Input Controls for Entering Uncertain Data: Probability Distribution Sliders

MIRIAM GREIS, HENDRIK SCHUFF, MARIUS KLEINER, NIELS HENZE, and ALBRECHT SCHMIDT, University of Stuttgart

Although more and more data is collected automatically, many interfaces still require manual input. When we, for example, enter our daily calorie intake or calculate our ecological footprint, we often have to guess the weight of the food or what distance we have covered with our car. In this paper, we propose a solution to overcome the problem of forcing users to enter a single value when they are unsure about the actual input. On the basis of a slider, we designed four input controls which allow the input of uncertain data in the form of probability distribution functions. To evaluate our input controls, we conducted two studies collecting subjective and objective feedback. Based on the evaluation, we derived implications for their usage. We additionally provide an open-source toolkit with the evaluated input controls that can be included in web applications and customized for different contexts and tasks.

CCS Concepts: • **Human-centered computing → User interface toolkits**;

Additional Key Words and Phrases: Input control, probability distribution, slider, uncertainty, toolkit

ACM Reference format:
Miriam Greis, Hendrik Schuff, Marius Kleiner, Niels Henze, and Albrecht Schmidt. 2017. Input Controls for Entering Uncertain Data: Probability Distribution Sliders. *Proc. ACM Hum.-Comput. Interact.* 1, 1, Article 2 (June 2017), 17 pages.
https://doi.org/http://dx.doi.org/10.1145/3095805

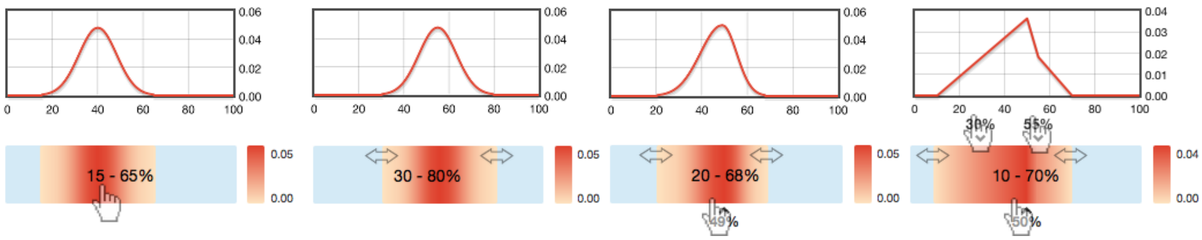
1 INTRODUCTION

Data uncertainty is a concept that everyone has to deal with in everyday life. Everyone makes decisions based on the weather forecasts or the bus schedule, many people carry around devices that count their steps or track their current activity, and nearly everyone uses navigation systems that recommend the best route. Most of this data, which is used as a basis for decisions, is uncertain and can be wrong.

Although uncertain data is present everywhere, its uncertainty is seldom taken into account when developing interactive systems, especially on the input side. Research has so far focused on the visualization of uncertain data. Different visualizations for experts such as glyphs [35] have been analyzed or compared [10, 24, 37]. For laymen, quantitative and qualitative methods can be used to communicate uncertain data, but both methods have their drawbacks. Even for well-educated adults it can be difficult to understand quantitative information such as easy probability questions [17], and qualitative formulations such as low risk or low uncertainty can be misleading due to different perceptions of qualitative terms [33].

One of the main problems of uncertainty visualization lies in the quantification of uncertainty. Before the uncertainty can be visualized, it has to be captured and modeled in a way that ensures data quality and correctness [1]. This also includes the quantification of uncertainty introduced through user input, e.g., uncertainty due to

2573-0142/2017/6-ART2 $15.00
https://doi.org/http://dx.doi.org/10.1145/3095805



(a) Fixed Range Slider (b) Flexible Range Slider (c) Flexible Range Best Estimate Slider (d) Advanced Flexible Range Best Estimate Slider

Fig. 1. Four slider controls each enabling users to enter a probability distribution function. The sliders have ascending degrees of freedom (left to right). The interaction cues shown in the pictures are displayed when hovering over an interactive element (drag handle) of the slider control.

technical inaccuracy for touch interfaces or forms [28]. What is not captured by interactive systems so far is the uncertainty directly introduced by the user. When tracking their calorie intake, for example, users have to manually enter the weight of the food that they have eaten. But without a scale, it is hard to enter the exact value. Another example is the calculation of the ecological footprint, where users have to specify how many kilometers they have driven with their car each month. This could be very different from month to month. Current interfaces offer number fields to enter such data which forces users to enter a single value, although they probably do not know this value. The uncertainty introduced through such an input can neither be quantified nor visualized or in any other way be reflected in the output. Knowing whether a user is uncertain or even how much uncertainty a manual input of a user contains makes it possible for systems to reflect this in the output. This makes the output more reliable, transparent, and easier for users as they can see how their uncertainty influences the output.

In this paper, we take a first step to explore standard interface controls for entering uncertain data. We designed new user interface controls: sliders that allow the user to enter a probability distribution function (see Figure 1). The sliders provide different degrees of freedom. We evaluated the input controls in two steps; a subjective evaluation in an online survey followed by an objective evaluation in a controlled user study. Based on the results, we implemented a web-based toolkit containing the input controls, which allows researchers and developers to easily use and customize the input controls for their projects. One goal of our work is to tackle a source of uncertainty in interactive systems that is so far mostly overseen: the user input. The more specific goals of our work are to understand how people interact with input controls that allow them to specify uncertainty, and to enable other researchers and developers to use our input controls and insights in future work.

Based on these goals, our contributions include the following:

(1) **Identification of sources of uncertainty in interactive systems**.
(2) **Design and evaluation of probability distribution sliders** for the input of uncertain data collecting objective and subjective feedback resulting in **recommendations and implications** for when to use the specific slider controls.
(3) **Development of a web-based open source toolkit** for web developers and researchers that allows easy modification of the probability distribution sliders for different tasks and contexts.

Our work encourages researchers and developers designing interactive systems to take uncertainty in manual user interfaces into account.

2 BACKGROUND & RELATED WORK

Our work is mainly based on three topics: definitions of uncertainty, visualization of uncertainty, and sliders.

2.1 Definitions of Uncertainty

Most previous work about the classification and visualization of uncertain data was conducted in a specific domain which leads to manifold definitions and references to the term uncertainty. The term is used inconsistently across domains, because each domain focused on their research independently [18]. Specialized topologies could assist analysts to make better decisions, for example for geospatial information [32].

Skeels et al. [30] built a classification of uncertainty based on previous literature and by conducting interviews. The classification contains three levels of uncertainty: measurement precision, completeness, and inferences. Gershon [6] discusses the manifold sources of imperfection of information; distinguishing between corrupt, incomplete, inconsistent, complicated, and uncertain data. Clearly driven by the visualization domain, Pang et al. focus [24] on a classification of methods for uncertainty visualization by data type and visualization extent.

As seen, most of this work focuses on different aspects when building classifications of uncertainty, but do not fully address or apply to interactive systems.

2.2 Visualization of Uncertainty

The visualization of uncertain data has been explored in different research areas. Pang et al. [24] and Zuk and Carpendale [37] worked on glyph visualizations to visualize uncertainty in vector fields and surfaces for experts. Other work in the visualization domain focuses on a specific type of uncertainty (e.g., bounded uncertainty [22]) or a specific visualization such as line graphs [31], box plots [26], or bar charts [3].

Other work addresses domain-specific problems such as weather forecasting, in which the research focus lies on the visualization of uncertainty information in weather forecasts and its influence on both meteorologists and laymen. A study by Morss et al. [21] indicates that people are aware of the uncertain nature of deterministic weather forecasts, although the perceived range of this uncertainty differs between people. This and further studies clearly indicate that showing uncertainty information is not only preferred by people [21], but also helps them to make better decisions [11, 27]. The same studies as well indicate that showing uncertainty increases the perceived transparency and reliability due to the increased trust in a forecast [11, 27]. This does not only hold for weather forecasts but can be applied in many situations, for example for body weight measurements where missing uncertainty information decreases trust [14]. All these positive findings suggest that showing uncertainty information should also be considered in interactive systems.

A large number of studies compared uncertainty visualizations for different groups of people. Pappenberger et al. [25] identified quantiles as the representation most used by experts in meteorology. Ibrekk et al. [8] conducted a study with laymen who had to find the mean of a probability distribution with nine different visualizations. They suggest displaying a normal probability distribution function and a cumulative probability function on top of each other. This is in line with findings of Greis et al. [7], who found that laymen performed best in a game when supported by a probability distribution function plot. Color gradients also proved suitable for visualizing probability distributions [9], and have even been suggested as an alternative to bar charts with error bars [3]. Overall, visualizing uncertainty as a probability distribution function or gradient seems a promising direction to communicate uncertainty both to experts and laymen.

Research in HCI has started to focus on uncertainty visualization and communication recently as well, for the exploration of personal genomics data [29], data analysis [5], machine learning [15, 36], bus arrival predictions [13], range anxiety in electric cars [12] and other applications.

2.3 Sliders

Sliders are flexible input controls used for a variety of tasks. As visual analogue scales, they are for example used in clinical trials and research [19]. Their design in this context has been studied thoroughly revealing that, for example, tick marks introduce a bias while a banded design or dynamic feedback does not [20]. Sliders have an even longer history of being used for data exploration. With the help of the alpha slider [23] words, phrases, or names from textual lists can be selected. In this context, sliders are used as a selection input control.

Eick et al. [4] developed enhanced sliders that show the distribution of the data as a density plot in the slider bar, thus the sliders are used as a filtering mechanism for existing data. A similar approach is suggested by Willett et al. [34] with their scented widgets, in which the sliders incorporate visual elements that help to select and explore data. Sliders therefore can support filtering and exploration of data by providing visual feedback of their selection. Lasram et al. [16] visually showed the effect of a slider control on an image by enhancing the slider bar. This served as a preview to allow users to intuitively use the sliders. Here, visual elements helped to understand the outcome of the interaction while adjusting the sliders. Overall, sliders offer a great potential to integrate visual feedback on the selection.

3 DESIGN PROCESS

Research has shown that visualizing uncertainty information has many advantages, and the topic is also recently gaining attention in HCI. In the following, we identify potential sources of uncertainty in interactive systems that need to be addressed to quantify uncertainty. We observe that user input plays an important role and is so far seldom taken into account when quantifying uncertainty. We therefore focus on designing for uncertain input and present our design rationale for our input controls, the probability distribution sliders.

3.1 Sources of Uncertainty in Interactive Systems

Based on the related work we have presented and the architecture of interactive systems, we derived potential sources of uncertainty for interactive systems based on real world models (see Figure 2). The following sources should be taken into account: (1) *Structural uncertainty* is introduced because a model can only partially copy the real world. During the modeling step, parts of the real world have to be left out deliberately and not considered in the model. (2) *Algorithmic uncertainty* could be introduced when converting a model to source code as an identical transformation can be difficult. (3) Due to a *lack of knowledge* or *imprecise measurements*, e.g., of the amount of consumed food, a user may enter wrong data when using an interactive system. (4) Additionally, the user might have a wrong *understanding of the input methods* provided by the system or face technical *input method limitations* (for example when using a slider with a fixed number of pixels). (5) The input method itself might not support enough *degrees on freedom* (e.g., a number field instead of allowing probabilistic input) or require *data transformation* (e.g., the input data could be transformed with an outdated measure such as an outdated currency exchange rate). (6) As input methods, output methods can have restricted degrees of freedom (e.g. not supporting uncertainty information). (7) The potentially wrong *understanding of the output methods* and potential *output method limitations* apply, too. (8) Lastly, the user can *misinterpret* the presented data and understand the output in a different way.

User input plays an important role in interactive systems. So far, input controls do not support the degrees of freedom necessary to support the input of uncertain data. In the following, we make a first step to explore input controls for uncertain data.

3.2 Designing for Uncertain Input

In a prototyping session we designed low fidelity prototypes for the input of uncertain data. We realized that designs based on different input controls such as number fields, radio buttons, or sliders lead to completely

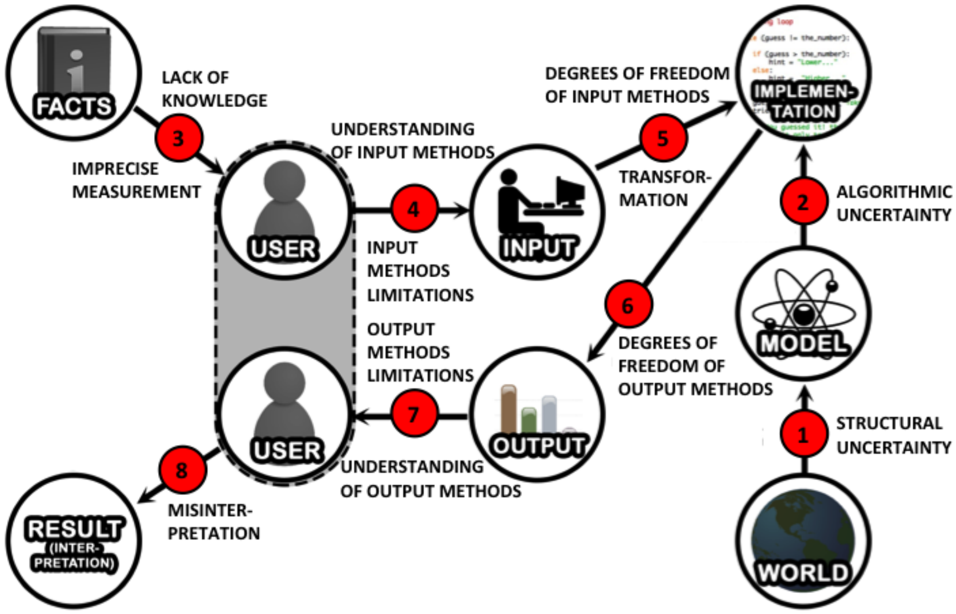


Fig. 2. Schematic representation of an interactive system that includes a real-world model. Each number on an edge represents a step in the process where uncertainty could be introduced specifying different sources of uncertainty that have to be taken into account.

different prototypes that do not have any common features and therefore cannot be compared in a meaningful way. As there is no former research which could be used as a baseline, we decided to focus on one specific input control only, allowing us to compare the newly designed input controls against a meaningful baseline. Additionally, we aimed to design input controls that had the greatest possible transparency by giving users feedback on how their input would be interpreted by the system.

One input control that has a huge amount of flexibility and fulfills these requirements is a slider. Sliders are often combined with enhancing visualizations, which can be easily adapted for the input of uncertain data. Adding enhancing visualizations to other input controls would be unusual for people. Additionally, people already use different versions of sliders for searching and filtering data; sometimes even specifying a range instead of a single number (e.g., on websites of online shops). Related work shows that probability distribution functions either as function or gradient plots are a promising way to communicate uncertain data to non-technical people. Other representations such as box plots or quantiles are not suitable for laymen but mainly used by experts. The combination of sliders with probability function plots or gradient plots is therefore a promising first step to design input controls for uncertain input.

3.3 Probability Distribution Sliders

A common input control allows users to enter a single value (e.g. a number), which could be either the mode, the median, or the mean (expected value) of a probability distribution. Based on this, we first derived levels with varying degrees of freedom for probabilistic input. A probability distribution function can be specified by the following four properties: mode, standard deviation, skew, and kurtosis. Each of these properties can be either

	SD	**Skew**	**Kurtosis**
Level 0	not included	not included	not included
Level 1	fixed	fixed	fixed
Level 2	adjustable	fixed	fixed
Level 3	adjustable	adjustable	fixed
Level 4	adjustable	adjustable	adjustable

Table 1. Deriving levels with varying degrees of freedom for entering a probability distribution function.

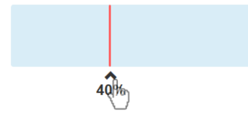


Fig. 3. The basic slider used as baseline input control for the probability distribution sliders.

not included, fixed, or adjustable by the users. As the properties have a rising complexity, we derived five levels with rising flexibility listed in Table 1.

As a baseline, we used a standard slider (depicted in Figure 3), which just allows the user to enter a single value (level 0). Based on the standard slider and the four additional levels, we developed four input controls (abbreviated with *IC* in the following) for specifying a probability distribution function. The number of the *IC* corresponds to its level.

IC1 allows the user to input a fixed range by moving a fixed part of the slider bar (see Figure 1a) which corresponds to a probability distribution with fixed standard deviation, fixed skew, and fixed kurtosis.

IC2 allows the user to drag at the two ends of the selection (see Figure 1b) to create a range with a flexible size.

IC3 offers an additional selection of the mode on top of the flexible range selection of IC2 (see Figure 1c). By selecting the mode of the probability distribution function, users can influence its skew to specify asymmetric distributions.

IC4 additionally provides the possibility to specify two more values (half as high as the mode) of the probability distribution function (see Figure 1d). This allows users to influence the kurtosis of the function.

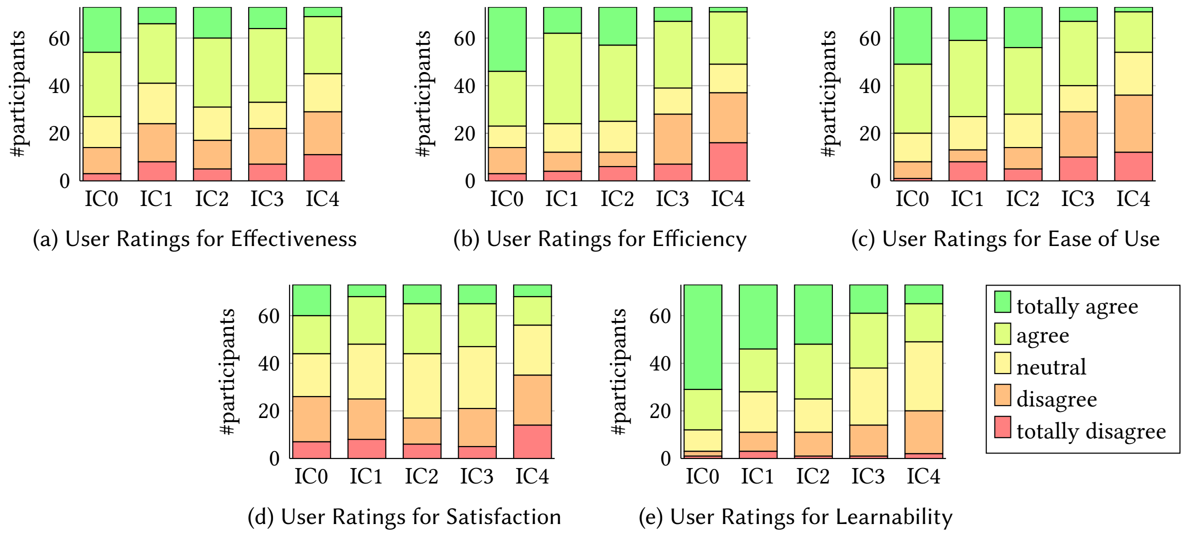
All input controls were designed to have analogous features allowing them to be evaluated and compared. For *IC4*, we could have chosen an arbitrary amount of additional values, but this would have introduced a completely new input control not comparable to the others. To achieve transparency of the input process we added three supportive visual elements: (1) A gradient plot providing the height information of the probability distribution function to support the intuitive feeling of the user without understanding the details of a probability distribution function; (2) A gradient height legend to determine the height of different points on the function or at least the height of the peak, which can be directly read from the top of the legend; (3) A plot of a probability distribution function where the height of a certain point can be easily determined, although basic mathematical knowledge might be required to fully understand the plot. In addition to the supportive visual elements, we added interaction cues and tooltips. Tooltips are shown before users interact with a control while the interaction cues are displayed when hovering over the interactive elements of a control.

4 SUBJECTIVE ASSESSMENT

To evaluate the probability distribution sliders, we first conducted an online survey. We collected subjective feedback according to perceived effectiveness, efficiency, ease of use, satisfaction, and learnability of the controls.

4.1 Tasks & Procedure

After a general introduction, participants were first asked to provide demographic information and assess their knowledge about stochastics, statistics, probability theory, and probability distributions. We then presented the base line slider and the probability distribution sliders randomizing the order of the ICs across participants to reduce sequence effects. For each IC, we presented an exemplary task displaying a table that showed how often a car was used over 36 months by "Sam Sample" to get to work, and provided the following question: "How many



(a) User Ratings for Effectiveness (b) User Ratings for Efficiency (c) User Ratings for Ease of Use

(d) User Ratings for Satisfaction (e) User Ratings for Learnability

Fig. 4. Results of the online survey showing the agreements of 73 participants on a five-point Likert scale about the input controls (IC0 to IC4). The exact formulation of the statements is provided in Section Tasks & Procedure.

	IC0		IC1		IC2		IC3		IC4	
	M	SD	M	SD	M	SD	M	SD	M	SD
Effectiveness	3.66	1.15	3.10	1.18	3.45	1.17	3.27	1.20	2.89	1.19
Efficiency	3.82	1.21	3.60	1.05	3.63	1.16	3.07	1.18	2.63	1.21
Ease of Use	3.93	1.00	3.53	1.20	3.59	1.18	3.00	1.24	2.63	1.10
Satisfaction	3.12	1.26	2.96	1.11	3.19	1.09	3.11	1.09	2.63	1.17
Learnability	4.38	0.91	3.79	1.18	3.84	1.09	3.44	1.01	3.14	1.00

Table 2. Mean values and standard deviations for all Likert scale ratings of participants in the online survey.

times each month does Sam Sample use his car to go to work?". The description made clear that the task was an example and participants did not have to solve it. We additionally presented a short description of the control. Participants were then asked to try out the control and indicate their level of agreement on a five-point Likert scale with the following four items:

Effectiveness: "*I am confident that I am able to correctly enter data with this input method.*"
Efficiency: "*I was able to quickly enter data using this input method.*"
Ease of use: "*It was simple to use this input method.*"
Satisfaction: "*I liked using this input method.*"

Additionally, we asked participants to judge the learnability of the input control by using a five-point Likert scale ranging from "*I could use this input method intuitively (without reading the description)*" to "*I doubt I will ever be able to confidentially use this input method (even after training)*". Participants were then able to optionally leave positive and negative remarks, reasons for their judgments, and further suggestions, questions, or comments in a text field. In the end, we asked participants to rank the ICs according to how much they liked them and how useful they think they were. We additionally asked them to judge whether they understood, liked, and found the supportive visual elements (the probability distribution function plot, the gradient plot, and the gradient height legend) useful.

4.2 Participants

In total, 75 participants (34 female, 40 male, 1 preferred not to say) answered the online survey completely. They were recruited via social media and by e-mailing our list of volunteers. We decided to exclude two male participants from the analysis: One asked us to remove his data in one of the comment fields; the other took less than five seconds to click through each page of the survey, which deviated too much from the average time participants needed to answer the questions appropriately.

Finally, we analyzed the data from 73 participants who had an average age of 25.97 years ($SD = 5.97$) and who were mostly students and employees with different subjects and fields of work. More than 90 % had a high school degree or an even higher degree, such as a bachelor (31.51%) or master degree (17.81%). We also asked participants about their previous stochastic and statistical knowledge and presented six different levels from which they should select: from "*no knowledge at all*" to "*knowledge about stochastic, statistics, probability theory, and probability distributions*". We converted the answers to a knowledge level with rising knowledge from 1 to 6. Participants on average reported to have a knowledge level of 4.55 ($SD = 1.85$), which correlates to them having some stochastic and statistics knowledge.

4.3 Results

For the analysis, we converted all Likert scale ratings of the subjective assessment to numbers, associating totally disagree with the number 1 and totally agree with the number 5. We also converted the items for learnability to the same scale. We then analyzed all of the statements independently.

4.3.1 Metrics for Input Controls. All Likert scale ratings for all statements are depicted in Figure 4 and for easier comparison, means and SDs are depicted in Table 2. For each statement, we conducted a Friedman test to show that there is a statistically significant difference in the perceived effectiveness, efficiency, ease of use, satisfaction, and learnability of the input controls with a significance level of $\alpha = 0.05$. As a post hoc analysis, we conducted Wilcoxon signed-rank tests with an applied Bonferroni correction, resulting in a significance level of $p < 0.005$ for each statement. In the following, we only report significant results.

Effectiveness. We found a significant difference in terms of perceived confidence to be able to correctly enter data, $\chi^2(4) = 23.94, p < 0.001$. IC0 and IC2 were rated significantly better than IC1 and IC4.

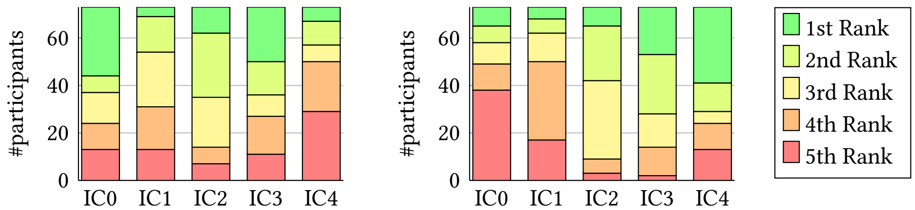
Efficiency. We found a significant difference in terms of perceived ability to quickly enter data, $\chi^2(4) = 61.56, p < 0.001$. IC0, IC1, IC2, and IC3 were rated significantly better than IC4. IC0, IC1, and IC2 were also rated significantly better than IC3.

Ease of Use. We found a significant difference in terms of perceived ease of use, $\chi^2(4) = 63.85, p < 0.001$. IC0, IC1, and IC2 were rated significantly better than IC3 and IC4.

Satisfaction. We found a significant difference in terms of perceived satisfaction, $\chi^2(4) = 16.48, p = 0.002$. IC2 and IC3 were rated significantly better than IC4.

Learnability. We found a significant difference in terms of perceived learnability, $\chi^2(4) = 73.74, p < 0.001$. Participants thought that IC0 would be significantly easier to learn than all other input controls. IC1 and IC2 were also rated significantly better than IC3 and IC4.

4.3.2 Rankings. For the rankings, we assigned the number of the rank to the input controls (see Figure 5 for an overview of the ranking results). A lower number therefore corresponds to a better rank. We then performed the same analysis as for the metrics.



(a) Ranking on Likeability

(b) Ranking on Usefulness

Fig. 5. Results from our online survey showing the rank that participants gave to the input controls (IC0 to IC4) in terms of likeability and usefulness.

Likeability. We found a significant difference in the ranking of the input controls for likeability , $\chi^2(4) = 31.46, p < 0.001$. Participants ranked IC0, IC2, and IC3 significantly better than IC4. Additionally IC2 was ranked significantly better than IC1.

Usefulness. We also found a significant difference in the ranking of the input controls for usefulness , $\chi^2(4) = 62.18, p < 0.001$. Participants found IC2, IC3, and IC4 significantly more useful than IC0 and IC1.

4.3.3 Visual Elements. The majority of participants agreed to the statement of "*I understood how ... works*" for all graphical elements. The probability distribution function plot was rated best with an average of 4.16 (SD = 0.90), followed by the gradient plot with an average of 3.64 (SD = 1.10) and the gradient height legend with an average of 3.53 (SD = 1.17). We found that the same order applied for the statement of "*I liked ...*". The function plot got an average rating of 3.70 (SD = 1.02), followed by the gradient plot with an average of 3.12 (SD = 1.00) and the gradient height legend with an average of 2.93 (SD = 1.08). For all elements, participants agreed on the statement that they are useful with an average agreement of 3.92 (SD = 0.88) for the function plot, 3.36 (SD = 0.98) for the gradient plot, and 3.22 (SD = 0.99) for the gradient height legend.

4.3.4 Qualitative Feedback. We analyzed the qualitative feedback for each input control and the overall feedback related to the rankings of the input controls.

According to four participants, IC0 was easy to use. Although we neither assumed nor expected participants to calculate the mean, three of them did calculate it. They complained that they were not able to enter the exact value (15.67), because the slider only allowed integer input. Nevertheless, one participant found that "*on the other hand, Sam can only go by car 15 or 16 times, not 15,6666...so the input is alright*".

On the one hand, IC1 was described as easy to use (three participants) and fast (two participants). On the other hand, three participants complained about the low number of degrees of freedom that the input control offers. Three others also mentioned that it could be a problem that the width of the range of the probability distribution could not be adapted, because "*[...] this restricts the input options and one is unable to vary the standard deviation [...]. This could be either positive or negative depending on the purpose.*"

Participants described IC2 as "*awesome for this type of question*", visually appealing, good for entering the lowest and highest value, and self-explaining. Two others stated that it was difficult for them to figure out how to use the control. One participant complained about the degree of freedom, which was perceived too low in comparison with other input controls.

Participants had very diverse opinions about IC3. Whilst one described that it was easy to enter data and that usability and accuracy of the control were pleasing, another stated that "*my stochastic knowledge is weaker than the form of presentation. I guess it's adequate and useful for people in the field.*"

Participants had even more diverse opinions about IC4. On one hand, three liked it because of its degrees of freedom and one participant each found it effective, a nice visualization, and a "*very good input control*". On the other hand, three did not understand it, two thought that there were too many inputs, and one participant each thought that it was too complicated, overloaded, and that it would be faster to draw the graph.

For all input controls, participants specified that the tooltips helped them to understand how to operate the sliders. Few of them provided us with specific feature requests such as directly clicking in the slider bar to set the value, placing the numbers showing the value not below the bar, making more parts of the slider interactive, including the ability to pull the left border across the right border to make the input easier, or styling the slider more like a standard web slider.

For the graphical element, participants specified that the function plot was very useful and better to judge than the color gradient. The color gradient was found to be useful by two participants, but three mentioned that it was hard to interpret. One stated that this is "*only nice to have*". The transition legend was not considered as useful and two participants specified that they did not even understand it. Also the unit of the legend and the plot was hard to understand according to participants.

Participants' comments about the rankings gave more insights in why they liked specific controls and which controls they liked best. They stated that they ranked for different properties such as provided details, number of settings that can be adjusted (degrees of freedom), simplicity and straightforwardness, easy to handle and understand, or rating from easiest to hardest. One participant again mentioned that "*[...] the first two methods [(IC0, IC1)] didn't specify what should be the input. I assumed mean, which I can't do in my head.*" Another one specified that he "*[...] needed some time to find out how to adjust the different sliders and which function they had.*" Following the ranking, participants mostly disliked IC1. They found that "*[IC1] is horrible, because the plot is presented with an amount of detail (the pdf, which includes all kinds of assumptions on the shape) that is not at all supported by the data that's put in.*" and "*[IC1] and [IC2] were somehow in between (too simple and too unspecific)*". By contrast participants specified that they liked IC3, which offered the ability to enter a flexible range and the peak of the probability distribution function. One participant stated that "*The more power the methods hold, the more complex and annoying the interaction got. Number [IC3] was a good compromise. Easy sliding and the possibility to change the width.*" and another stated that "*[... IC3] is where you provide the most information, while not being very cumbersome at the same time.*" One participant also summarized all these findings by highlighting that he "*[...] liked [IC3] because you can adjust some features, but it's still relatively quick. Method [IC2] and [IC1] give less possibilities. But in [IC4] it's too much you have to enter.*"

4.4 Discussion

We found that IC2 got the highest agreement for the item on satisfaction although IC0 got higher agreement for all other statements. Participants realized that they could not provide a good answer for the task using IC0. Additionally, IC2 outperformed IC1 for all statements. Participants did not like the fixed bar restricting them in their input, making it more difficult to adjust. Participants agreed less with the satisfaction statements for IC1 than for IC2, IC0, and IC3. To improve IC1, an explanation about why this range was chosen could help to make users more aware of good reasons for the choice and less miserable about the fact that the size could not be changed. IC4 got the lowest agreement on all five statements. The control was perceived too cumbersome and difficult to handle. Although IC3 did not get good single ratings, it was ranked high and comments about the ranking showed that the majority of participants liked it and saw it as the best compromise. Nevertheless, it was already difficult to handle for participants with low statistical knowledge. Although statistical knowledge was needed to understand it, participants surprisingly perceived the probability distribution function plot to be the most useful graphical support element. The gradient height legend was more difficult to understand.

The subjective assessment showed that participants in general liked the idea of using input controls for uncertain data. They especially liked IC2 and IC3, while finding IC1 too restrictive and IC4 to cumbersome. One of the findings was that although we carefully formulated our question, three participants were not sure what to enter: mean, median, or mode. This is also a problem of current interfaces that provide a single value input. In most cases, there is no specification whether the mean, the median, or the mode should be entered.

In addition to a subjective assessment, an objective assessment of the input controls is necessary to assure the correctness of the input and get more insights in when to use which input control.

5 OBJECTIVE ASSESSMENT

As a second step of the evaluation, we conducted a user study as objective assessment of the five input controls. We therefore focused on how long participants need to make an input, how much help is necessary, and how well they were able to provide the required input.

5.1 Method

We invited prospective participants by sending out e-mails to a mailing list that volunteers for user studies could subscribe to. When arriving, prospective participants were informed about the context and were asked to sign a consent form. We used a within-subjects design, thus participants had to solve three tasks and complete a System Usability Scale (SUS) questionnaire [2] for each input control. To minimize sequence effects, we randomly assigned an order of input controls to each participant. The study instructor manually started a new task by key press.

We used the following three tasks for each input control:

(1) For the *first task*, we modified the question from the online survey learning from previous findings. We adapted the question to "*What is the most likeliest value ...*" to make sure that participants knew that we expected them to enter the mode. We prepared five different tables and randomly assigned them to the input controls to make sure that participants did not know the solution beforehand and were not influenced by the values. We used this task to compare the actual input with the underlying probability distribution function that was used for generating the tables.
(2) In the *second task*, we asked participants to specify how much money they spent when doing their grocery shopping. This task allowed participants to enter information that they knew without needing to process a table. The goal of the task was to give participants a feeling how to interact with the input control on a free task.
(3) In the *third task*, we asked about the possible outcome of dice rolls. Multiple rolls produce a probability distribution which can be well specified by using the input controls.

For the study, we added the input controls to a web page containing a description of the task, the question that needed to be answered with the input control, and five buttons where participants could judge their confidence about the correctness of their input with five options: "*My input is correct*", "*My input is nearly correct*", "*I'm not sure whether my input is correct*", "*I doubt that my input is correct*", and "*I don't understand it*". The web page with one of the input controls each time was opened in Firefox.

We recorded all clicks on the input controls, the input time, the clicks on the help button and the confidence buttons, and how long the help information was opened. Additionally, we logged the input and did a screen recording.

5.2 Participants

30 participants (11 female, 19 male) with an average age of 26.03 ($SD = 7.15$) took part in the study. We recruited them in a university setting, so the majority of them were undergraduate students with different subjects or university employees. None of them participated in the online survey. As in our online survey, we asked participants for their knowledge of stochastics, statistics and probability theory. Participants reported to have an

average knowledge level of 4.73 ($SD = 1.68$). As the majority had a high school degree as their highest educational degree, the knowledge related to the knowledge acquired in school.

5.3 Results

For all efficiency, confidence, effectiveness, and usability metrics, we calculated Pearson's r to detect possible correlations between the statistical knowledge of participants and their performance in the study. We only found moderate positive correlations ($0.40 < r < 0.45$) in terms of how often and how long the help was consulted for task 2 and 3 of IC2 and IC4. The statistical knowledge did not significantly correlate with any other metrics ($-0.35 < r < 0.35$) for any other input control and task, thus no distinction between participants with low and high statistical knowledge was made in the analysis. We analyzed the task separately, but as we did not find any significant differences between tasks, we merged them for the analysis.

For all metrics, we conducted a Friedman test to test for statistical significance with a significance level of $\alpha = 0.05$. As a post hoc analysis, we conducted Wilcoxon signed-rank tests with an applied Bonferroni correction, resulting in a significance level of $p < 0.005$ for each metric.

5.3.1 Efficiency. To compare the efficiency of the input controls, we recorded all clicks on the input controls and therefore how often the participants interacted with the controls. We divided the total number of clicks per input control by the number of interactive elements (drag handles) to achieve a comparable value. The Friedman test showed a significant difference in how often participants clicked on the input controls, $\chi^2(4) = 43.54, p < 0.001$. Most clicks per element were recorded for IC1 with 2.89 clicks, which was significantly clicked more often than IC0, IC3 and IC2. IC1 was followed by IC4, which was also clicked significantly more often than IC2 and IC3.

We additionally analyzed the total input time. The Friedman test showed a significant difference in how long the participants interacted with the input controls, $\chi^2(4) = 98.05, p < 0.001$. Participants were fast using IC0, IC1, and IC2. With IC3, it took them significantly more time to enter data than with IC0 and IC2. Entering data with IC4 took significantly longer than with all other input controls.

5.3.2 Confidence. We recorded how many times the help button was clicked to open the help information. Most help was needed for IC0, where in 20 of 90 tasks, participants opened the help information. For all other input controls, help information was only opened for 11 up to 14 out of 90 tasks. How long participants opened the help information was also different, but the Friedman test showed no significant difference. For whether the help information was used or not, we did also not observe any ordering effects.

At the end of each task, participants rated their confidence of the correctness of their input. We converted the statements to numbers from 1: "*I don't understand it*", to 5: "*My input is correct*". The Friedman test showed a significant difference in the confidence, $\chi^2(4) = 16.85, p = 0.002$. Participants were significantly more convinced of the correctness of their input when using IC3 than when using IC0 and IC1.

5.3.3 Effectiveness. We calculated the absolute deviation of the answers for task 1 and task 3 by calculating the mean deviation of all interactive elements, but the Friedman test showed no significant differences for task 1 for task 3.

5.3.4 Usability. Each participant had to answer a SUS questionnaire for each input control which was adapted by replacing the term system with the term input control. We found a significant difference in the reported usability, $\chi^2(4) = 33.64, p < 0.001$. IC4 was rated significantly worse than all other input controls with a SUS score of 47.83.

5.4 Discussion

Recording the number of clicks per interactive element revealed that participants used most clicks for IC1 and least clicks for IC2. This happened probably because participants were more unsure what to enter with IC1 and

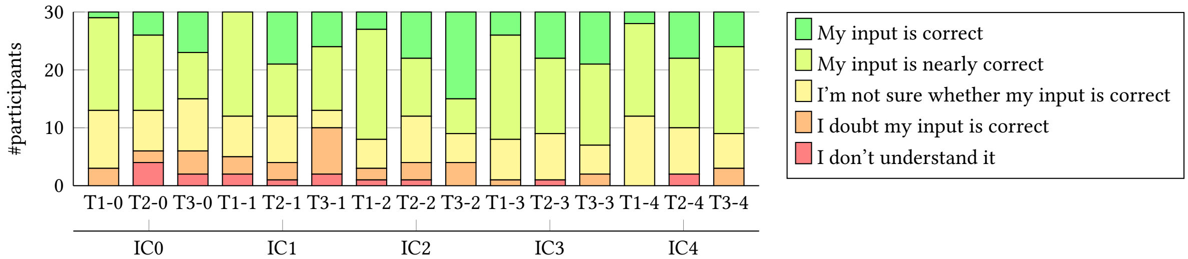


Fig. 6. Results of the user study showing the confidence of 30 participants about the correctness of their input sorted by input control (IC0 to IC4) and task (T1 to T3).

	IC0		IC1		IC2		IC3		IC4	
	M	SD	M	SD	M	SD	M	SD	M	SD
Clicks per Element	1.86	1.63	2.89	2.81	1.43	0.98	1.84	1.72	2.53	1.56
Total Input Time (s)	33.2	32.4	35.5	30.1	31.4	21.8	45.6	33.7	63.3	32.5
Help Time (s)	8.90	26.67	5.71	17.24	5.71	17.24	7.18	27.49	10.99	27.81
Perceived Correctness	3.44	1.06	3.49	1.11	3.82	1.03	3.91	0.83	3.76	0.87
Deviation of Answers T1	2.37	1.90	2.63	1.97	2.59	2.08	2.64	2.14	2.39	1.82
Deviation of Answers T3	53.17	85.00	42.17	36.31	29.44	54.47	40.56	62.91	30.78	46.9
SUS score	65.25	20.05	65.67	16.00	70.92	17.43	72.5	16.19	47.83	21.52

Table 3. Results of the user study showing the means and standard deviations of all 30 participants for each of the used metrics.

needed to correct their input several times. Additionally, participants on average needed the shortest input time to enter results with IC2 although IC2 had one more interactive control than IC0 and IC1. For the participants, IC2 was easier to use intuitively.

We additionally experienced that for IC0, the standard slider input control, help information was opened more often than for all other input controls. Probably participants were unsure about what value to enter although we changed the question after the online survey and did an informal pre-test with five potential participants to see whether they could understand the intention of the question. The help time was also longer for IC0 than for nearly all other input controls.

Regarding confidence, participants were mostly convinced of the correctness of their input for IC3 and IC2, although analyzing the effectiveness showed that IC3 had the highest deviation of the exact value. This could be because participants with less statistical knowledge had more difficulties in understanding and using IC3, which is also indicated by the highest standard deviation. Nevertheless, IC3 had the highest usability score, followed by IC2.

6 RECOMMENDATIONS FOR USAGE

In the following, we outline implications and recommendations for the usage of the five input controls based on the results of our evaluation.

Basic Slider (IC0): The basic slider can be used when the other input methods are not applicable or the users have to be very fast. However, based on the question and the input data, an ordinary number input box might provide similar or better results. Despite of what input control is used, it is very important to clarify whether the interface expects users to put in the mean, median, or mode.

Fixed Range Slider (IC1): The fixed range slider was rated negative in the survey and did also not compete very well in the user study. Participants felt unsatisfied with the interaction and were unconfident about the range size, because it did not necessarily match their expectations. We suggest to preferably use the flexible range slider, except for very specific situations where the range is known beforehand. Nevertheless, the choice of range should be explained to participants to avoid confusion.

Flexible Range Slider (IC2): The flexible range slider is applicable in most scenarios and even people without much knowledge of statistics were able to use it and correctly enter data. Specifying a minimum and maximum value was an intuitive task for most participants that was easy to solve. In addition, they were quite confident about their input and rated the usability high.

Flexible Range Best Estimate Slider (IC3): The dynamic range slider was seen as a good compromise for our type of tasks by most of the participants. They stated in their comments that they preferred using it, which also showed the good ranking for likeability and usefulness in the online survey as well as the good usability score in the user study. However, the input control could be demanding for users without statistical knowledge and should therefore only be used if it is very likely that users have basic statistical knowledge.

Advanced Flexible Range Best Estimate Slider (IC4): The advanced flexible range best estimate slider got the most negative ratings and also a very bad SUS score in the user study. We do not recommend this input control for laymen. Nevertheless, participants with previous knowledge about statistics liked it, which makes it usable for people with a high amount of statistical knowledge.

6.1 Limitations

Our input controls are limited to unimodal probability distribution functions. Other shapes are not supported and we also use clipped probability functions. However, our input controls are clearly directed to people with an average amount of statistical knowledge, not to experts of statistics who would probably like to enter bimodal and more complex distributions. The development of more complex input controls for experts has to be analyzed separately. We selected our sample of participants to be well-educated adults with a basic statistical knowledge acquired in school. This is also the target group for our input controls.

We evaluated the sliders with three different tasks, but only one task was related to real usage behavior. As tasks related to user behavior are very difficult to evaluate according to effectiveness, we decided to introduce two other tasks to better understand how well people can enter data based on given data. We see this as one of the first evaluation steps to compare the input controls against each other. Future work should also provide an empirical investigation that compares real data of people with their input.

7 TOOLKIT

We implemented a toolkit that contains all five web-based input controls. The toolkit consists of a JavaScript and a CSS file that can be easily embedded in a web application or web page. To include one of the input controls, developers have to create an HTML div-element (see Figure 7). By specifying parameters, the look and feel of the control can be influenced. Figure 8 shows an overview about the basic parameters that can be chosen by developers for IC3. Besides specifying the minimum value, the maximum value, the initial values, and the colors of the slider bar, developers can also decide whether they want to show the color gradient, its legend, and the distribution plot. Developers can also influence the type of distribution by specifying a distribution type (see Figure 9). So far, we implemented four different distribution types: triangular distribution, normal distribution, wigner semicircle distribution and laplace distribution. In addition to the shown parameters, the control can also be disabled and developers can enable or disable tooltips that provide supportive information on how to use the input control. The documentation of the toolkit provides additional information on how to import the toolkit,

```html
<div show_plot="true" distribution_type="1"
gradient_mode="chart-legend" value_upper="40" value_lower="10"
value="24" value_max_sel="50" value_min_sel="0" value_max="50"
value_min="0" value_desc_post="%" value_desc_pre=""
color_center="#e34a33" color_border="#fee8c8"
text_color="#000000" graph_color="#e34a33" class="PDS-uist-im-
fberca-container" id="slider-1">
```

Fig. 7. HTML code for including a probability distribution slider in a web application. The slider is specified by a div-element that contains all parameter information.

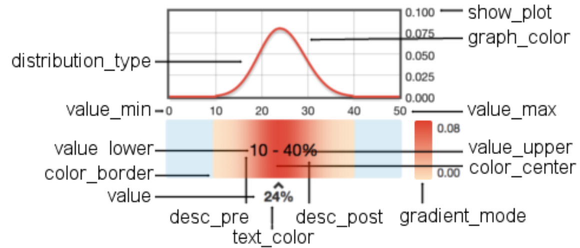
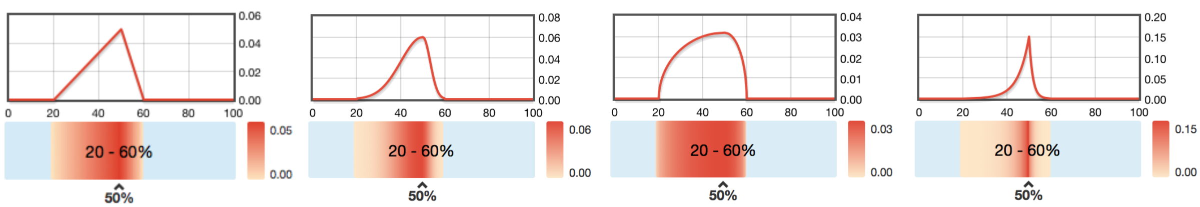


Fig. 8. Parameters for input controls.



(a) Triangular Distribution (b) Normal Distribution (c) Wigner Semicircle Distribution (d) Laplace Distribution

Fig. 9. Different types of probability distribution functions that can be used to describe the distribution of uncertain data.

include the input controls in a web application, how to choose the parameters for customizing the input control, and how to fetch the input of a user.

The toolkit and its documentation is provided open-source and free to use as supplementary material and at the following URL: http://probability-distribution-sliders.hci.simtech.uni-stuttgart.de/.

8 CONCLUSION

This paper addresses the issue of entering uncertain data in graphical user interfaces. With the help of input controls that allow users to input such data, the uncertainty in the input can be quantified and taken into account in the calculation process of an application. We showed that non-probabilistic input controls are perceived less usable and useful for tasks that force users to enter uncertain data. The flexible range slider proved to be suitable for probabilistic input for users that have no or limited knowledge about statistics. This is supported by high rankings in the online survey and a good usability score in the user study. Input controls with more degrees of freedom did perform well for participants with basic statistical knowledge. The results of the user study also indicate that participants are more confident about the correctness of their input when using input controls for probabilistic input. We therefore implemented a toolkit containing the five input controls developed in the context of these studies to make probabilistic input available for web developers. By including the controls in their web applications, they can empower users to enter better input values. Based on the usage patterns, adaption mechanisms and help mechanisms for untrained users could be established. The toolkit is additionally available for other researchers which gives them the opportunity to empirically explore probabilistic user input in different contexts.

ACKNOWLEDGMENTS

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] Nadia Boukhelifa and David J. Duke. 2009. Uncertainty Visualization: Why Might It Fail?. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, 4051–4056. https://doi.org/10.1145/1520340.1520616

[2] John Brooke. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[3] Michael Correll and Michael Gleicher. 2014. Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error. *Visualization and Computer Graphics, IEEE Transactions on* 20, 12 (Dec 2014), 2142–2151. https://doi.org/10.1109/TVCG.2014.2346298

[4] Stephen G. Eick. 1994. Data Visualization Sliders. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology (UIST '94)*. ACM, New York, NY, USA, 119–120. https://doi.org/10.1145/192426.192472

[5] Nivan Ferreira, Danyel Fisher, and Arnd Christian Konig. 2014. Sample-oriented Task-driven Visualizations: Allowing Users to Make Better, More Confident Decisions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 571–580. https://doi.org/10.1145/2556288.2557131

[6] Nahum Gershon. 1998. Visualization of an imperfect world. *Computer Graphics and Applications, IEEE* 18, 4 (Jul 1998), 43–45. https://doi.org/10.1109/38.689662

[7] Miriam Greis, Thorsten Ohler, Niels Henze, and Albrecht Schmidt. 2015. Investigating Representation Alternatives for Communicating Uncertainty to Non-experts. In *Human-Computer Interaction - INTERACT 2015*. Vol. 9299. 256–263. https://doi.org/10.1007/978-3-319-22723-8_21

[8] Harald Ibrekk and M. Granger Morgan. 1987. Graphical Communication of Uncertain Quantities to Nontechnical People. *Risk Analysis* 7, 4 (1987), 519–529. https://doi.org/10.1111/j.1539-6924.1987.tb00488.x

[9] Christopher H. Jackson. 2008. Displaying uncertainty with shading. *The American Statistician* 62, 4 (2008), 340–347. https://doi.org/10.1198/000313008X370843

[10] Chris. R. Johnson and Allen R. Sanderson. 2003. A Next Step: Visualizing Errors and Uncertainty. *Computer Graphics and Applications, IEEE* 23, 5 (Sept 2003), 6–10. https://doi.org/10.1109/MCG.2003.1231171

[11] Susan L. Joslyn and Jared E. LeClerc. 2012. Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *Journal of Experimental Psychology: Applied* 18, 1 (2012), 126–140. https://doi.org/10.1037/a0025185

[12] Malte F. Jung, David Sirkin, Turgut M. Gür, and Martin Steinert. 2015. Displayed Uncertainty Improves Driving Experience and Behavior: The Case of Range Anxiety in an Electric Car. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2201–2210. https://doi.org/10.1145/2702123.2702479

[13] Matthew Kay, Tara Kola, Jessica R. Hullman, and Sean A. Munson. 2016. When (Ish) is My Bus?: User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5092–5103. https://doi.org/10.1145/2858036.2858558

[14] Matthew Kay, Dan Morris, M.C. Schraefel, and Julie A. Kientz. 2013. There's No Such Thing As Gaining a Pound: Reconsidering the Bathroom Scale User Interface. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 401–410. https://doi.org/10.1145/2493432.2493456

[15] Matthew Kay, Shwetak N. Patel, and Julie A. Kientz. 2015. How Good is 85Classifier Evaluation to Acceptability of Accuracy. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 347–356. https://doi.org/10.1145/2702123.2702603

[16] Anass Lasram, Sylvain Lefebvre, and Cyrille Damez. 2012. Scented Sliders for Procedural Textures. In *EUROGRAPHICS short papers (Proceedings of the Eurographics conference (short papers))*. Cagliari, Italy. https://doi.org/10.2312/conf/EG2012/short/045-048

[17] Isaac M. Lipkus, Greg Samsa, and Barbara K. Rimer. 2001. General Performance on a Numeracy Scale among Highly Educated Samples. *Medical Decision Making* 21, 1 (2001), 37–44. https://doi.org/10.1177/0272989X0102100105 arXiv:http://mdm.sagepub.com/content/21/1/37.full.pdf+html

[18] Alan M. MacEachren, Anthony Robinson, Susan Hopper, Steven Gardner, Robert Murray, Mark Gahegan, and Elisabeth Hetzler. 2005. Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science* 32, 3 (2005), 139–160. https://doi.org/10.1559/1523040054738936 arXiv:http://dx.doi.org/10.1559/1523040054738936

[19] Dawn M. Marsh-Richard, Erin S. Hatzis, Charles W. Mathias, Nicholas Venditti, and Donald M. Dougherty. 2009. Adaptive Visual Analog Scales (AVAS): A modifiable software program for the creation, administration, and scoring of visual analog scales. *Behavior Research Methods* 41, 1 (2009), 99–106. https://doi.org/10.3758/BRM.41.1.99

[20] Justin Matejka, Michael Glueck, Tovi Grossman, and George Fitzmaurice. 2016. The Effect of Visual Appearance on the Performance of Continuous Sliders and Visual Analogue Scales. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5421–5432. https://doi.org/10.1145/2858036.2858063

[21] Rebecca E. Morss, Julie L. Demuth, and Jeffrey K. Lazo. 2008. Communicating Uncertainty in Weather Forecasts: A Survey of the U.S. Public. *Weather and forecasting* 23, 5 (2008), 974–991. https://doi.org/10.1175/2008WAF2007088.1

[22] Chris Olston and Jock D. Mackinlay. 2002. Visualizing data with bounded uncertainty. In *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on.* 37–40. https://doi.org/10.1109/INFVIS.2002.1173145

[23] Masakazu Osada, Holmes Liao, and Ben Shneiderman. 1993. Alpha Slider: Searching Textual Lists with Sliders. (1993).

[24] Alex T. Pang, Craig M. Wittenbrink, and Suresh K. Lodha. 1997. Approaches to uncertainty visualization. *The Visual Computer* 13, 8 (1997), 370–390. https://doi.org/10.1007/s003710050111

[25] Florian Pappenberger, Elisabeth Stephens, Jutta Thielen, Peter Salamon, David Demeritt, Schalk Jan van Andel, Fredrik Wetterhall, and Lorenzo Alfieri. 2013. Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication. *Hydrological Processes* 27, 1 (2013), 132–146. https://doi.org/10.1002/hyp.9253

[26] Kristin Potter, Joe Kniss, Richard Riesenfeld, and Chris R. Johnson. 2010. Visualizing Summary Statistics and Uncertainty. *Computer Graphics Forum* 29, 3 (2010), 823–832. https://doi.org/10.1111/j.1467-8659.2009.01677.x

[27] Mark S. Roulston, Gary E. Bolton, Andrew N. Kleit, and Addison L. Sears-Collins. 2006. A laboratory study of the benefits of including uncertainty information in weather forecasts. *Weather and Forecasting* 21, 1 (2006), 116–122. https://doi.org/10.1175/WAF887.1

[28] Julia Schwarz, Scott Hudson, Jennifer Mankoff, and Andrew D. Wilson. 2010. A Framework for Robust and Flexible Handling of Inputs with Uncertainty. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology (UIST '10).* ACM, New York, NY, USA, 47–56. https://doi.org/10.1145/1866029.1866039

[29] Orit Shaer, Oded Nov, Johanna Okerlund, Martina Balestra, Elizabeth Stowell, Lauren Westendorf, Christina Pollalis, Jasmine Davis, Liliana Westort, and Madeleine Ball. 2016. GenomiX: A Novel Interaction Tool for Self-Exploration of Personal Genomic Data. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16).* ACM, New York, NY, USA, 661–672. https://doi.org/10.1145/2858036.2858397

[30] Meredith Skeels, Bongshin Lee, Greg Smith, and George G. Robertson. 2010. Revealing Uncertainty for Information Visualization. *Information Visualization* 9, 1 (2010), 70–81. https://doi.org/10.1057/ivs.2009.1 arXiv:http://ivi.sagepub.com/content/9/1/70.full.pdf+html

[31] Susanne Tak, Alexander Toet, and Jan van Erp. 2014. The Perception of Visual Uncertainty Representation by Non-Experts. *Visualization and Computer Graphics, IEEE Transactions on* 20, 6 (June 2014), 935–943. https://doi.org/10.1109/TVCG.2013.247

[32] Judi Thomson, Elizabeth Hetzler, Alan MacEachren, Mark Gahegan, and Misha Pavel. 2005. A typology for visualizing uncertainty. *Proc. SPIE* 5669, 146–157. https://doi.org/10.1117/12.587254

[33] Thomas S. Wallsten, David V. Budescu, Amnon Rapoport, Rami Zwick, and Barbara Forsyth. 1986. Measuring the vague meanings of probability terms. *Journal of Experimental Psychology: General* 115, 4 (1986), 348–365. https://doi.org/10.1037/0096-3445.115.4.348

[34] Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. 2007. Scented Widgets: Improving Navigation Cues with Embedded Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (Nov. 2007), 1129–1136. https://doi.org/10.1109/TVCG.2007.70589

[35] Craig M. Wittenbrink, Alex T. Pang, and Suresh K. Lodha. 1996. Glyphs for visualizing uncertainty in vector fields. *Visualization and Computer Graphics, IEEE Transactions on* 2, 3 (Sep 1996), 266–279. https://doi.org/10.1109/2945.537309

[36] Jianlong Zhou, Constant Bridon, Fang Chen, Ahmad Khawaji, and Yang Wang. 2015. Be Informed and Be Involved: Effects of Uncertainty and Correlation on User's Confidence in Decision Making. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15).* ACM, New York, NY, USA, 923–928. https://doi.org/10.1145/2702613.2732769

[37] Torre Zuk and Sheelagh Carpendale. 2006. Theoretical analysis of uncertainty visualizations. *Proc. SPIE* 6060, 606007–606007–14. https://doi.org/10.1117/12.643631