The Biases of Thinking Fast and Thinking Slow

Dirk Streeb* University of Konstanz Min Chen[†] University of Oxford Daniel A. Keim[‡] University of Konstanz

ABSTRACT

Visualization is a human-centric process, which is inevitably associated with potential biases in humans' judgment and decision making. While the discussions on humans' biases have been heavily influenced the work of Daniel Kahneman as summarized in his book "Thinking, Fast and Slow" [8], there have also been viewpoints in psychology in favor of heuristics (e.g., [6]). In this paper, we present a balanced discourse on the humans' heuristics and biases as the two sides of the same coin. In particular, we examine these two aspects from a probabilistic perspective, and relate them to the notions of global and local sampling. We use three case studies in Kahneman's book to illustrate the potential biases of human- and machine-centric decision processes. Our discourse leads to a concrete conclusion that visual analytics, where interactive visualization is integrated with statistics and algorithms, offers an effective and efficient means to overcome biases in data intelligence.

1 INTRODUCTION

Daniel Kahneman's "Thinking, Fast and Slow" [8] is a best-selling book about the biases of human intuition. The book provides an insightful and enjoyable omnibus tour of human mind in decision making, drawing from empirical findings in a wide range of research publications in cognitive sciences and especially from the research activities of the author, Amos Tversky, and their colleagues and students. As one of the most popular non-fiction books in this decade, it has generated a profound impact on politics, business, healthcare, and many scientific and scholarly fields.

In "Thinking, Fast and Slow", Kahneman describes two systems of decision making. "System 1 operates automatically and quickly, with little or no effort and no sense of voluntary control. System 2 allocates attention to the effortful mental activities that demand it, including complex computations." [8, p. 20]. A major contribution of Kahneman's scientific work was to point out numerous differences between prescriptions of normative theory and observed human behaviors. The book highlights more than twenty types of such differences as weaknesses of System 1 thinking, e.g., priming effect (§4), overestimating and overweighting rare events (§30), inference without statistics (§10, §16) or algorithms (§21). As many decisions made by System 2 may feature some shortcomings as well, Kahneman attributes these to System 2 being busy, depleted or lazy (§3).

In the book, Kahneman champions for statistics and algorithms (e.g., §16, §21). This view is often used as a supporting evidence for minimizing humans' role in data intelligence. From the perspective of visual analytics, which advocates the integration of machine-centric processes (i.e., statistics, algorithms, and machine learning) with human-centric processes (i.e., visualization and interaction), the book appears to lean heavily towards the former.

The labeling of Kahneman's two systems is meant to be a characterization rather than a panacea [20]. Similar dichotomies can be found in other psychological literature (cf. [10]). There are also suggestions for a one system or *k*-system models (e.g., [10]). In this short paper, we follow the two-system discourse by considering two types of human-centric processes: (A) intuition-based decision making and (B) analysis-based decision making using, e.g., logical, statistical, rule-based, and procedural reasoning. Later in Section 4, we add two additional types involving machine-centric processes: (C) fully automated decision making using only machine-centric processes, and (D) *visual analytics* where machine- and human-centric processes are integrated together for decision making.

In the remainder of this paper, we first summarize the scholarly discourse on heuristics and biases in the literature (Section 2). We then examine three case studies in Kahneman's "Thinking, Fast and Slow" [8], pointing out the potential biases that Kahneman may have unwittingly introduced (Section 3). Finally we provide an alternative interpretation to the empirical findings in the book, making a case that the visual analytics paradigm remains to be a necessary approach to many complex decision processes (Section 4).

2 HEURISTICS AND BIASES

In statistics *bias* has a relatively clear definition, i.e., the expected value of a procedure A used for fitting or predicting a random variable B is not equal to the expected value of B. On the other hand, the notion of bias as used in the context of human decision making is much looser. Here, bias describes the deviation of human decision making from taking some option an experiment designer proposes as being optimal. Thereby, bias is attributed a clearly negative meaning, which is not as prominent in statistics, where usually other measures are minimized, e.g., mean squared error.

In psychology and decision sciences, *heuristics* are commonly considered as decision strategies that enable fast and frugal judgments by ignoring some of the available information, typically under the conditions of limited time and information [6]. This notion characterizes the type (A) human-centric processes. One earlier and different notion of *heuristic* is that they are useful strategies to cope with problems that are intractable using exact computation such as playing chess [18]. This implies the type (B) human-centric processes, and corresponds to some machine-centric processes such as the concept of *heuristic algorithms* in computer science. Regardless which notions, psychologists all agree that heuristics may lead to *biases*, resulting in various types of errors in judgments and decisions [8]. Hence heuristics and biases are the two sides of the same coin.

The research and discourse on heuristics and biases has attracted much attention in the literature. Fiedler and von Sydow [5] provide an overview of the research in psychology since Tversky and Kahneman's 1974 paper [21]. There are papers that focus on specific examples in Kahneman's book [8], such as on the hot-hand fallacy [7, 13], the priming effect [24], and experiment reproducibility [9, 14, 17, 22].

In this paper we examine the sources of biases from a probabilistic perspective. As illustrated in Fig. 1, let's consider a domain of entities (e.g., objects, phenomena, or events) that is referred to as the *global domain*. Any systematic sampling of this domain is usually limited to a pre-defined set of variables. To maintain the statistical validity of the sampling process, the required sample size is exponentially related to the number of variables, which is understandably kept low in most situations. Meanwhile, an individual most likely

^{*}e-mail: streeb@dbvis.inf.uni-konstanz.de

[†]e-mail: min.chen@oerc.ox.ac.uk

[‡]e-mail: keim@dbvis.inf.uni-konstanz.de



Figure 1: The two types of human-centric processes typically correspond to the information acquired from, and knowledge related to, two sampling domains.

encounters some of these entities in a sub-domain (referred to as a *local domain*). The observations made in these encounters typically include more variables (as illustrated by 3-D features and shape variations in Fig. 1) than what a systematic sampling may cover. To process a sequence of input data, Type (A) human-centric processes rely on the experience built on accumulated information about the local domain, while Type (B) human-centric processes rely on the summary resulting from systematic sampling in the global domain. In addition, both will make use of some a priori knowledge, though the two will likely utilize or focus on different forms of knowledge.

Many shortcomings of human heuristics highlighted by Kahneman [8] represent attempts of using experience about a local domain to make a decision about an input from the global domain. In such situations, it is often reasonable to think that using the global statistics would be more appropriate. However, if a decision is to be made about an input that is known to be in a local domain and feature more variables than those covered by the global sampling, using the experience accumulated in this local domain could be more appropriate than blindly using the global statistics.

Both, Kahneman [8, p. 222] and Gigerenzer [6, p. 4] point out that more information is not necessarily the key to better decisions. Instead, it is essential to choose the more appropriate global and/or local information in a given situation. In the next section we use some of Kahneman's examples to illustrate that biases can result from inappropriate use of global or local information.

3 CASE STUDIES

We appreciate that most of the case studies in Kahneman's book [8] are used to illustrate the biases of human intuition, and the absolute necessity for anyone working in data science and decision science to be aware of the limits posed on decision making processes. We find that most of his case studies are well designed and presented. Here we only show three case studies that feature biases unaccounted for in the book. We flag out these cases to illustrate that the sources of biases may come from both sampling domains as shown in Fig. 1. One should not read the discourse on these case studies as an argument against the main thesis of Kahneman's book.

3.1 Causes Trump Statistics

In §16 Causes Trump Statistics, Kahneman describes the following synthetic scenario:

"A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data: • 85% of the cabs in the city are Green and 15% are Blue. • A witness identified the cab as Blue. The court tested the liability of the witness under the circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time." [8, p. 166, ch. 16]

Using Bayesian inference, Kahneman concludes that the probability of the guilty cab being Blue is 41%, though most people would instinctively ignore the base-rate of 15% and go with the witness for 80%. This example is meant to illustrate a failure of human reasoning.

It is necessary to notice the difference between the sampling domains involved. The overall statistics of cabs in the city are likely compiled in the domain of the whole city, and day and night. However, the incident happened in a particular area in the city and at a night. There is no reason to ascertain that the global statistics would be the same as the local statistics. This is of the same logic that the local sample cannot simply be assumed to be representative of the global population. Meanwhile, the court is likely to test the reliability of the witness indoors using two-dimensional photos. This sampling domain is different from the outdoor environment where the witness observed a real three-dimensional car. In addition, when one watches a real world event unfolding, one observes many variables, such as the color and shape of the cab and the telephone number and the font in which it is written on the cab. As memorization and recall involve an abstraction process, naturally one may use "Blue" as a cue for the memory representation. Hence, the witness is likely to remember more than just a color. Those additional variables are not part of the uni-variate global statistics about cab colors.

While it is reasonable to suggest that the global statistics about cabs should be considered at court, it is necessary to examine if it is possible for the local statistics to deviate from the global statistics. It is inappropriate to apply the indoor testing results (as in one local domain) to an outdoor scenario (a different local domain) without any moderation. Therefore, it is unsafe to simply combine the two pieces of statistics using Bayes' rule, and to conclude that the witness is likely to be wrong. Indeed, those people who trusted the witness may have taken the factors of sampling domain and additional variables into account. As a result this example is not necessary a case of "biases trump powerful statistics" but a case of "humans' heuristics trump biased statistics".

The potential deviation of statistical measures of a local sample from those of a global sample created by amalgamating many local samples is understood by statisticians. The well-known Simpson's paradox is such an example in a manifest form [16, 19, 23]. There are many similar real-world examples, one of which is a report by Bickel et al. on sex bias in university admission [2]. Furthermore, Birnbaum [3] points out that a normative analysis following Kahneman is incomplete due to its ignorance to signal detection theory and judgement theory. Krynski and Tenenbaum [11] argue that the problem features a high false-positive rate and a lack of a causal structure. In a later work [12], they highlight the overlooked value of humans' "causal knowledge in real-world judgment under uncertainty" [15].

While there are also papers (e.g., [1]) that support Kahneman's conclusion that Bayesian inference is superior to humans' intuition, they are all based on the crucial but subjective assumptions that the problem did not feature a partial Simpson's paradox and the witness did not observe extra information in addition to colors.

Taking all these points together, we can observe that while the case study was presented as an illustration of humans' bias in decision making, its Bayesian inference also features biased assumptions. Were these biased assumptions properly taken into consideration, the conclusion could be very different.

Relevance to Visualization. Interactive visualization typically provides indispensable support to spatiotemporal data analysis. It is usually difficult to capture all necessary variables that describe the characteristics of each location and each path between two locations. Any statistical aggregation across many locations and routes can thus be very sensible to uncaptured confounding variables. Visualization stimulates human analysts' knowledge about these uncaptured variables, while interaction allows analysts to explore different spatial regions and temporal ranges of the data according to both global statistics as well as local characteristics.

3.2 Tom W's Specialty

In the book, Kahneman presents another case study in favor of a base-rate (i.e., a piece of global statistics).

"Tom W is a graduate student at the main university in your state. Please rate the following nine fields of graduate specialization in order of the likelihood that Tom W is now a student in each of these fields. [...] • business administration • computer science [...] • social science and social work" [8, p. 146, ch. 14]

Kahneman further describes a personality sketch of Tom W. He then states that the variables featured in the personality sketch (e.g., nerdiness, attention to details, sci-fi type, etc.) are irrelevant to the question, and only the base-rate of enrollments in these subjects provides the solution.

Most of us agree that for Tom W to choose a subject to study is a complicated decision process. This may depend on many variables, including some captured in his personality sketch, as well as many others unknown to readers. As illustrated in Fig. 1, the base-rate is a piece of global statistics based on the variable of enrollment outcomes, while the personality sketch define a local sampling domain for those fitting the sketch. Ideally one would also wish to have other global statistics, such as enrollment outcomes according to some twenty variables mentioned in his personality sketch (e.g., gender, intelligence, creativity, nerdiness, attention to details, writing skills, and so on), since many of these are confounding variables that influence the base-rate.

Understandably, most organizations do not have all these global statistics or feel inappropriate to disclose them. Nevertheless, using only the base-rate of one variable to infer a clearly multi-variate decision will introduce a huge amount of bias against the conditional probability constrained by the local sampling domain. Such bias is alarmingly dangerous, for instance, by imagining that personality sketches were ignored in criminal investigation, business management, and the media and entertainment industries. In practice, the lack of comprehensive global statistics and the biases of uni-variate approximation of a multi-variate problem is usually compensated by human heuristics that have captured some multi-variate information of a local sampling domain.

Notably, Kahneman himself points out that the probability of Tom W's subject is complex. "Logicians and statisticians disagree about its meaning, and some would say it has no meaning at all." [8, p. 150, ch. 14].

Let "Tom studies $a \in X$ " be a truth statement, where X is a variable of subjects and a is one particular subject. The probability function P(X | Tom) is an approximation of this truth statement. Since a personality sketch is used to describe a set of variables of a person, V_1, V_2, \ldots, V_n , $P(X | V_1, V_2, \ldots, V_n)$ is a further approximation. While ignoring the base-rate P(X) is clearly a naive bias, ignoring the impact of variables V_1, V_2, \ldots, V_n is also a dangerous bias. For example, it is well known that, nowadays, women are overrepresented in psychology and underrepresented in computer science. Only by considering that Tom is male one can correct the bias of the global sample towards psychology and away from computer science. Thus, in contrast to Kahneman, we consider P(X | Tom) to be the best approximation of the truth statement, not P(X).



Figure 2: (a) Illusion as presented by Kahneman [8, p. 100, ch. 9], (b) examples of the global domain, (c) Kahneman's example with additional information, associating the three figures with a 2-D plane.

Relevance to Visualization. Multivariate data visualization techniques, such as parallel coordinates plot and glyph-based visual encoding, enable human users to observe many variables simultaneously. These techniques complement major statistical measures (e.g., base-rate) by preventing any complex decision process (e.g., ranking countries' healthcare quality) from simply depending on the base-rate of one variable (e.g., mortality rate).

3.3 The 3-D Heuristic

Let us go back to a visual case study that Kahneman discussed in the early part of the book. After presenting a Ponzo illusion as shown in Fig. 2 (a), Kahneman asks:

"As printed on the page, is the figure on the right larger than the figure on the left?" [8, p. 100, ch. 9]

Kahneman considers that the question is not misleading. "The bias associated with the heuristic is that objects that appear to be more distant also appear to be larger on the page." [8, p. 101]

We do agree that the Ponzo illusion demonstrates a bias in human perception. However, such bias does not in any way suggest that the humans' 3-D heuristic is ineffective. The humans' 3-D heuristic, in fact, is a form of reasoning using a base-rate, which is obtained by observing a huge number of instances in global sampling as illustrated in Fig. 2 (b). Illusions are outlier events, most of which were created purposely to illustrate the existence of instances contradicting the perception based on the base-rate.

In many cases, biases due to global statistics occur to machinecentric processes. In this case study, the base-rate biases occur to a human-centric process. Instead of considering this being a misled application of global statistics to a very specialized local domain, Kahneman somehow attributes the problem to human heuristics in general.

When applying global statistics to a local sampling domain, biases can often be alleviated by introducing additional variables that humans can reason with. In Kahneman's question, the phrase "As printed on the page" is designed to be misleading in order to maximize the illusion. Hence, the variable about whether the three figures are meant to be measured in their 2-D projections or the perceived 3-D reality is stated ambiguously. If the requirement of 2-D projections was clearly defined, e.g., as 2-D stickers on a piece of glass as illustrated in Fig. 2 (c), the illusion would become much less effective.

Relevance to Visualization. Gestalt grouping, which is a major cause of illusion, holds an important key to the power of visualization. If the human vision system was not equipped with Gestalt grouping, our ability to spot patterns, clusters, anomalies, and changes in visualization would be significantly degraded. But, it is a prerequisite for Gestalt grouping that humans read beyond the data provided by their eyes. Visualization designers should not worry about the existence of false readings. Instead, they should reckon how large negative effects might be and how difficult it is to become conscious about them.

4 IMPLICATIONS FOR VISUALIZATION AND VISUAL ANALYT-ICS

The discourse in Sections 2 and 3 shows that both types of humancentric processes, namely (A) human heuristic and (B) structured reasoning using logic, statistics, rules, and procedures, can lead to biases. If a local domain has the same variables and statistics as the global domain, i.e., being representative, both types of processes will likely function well in both domains. When they have different variables and statistics, for humans to perform decision making tasks in the global domain will lead to biases if there is not enough context knowledge to build a bridge. Similarly, decision making based on global statistics in a local domain can lead to biases if it does not encapsulate the variations and variables necessary for sound decisions in the local domain [4]. Hence, both approaches may suffer from biases.

A common desire is to replace both human-centric processes by (C) a fully automated decision system, which is equipped with many types of global statistics, logic, rules, and algorithms. Kahneman's book [8] naturally inspires such desires. The main challenges for such machine-centric processes to be effective are the difficulty to collect enough data in every local domain to know if its sample features the same statistical measures as the global sample, the difficulty to sample a variety of variables as humans would observe in a real-world environment, and the difficulty to determine what variables should be used in a decision process dynamically according to contextual conditions. With rapidly increasing capabilities of data collection, some machine-centric processes will become more effective decision tools in individual local domains at an acceptable level of biases. However, we must beware of these biases due to global statistics and missing variables in global sampling.

For almost all complex decisions the most appropriate approach is (D), namely, to support human-centric processes with machinecentric processes and vice versa. Following this approach, the two types of human-centric processes (A) and (B) are combined with machine-centric processes based on statistics, logic, rules, and algorithms. But unlike (C), (D) does not intend to fully replace (A) and (B). Instead, (D) uses computers' extensive memory capacity and fast computation capability to assist humans and relieve (A) and (B) from computationally intense tasks. One most effective form of such help is *Visualization*, which enables humans to observe a vast amount of raw data, statistics, analytical results, and computed predictions and recommendations in a rapid fashion.

Meanwhile, the human in the loop of (D) is aware of its limitation in gathering all necessary variables and data for a decision, and its inability to judge whether it is biased towards local statistics. (D) allows human-centric processes to override automated decision recommendations. This is indeed the role of *Interaction* for humans to add their information and knowledge to decision processes.

(D) is what the field of visual analytics has been advocating. With (D) or *Visual Analytics*, semi-automated analyses (including statistics, logic, rules, and algorithms) provide human analysts with (i) numerically aggregated information based on large volumes of data, that humans themselves cannot store, (ii) consistent and rapid application of logic, rules, and algorithms, which humans cannot match in speed and consistency, (iii) visual representations of a variety of information, that humans cannot easily imagine in their minds, and (iv) systematically sequenced interactions, which help structure the operational workflow. At the same time this allows human analysts to inform (D) about information related to the tasks at hand and local variables known only to analysts as well as to guide (D) to process and display information that is most relevant to tasks and local contexts, alleviating the biases of global statistics and the logic, rules, and algorithms designed for common scenarios.

In summary, the studies of heuristics and biases in psychology have provided a rich collection of empirical evidence that enables us to appreciate the benefit of heuristics and the common existence of biases. We examine the sources of biases from a probabilistic perspective, revealing that the fundamental cause is the difference between the two sampling domains, i.e., the domain that human heuristics are based on or statistical samples are obtained from, and the domain heuristics and statistics are applied to for decision making. Biases can be caused by both heuristics and statistics, or in general, by the information and knowledge related to both sampling domains as shown in Fig. 1. All this points to the conclusion that visual analytics, or (D), is the most sensible approach to complex decision making.

ACKNOWLEDGMENTS

Part of this work was funded by the German federal state's Initiative of Excellence via the Graduate School of Decision Sciences at the University of Konstanz.

REFERENCES

- M. Bar-Hillel. The base-rate fallacy in probability judgements. Ars Psychologica, 44(3):211–233, 1980. doi: 10.1016/0001-6918(80)90046-3
- [2] P. J. Bickel, E. A. Hammel, and J. W. O'Connell. Sex bias in graduate admissions: Data from Berkeley. *Science*, 187(4175):398–404, 1975. doi: 10.1126/science.187.4175.398
- [3] M. H. Birnbaum. Base rates in Bayesian inference: Signal detection analysis of the cab problem. *American Journal of Psychology*, 96(1):85– 94, 1983. doi: 10.2307/1422211
- [4] K. Fiedler. The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1):186–203, 2008. doi: 10.1037/0278-7393.34.1. 186
- [5] K. Fiedler and M. von Sydow. Heuristics and biases: Beyond Tversky and Kahneman's (1974) judgment under uncertainty. In M. W. Eysenck and D. Groome, eds., *Cognitive Psychology: Revisiting the Classic Studies*, chap. 12, pp. 146–161. Sage Publications, 2015.
- [6] G. Gigerenzer, R. Hertwig, and T. Pachur, eds. *Heuristics: The foundations of adaptive behaviour*. Oxford University Press, New York, 2011.
- [7] B. S. Green and J. Zwiebel. The hot-hand fallacy: Cognitive mistakes or equilibrium adjustments? Evidence from Major League baseball. *Management Science*, forthcoming. doi: 10.2139/ssrn.2358747
- [8] D. Kahneman. *Thinking, fast and slow.* Penguin Books Ltd., London, 2011.
- [9] D. Kahneman. Comment in response to Schimmack et al. online, 2017. accessed in June 2017.
- [10] G. Keren and Y. Schul. Two is not always better than one. *Perspectives on Psychological Science*, 4(6):533–550, 2009. doi: 10.1111/j.1745 -6924.2009.01164.x
- [11] T. R. Krynski and J. B. Tenenbaum. The role of causal models in reasoning under uncertainty. In *Proceedings of the 25th Annual Conference* of the Cognitive Science Society. Mahwah, NJ: Erlbaum, 2003.
- [12] T. R. Krynski and J. B. Tenenbaum. The role of causality in judgment under uncertainty. *Journal of Experimental Psychology: General*, 136(3):430–450, 2007. doi: 10.1037/0096-3445.136.3.430
- [13] J. B. Miller and A. Sanjurjo. Surprised by the gambler's and hot hand fallacies? A truth in the Law of Small Numbers. *IGIER Working Paper No. 552*, 2016. doi: 10.2139/ssrn.2627354
- [14] Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*, 349(6251), 2015. doi: 10.1126/science. aac4716
- [15] J. Pearl. Causality: Models, Reasoning and Inference. Cambridge University Press, 2 ed., 2009. doi: 10.1017/CBO9780511803161
- [16] J. Pearl. Understanding Simpsons Paradox. Technical report, 2013.
- [17] U. Schimmack, M. Heene, and K. Kesavan. Reconstruction of a train wreck: How priming research went off the rails. online, 2017. accessed in June 2017.

- [18] H. A. Simon. Invariants of human behavior. Annual Review of Psychology, 41:1–20, 1990. doi: 10.1146/annurev.ps.41.020190.000245
- [19] E. H. Simpson. The interpretation of interaction in contingency tables. Journal of the Royal Statistical Society. Series B (Methodological), 13(2):238–241, 1951. doi: 10.2307/2984065
- [20] S. A. Sloman. The empirical case for two systems of reasoning. *Psychological Bulletin*, 119(1):3–22, 1996. doi: 10.1037/0033-2909.119.1.3
- [21] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131, 1974. doi: 10.1126/science .185.4157.1124
- [22] E.-J. Wagenmakers, ..., and R. A. Zwaan. Registered replication report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science*, 11(6):917–928, 2016. doi: 10.1177/1745691616674458
- [23] C. H. Wagner. Simpson's paradox in real life. *The American Statistician*, 36(1):46–48, 1982. doi: 10.2307/2684093
- [24] E. Yong. Nobel laureate challenges psychologists to clean up their act: Social-priming research needs "daisy chain" of replication., 2012. Including a public letter by Daniel Kahneman. doi: 10.1038/nature. 2012.11535