Trust Calibration as a Function of the Evolution of Uncertainty in Knowledge Generation: A Survey

Joshua Boley*
Northern Illinois University

Maoyuan Sun†
Northern Illinois University

ABSTRACT

User trust is a crucial consideration in designing robust visual analytics systems that can guide users to reasonably sound conclusions despite inevitable biases and other uncertainties introduced by the human, the machine, and the data sources which paint the canvas upon which knowledge emerges. A multitude of factors emerge upon studied consideration which introduce considerable complexity and exacerbate our understanding of how trust relationships evolve in visual analytics systems, much as they do in intelligent sociotechnical systems. A visual analytics system, however, does not by its nature provoke exactly the same phenomena as its simpler cousins, nor are the phenomena necessarily of the same exact kind. Regardless, both application domains present the same root causes from which the need for trustworthiness arises: Uncertainty and the assumption of risk. In addition, visual analytics systems, even more than the intelligent systems which (traditionally) tend to be closed to direct human input and direction during processing, are influenced by a multitude of cognitive biases that further exacerbate an accounting of the uncertainties that may afflict the user's confidence, and ultimately trust in the system.

In this article we argue that accounting for the propagation of uncertainty from data sources all the way through extraction of information and hypothesis testing is necessary to understand how user trust in a visual analytics system evolves over its lifecycle, and that the analyst's selection of visualization parameters affords us a simple means to capture the interactions between uncertainty and cognitive bias as a function of the attributes of the search tasks the analyst executes while evaluating explanations. We sample a broad cross-section of the literature from visual analytics, human cognitive theory, and uncertainty, and attempt to synthesize a useful perspective.

1 INTRODUCTION

In the age of "big data", trustworthiness in computing has become a complex, multi-faceted challenge for system designers and developers. The rise and advancement of intelligent sociotechnical systems in particular has compounded the trust issues facing responsible, ethical software vendors; as modernized high-traffic and high-demand computational ecosystems increasingly rely on intelligent agents to aid human decision makers. The most visible and controversial of these systems, such as those used in recidivism and credit scoring [4, 28], often have a high degree of impact and carry significant implications and consequences for society [2, 41]. The issue of user trust in the data, models and processing of increasingly large datasets—particularly when using opaque models such as deep neural networks—has its origins within the context of automation, where the human is generally excluded or only minimally "in the loop". But as the sophistication of intelligent systems and the scope of the problems they solve have evolved, so too has the need to

---

*e-mail: jboley2@niu.edu
†e-mail: smaoyuan@niu.edu

bring humans into decision-making processes originally belonging entirely to the automated system. At the same time, the HCI community has largely concerned itself with enabling more effective communication between human and machine, introducing more sophisticated interactive methods and ideas which have brought about several frameworks describing human-AI collaboration [6].

The themes of accountability and trustworthiness have resonated quite clearly within the visual analytics community, where both the human and the machine intersect to cooperatively steer analysis in increasingly collaborative settings (e.g., [24]). At the high level, the user or analyst guides the process by requesting transformations of data into useful information presented as visualizations, from which hypotheses can be formulated and tested. At the low level, the machine responds to the user, performing the transformations of raw or preprocessed data, and also the mapping of processed results to a graphical syntax and grammar that it can use to construct meaningful visualizations. The software industry is rife with examples of systems that fail because they are unreliable or otherwise unable to meet the needs of their users. Collaborative, AI-backed systems are as susceptible to poor design and implementation decisions as any, if not more so because of their complexity. But in the visual analytics domain there is even more at stake. Responsible developers do not *only* need to know when their systems are failing the users; the analyst's trust in the tools they are expected to use is an absolutely crucial component of knowledge generation, and any visual analytics tool ignored and excluded from knowledge generation has little value. Worse, they can be harmful if the analyst actually trusts them when they should not or cannot avoid using them [32].

Circumventing the limitations of human cognition is at the very heart of both visual analytics and the use of visualizations in general. Hilbert et al. supply an extremely informative discussion in [16] concerning noisy mechanisms in human cognitive processing, which we will later outline. Prior to their work, over half a century of scholarship had already revealed an impressive catalog of both biases and heuristic shortcuts in human cognition, which Tversky and Tahneman compile and discuss at length in [38]. Those works collectively describe and illustrate the built-in heuristics used in low-level cognition during human reasoning tasks, all designed to work around cognitive limitations. For example, Simon's "bounded rationality" [35, 36] was an early attempt to describe low-level cognitive shortcuts the human brain uses to overcome limited short-term memory and imperfect recall. Several works have since enhanced the discussion to include computational aspects such as limited short-term storage capacity and epistemic "noise" introduced by imperfect long-term storage and recall mechanisms [12, 20, 34].

We can therefore expect that such built-in cognitive mechanisms will lead to specific patterns or tendencies in how visual information is processed with some degree of consistency across users. Hilbert et al. in [16] compellingly argue that *all* information processing in human reasoning—from extraction to interpretation, and then clear on from reasoning with abstractions to the formulation of conclusions—is inherently noisy as an inevitable consequence of what is ultimately an ecologically competitive, but not perfect, evolutionary design. It is reasonable to suppose that biased information processing likewise exerts biases on hypothesis formation and knowledge generation. If low-level cognitive bias generalizes across users then so too may

the ways in which a user's conclusions are influenced as low-level, attentional bias maps to high-level reasoning bias in what, to use a computation-inspired metaphor, we could consider a complete human cognitive pipeline for information discovery and symbolic reasoning.

Within the machine learning domain, Bhatt et al. [3] contend that the measurement and communication of uncertainty is essential in building models that a user can usefully incorporate into their decision-making processes. At first glance there may seem to be some barriers to translating their conclusions directly to the visual analytics space. First, their discussion is constrained to uncertainty of outcomes as singletons—that is, one-time transformations of inputs to outputs—as a result of machine learning processing, instead of a sequence of interactions where the machine responds directly to user feedback to intermediate results. Second, users of visual analytics typically have considerably more freedom to select the parameters of the analyses (e.g., type of processing or algorithm, type of graph, etc.) through which they can explore a data space. Finally, there is typically a mismatch between the types of tasks being performed; visual analytics is *explorative* while traditional machine learning algorithms simply *make decisions*; for example, deduce a class or make a prediction based on features in the data learned solely through training algorithms.

However, over the past several years a similar concern for the effects of uncertainty in data and processes has emerged in the visual analytics and visualization communities, arriving at similar conclusions [8, 32]. It is additionally worth noting that while Bhatt et al.'s work is rooted in the explainable AI (or XAI) context, visual analytics and machine learning have become increasingly intertwined over the past several years in general [10, 25, 31, 37, 42]. Algorithmic, transparency and explainability certainly have their place in minimizing the user's uncertainty in the transformations performed during the analytic process, *particularly* where machine intelligence is woven into the fabric of the visual analytics process, be it as a simple data source or analysis tool provider.

In this paper we propose that trust in the knowledge generation process is influenced by uncertainty that propagates from data into knowledge discovery. Noise and errors from biased data and potentially biased processing are incorporated on the machine side into the visualizations the analysts explore to discover gain new insights. We also take the position that cognitive bias in both foraging and the formation of explanatory associations will tend to degrade the usefulness of the analytic model—the collective interactive processes and strategies executed by the analyst to generate insight [32]. In regards to trust calibration, we additionally maintain that the analyst's overall perception of the utility of their models within the context of the problems they solve greatly influences when and to what degree a visual analytics system is engaged, and that a lack of useful or correct results harms user trust.

We begin by presenting a survey of the literature that covers a broad range of relevant topics in sections 2, 3, and 4. In section 5 we present research specific to models of human cognition through which we build our case for the exploration of how uncertainty propagates from visualizations on the machine side into human reasoning about the data being explored, through the mechanisms of cognitive bias. We briefly discuss the implications and research questions introduced by our work in section 6 and conclude in section 7.

2 TRUST CALIBRATION IN THE KNOWLEDGE GENERATION PROCESS

Definitions of trust vary widely within the fields of computer science, sociology, philosophy, and beyond. While traditionally such discussions center on trust between or among humans, within the past few decades, human-AI trust has come into the spotlight not only for researchers, but also for professional and casual audiences interacting with AI in their work or leisure activities. Before we can examine what trust means in the context of visual analytics (to say nothing of AI-driven visual analytics), we must consider different conceptualizations of human-AI trust as a whole.

In a survey of research concerning trust in AI [39], Vereschak et al. highlight three aspects central to many definitions of human-AI trust throughout the literature. The first is a conceptualization of trust as an *attitude* rather than a behavior. The second is a sense of *vulnerability* as the user accepts that the AI might not yield a favorable result. The third is the presence of *positive expectations*; the user believes that the AI will produce a beneficial outcome.

Vereschak et al. [39] additionally identify two main perspectives on the distinction between trust and distrust in the literature. From one viewpoint, trust can be conceptualized as a single continuum from trust to distrust. A recent version of such a continuum can be found in [14], where distrust is placed on the spectrum adjacent to "undistrust" (that is, suspicion, or a lower level of distrust with an overall negative expectation for the AI's behavior), followed by the more positive notions of "untrust" and trust, respectively. From another standpoint, trust and distrust exist on separate axes; high levels of trust do not imply low levels of distrust, and vice versa. This perspective suggests a more granular view of trust as a multidimensional phenomenon; a user may have faith in one aspect of an AI system, for example, but not another. Jacovi et al.'s recent formalization of (multi-)contractual trust extends these concepts into a space defined by user objectives [19]; each *contract*, or "useful" AI functionality, carries its own trust-distrust axis.

Trust may be thought to calibrate on a continuum by the degree to which outcomes conform to user expectations, as the user engages with the system over time. Positive contributions to the perception of trustworthiness calibrate trust towards the positive end of the spectrum, and vice versa negative perceptions quickly move calibration in the opposite direction. We may further conceptualize zones on the ends of the spectrum beyond which trust is more difficult to calibrate towards the opposing end [14, 27]. A *threshold of cooperation* and *limit of forgivability* demarcate the boundaries of these zones, in which the user trusts or distrusts the AI system to the point where the user is either willing to engage with the system cooperatively or unwilling to believe that the system is working towards their objectives, respectively.

Several works have tackled the processes at work in knowledge generation, in both the general sense and specific application domains [1, 30, 32, 33]. A general model of knowledge generation within the visual analytics domain was developed in [33], proposing a complex interplay between processing on the machine side and knowledge generation on the human side. The essential components of this model are the loops through which the analyst gathers information and tests hypotheses:

1. **Exploration loop**, where the analyst interacts with the system to generate observable outcomes to elicit insights leading to the formation of hypotheses.

2. **Verification loop**, where the analyst confirms or rejects hypotheses, steering the direction of the exploration loop.

3. **Knowledge generation loop**, where the analyst generates new knowledge through information foraging and hypothesis testing.

Outside of trivial tasks such as initialization and interface configuration, these loops capture the majority of the analyst's interaction with the system, during which a two-way flow information is conducted. Beyond "one-off" initial calibrations of trust, such as from favorable impressions of the developer or interface design, calibration of the analyst's trust predominantly occurs during the execution of activities within these loops. Indeed, according to Sacha et al.

in [32], trust calibration occurs primarily during the knowledge generation process (i.e., the knowledge generation loop), along several dimensions originally discussed by Muir in [29].

Muir additionally proposes three essential components situated outside the knowledge generation process that heavily influence trust-building: 1) Expert knowledge (i.e., domain expertise), 2) Technical facility, and 3) Routine performance (i.e., development of rote proficiency with a system). As evidenced by accounts of expert problem-solving in e.g. [23], there is a significant difference between expert and novice problem-solving knowledge; the expert has background and contextual information outside of the pipeline that helps them steer their analysis around uncertainties that hinder the less-experienced analyst. Expertise with a system, or visual analytics systems in general, also affects search behaviors though mostly in the form of stream-lining of the overall analysis.

3 UNCERTAINTY IN VISUAL ANALYTICS

The study of uncertainty in both computation and human cognition has a long and varied history. To date, the majority of the work has focused on the computational aspects.

3.1 Sources of Uncertainty

From [43] (as cited in [32]), uncertainty may be understood as a "composition" of several different concepts:

- **Error.** Deviation of data points from their true values (e.g., outliers).

- **Imprecision.** Lower resolution of datapoints relative to that needed for the analysis task.

- **Accuracy.** Tightness of the ranges within which values tend to fall relative to true values.

- **Lineage.** Origin and evolution of the data (i.e., provenance).

- **Subjectivity.** Qualitative judgement/interpretation injected into the data.

- **Non-specificity.** Vagueness in the values assigned to an attribute of a data point. We alternately propose that, qualitatively, the concept includes the degree to which attributes of data points fail to meaningfully distinguish one from another as related to the performance of a task (i.e., unsuitability of selected features).

- **Noise.** Background clutter created by random processes that are either not explainable or not relevant to the task

Based on Sacha et al.'s discussion in [32], we infer that imprecision, subjectivity and non-specificity especially confound user mental models, as a deficit of definitive incorporation of information into hypothesis testing. This is further complicated by the fact that the uncertainty in data sources has probabilistic characteristics, typically falling within distributions [18, 21] that are not readily apparent and at worst, may give the appearance of concealing deeper information when in fact laying false trails.

Frederico et al. in [11] give a more exacting account of where in the knowledge generation process uncertainty is most likely to manifest or become injected, including 1) Automatic analysis (machine-side processes), visualization, and the interaction methods. Automatic analysis is not only subject to potential algorithmic bias, but in general statistical errors carried through from the raw datapoints is amplified. The communication of information can also confound uncertainty when visualizations are inappropriately used or poorly constructed (as discussed at length by Lo et al. in [26]) distracting from or obscuring important/salient information. Finally, interaction methods introduce uncertainty when they misalign with potential "paths" of investigation that would efficiently lead to useful new information, i.e., complicate the means of investigation (e.g., added steps), divert the user from identifying useful patterns, or otherwise hinder the analyst's mental processing.

Looking to such works as [26], we may also find consideration of the data sources, where both raw and preprocessed data are recognized to introduce bias, as well as other types of statistical error. More subtly, views of data , or filters applied to raw or derived data that e.g. conform data to a model (i.e., dropping irrelevant features), or filter samples where one or more features do not fall into specific ranges of values, may potentially obscure relevant patterns [17].

3.2 Propagation of Uncertainty

Accounting for uncertainty is by far simplest in the analytic system: Uncertainty enters via data sources and is amplified via processing that first incorporates data into a model and then interprets models into a visualization [32]. A pipeline describing the capture of uncertainty in visualization is presented by Griethe and Schumann in [13], however this pipeline also captures subsequent entry points of uncertainty during the filtering, mapping and rendering stages. As implied in [13], each is influenced by a set of parameters that are required for meaningful, task-relevant interpretation of the data, collectively encoding a set of assumptions and/or knowledge artifacts. Derived from prior propositions, expertise and information, all of which carry over some uncertainty introduced during prior processing. Assumptions tend to rely on heuristic generalizations of human knowledge constructs, making them subject to bias.

Uncertainty is somewhat more difficult to account for on the human side of the process. Obviously, there is no direct or practical way to measure the analyst's uncertainty on a step-by-step basis. However, it is possible to identify where uncertainty manifests. Zuk and Carpendale's work in [43] presents knowledge constructs in the context of uncertainty visualization, citing three primary types of constructs as discussed in [7]:

1. **Arguments.** In the words of Thomas and Cook (in [7]), the "logical inferences linking evidence and other reasoning artifacts into defensible judgments of greater knowledge value". As set forward in [43], uncertainties contribute to ill-defined and poorly structured problems, creating difficulties with devising useful or valid argumentation by which to solve them as problem specifications may be unclear relevance of data and representations become nebulous, among other problems (the reader may refer to [5] for a deeper discussion).

2. **Causality.** The cause-effect relationship between events, where actions contribute to resulting states and responsive actions. Perception of causality is highly temporalized such that co-occurrent events may be misconstrued as causal. Building on [43], we additionally note that causality may be falsely implied by temporal characteristics of a visualization chain, creating false leads and maladaptively distorting the analytic model. User preconceptions may additionally bias towards assumptions of causality even when diagnostic information establishes otherwise [15].

3. **Models of Estimation.** Allow for the mapping of concrete data to useful abstractions. Uncertainties may be introduced during.

3.3 Effects of Uncertainty on Trust Calibration

Sacha et al. posit in [32] that trust calibration is influenced by the analyst' *awareness* of uncertainty, indicating that recognized uncertainty tends to correctly calibrate trust, while that remaining undetected will generally miscalibrate trust. However, mistaken uncertainty, whether as a result of uncertainty bias or miscommunication is also possible, and will tend to miscalibrate trust. Miscalibrated trust

ultimately leads to greater bias and error in the results, detrimentally impacting the analyst's trust in the entire analytic process over time. Miscalibration disrupts the knowledge generation process in several ways, not least among them the distrust for the system as analyses tend to generate results that are incorrect or have low utility, and possibly even degradation of the user's trust in their own competence in the worst case, depending on their domain expertise.

According to Sacha et al. [32], user trust will tend to be high in two cases: 1) When there are little to no uncertainties and the analyst is aware, in which case trust is properly calibrated; and 2) When there are uncertainties and the analyst is unaware, in which case trust is micalibrated. The analyst may also mistakenly perceive uncertainties when there are none, in which case trust is lower than it should be and therefore miscalibrated. If there are uncertainties but the analyst is *mistaken* about what they are, whether as a result of bias or miscommunication, trust will likewise become miscalibrated and overall trust will be lower should the problems persist. It is important to note, however, that trust may not *necessarily* be miscalibrated in the short term, as the overall level of calibration might be appropriate. Unfortunately, the nature of the bias and errors introduced may surprise the user, reducing overall trust in the system.

4 COGNITIVE BIAS IN VISUAL ANALYTICS

Despite a common negative connotation, cognitive bias has the adaptive quality of enabling rapid decision-making in the absence of detailed information or when time constraints preclude rigorous analysis [9, 40]. However, in the analytic space, cognitive bias often works against the analyst when superficial patterns in the data obscure those that are deeper and more meaningful. The effects of cognitive biases are both fairly consistent and predictable, and manifest as traceable interactive behaviors [40]. Within the context of intelligence analysis, Heur in [15] identifies several types of cognitive bias that emerge during evaluation:

1. **Vividness criterion.** The analyst attends to data and information that is more salient (i.e., specific, personally relevant) than that which is vague or nebulous.

2. **Absence of evidence.** Evidence that is missing tends to be overlooked in favor of evidence that is present.

3. **Oversensitivity to consistency.** Analysts build or otherwise favor hypotheses that are (superficially) supported by the most data.

4. **Coping with evidence of uncertain accuracy.** Analysts tend to overlook partial accuracy of a hypothesis, wholly accepting or rejecting it instead.

5. **Persistence of impressions based on discredited evidence.** Analysts tend to carry over conclusions reached prior to the discrediting of hypotheses, regardless of its validity.

Dimara et al. extend the discussion of cognitive bias in [9], introducing a taxonomy of documented bias grouped by task type. They discuss an *estimation task*, in which users engage in the assessment of some value relative to a question or problem-solving step, as well as a *decision task*, which consists of the selection of an option among many. Also discussed is a *hypothesis assessment task*, in which the user evaluates the hypotheses formulated during the knowledge generation loop. Finally, a *causal attribution task* is introduced, where users attempt to infer a cause-effect relationship among events or pieces of evidence.

Dimara et al. further observe that types of cognitive bias coalesce around a set of fundamental characteristics (or "flavors"), regardless of the type of task: Association, baseline, inertia, outcome, and self-perspective. Associations are the perceived connections between evidence which naturally predispose the user towards assumptions incorporating those connections. Assumed baselines (with or without supporting evidence) predispose the user towards a specific standard for comparison that influences further analysis. In the group of biases grouped under outcome, the central theme is expectation creating cognitive influence that steers the analysis. In the self-perspective category, analysis in influenced (or constrained) by the particular viewpoint of the analyst, whether as a result of domain expertise or uninformed assumptions.

From the preceding discussion it also emerges, though indirectly, that prior patterns of interaction (i.e., habits) and reinforced perceptions of the relevance of subsets of data points or attributes may also contribute to observable higher-level forms of bias; in other words, biases can be learned. These manifest as patterns in what types of data are attended to most frequently or what types of analytic tasks are most frequently undertaken.

5 THE INTERPLAY BETWEEN UNCERTAINTY AND COGNITIVE BIAS IN SENSEMAKING

Based on the conclusions offered by Sacha et al. in [32], we expect that cognitive biases may misdirect the user into information seeking behaviors that may not be consistently effective or useful. Those behaviors may either create incorrect perceptions of uncertainty as the analytic model degrades or deflect the analyst's attention from uncertainties that are important to the proper calibration of trust. This directly implies that cognitive bias may have detrimental effect on the analyst's overall trust in the system or knowledge generation process.

Klein et al.'s view of sensemaking as a data-framing problem from [23] potentially offers a means to describe user analytical reasoning in the context of a network of relationships between propositions and premises. Klein et al. describe these networks as "frames", where information extracted from visualizations is used to formulate propositions and gather supporting (or contradictory) evidence that strengthen or weaken inferred relationships. Klein et al. do not provide a clear definition of the term "data"; we therefore adapt their term to describe any discrete piece of information derived from actual data (e.g., a trend in a line graph, clustering in a scatterplot, etc.) that the user incorporates into their reasoning.

Similar to Correa et al.'s conception of uncertainty in data sources from [8], we propose that it is possible to capture uncertainty within a user's frame through Gaussian mixture models. This would be convenient, as it would allow us to think of "Kleinian" frames as a type of Gaussian network that changes over time, presumably with a central set of propositions forming more stable regions of the network that would tend to converge relatively early on. Restructuring would occur as new relationships are inferred, old relationships are discarded, and new propositions are discovered.

Klein additionally reviewed 120 case studies in [22], noting that the "anchors", or central propositions of an explanatory framework shift as new insights are generated. These anchors directly steer the analysis and guide the analyst's attention in the subsequent data and information foraging activities, until such a time as contrary or unaccounted-for evidence spurs further investigation, potentially leading to a shift or restructuring in the framework. We additionally observe that under ideal conditions, the frequency and degree of shifts over time should stabilize as the analyst converges upon a solution.

It therefore seems reasonable to propose that 1) the trend of such convergence reveals much about the degree of uncertainty in the analyst's reasoning, and 2) the characteristics of the information foraging activities (e.g., number of steps, number of unrelated data sources, the amount of time spent using a visualization before it is discarded, etc.) reflect the analyst's uncertainty regarding a proposition under investigation. We can think of this in terms of the parameters of the search: If the analyst reuses a significant portion

of the parameters used in previous searches regarding the subset of entities of the frame targeted by investigation, then investigation is likely confirmatory. However, if relatively few are employed, then it is likely that the investigation either seeks to resolve conflicting new information, is following a loosely-related parallel, supportive line of reasoning, or the analyst is pursuing a new conception (i.e., expanding upon the frame) that happens to include some common parameters. For our purposes, we are most interested in the first case. We therefore propose that a reasoning process can be approximately described as a series of parameter selection problems, and that the correlations between the frequencies with which parameters are selected across problem instances reveal useful information in regards to the degree of uncertainty in the analyst's reasoning.

6 So Where Does This Leave Us?

The problem of tracing uncertainties in visualization through the analysis process is extremely difficult, especially given that cognitive bias is not likely to present in exactly the same way under all conditions by all users. The authors are unaware of any significant, direct undertaking to rigorously investigate how cognitive biases affect the uncertainties that propagate from the machine into the analyst's decision making.

However, we can make some educated guesses. Based on the discussion in [32], we should expect that cognitive biases likely misdirect the user into information seeking behaviors that are only partially effective at best. We can suppose those behaviors could also create incorrect perceptions of uncertainty, assuming the system attempts to communicate them (e.g., [8, 18, 43], etc.). In any case, it is likely to deflect the analyst's attention from important information, possibly calibrating user trust downwards. If the analyst remains unaware of it then their trust will likely not be appropriately calibrated, with potentially serious consequences.

While it is possible to model uncertainty as it propagates and aggregates during analysis on the machine side, we lack a unified model of the propagation of uncertainty through analytic reasoning. As noted in [40], the purpose of the visual analytic process is to reduce the uncertainty introduced by the kinds of reasoning tasks discussed in [16]. While visual analytics has proven to handily mitigate many of the errors introduced through human reasoning by externalizing both memory channels and significant transformations of information, human cognitive bias is still a significant factor [40]. It is our belief that a better understanding of how error transmitted to the human *through* the visualizations influences information gathering behaviors, is required.

A related question is how the influence of such biases could be measured. Our proposal ultimately assumes that observable characteristics of the interactions reveal useful information about the analyst's internal mapping of information to evidence for or against a hypothesis, and that there is a definite correlation between the analyst's overall trust over time and their specific foraging behaviors. Wall et al.'s work in [40] does handily establish that cognitive biases can be detected through interaction traces, but how would the data recorded in these traces correlate with the analyst's foraging strategies, and what can be inferred about user trust from them? These and similar questions must be assessed before we can even begin to determine whether this is a line of inquiry worth pursuing.

7 Conclusion

In this article we have argued the need for an accounting of how uncertainty propagates from data sources through to the extraction of information and hypothesis testing as a requirement to a deeper understanding of how user trust in a visual analytics system evolves over its lifecycle. We presented a survey of a segment of the literature regarding uncertainty and cognitive bias in visual analytics, focused on those works which seem to collectively offer the best clues towards the creation of a useful theoretical framework for

investigating the propagation of mechanical uncertainty through cognitive bias into resolution of the analyst's uncertainty in their reasoning. Our work continues as we gather more information from the literature and refine and extend our models.

References

[1] C. Antonelli and A. Colombelli. External and internal knowledge in the knowledge generation function. In *The Evolutionary Complexity of Endogenous Innovation*. Edward Elgar Publishing, 2018.

[2] V. Bellotti and K. Edwards. Intelligibility and accountability: human considerations in context-aware systems. *Human–Computer Interaction*, 16(2-4):193–212, 2001.

[3] U. Bhatt, J. Antorán, Y. Zhang, Q. V. Liao, P. Sattigeri, R. Fogliato, G. Melançon, R. Krishnan, J. Stanley, O. Tickoo, et al. Uncertainty as a form of transparency: Measuring, communicating, and using uncertainty. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 401–413, 2021.

[4] J. B. Biddle. On predicting recidivism: Epistemic risk, tradeoffs, and values in machine learning. *Canadian Journal of Philosophy*, pp. 1–21, 2020.

[5] J. M. v. Bruggen, H. Boshuizen, and P. A. Kirschner. A cognitive framework for cooperative problem solving with argument visualization. In *Visualizing argumentation*, pp. 25–47. Springer, 2003.

[6] C. J. Cai, E. Reif, N. Hegde, J. Hipp, B. Kim, D. Smilkov, M. Wattenberg, F. Viegas, G. S. Corrado, M. C. Stumpe, et al. Human-centered tools for coping with imperfect algorithms during medical decision-making. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2019.

[7] K. A. Cook and J. J. Thomas. Illuminating the path: The research and development agenda for visual analytics. Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2005.

[8] C. D. Correa, Y.-H. Chan, and K.-L. Ma. A framework for uncertainty-aware visual analytics. In *2009 IEEE symposium on visual analytics science and technology*, pp. 51–58. IEEE, 2009.

[9] E. Dimara, S. Franconeri, C. Plaisant, A. Bezerianos, and P. Dragicevic. A task-based taxonomy of cognitive biases for information visualization. *IEEE transactions on visualization and computer graphics*, 26(2):1413–1432, 2018.

[10] A. Endert, W. Ribarsky, C. Turkay, B. W. Wong, I. Nabney, I. D. Blanco, and F. Rossi. The state of the art in integrating machine learning into visual analytics. In *Computer Graphics Forum*, vol. 36, pp. 458–486. Wiley Online Library, 2017.

[11] P. Federico, M. Wagner, A. Rind, A. Amor-Amorós, S. Miksch, and W. Aigner. The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics. In *2017 ieee conference on visual analytics science and technology (vast)*, pp. 92–103. IEEE, 2017.

[12] D. G. Goldstein and G. Gigerenzer. Models of ecological rationality: the recognition heuristic. *Psychological review*, 109(1):75, 2002.

[13] H. Griethe, H. Schumann, et al. The visualization of uncertain data: Methods and problems. In *SimVis*, pp. 143–156, 2006.

[14] W. Han and H.-J. Schulz. Beyond trust building—calibrating trust in visual analytics. In *2020 IEEE workshop on trust and expertise in visual analytics (TREX)*, pp. 9–15. IEEE, 2020.

[15] R. J. Heuer. *Psychology of intelligence analysis*. Center for the Study of Intelligence, 1999.

[16] M. Hilbert. Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychological bulletin*, 138(2):211, 2012.

[17] J. Huang. From big data to knowledge: Issues of provenance, trust, and scientific computing integrity. In *2018 IEEE international conference on big data (big data)*, pp. 2197–2205. IEEE, 2018.

[18] J. Hullman, X. Qiao, M. Correll, A. Kale, and M. Kay. In pursuit of error: A survey of uncertainty visualization evaluation. *IEEE transactions on visualization and computer graphics*, 25(1):903–913, 2018.

[19] A. Jacovi, A. Marasović, T. Miller, and Y. Goldberg. Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in ai. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, pp. 624–635, 2021.

[20] D. Kahneman and A. Tversky. The simulation heuristic judgment under uncertainty: Heuristics and biases, 1982.

[21] A. Kale, M. Kay, and J. Hullman. Visual reasoning strategies for effect size judgments and decisions. *IEEE transactions on visualization and computer graphics*, 27(2):272–282, 2020.

[22] G. Klein. *Seeing what others don't: The remarkable ways we gain insights*. Public Affairs, 2013.

[23] G. Klein, J. K. Phillips, E. L. Rall, and D. A. Peluso. A data–frame theory of sensemaking. In *Expertise out of context*, pp. 118–160. Psychology Press, 2007.

[24] D. Liu, S. Alnegheimish, A. Zytek, and K. Veeramachaneni. Mtv: Visual analytics for detecting, investigating, and annotating anomalies in multivariate time series. *arXiv preprint arXiv:2112.05734*, 2021.

[25] S. Liu, X. Wang, M. Liu, and J. Zhu. Towards better analysis of machine learning models: A visual analytics perspective. *Visual Informatics*, 1(1):48–56, 2017.

[26] L. Y.-H. Lo, A. Gupta, K. Shigyo, A. Wu, E. Bertini, and H. Qu. Misinformed by visualization: What do we learn from misinformative visualizations? *arXiv preprint arXiv:2204.09548*, 2022.

[27] S. Marsh and P. Briggs. Examining trust, forgiveness and regret as computational concepts. In *Computing with social trust*, pp. 9–43. Springer, 2009.

[28] M. Mirtalaie, M. Azadeh, M. Saberi, and B. Ashjari. A trust-based credit scoring model using neural network. *Advances in Industrial Engineering*, 46(1):91–104, 2012.

[29] B. M. Muir. Trust between humans and machines, and the design of decision aids. *International journal of man-machine studies*, 27(5-6):527–539, 1987.

[30] M. Pohl, M. Smuc, and E. Mayr. The user puzzle—explaining the interaction with visual analytics systems. *IEEE transactions on visualization and computer graphics*, 18(12):2908–2916, 2012.

[31] D. Sacha, M. Kraus, D. A. Keim, and M. Chen. Vis4ml: An ontology for visual analytics assisted machine learning. *IEEE transactions on visualization and computer graphics*, 25(1):385–395, 2018.

[32] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim. The role of uncertainty, awareness, and trust in visual analytics. *IEEE transactions on visualization and computer graphics*, 22(1):240–249, 2015.

[33] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. A. Keim. Knowledge generation model for visual analytics. *IEEE transactions on visualization and computer graphics*, 20(12):1604–1613, 2014.

[34] A. K. Shah and D. M. Oppenheimer. Heuristics made easy: an effort-reduction framework. *Psychological bulletin*, 134(2):207, 2008.

[35] H. A. Simon. A behavioral model of rational choice. *The quarterly journal of economics*, 69(1):99–118, 1955.

[36] H. A. Simon. Rational choice and the structure of the environment. *Psychological review*, 63(2):129, 1956.

[37] T. Spinner, U. Schlegel, H. Schäfer, and M. El-Assady. explainer: A visual analytics framework for interactive and explainable machine learning. *IEEE transactions on visualization and computer graphics*, 26(1):1064–1074, 2019.

[38] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, 185(4157):1124–1131, 1974.

[39] O. Vereschak, G. Bailly, and B. Caramiaux. How to evaluate trust in ai-assisted decision making? a survey of empirical methodologies. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–39, 2021.

[40] E. Wall, L. M. Blaha, L. Franklin, and A. Endert. Warning, bias may occur: A proposed approach to detecting cognitive bias in interactive visual analytics. In *2017 ieee conference on visual analytics science and technology (vast)*, pp. 104–115. IEEE, 2017.

[41] D. Wang, Q. Yang, A. Abdul, and B. Y. Lim. Designing theory-driven user-centric explainable ai. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–15, 2019.

[42] J. Yuan, C. Chen, W. Yang, M. Liu, J. Xia, and S. Liu. A survey of visual analytics techniques for machine learning. *Computational Visual Media*, 7(1):3–36, 2021.

[43] T. Zuk and S. Carpendale. Visualization of uncertainty and reasoning. In *International symposium on smart graphics*, pp. 164–177. Springer, 2007.