A Visual Analytics Approach to Combat Confirmation Bias for a Local Food Bank

Ritson Delpish(✉), Steven Jiang, Lauren Davis, and Kehinde Odubela

Department of Industrial and Systems Engineering, North Carolina A&T State University, Greensboro, NC 27411, USA
{rkdelpis,kaodubela}@aggies.ncat.edu,
{xjiang,lbdavis}@ncat.edu

Abstract. In the fight against hunger, Food Banks must routinely make strategic distribution decisions under uncertain supply (donations) and demand. One of the challenges facing the decision makers is that they tend to rely heavily on their prior experiences to make decisions, a phenomenon called cognitive bias. This preliminary study seeks to address cognitive bias through a visual analytics approach in the decision-making process. Using certain food bank data, interactive dashboards were prepared as an alternative to the customary spreadsheet format. A preliminary study was conducted to evaluate the effectiveness of the dashboard and results indicated dashboards reduced the amount of confirmation bias.

Keywords: Human factors · Human-Systems integration · Cognitive bias
Visual analytics

1 Introduction

Food Banks operate in a complex environment. Their distribution networks are dynamic and consist of multiple configurations (e.g. hub and spoke) with many charitable agency partners that can receive donated food. There is a high degree of uncertainty on both the supply and demand side making it difficult to understand available supply and food need. Their food distribution decisions have to be made in a way that (i) provides food equitably to those in need, (ii) efficiently maximizes the yield of the donated supply and minimizes waste, and (iii) is distributed in a cost effective manner. Decision-makers in this environment rely on information systems to manage their operational processes. They must interact with large amounts of unstructured and complex data on food collection, inventory management, and distribution activities.

The process of translating that data to a meaningful format is a cumbersome process typically performed using spreadsheets. In addition, existing information systems of the food bank lacks the capabilities to interpret this large-scale data to real-time policy recommendations that support operational and strategic decision-making. Models that could assist with better understanding of uncertainty and how to make decisions given uncertainty are not reflected in the systems and often not directly considered during the decision-making process.

R. L. Boring (Ed.): AHFE 2018, AISC 778, pp. 13–23, 2019.
https://doi.org/10.1007/978-3-319-94391-6_2

It is known that managers employ subjective judgement in formulating strategy and that heuristics are often used in the presence of uncertainty [1]. In many instances, this may be useful and appropriate. However, information overload can often cause a reliance on such strategies making decisions susceptible to cognitive bias and leading to errors and missed opportunities. In this preliminary study, visual analytics tools were explored to help decision-makers combat cognitive biases. Data collected from a local food bank were used in this study.

2 Background

2.1 Food Bank of Central and Eastern North Carolina

The Food Bank of Central and Eastern North Carolina (FBCENC) is the largest food bank in North Carolina and serves 34 of the 100 counties. FBCENC has six branches: Wilmington, Durham, Raleigh, Sandhills, Greenville, and New Bern. Raleigh is the headquarters of the FBCENC which is part of Wake County. The FBCENC receives food donations from different sources and distributes it through soup kitchens, food pantries, homeless shelters, elderly nutrition programs and churches. In 2014-2015, FBCENC distributed more than 57.2 million pounds of food related aid through over 800 partner agencies across 34 counties.

2.2 Food Bank Decision-Making Challenges

The primary goal of Food Banks and other nonprofit hunger relief organizations (NHRO) is to provide access to nutritious food to populations who do not have sustained access to it. These organizations achieve this goal by obtaining food donations from a variety of sources to create large food inventories. The supplied food donation must be screened to ensure quality and safety, sorted, packaged and labeled prior to distribution to charitable agencies where food insecure populations are served. NHRO decision-makers are charged with making strategic and tactical decisions affecting food aid while simultaneously meeting core objectives of equitable food sharing, efficient management of donations, and cost-effective distribution. These decision makers face several challenges such as:

- Uncertainty in the source, frequency and magnitude of supply and demand for food
- Food shelf life, safety and quality of donated food
- Resource constraints such as availability of warehousing, volunteers, transportation, and technology
- Lack of optimal methods and processes for determining true food need
- Compliance with federal and Feed America rules

In this study, historical distribution data from the FBCENC's *fair share* policy was examined. Fair share is a primary program that uses local county poverty rates to identify how to allocate food equitably. The fair share distribution determines how much food each branch and county should receive relative to total food donations across the entire food network. A fair share rate is associated with each branch and that rate further divided to a rate for each county. Using this rate, the theoretical fair share pounds of food is calculated. The theoretical fair share amount is compared to the actual food distribution or *pounds per person in poverty* (PPIP) to determine the effectiveness of the policy. Additionally, Feeding America guidelines recommend that distributions meet 75% of PPIP at a minimum.

Given such vast amounts of information, constraints and uncertainties, Food Bank decision makers often look to past approaches to make strategic planning decisions. This exposes the decision-making process to cognitive bias by using heuristic approaches to simplify problems that are mentally difficult rather than taking a rational problem-solving approach considering all available information [1].

2.3 Cognitive Bias

Cognitive bias or psychological bias is a tendency for individuals to draw conclusions or adopt beliefs given insufficient or no rational evidence that logically support it [2]. It can lead to missed opportunities and poor decision-making. Much has been written about the effect of cognitive bias in individual decision-making [3]. In an organizational setting, individual decision-making introduces bias during negotiations where each party brings their own particular beliefs to the table [1, 3–5]. A particularly troublesome example is one where a particular individual viewpoint continuously influences and dominates others leading to groupthink. Methods to combat cognitive bias have focused on increasing awareness and training [2, 6]. However, this might not be sufficient in cases when information overwhelms the capacity limits of human working memory [7].

Considering the pressures of time (food shelf life) and large data, the decision maker's ability to follow rational approaches is often challenged leading to a reliance on heuristic approaches [7, 8]. A great number of cognitive biases are well documented in the literature [2]. Recent work has begun to look at design considerations for reducing cognitive bias [9, 12]. In this preliminary study, confirmation bias was selected as a test case of bias that might be addressed through visual analytics. Confirmation bias, also known as confirmatory bias describes the tendency a person to look for and recall information that confirms his or her preexisting beliefs. Task analysis done with the FBCENC operations managers reveals that they do rely on their prior experience to support their preconceived hypotheses. Hence, conformation bias does exist in their decision-making process. To combat such bias, new tools such as visual analytics are needed that expands the decision-makers' attention more broadly include dissenting evidence as well [7].

2.4 Visual Analytics

This study utilizes the power of visual analytics to provide analysts a structured approach to combat the cognitive bias in their decision-making. Visual analytics is a process that includes information gathering, data preprocessing, knowledge representation, and interaction to facilitate analytical reasoning via interactive visual interfaces [10, 11]. It is a direct response to the ease and scale of storing big data and the inherent limited capacity of human attention and working memory. There are few recent research writings found that directly address cognitive bias detection in visual analytics environments [13–15].

The focus of those papers involved mitigating biases using approaches such as sensemaking to increase awareness when using interactive intelligence visualization. No research up to this writing was found that addressed utilization and design of visual analytics to reduce cognitive bias in humanitarian supply chain decision-making.

In this study, interactive dashboards were developed to enhance the decision makers' tasks allowing them to see visualized information and get prompt feedback to assess the effects of alternative approaches. An evaluation of dashboard effectiveness using structured responses and feedback from participants indicated dashboards do reduce confirmation bias. This preliminary work serves as a starting point for defining what approaches and evaluation techniques can support future design efforts for Food Bank decision-making.

3 Methodology

3.1 Participants

Participants were recruited from a pool of Industrial & Systems Engineering students with at least senior standing. A diverse group of six volunteers (50% females) between the ages of 25 and 44 years took part in the study. All participants held graduate degrees and experience with frequent use of electronic records and documents such as spreadsheets. All participants reported having advanced computer skills and two-thirds had experience interacting frequently with information systems platforms (the remainder reported occasional interaction). The duration of the exercise ranged from 30–50 min.

3.2 Stimulus Materials

A desktop computer and two 26″ monitors were used to present the task materials (questionnaires and task descriptions) and the stimulus (spreadsheet data file and interactive dashboard). Tasks analysis results revealed that currently, the FBCENC uses spreadsheet in their operations. Each participant performed their assigned task in a quiet location away from the gaze and distraction of others. The spreadsheet data file comprised of historical fair share data for a specific month over a five-year period. A screenshot of the spreadsheet is provided in Fig. 1. Each year was presented in its separate sheet and showed the distribution of food (in pounds) to all the counties served by the food bank network. The interactive dashboard created with Tableau software is shown in Fig. 2. It displays line plots using the same data presented in the spreadsheet filtered by county, month and year. A linear fit trend line was included with each plot for the particular month. Only data from the first four years was presented in the tasks. Using both the spreadsheets and the dashboard, participants were asked to determine whether the fair share distribution would be met (or exceeded) or fall short in year five.

County/ Branch	*PPIP Poverty Pop (a)*	*Fairshare %*	*Theoretical Fairshare*	Total lbs	Difference lbs	Difference %	% of FA Median
Brunswick	12,055	3.27%	1,617,701	1,533,251	(84,450)	-5.22%	161.71%
Columbus	11,966	2.04%	1,005,849	978,577	(27,272)	-2.71%	103.98%
New Hanover	26,855	5.91%	2,922,321	2,258,825	(663,496)	-22.70%	106.94%
Pender	7,466	1.62%	800,182	2,113,471	1,313,289	164.12%	359.92%
Wilmington Total	**58,342**	**12.84%**	**6,346,053**	**6,884,124**	538,071	**8.48%**	**150.03%**
Carteret	7,279	1.76%	867,901	894,245	26,344	3.04%	156.20%
Craven	13,498	2.85%	1,408,862	1,394,094	(14,768)	-1.05%	131.32%
Greene	3,598	0.62%	306,775	265,993	(40,782)	-13.29%	94.00%
Jones	1,798	0.33%	160,791	219,630	58,839	36.59%	155.31%
Lenoir	12,999	2.11%	1,043,511	1,003,223	(40,288)	-3.86%	98.13%
Onslow	19,570	3.99%	1,969,003	1,437,532	(531,471)	-26.99%	93.40%
Pamlico	1,635	0.37%	181,306	283,814	102,508	56.54%	220.71%
Pitt	35,094	5.86%	2,895,906	3,713,072	817,166	28.22%	134.52%
Greenville Total	**95,471**	**17.88%**	**8,834,056**	**9,211,603**	**377,547**	**4.27%**	**122.68%**
Chatham	7,273	1.18%	583,435	627,618	44,183	7.57%	109.72%
Durham	35,936	7.39%	3,652,641	2,449,695	(1,202,946)	-32.93%	86.67%
Granville	6,222	1.53%	755,717	504,887	(250,830)	-33.19%	103.17%
Orange	18,639	2.53%	1,248,645	1,199,421	(49,224)	-3.94%	81.82%
Person	5,596	1.20%	591,919	628,650	36,731	6.21%	142.83%
Vance	12,141	1.90%	939,502	1,583,716	644,214	68.57%	165.85%
Durham Total	**85,807**	**15.73%**	**7,771,860**	**6,993,987**	**(777,873)**	**-10.01%**	**103.63%**
Duplin	6,000	0.95%	469,325	985,320	515,995	109.94%	208.80%
Franklin	8,110	1.75%	864,003	971,707	107,704	12.47%	152.34%
Halifax	13,437	2.14%	1,055,665	1,172,951	117,286	11.11%	110.99%
Harnett	8,576	1.74%	859,718	1,332,233	472,515	54.96%	197.51%
Johnston	22,053	4.88%	2,411,532	1,772,099	(639,433)	-26.52%	102.17%
Nash	13,646	3.11%	1,534,592	1,322,986	(211,606)	-13.79%	123.27%
Sampson	6,401	1.07%	529,090	810,367	281,277	53.16%	160.97%
Wake	75,518	21.07%	10,409,116	7,323,499	(3,085,617)	-29.64%	123.30%
Warren	5,102	0.68%	334,643	529,895	195,252	58.35%	132.05%
Wayne	19,864	3.57%	1,763,481	1,584,297	(179,184)	-10.16%	101.41%
Raleigh Total	**178,707**	**40.95%**	**20,231,165**	**17,805,354**	**(2,425,811)**	**-11.99%**	**126.68%**
Moore	10,046	2.23%	1,100,669	1,517,108	416,439	37.84%	192.01%
Richmond	11,351	1.85%	913,538	1,956,869	1,043,331	114.21%	219.19%
Scotland	9,839	1.54%	762,391	1,492,056	729,665	95.71%	192.81%
Sandhills Total	**31,236**	**5.62%**	**2,776,599**	**4,966,033**	**2,189,434**	**78.85%**	**202.14%**
Edgecombe	12,200	2.18%	1,075,059	1,457,617	382,558	35.58%	151.91%
Lee	8,452	1.90%	938,072	808,180	(129,892)	-13.85%	121.58%
Wilson	15,171	2.91%	1,436,410	1,282,376	(154,034)	-10.72%	107.47%
Shared County Total	**35,823**	**6.98%**	**3,449,542**	**3,548,173**	**98,631**	**2.86%**	**125.93%**
GRAND TOTALS	***485,386***	***100.00%***	**49,409,274**	**49,409,274**			**129.43%**

Fig. 1. A screenshot of the spreadsheet stimulus showing fair share distribution of pounds of food

3.3 Task Procedure

After describing the exercise and gaining consent, participants were asked to complete a questionnaire to gather demographic and experiential information. Next, the first stimulus was presented on one monitor and a form to collect responses was loaded on the second monitor. A brief explanation of the fields or a description of the dashboard and controls was provided at the beginning of each task. Both stimuli sets (spreadsheets or dashboard) were presented in random order across participants and tasks. Five items made up of branches and counties were presented in random order across both stimuli and participant. For each stimulus, the participant looked up a branch or county and reviewed the distribution history over four years before making a determination about distribution performance in year five.

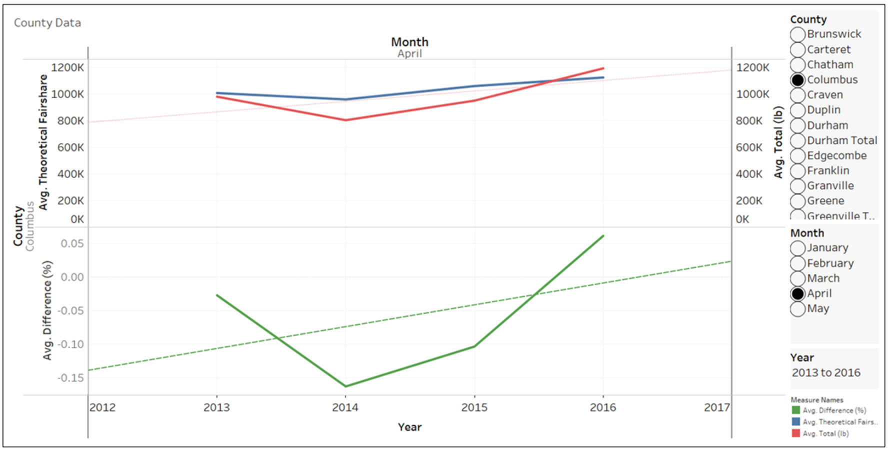


Fig. 2. The trends of avg. theoretical fair share, avg. total (lbs.) and avg. difference (%) over a 4 year period for the month of April of a selected County. Plots may be filtered on year, county and month as available in the dataset using the control panels to the right side.

Using the provided information, determine if the branch or county will meet/exceed its fair share or have a shortfall in 2017. *

	> 20% short	10 - 20% short	0 - 10% short	0 - 10% exceed	10 - 20% exceed	> 20% exceed	High Confidence	Moderate Confidence	Low Confidence	No Confidence
Durham Branch	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Craven County	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Wake County	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Raleigh Branch	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Orange County	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Fig. 3. Instrument used to collect participants' task response and confidence for decision about year five.

Using the provided response form shown in Fig. 3, the participant entered a check in the box matching their decision of shortfall or goal exceeded. On the same line, another check mark was entered to indicate the level of confidence about the decision just made.

4 Analysis and Results

An electronic form was used to collect participant responses as they performed the decision-making task. The order of the stimuli presentation was alternated to minimize the likelihood of order effects. Participants made two selections for each county/branch task using a composite of decision choices and confidence ratings - 6 choices for the decision, and four choices for confidence rating. Weights were assigned to the responses according to its proximity to the true year five decision to allow the information to be plotted as shown in this section. The weights assigned are shown in Table 1.

Table 1. Weights assigned to participant decision and confidence rating responses

Response scale (Reference)	Decision responses	Weight	Weight	Confidence scale
>20% exceed	Correct	9	6	High
10–20% exceed	10% away	6	3	Moderate
0–10% exceed	20% away	1	1	Low
0–10% short	Over 20%	0	0	No confidence
10–20% short				
>20% short				

4.1 Task Responses

The objective of the assigned tasks was to determine whether the fair share distribution goal was met in year five using data from the previous four years. Additionally, the answers were to indicate the scope of surpassing or falling short using discrete categories. The five tasks are summarized in Table 2.

Table 2. Year 5 distributions performance categories for counties and branches used in decision tasks

Response scale						
County/Branch	>20% exceed	10–20% exceed	0–10% exceed	0–10% short	10–20% short	>20% short
Craven				O		
Orange						O
Wake						O
Raleigh Br				O		
Durham Br					O	

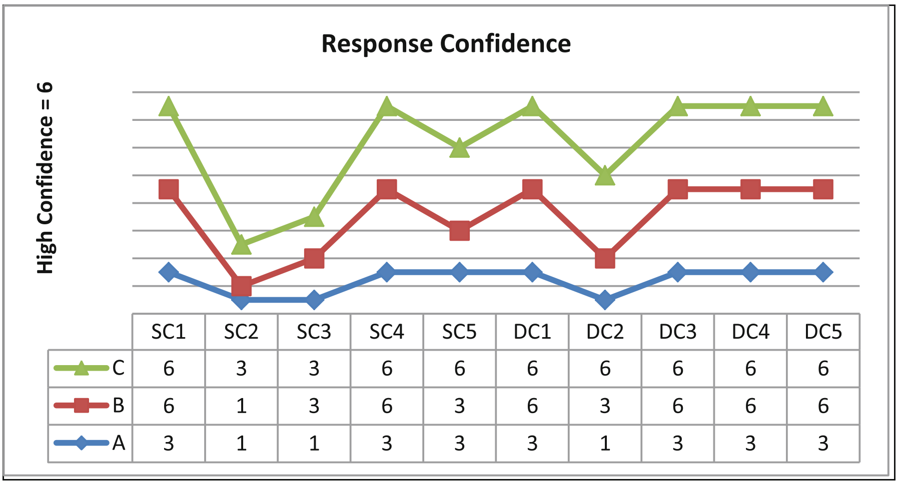
Response from the participants were analyzed to study whether two modes yielded any statistically different average responses. Model adequacy check was conducted on the data and normality assumption was met (Shapiro-Wilk $W = 0.883219$, $p = 0.2842$). A paired t test was conducted and results indicated no statistical significant difference in scores between the two presentation modes (spreadsheet vs. dashboard) ($t_5 = 1.54$, $p = 0.1850$). Concerning the small sample size, a non-parametric Wilcox Signed Rank test was also done and results revealed no statistical significant difference ($S = -4.5$, $p = 0.3125$) in scores between the two presentation modes (spreadsheet vs. dashboard).

4.2 Concurrent Confidence Ratings

Participants provided a rating of their decision confidence (see Table 1 and Fig. 3) for each sub task they performed. Participants were instructed to rate their confidence upon committing to a decision to ensure that they implicitly related confidence to decision response. Their responses can be seen in Figs. 4 and 5.

Concurrent confidence ratings were analyzed to study whether two modes yielded any statistically different average responses. Model adequacy check was conducted on the data and normality assumption was met (Shapiro-Wilk $W = 0.913362$, $p = 0.4589$). A paired t test was conducted and results indicated a statistical significant difference in ratings between the two presentation modes (spreadsheet vs. dashboard) ($t_5 = 5.22$, $p < 0.01$). Concerning the small sample size, a non-parametric Wilcox Signed Rank test was also done and similar results were obtained ($S = 10.5$, $p < 0.05$). Concurrent confidence ratings of the dashboard are significantly higher than those of the spreadsheet.

All participants reported at least moderate confidence following the task1: subtask 1, with 80% giving a high rating, but then confidence dropped across the board on subtask 2. A similar pattern occurred for task 2: subtask1 to subtask 2. At the end of task 1 and task 2, confidence was at least moderate and rated higher than confidence at each subtask 2. A noted observation for the dashboard in both cases was that confidence rose



	SC1	SC2	SC3	SC4	SC5	DC1	DC2	DC3	DC4	DC5
C	6	3	3	6	6	6	6	6	6	6
B	6	1	3	6	3	6	3	6	6	6
A	3	1	1	3	3	3	1	3	3	3

Fig. 4. Confidence rating of responses from spreadsheet task (SR#) followed by dashboard task (DR#). High = 6, Moderate = 3, Low = 1, No confidence = 0.

	DC1	DC2	DC3	DC4	DC5	SC1	SC2	SC3	SC4	SC5
F	3	6	6	6	6	6	3	3	6	6
E	6	3	6	6	6	6	1	3	6	3
D	6	1	3	3	3	3	1	1	3	3

Fig. 5. Confidence rating of responses from dashboard task (DR#) followed by spreadsheet task (SR#). High = 6, Moderate = 3, Low = 1, No confidence = 0.

after the task 2 and remained steady until dashboard tasks were complete. This suggests that a judgement policy was adopted and maintained whereas in the spreadsheet task, confidence appeared to remain steady over just the last two tasks for four of the six participants.

Given respondents knew very little about fair share distributions or food bank decision-making, lower ratings were anticipated. However, this reporting of confidence fits with studies that have shown people do think they can estimate values with greater precision than can actually be done and are typically very confident about judgements they make based on heuristics [1]. This is because they tend to be insensitive to the unsubstantiated nature of the underlying assumptions for their judgement [1].

4.3 Feedback from Participants

Most participants expressed a preference for the interactive dashboard and rated it as relatively easier to use than the spreadsheet. Participants also reported the dashboard as helpful on providing relevant information, helping to make unbiased decisions, reducing time to decision, guiding to the right information, and providing visual information.

5 Discussion and Conclusion

The intent of this study was to explore visual analytics as a tool to address cognitive bias. Ultimately, the objective is to support a broader design and research effort to develop a smart decision support system for Food Bank decision-makers working with big data. This work is in the preliminary stages and results are limited to the applications as described. It does contribute an approach on which future work may be built and qualitative evidence that will guide future design efforts.

The current study provided limited information to the participants, resulting a statistical non-significant difference in average response scores between the two presentation modes. In the future, more complex and unstructured data will be utilized to assist decision makers to combat their cognitive bias with the tool of visual analytics. For instance, the economic data such as unemployment rate, number of people applied for unemployment benefits will be visualized and more interactivity will be developed to allow participants to maneuver different visualized data in an easy way. Even with this limitation, the current study revealed participants had statistically significant higher confidence ratings for the dashboard. This was a very promising result.

Due to the small sample size, results are not generalizable to other applications. Other limitations included, limited pool of participants selected and the limits of qualitative methods chosen using subjective feedback that is subject to response bias. The subject pool comprised of experienced analytical thinkers suitable to the tasks assigned. However, the unfamiliarity of the visual analytics platform did initially appear to pose a challenge suggesting the need for longer orientation time. A more elaborate study is planned to include quantitative data collection methods such as eye gaze tracking, and utilization of more in-depth qualitative approaches using a larger sample size.

Another limitation is the limited functionality of the dashboard. In future research, proficiency with visual analytics information presentation will be increased to improve the intuitiveness to subjects. The design of survey instruments is also to be reviewed based on the collected responses to identify where experimenter bias can be eliminated for future studies. There is little known about the role that software can play in mitigating cognitive bias and insufficient evidence on the effectiveness of cognitive bias trainings. Moreover, measurement techniques for cognitive bias remain a challenge. It is expected that future research will contribute to the knowledge in these areas.

Acknowledgments. The study is supported in part by a National Science Foundation grant - PFI: Flexible Equitable Efficient Effective Distribution (FEEED) (Award number: 1718672).

References

1. Barnes Jr., J.H.: Cognitive biases and their impact on strategic planning. Strateg. Manage. J. **5**(2), 129, (1984)
2. Haselton, M.G., Nettle, D., Murray, D.R.: The Evolution of Cognitive Bias. In: The Handbook of Evolutionary Psychology (2005)
3. Caputo A.: A literature review of cognitive biases in negotiation processes. Int. J. Confl. Manag. **24**(4), 374–98 (2013)
4. Bazerman, M.H., Moore, D.: Judgment in Managerial Decision Making. Wiley, New York (2008)
5. Caledonia: Wiley (2009)
6. Lewicki, R.J., Saunders, D., Barry, B.: Negotiation. McGraw Hill, New York (2005)
7. Morewedge, C.K., Yoon, H., Scopelliti, I., Symborski, C.W., Korris, J.H., Kassam, K.S.: Debiasing decisions: improved decision making with a single training intervention. Policy Insights Behav. Brain Sci. **2**(1), 129–140 (2015)

8. Pirolli, P., Card, S.: The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In: Proceedings of International Conference on Intelligence Analysis, vol. 5, pp. 2–4, May 2005
9. Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. Science **185**, 1124–1131 (1974)
10. Liedtka, J.: Perspective: linking design thinking with innovation outcomes through cognitive bias reduction. J. Prod. Innov. Manag. **32**(6), 925–938 (2015)
11. Keim, D.A., Mansmann, F., Schneidewind, J., Thomas, J., Ziegler, H.: Visual analytics: scope and challenges. In: Visual Data Mining, pp. 76–90. Springer, Heidelberg (2008)
12. Cook, K.A., Thomas, J.J.: Illuminating the path: the research and development agenda for visual analytics (2005)
13. Denmark, D., Harker, D., McCollough, A.: Interliminal design: mitigating cognitive bias and design distortion (2013)
14. Nussbaumer, A., Verbert, K., Hillemann, E.C., Bedek, M.A., Albert, D.: A framework for cognitive bias detection and feedback in a visual analytics environment. In: European Intelligence and Security Informatics Conference (EISIC), pp. 148–151. IEEE, August 2016
15. Pohl, M., Winter, L.C., Pallaris, C., Attfield, S., Wong, B.W.: Sensemaking and cognitive bias mitigation in visual analytics. In: 2014 IEEE Joint Intelligence and Security Informatics Conference (JISIC), pp. 323–323. IEEE, September 2014
16. Wall, E., Blaha, L.M., Franklin, L., Endert, A.: Warning, bias may occur: a proposed approach to detecting cognitive bias in interactive visual analytics. In: IEEE Conference on Visual Analytics Science and Technology (2017)