Investigating Representation Alternatives for Communicating Uncertainty to Non-experts

Miriam Greis, Thorsten Ohler, Niels Henze, Albrecht Schmidt

Investigating Representation Alternatives for Communicating Uncertainty to Non-Experts

Miriam Greis, Thorsten Ohler, Niels Henze, Albrecht Schmidt

VIS, University of Stuttgart, Stuttgart, Germany
firstname.lastname@vis.uni-stuttgart.de

Abstract. Non-experts are confronted with uncertainty of predictions everyday when, e.g., using a navigation device or looking at the weather forecast. However, there are no standards for representing uncertain information and representations could be easily misleading. Thus, we selected twelve representations that provide different levels of uncertainty information. We compared the representations in an online survey with 90 participants where we asked participants to judge their support in decision-making, familiarity, easiness to understand, and visual appeal. We further evaluated the four most promising representations in a turn-based online game. Players had to make decisions in a farming scenario based on a displayed weather forecast. The results of the survey and the game indicate that a function graph of a probability distribution function is the best way to communicate uncertain information. Nevertheless, our results also show that presenting more uncertainty information does not necessarily lead to better decisions.

Keywords. Uncertainty · Representations · Visualizations · Non-Experts

1 Introduction

Simulation is a very powerful technique used in different fields to explore the behavior of complex systems as, e.g., the flow of ground water, the human walking behavior, and the world's climate. Due to its applicability to many problems, simulation is one of the most used techniques in research and management sciences [4].

The results of simulations are uncertain due to, e.g., assumptions made in the modeling process or the parameter choice. Non-experts, who have no specialized knowledge about simulations, are confronted with simulation results and uncertain data everyday, e.g., when looking at the weather forecast that shows the possible temperature for the next day. They rely on uncertain data to implicitly predict the future, plan activities, and make decisions. These decisions could be even unintentionally manipulated by choosing a specific representation.

Although several aspects of the communication and visualization of uncertain data were examined before, still work has to be done to get more insights in how different representations influence decision-making. From existing work, we selected twelve representations (three textual and nine graphical representations) used for uncertainty communication for experts and non-experts. This leads to a wide range of used representations with different degrees of included uncertainty information.

First, we present the results of an online survey where participants had to rate the representations according to the perceived help for decision support in a farming scenario, familiarity, easiness to understand, and visual appeal. We found significant differences in the rating of the representations and also correlations between ratings. Surprisingly, we did not find any strong correlations between subjective ratings and the degree of uncertainty information included in a representation.

Second, we present the results of a small experiment with a turn-based online game. The four representations that performed best in the online survey were used to display a weather forecast. Decisions of players were analyzed on the basis of optimal decisions. The function graph of a probability distribution function performed best, but the representation with no information about the included uncertainty did not perform dramatically worse.

The contribution of our work is twofold. First, we show that people do not judge representations for the degree of uncertainty information they show. Other factors are more important for their judgment. Second, we show that a higher degree of uncertainty information does not necessarily lead to better decisions and that people are able to make good decisions with the help of a probability density function.

2 Related Work

The topic of uncertainty visualization is well explored for experts. Multiple areas such as vector field, surface and glyph visualizations are for example explored by Pang et al. [9] and Zuk and Carpendale [15]. Additionally, new versions of basic representations (e.g., box plots) are for example developed by Potter et al. [11]. In our paper, we do not take into account such visualizations and focus on basic representations developed for experts and non-experts.

For communicating uncertainty to non-experts, usually quantitative information, especially probabilities are used. One problem when using quantitative information is the inability of even well-educated adults to solve easy numeracy probability questions [5]. But also qualitative information, e.g., labels such as low uncertainty or low risk, could be misleading, as already examined by Wallsten et al. [14]. Additionally, the formulation of a risk or uncertainty, whether negative or positive, has a huge influence on decision-making [6].

One strand of work investigated uncertainty information in weather forecasts. Morss et al. [7] found that most people are aware of the uncertainty in deterministic weather forecasts, although the range of this uncertainty was perceived very differently. Additionally, 70 % of the people preferred forecasts that contained information about the uncertainty of the forecast. Studies by Roulston et al. [12] and Joslyn et al. [3] showed that people make better decisions when having information about the uncertainty of a forecast and that information about the uncertainty also increases the trust in a forecast. They used a small number of alternative representations in a decision task.

Another strand of work investigated and compared visualizations including uncertainty information. Ibrekk et al. [2] compared nine visualizations for uncertainty by giving non-experts specific tasks (e.g., finding the mean). They suggest displaying a normal probability distribution function together with a cumulative probability distribution function. Pappenberger et al. [10] made a study with experts in meteorology and asked them about their preferred representation for a probabilistic forecast. The most used representation were quantiles.

Additional studies focus on one specific visualization or aspect. Olston et al. [8] examined visualizations for presenting bounded uncertainty by adjusting the visual elements and including transparency in the visualizations and Tak et al. [13] used seven variations of a line graph to investigate the perceived certainty. Correll and Gleicher [1] proposed a redesign of bar charts and found that less well-known visualizations improve performance for interferential tasks.

Previous work mainly focused either on a small number or representations, variations of one representation, or a very specific task (e.g., finding the mean). They found that uncertainty information leads to better decisions, but it is not clear how this conclusion relates to different degrees of uncertainty information and how aggregated uncertainty information influence decision-making. In contrast to previous work, we decided to take a wide range of basic representations and concentrate our research on the degree of uncertainty information that is included in these representations.

3 Online Survey

To understand if the presented degree of uncertainty changes the perceived value of a representation for decision support and the easiness to understand the representation, we conducted an online survey. Building upon prior research, we selected twelve representations (see Figure 1) with different properties for communicating uncertainty information. All representations show the expected rainfall for the next three days. Three representations use a textual representation of the information, whilst the other nine representations are graphical. We use a line chart, a box-and-whisker plot, bar charts, stacked bar charts, stacked area diagrams, shaded bars, and function graphs. The representations also communicate different degrees of uncertainty information, from no information about uncertainty at all up to detailed information. The displayed degree of uncertainty information for each representation is depicted in Table 1.

3.1 Questionnaire

On the first page of the questionnaire, we asked participants for demographic information: age, gender, their highest degree, and their field of work.

We then displayed a scenario description that told participants to imagine that they are a farmer and want to grow plants. The plants need a certain amount of water to grow and survive. A weather forecast will be available to support participants in making a decision, but it will be uncertain.

Participants then had to navigate through twelve pages with one representation on each page. The order of the representations was randomized across participants. For each representation, participants had to indicate their level of agreement on a five-level Likert scale from totally disagree to totally agree with four statements:

1. The representation supports me in making a decision.
2. I am familiar with the representation.
3. The representation is easy to understand.
4. The representation is visually appealing.

Fig. 1. All 12 representations compared in the online survey. Explanations: 1 – 3 Textual representations, 4 – Line chart with area diagram, 5 – Box-and-whisker plot, 6 – 7 Bar charts, 8 – Stacked bar charts, 9 – Area chart, 10 – Shaded horizontal bars, 11 – 12 Function graph

Table 1. Degree of uncertainty information included in the representations.

	Textual Representation	Graphical Representation
No Uncertainty Information	REPR 1: Expected values	/
Aggregated Uncertainty Information	REPR 2: Expected values and standard deviation REPR 3: Quantiles	REPR 4: Expected values and confidence interval REPR 5: Quantiles REPR 6: Expected values and standard deviation
Detailed Aggregated Uncertainty Information	/	REPR 7 – 9: Aggregated probability density function REPR 10: Color-coded probability density function
Detailed Uncertainty Information	/	REPR 11: Probability Density Function REPR 12: Cumulative Probability Density Function

3.2 Participants

We invited participants through social networks to participate in our survey. In total, 90 participants (36 female, 54 male) fully answered our online survey. Participants' age ranged from 18 to 82 years with a mean of 31 years (sd: 12.6). 45 % of the participants had a university degree, 28 % had a high school diploma, further 20 % finished a vocational training, and all other participants had a minor degree or no degree at all. Participants worked in different field such as computer science, economics, commerce, teaching, mechanics, services, and others. They had no specialized knowledge about simulations.

3.3 Results

For each representation, we calculated the mean for each of the four statements and the overall mean (see Table 2). For each statement, we also conducted a Friedman test and for the post hoc analysis Wilcoxon signed-rank tests with an applied Bonferroni correction. The Friedman test showed that there is a statistically significant difference between the twelve representations for each statement.

The Wilcoxon singed-rank tests showed that representation 1, 4, 7, and 11 performed significantly better than the majority of other representations in at least one judgment each. Representation 3 was rated significantly worse than the majority of representations on three scales.

We ran a Spearman's rank-order correlation to determine relationships between our 1080 Likert items and the degree of uncertainty information of the different representations. As expected, we found strong, positive correlations between all pairs of Likert items, which were all statistically significant ($p < 0.0005$), see Table 3 for detailed values. Surprisingly, we did not find any significant positive or negative correlation between the Likert items and the degree of uncertainty information of the representations, except one moderate positive correlation with the Likert items for visual appeal (see Table 3). We assume that this correlation occurred because we used textual representations with low degrees of uncertainty information.

Table 2. Calculated mean values for the level of agreement on a five-level Likert scale from totally disagree (1) to totally agree (5) with the statements: S1 – The representation supports me in making a decision., S2 – I am familiar with the representation., S3 – The representation is easy to understand., S4 – The representation is visually appealing., and O – the overall mean values for all 12 representations.

REPR	S1	S2	S3	S4	O
1: Expected values	3.63	**4.29**	**4.28**	2.30	**3.63**
2: Expected values and standard deviation	3.68	3.90	3.34	2.13	3.26
3: Quantiles	2.86	2.79	2.37	1.70	2.43
4: Line chart with confidence interval	**4.12**	3.74	4.12	**4.01**	**4.00**
5: Boxplot	3.22	2.73	2.52	2.48	2.74
6: Bar chart with error bars	3.48	3.11	3.08	2.98	3.16
7: Histograms as bar charts	3.93	4.21	3.67	3.67	**3.87**
8: Histograms as stacked bar charts	3.47	3.49	3.16	3.52	3.41
9: Histograms as area chart	3.16	2.84	2.74	3.42	3.04
10: Shaded horizontal bars	3.73	2.31	3.59	3.59	3.31
11: Probability distribution function	3.89	3.88	3.46	3.60	**3.71**
12: Cum. probability distribution function	3.44	3.50	2.88	3.52	3.34

Table 3. Spearmen's rho for a Spearman's rank-order correlation between Likert scale items of the online survey and the degree of uncertainty information of the presented representations. Statistically significant values are marked with asterisk(s). $^{**}p < 0.01$, $^{*}p > 0.0005$

	Decision Support	Familiarity	Easiness to Understand	Visual Appeal
Degree of Uncertainty	0.048	-0.040	-0.084**	0.365*
Decision Support	-	0.505*	0.670*	0.527*
Familiarity	-	-	0.609*	0.530*
Easiness to Understand	-	-	-	0.530*

4 Experiment

Based on the results of the online survey, we implemented a turn-based online game that displayed a weather forecast using the four representations that performed best in our online survey. With the help of the weather forecast for the next three days, players had to decide which crops they want to plant. Crops needed specific values for rainfall, wind, and sun that were always displayed and gave different amounts of money when fully grown and harvested. Players had the goal to get as much money as possible. If one weather condition for a planted crop was violated, it withered and it gave no money at all.

We conducted a controlled experiment with 12 participants (4 female, 8 male), recruited with the help of social networks, who had to play the game with each of the four representations. The order of the representations was changed for each participant and randomly assigned to them. Each game lasted for 10 rounds. Participants were not trained and did not have any specialized knowledge about simulations. For the analysis, we logged all relevant information, which included the forecasted weather, the real weather, and all crops that were planted.

Participants played in total 480 rounds of the game, 120 rounds with each representation. For our analysis, we only considered rounds in which the players at least clicked on the button to open the weather forecasts once. This resulted in 442 rounds.

For each round, we calculated the optimal decision based on the weather forecast displayed and compared it with players' decisions. With representation 11, participants made the most optimal decisions, in 69 % of the rounds, with representation 7 in 64 % of the rounds, with representation 1 in 60 % of the round and with representation 4 in 57 % of the rounds.

At the end, we asked participants which representation they liked most. 8 participants selected representation 11, three participants selected representation 7, 1 participant selected representation 4, and no participant selected representation 1.

5 Discussion & Conclusion

The results of our online survey show that the representations were judged very differently. Surprisingly, the four best-rated representations taking the overall mean show a different degree of uncertainty information each. Correlating the judgments and the degree of uncertainty information provided by the representations, we found that there are no significant correlations between the perceived support in decision-making and the degree of uncertainty information. Thus, we showed that participants do not judge the perceived support of decision making for a representation in regard to the degree of uncertainty information presented. In contrast, factors such as familiarity, easiness to understand, and visual appeal have a huge influence on the judgment. Our work indicates that other factors besides the presented degree of uncertainty information have to be considered when displaying uncertain data.

Comparing results from the online survey and the experiment, we found that the line chart with the highest rating in the survey performed worse than the other representations in the experiment. Although other studies showed that participants make better decision when having uncertainty information, our experiment suggests that aggregated uncertainty information do not provide enough details to make better decisions than only information about the expected values. We assume that the aggregated uncertainty information, the confidence interval, was not enough information to make a better decision but instead alienated participants. Nevertheless, most participants in our experiment preferred using the representation of the probability distribution function. Additionally, participants made very good decisions using this probability density function although earlier studies suggest that people are not very good at intuitively interpreting statistical information. This indicates that a probability function should be used to communicate uncertain data for non-experts and that too much aggregation of uncertainty information should be avoided.

Our experiment is clearly limited by the small number of participants and the scenario, but nevertheless shows interesting factors that are relevant for uncertainty communication. Future work on different scenarios with more participants will help to generalize the findings.

Acknowledgements. The authors would like to thank the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/2) at the University of Stuttgart.

References

1. M. Correll and M. Gleicher, "Error bars considered harmful: Exploring alternate encodings for mean and error," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, 2014.
2. H. Ibrekk and M. G. Morgan, "Graphical communication of uncertain quantities to nontechnical people," *Risk Anal.*, vol. 7, no. 4, 1987.
3. S. L. Joslyn and J. E. LeClerc, "Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error.," *J. Exp. Psychol. Appl.*, vol. 18, no. 1, 2012.
4. A. M. Law, W. D. Kelton, and W. D. Kelton, *Simulation modeling and analysis*, 1991.
5. I. M. Lipkus, G. Samsa, and B. K. Rimer, "General performance on a numeracy scale among highly educated samples," *Med. Decis. Mak.*, vol. 21, no. 1, 2001.
6. B. J. McNeil, S. G. Pauker, H. C. Sox Jr, and A. Tversky, "On the elicitation of preferences for alternative therapies.," *N. Engl. J. Med.*, vol. 306, no. 21, 1982.
7. R. E. Morss, J. L. Demuth, and J. K. Lazo, "Communicating uncertainty in weather forecasts: A survey of the US public," *Weather Forecast.*, vol. 23, no. 5, 2008.
8. C. Olston and J. Mackinlay, "Visualizing data with bounded uncertainty," in *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*, 2002.
9. A. Pang, C. Wittenbrink, and S. Lodha, "Approaches to Uncertainty Visualization," *Vis. Comput.*, vol. 13, no. 8, 1997.
10. F. Pappenberger, E. Stephens, J. Thielen, P. Salamon, D. Demeritt, S. J. Andel, F. Wetterhall, and L. Alfieri, "Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication," *Hydrol. Process.*, vol. 27, no. 1, 2013.
11. K. Potter, J. Kniss, R. Riesenfeld, and C. R. Johnson, "Visualizing Summary Statistics and Uncertainty," *Comput. Graph. Forum*, vol. 29, no. 3, 2010.
12. M. S. Roulston, G. E. Bolton, A. N. Kleit, and A. L. Sears-Collins, "A laboratory study of the benefits of including uncertainty information in weather forecasts," *Weather Forecast.*, vol. 21, no. 1, 2006.
13. S. Tak, A. Toet, and J. Van Erp, "The perception of visual uncertainty representation by non-experts," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, 2014.
14. T. S. Wallsten, D. V Budescu, A. Rapoport, R. Zwick, and B. Forsyth, "Measuring the vague meanings of probability terms.," *J. Exp. Psychol. Gen.*, vol. 115, no. 4, 1986.
15. T. Zuk and S. Carpendale, "Theoretical analysis of uncertainty visualizations," *Proc. IS&T/SPIE Electron. Imaging*, vol. 6060, 2006.