) shows the number of crimes as bars according to a classification scheme. Discrimination factors include crime types, districts or resolving state. According to such discriminators the numbers of crimes are shown on a bar chart sorted by the number of crime incidents. Clicking on a particular bar limits the data set and synchronizes the other tools accordingly

The *List tool* presents a list of the currently selected crimes including their details. These details consist of metadata (time, location, etc.) and their descriptions.

The *Similarity Space Selector* (S3) tool builds and visualizes clusters of crime incidents. Clusters of crime incidents are built by analyzing the similarity between them. For this, crime descriptions are analyzed according to a pre-defined set of relevant terms and by calculating distances between the descriptions. The clusters are presented on a 2D plane, where they are represented as polygons and located according to their distances to each other.

Similarly, the *Crime Classification Table* (CCT) tool provides a matrix representation of the clusters and terms, where the matrix elements are the related crime documents.

3 SELECTION OF COGNITIVE BIASES IN VALCRI

A large number of cognitive biases have been suggested and described in the literature. However, in the course of the VALCRI project, the following set of eight cognitive biases has been selected, based on their significance for the daily routines of analysts:

Confirmation Bias, where pieces of information that support the initial expectation are disproportionally considered and selected [8].

Anchoring, which is the tendency to rely too heavily upon or to "anchor" on a past reference or on one trait or piece of information when making decisions [7].

Clustering Illusion, which is a tendency to “see patterns where no patterns exist”, e.g. interpreting patterns or trends in random distributions [4].

Framing Effect, which is the tendency to draw different conclusions from the same information, depending on how that information is presented [12].

Availability Bias, where likelihood-estimations of something to happen is “*by the ease with which instances of occurrences can be brought to mind*” [11], (p. 1127).

Base Rate Fallacy: The tendency to base judgment on specifics, ignoring general statistical information [3].

Selective Perception: Selective perception occurs when people pay particular attention to some parts of their environment to the point where it distorts the reality of the situation [1].

Group-think: A deterioration of mental efficiency, reality testing and moral judgment resulting from group pressure [5].

4 DISCOVERING COGNITIVE BIASES

To discover cognitive biases or to measure if an analyst is affected by at least one of them *while* being engaged with the VALCRI platform we applied three methodological approaches in parallel: a theory-driven approach, a behavioral observation approach and a data-driven approach.

4.1 Theory-driven Approach

This approach is called “theory-driven” because it is solely carried out by domain experts. Cognitive scientists operationalize cognitive bias based on their descriptions and map them to VALCRI tools and its features.

4.1.1 Process-oriented Operationalization

Based on the description of the cognitive biases, they have been operationalized by identifying behavioral indicators, such as actions and interactions that may differentiate biased and unbiased usage of the tools and the platform.

In the following we will focus on the example of *Selective Perception*. As mentioned above, this cognitive bias is defined as being focused on a particular area of the information space. A similarity measurement between the key words entered into the *Search tool*, between the documents and crime reports further examined via the *List tool*, or between the parameters of the visualizations (location, crime-types, peoples, time) of the *Location tool* can be computed. A high similarity between the key words, the selected documents and the visualization parameters over a longer period of time is considered as an indication for the user being focused on a particular area of the information space, i.e. for *Selective Perception*.

In the context of criminal or intelligence analysis, it is important to distinguish between different kinds of search, e.g. explorative, investigative, hypothesis- or question-driven, etc. The validity of the operationalization of any cognitive bias can be improved when taking such *context* information into account. For example, in case of a hypothesis-driven search, an analyst who is engaged within a small area of the information space shouldn´t be diagnosed as being affected by selective perception (but potentially from confirmation bias).

4.1.2 Tool x Bias Matrix

A less fine-tuned theory-driven approach is to identify the extend by which different tools of the platform might lead to biased or unbiased decisions. Three cognitive scientists and psychologists filled out a simple cross table or "tool x bias matrix" and evaluated if a certain cognitive bias is either induced or hindered by features of a particular tool. In the following we will briefly outline two examples:

The *Time tool* (see Figure 1) or to be more precise, the "connecting lines" of the line chart might give the impression of "patterns" (such as trends) even if these patterns don´t exist. This is due to the fact that only aggregated values are shown (e.g. on a week-base level). Seeing such patterns can be interpreted as *Clustering Illusion*. As a suggestion for avoiding such a misinterpretation, a user should have the opportunity to switch to a histogram visualization.

The *Location tool* (see Figure 1) might induce a base rate fallacy if the user doesn´t look at the actual numbers but just consider the size of the filled rectangles in in the selected areas. As mentioned above, the size of the filled rectangles indicates the number of incidents, relative to the area with the most incidents, rather than absolute values. A *Base Rate Fallacy* might be avoided if users could also switch a different view, e.g. showing the sizes of the filled rectangles relative to the maximum value of the overall city or the particular city district.

4.2 Behavioral Observation Approach

In this section we describe the procedure and goal of a behavioral observation. Nine experienced law-enforcement analysts worked on a task for around 2 hours, separately from each other. While working on the task, they were asked to "think aloud" on their reasoning, ideas and conceptions. Their activities while working on the task have been video- and audio recorded and a screen capturing took place. The participating analyst´s task was to analyze a particular crime type (burglary) in a certain city district over a certain period of time and the main question for them was if more patrols should be sent to this city district or not. Afterwards, a qualitative interview has been carried out.

While working on the task, they have been observed by at least one expert on cognitive biases who didn´t intervene during this exercise. The observer filled out a prepared form, indicating the time, which cognitive bias he or she observed, the tools that have been used by the analyst, and if necessary, further explanations on this observation in an open format. These observations have been validated and enriched by two other experts who used the video and audio recordings.

On the one hand, the outcome of this exercise was a validation and enrichment of the theory-driven tool x bias mapping described in the previous section, as well as the elaboration of new ideas for potential process-oriented indicators. On the other hand, compared to the purely theory-driven elaboration of the "tool x bias matrix" described in section 4.1.2 above, the outcome of this exercise resulted in a mapping between *sets of tools* and cognitive biases. The reason for this is that for certain, more complex workflows and processes the analysts used a combination of tools simultaneously.

An example would be the combination of the *Time tool*, the *SPC tool* and the *Location tool* when searching for "peaks in the noise", for a certain area and period of time. In many case the search for such peaks was focused on the maximum values and quite often, the analysts were not trying to falsify their initial hypothesis (e.g. by checking also for other periods of time or other city districts). In other words, this particular work process resulted in many cases to vastly overlapping combinations of certain cognitive biases: the confirmation bias, the framing effect, the base rate fallacy and the clustering illusion.

4.3 Data-driven Approach

The data-driven approach makes use of the user´s interaction data while being engaged with a task. This approach tries to identify indicators that predict the occurrence or strength of a cognitive bias. The prerequisite for such an approach is the availability of an "objective" measurement. This is what we call outcome-oriented operationalization.

4.3.1 Outcome-oriented Operationalization

Compared to the large number of cognitive biases mentioned in the literature, for only a few of them an "objective" measurement, such as a questionnaire or test, has been suggested. One example is the "Selective Exposure Paradigm" [2] to measure the confirmatory search tendencies, a main indicator of the confirmation bias. In an experiment, participants are confronted with two alternatives (e.g. 2 different supermarkets) and they have to make a decision (e.g. in which supermarket they would buy some food). After a preliminary decision is made, the participants are then exposed to various pieces of information that either confirm or disconfirm the initial decision. A tendency for confirmatory search can be identified if a participant doesn´t change his or her initial decision, even if overwhelmed by a large number of disconfirming pieces of information.

Such a well-established measurement procedure is required to apply data-mining and statistical analysis since it serves as criteria or for training purposes of the classification algorithm.

4.3.2 Data-mining and Statistical Analysis

In the case of the statistical analysis we expect likelihoods rather than certainty that a user is affected by a certain cognitive bias.

The statistical method compares interaction behavior of biased and non-biased users. In order to classify users with regards to a confirmation bias, they will participate in the selective exposure experiment. This experiment discriminates biased users from non-biased users. Then they complete some tasks and their interaction data is classified according to the bias state of the respective user. This can be seen as training data for the detection algorithm. An appropriate detection algorithm has to be selected and adapted, so that it is capable of classifying new interaction data without the results of the selective exposure experiment. Candidates of such algorithms will be chosen from the machine learning field. For example the Vector Space Model for similarity measures between documents [10] could be modified that it measures similarities between interaction data.

At the current stage this method has not been applied in the VALCRI project. However, in principle, a validated detection algorithm that makes use of interaction patterns that distinguish biased and non-biased users could be used to provide prompts or hints and recommendations to the user while being engaged with the platform (formative feedback) or at the end of a certain workflow (summative feedback).

5 Conclusion and Outlook

As implicitly mentioned in the previous section, the three methodological approaches are not mutual exclusive, but they encompass each other. Just a few examples, the theory-driven mapping between tools and cognitive biases is validated by the behavioral observation approach. Or the theory-driven operationalization of the cognitive biases, i.e. the process-oriented indicators, are validated by data-driven approach since the indicators should have some predictive power of the outcome-oriented operationalization (i.e. the indicators serve as predictors and the outcome-oriented operationalization serves as criteria in a regression analysis). Applying these different methods enables for a holistic mutual validation.

Our focus in the near future is to carry out a data-driven study for examining the *Clustering Illusion*. Different indicators have been defined, a procedure for an outcome-oriented operationalization has been elaborated and a data set, instruction and task have been selected and described.

Our impression is that the interplay between large data visualizations and their characteristics, state and trait variables of users as well as the context in which they have to make decisions, is a promising field to find new and compelling research questions to ask.

Acknowledgement

The research leading to these results of the VALCRI project has received funding from the European Union 7th Framework Programme (FP7/2007-2013) under grant agreement no FP7-IP-608142.

References

[1] DeWitt C. Dearborn and Herbert A. Simon, “A Note on the Departmental Identification of Executives”, Sociometry, vol.21, no. 2, pp. 140-144, 1958.

[2] Peter Fischer, Eva Jonas, Dieter Frey, and Stefan Schulz-Hardt, “Selective exposure to information: The impact of information limits”, European Journal of Social Psychology, vol. 35, no. 4, pp. 469–492, 2005.

[3] Gerd Gigerenzer “Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa)” in Subjective probability, G. Wright and P. Ayton, Eds. Chichester: Wiley, pp. 129-161, 1994.

[4] Thomas D. Gilovich, “How we know what isn’t so: The fallibility of human reason in everyday life”. New York: The Free Press, 1991.

[5] Paul Jones and Peter Roelofsma, “The potential for social contextual and group bias in team decision-making: biases, conditions and psychological mechanisms”, Ergonomics, vol. 43, no. 8, pp. 1129-1152, 2000.

[6] Gary Klein, “A naturalistic decision making perspective on studying intuitive decision making”, Journal of Applied Research in Memory and Cognition, vol. 4, no. 3, pp. 164-168, 2015.

[7] Thomas Mussweiler, “The Malleability of Anchoring Effects”, Experimental Psychology, vol. 49, no. 1, pp. 67-72, 2002.

[8] Raymond. S. Nickerson, “Confirmation bias: A ubiquitous phenomenon in many guises”, Review of general psychology, vol. 2, no. 2, pp. 175-220, 1998.

[9] Alexander Nussbaumer, Katrien Verbert, Eva-C. Hillemann, Michael A. Bedek, Dietrich Albert, “A Framework for Cognitive Bias Detection and Feedback in a Visual Analytics Environment”, 2016 European Intelligence and Security Informatics Conference, pp. 148-151, 2016.

[10] Gerard Salton, Andrew Wong, and Chungshu. S. Yang, “A vector space model for automatic indexing,” Communications of the ACM, vol. 18, no. 11, pp. 613–620, 1975.

[11] Amos Tversky and Daniel Kahneman, “Judgment under uncertainty: Heuristics and biases”, Science, vol. 185, no. 4157, pp. 1124–1131, 1974.

[12] Amos Tversky and Daniel Kahneman, “The framing of decisions and the psychology of choice” Science, vol. 211, no. 4157, pp. 453–458, 1981.