Value-Suppressing Uncertainty Palettes

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

ABSTRACT

Understanding uncertainty is critical for many analytical tasks. One common approach is to encode data values and uncertainty values independently, using two visual variables. These resulting bivariate maps can be difficult to interpret, and interference between visual channels can reduce the discriminability of marks. To address this issue, we contribute Value-Suppressing Uncertainty Palettes (VSUPs). VSUPs allocate larger ranges of a visual channel to data when uncertainty is low, and smaller ranges when uncertainty is high. This non-uniform budgeting of the visual channels makes more economical use of the limited visual encoding space when uncertainty is low, and encourages more cautious decision-making when uncertainty is high. We demonstrate several examples of VSUPs, and present a crowdsourced evaluation showing that, compared to traditional bivariate maps, VSUPs encourage people to more heavily weight uncertainty information in decision-making tasks.

ACM Classification Keywords

H.5.0. Information Interfaces and Presentation (e.g. HCI): General

Author Keywords

Uncertainty Visualization; Color Perception; Thematic Maps; Semiotics.

INTRODUCTION

Uncertainty is an inescapable component of collecting, analyzing, and presenting data. A common goal in the communication of uncertainty is promoting *uncertainty-aware decisions*: the audience should be aware of the risks and rewards of certain decisions, modulate their confidence in their conclusions, and perhaps refrain from making a decision at all if there is too much uncertainty. A way that designers can contribute to this goal is by ensuring that uncertainty information is *well-integrated* with the rest of the data. That is, it should be difficult to discount or ignore the uncertainty in a dataset.

Uncertainty is an inescapable component of collecting, presenting, and using data. A common goal in the communication

CHI'16, May 07–12, 2016, San Jose, CA, USA

ISBN 123-4567-24-567/08/06...$15.00

DOI: http://dx.doi.org/10.475/123_4

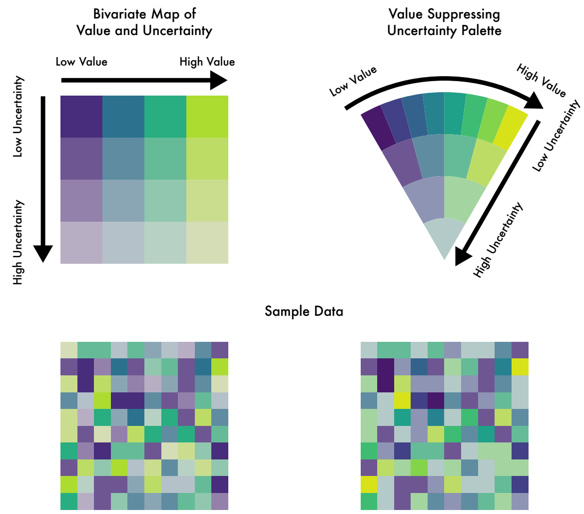


Figure 1: A standard bivariate map (left) and a VSUP (right), used to encode an identical 10x10 grid of random data. Both use the same visual channels to encode value (position along the Viridis [?] color map) and uncertainty (lightness and saturation). However, the VSUP uses a tree-like structure to allocate colors, defining more bins when uncertainty is low. This non-uniform budgeting affords better discrimination between values when uncertainty is low, even though the VSUP has fewer color bins (in this case, 15 to the bivariate map's 16). This tree-like structure also discourages analysis in regions where uncertainty may be unacceptably high.

of uncertainty is promoting *uncertainty-aware decisions*: the audience should be aware of the risks and rewards of certain decisions, modulate their confidence in their conclusions, and perhaps refrain from making a decision at all if there is too much uncertainty. A way that designers can contribute to this goal is by ensuring that uncertainty information is *well-integrated* with the rest of the data. That is, it should be difficult to discount or ignore the uncertainty in a dataset.

Simultaneous presentation of uncertainty and value necessitates the construction of a bivariate map — a relation, in terms of visual variables, between 2-tuples (value, uncertainty) and mark properties. Due to the interference and interplay between different visual variables, bivariate maps may suffer from limited discriminability.

In this paper, we contribute **Value-Suppressing Uncertainty Palettes** (VSUPs) for integrating data and uncertainty information in visualizations. VSUPs intentionally alias together data values with high uncertainty, affording greater discriminability as uncertainty decreases. Traditional bivariate maps might be thought of as a 2D square, with differing outputs for each combination of value and uncertainty. In contrast, VSUPs can be conceptualized as arcs: as uncertainty increases, values are mapped to smaller and smaller sets of outputs, culminating in a singularity where all inputs are mapped to an identical, highly uncertain mark regardless of data value. Figure **??** shows examples of both a traditional bivariate map and a VSUP.

We describe the motivations for VSUPs, provide examples of their utility for decision-making under uncertainty, and assess VSUPs in a crowdsourced experiment. Our results indicate that VSUPs create close integration between uncertainty and data, promoting rapid but cautious decision-making.

RELATED WORK

Despite the acknowledged importance of uncertainty in understanding data, explicit representations of uncertainty are often missing from visualizations [**?**]. This is partially due to the complexity of uncertainty as a concept. In typologies from Thomson et al. [**?**] and Buttenfield & Beard [**?**], the authors note that many, occasionally contradictory, concepts can fall under the category of "uncertainty," including data quality, sampling error, credibility, and provenance.

There are many methods for visualizing uncertainty. MacEachren et al. [**?**] presents a theoretical survey of uncertainty visualization generally. Relevant to our domain of heatmaps and thematic maps is work in GIS and cartography; for a survey of uncertainty visualization techniques in GIS, see Kinkeldey et al. [**?**, **?**]. Yet, despite these wide variety of methods, to our knowledge, there are no established standards for visualizing uncertainty in heatmaps.

Reviewing the state of the art in uncertainty visualization in the field, Greithe et al. [**?**] and Brodlie et al. [**?**] present lists of potential techniques for conveying uncertainty in concert with data. Some of these options (interaction, animation, and sonification) are not applicable for static charts; even for dynamic charts, many users do not interact with charts in sufficient detail to recover uncertainty information [**?**]. While we acknowledge the potential utility of these techniques in uncertainty visualization (for instance, the animation used to convey sampling error in Hullman et al.'s Hypothetical Outcome Plots [**?**]), we limit the scope of our discussion to techniques which center on static charts.

Visual Variables for Uncertainty

One challenge to conveying uncertainty information is that the decision to explicitly encode uncertainty increases the dimensionality of the data, and so requires the use of (at least one) additional visual channel [**?**]. Uncertainty therefore inherently increases the visual complexity of a visualization. When the data are already complex to convey, and many of the more common or accurate visual variables are in use, allocating an additional dimension is non-trivial. As the number of dimensions increases, finding visual variables that are both perceptually accurate (in either estimation of quantity or discrimination of category) as well as perceptually separable from all the other encoding channels, becomes more and more difficult.

A further hurdle is that not all visual variables are well-suited for conveying uncertainty. MacEachren et al. [**?**] evaluate a number of visual variables with respect to their *semiotic fit* for representing uncertainty. They observe that certain visual variables such as blurriness and transparency seem to have a more intuitive connection to uncertainty than other variables such as shape or hue. Unfortunately, Boukhelifa et al. [**?**] find that many visual variables habitually used for conveying uncertainty, such as blur and value, are also difficult to estimate. This results in a preference/performance gap where designers must choose between encoding uncertainty in a way that is intuitive but error prone, or use higher fidelity channels that may be more difficult to interpret.

Color Maps

For spatial visualizations like choropleth maps, heatmaps, and treemaps, visual channels such as position and length are reserved for data variables other than value (such as geographic location or relative size). In these situations, data value is often encoded using color. Colors in such univariate quantitative color maps should be sufficiently far apart as to be perceptually distinguishable [**?**], and vary in lightness as well as hue to afford an implicit ordering of value [**?**, **?**]. Different choices of color maps can highlight different features of the data, and should be chosen with care based on the data types and important features to be visualized [**?**].

A design consideration when constructing color maps is whether to quantize data into a discrete set of colors, or to encode data using a continuous mapping from value to color. While continuous color maps afford greater fidelity in presenting values [**?**], non-linearity in human color perception introduces errors in extracting numeric values from continuous colors [**?**]. Quantizing a color map is therefore an exercise in balancing *perceptual* error and *quantization* error [**?**]. Discrete maps offer a finer control over this balance, which can result in better performance in tasks involving heatmaps [**?**]. As such, we limit our discussion to color maps with discrete sets of colors.

Bivariate Maps

While uncertainty can be visualized entirely separately from the rest of the data (for instance, in a juxtaposed chart), this runs the risk that uncertainty information becomes *ignorable*[**?**]. It also complicates analysis, as interpreting the value and uncertainty of a given data point requires consulting two separate charts. For our goal on integrating uncertainty information into analysis, we focus instead on the construction of *bivariate maps*, where value and uncertainty information is displayed simultaneously in the same visualization.

Bivariate maps can, in principle, be constructed from the combination of any two visual variables (such as shape and color, or size and texture). E.g., Ware's "textons" overlay glyphs on top of regions to simultaneously encode two quantitative values [**?**]. However, in this work, our focus is on heatmaps and other sorts of thematic maps. In visualizations like heatmaps, common visual channels such as position and size are already in use to encode other information (such as spatial location and extent), and color is reserved for value. Bivariate maps in these settings therefore typically rely on a visual channel related to color for encoding a secondary variable, such as alpha [**?**] or pixel noise [**?**].

Due to the perceptual integrality of visual channels [**?**, **?**], the construction of color-based bivariate maps is difficult. Viewers need to be able to independently assess the variables of interest, but the resulting bivariate map should also be expressive and easy to interpret. Bivariate maps with these constraints are often limited to relatively few categories (say, a 4x4 matrix as in Figure **??**) [**?**, **?**]. For VSUPs, we adapt an insight from Dunn [**?**] that these limited categories ought to be assigned with regards to their importance to analytics tasks, rather than uniformly.

In general, the quality of bivariate color maps is a multivariate measure involving consideration of not just the component color channels, but also the interpolation scheme and the color of the surround [**?**]. However, even simple bivariate maps can be difficult for a general audience to interpret [**?**].

VALUE-SUPPRESSING UNCERTAINTY PALETTES

Value-Suppressing Uncertainty Palettes (VSUPs) are a technique for creating bivariate maps of data *value* and *uncertainty*. VSUPs make two central assumptions about bivariate maps:

1. There is a limited *budget* of potential outputs in a bivariate map.

2. Differences among *certain* data are more significant than differences among *uncertain* data.

In some cases, these assumptions are violated. For instance, an analyst might be interested in "long tail risks" or other "black swan events," where the impact of a value, no matter how uncertain, must be considered and planned for [**?**]. Other analysis tasks (such as filtering out outliers), require increased, rather than decreased, discriminability when uncertainty is high. However, for many information fusion tasks, the assumption is that uncertainty is related to data quality, or the uncertainty of data values [**?**].

However, if both of these assumptions hold, then it follows that the designer of a bivariate map should allocate more mark types to certain values, and fewer mark types to uncertain values. VSUPs codify this decision by *reducing the number of mark categories for representing value as uncertainty increases.* This means that data encoded using a VSUP will make visible only the largest of differences in uncertain data, but highlight comparatively small differences in value when uncertainty is low. VSUPs therefore act as both a filtering mechanism (in that values with too much uncertainty are all mapped to the same glyph), as well as an implicit test of effect size (in that smaller and smaller changes in value are visible as uncertainty decreases). This strategy of dampening the low quality or high uncertainty values in maps in order to focus on more informative regions has measurable benefits, including the removal of statistically spurious visual patterns, and the highlighting of regions of interest [**?**].

VSUPs rely on an underlying **quantization tree** that governs how values are discretized. Maximally uncertain values are all mapped to a singular "root" node. As uncertainty decreases, the tree branches into leaves, which evenly divide the data domain. The data value of a parent is the midpoint of all of its children. These leaves can then branch again, until a designer-specified stopping point. The layers act to the quantize the uncertainty domain, and the nodes the data domain. Figure **??**, for example, is a VSUP with a tree with a branching factor of 2, and 4 layers, resulting in a total of 15 different output colors.

Correll et al. [**?**, **?**] use a precursor of VSUPs in their LayerCake genomics visualization tool, where marks representing genomic data with increasing uncertainty are mapped to a smaller and smaller set of increasingly grey colors, creating the effect of uncertain values retreating into "confidence fog" while highly certain values remain prominent. Other bivariate visualizations implicitly alias together uncertain values. For instance, if uncertainty is encoded by transparency, a maximally uncertain glyph may be entirely transparent, and so impossible to distinguish from any other maximally uncertain glyph. Other channels with a semiotic connection to uncertainty, such as saturation, value, blur, or size, also have deleterious effects on the disambiguation of colors and shapes. In both cases, the property of aliasing is *ad hoc*, and places no guarantees on the discriminability of colors. The binning and degradation approach of VSUPs makes the choice to alias values explicit to both the designer and the viewer, and results in a bivariate mapping with known perceptual properties.

Javascript code for generating VSUPs for use in D3 charts is available at `URL-REMOVED-FOR-REVIEW`.

Design Considerations

There are multiple choices that designers must make before creating a VSUP. In particular, they must choose 1) which visual channels map to value and uncertainty, 2) structure of the underlying VSUP tree.

Often, the graphical conventions of a chart reserve variables such as position and area for other uses, and determine the visual channel to be used for value. In heatmaps, for instance, value is regularly encoded using color, while vertical and horizontal position are reserved for location. In this paper, we primarily deal with the case where value is encoded using color. While there are some similar conventions for representing uncertainty, generally designers have more freedom in choosing a visual channel for uncertainty [**?**]. We recommend channels with both a strong semiotic connection to uncertainty as well as a relatively large number of perceptually distinguishable levels. Here, we focus on value/saturation (or, equivalently, transparency against a white background). Correll & Gleicher [**?**] show that even audiences without statistical backgrounds can interpret uncertainty information encoded in these channels.

Examples

We present a set of examples showing how VSUPs can be applied to realworld datasets. The decision of VSUPs to allocate color bins asymmetrically amongst value and uncertainty

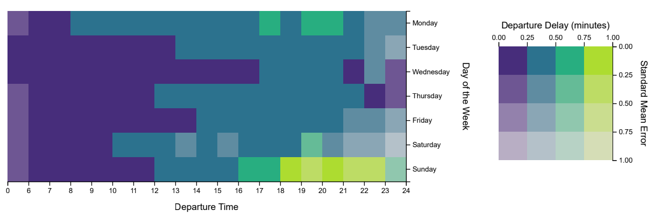


Figure 2: Average departure delay for different times of the day and days of the week visualized with a 2D uncertainty map. Horizontal position is the hour of scheduled departure, and vertical position is the day of the week.

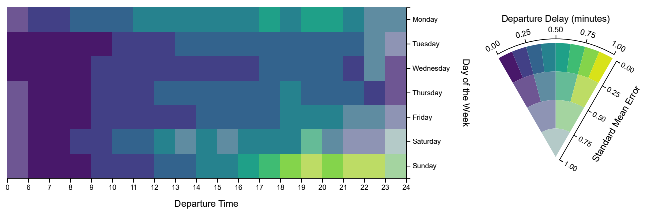


Figure 3: Flight delay data encoded with a VSUP.

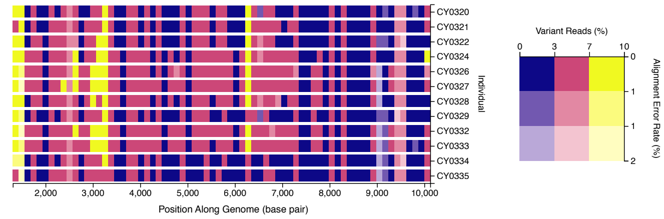


Figure 4: Variability for 12 different viral populations of the SIV virus, encoded with a traditional 2D bivariate map. Horizontal position denotes location along the SIV genome. Each row is the viral population of a different infected animal. Error in aligning short reads can lead to poor quality data, and so the amount of these errors is encoded as the uncertainty.

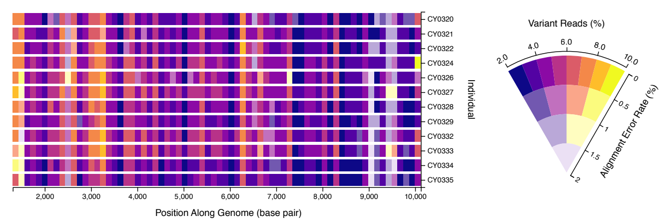


Figure 5: Variability for 12 different viral populations of SIV, encoded with a VSUP. The magnitude of mutation rate, as well as the generally poorer quality of data towards the end of the sequence, is more readily visible than with the traditional bivariate color map.

regions affords greater discriminability in regions of the chart that warrant closer scrutiny, while discouraging exploration of regions with noisy or unreliable signals.

Air Travel Delay

Using a public data set from the U.S. Bureau of Transportation Statistics [?], we created a heatmap showing average flight delay for different times of the day and days of the week for U.S. carriers in January 2017. We use standard error $(\sigma/\sqrt{n})$ as our measure of uncertainty. We created two alternate heatmaps of this delay information. Figure **??** shows a traditional bivariate map, whereas Figure **??** shows a VSUP generated under the same constraints.

Both maps illustrate a similar temporal trend: flight delays are shorter towards the beginning of the day, increasing on average over time. The traditional bivariate map affords only a coarse examination of this trend: two visible "blocks" of color, early in the day, and later in the afternoon. The VSUP, in contrast, has sufficient color resolution to show a quasi-linear trend. Uncertainty is high only for a few bins for which there are small number of flights, either on the weekends or late at night. The VSUP does not obscure any significant trends in the high uncertainty data, despite having fewer color categories overall.

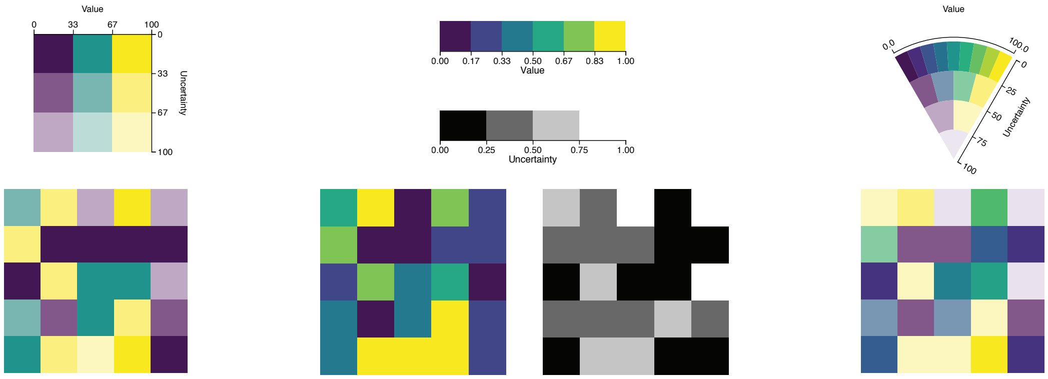
Viral Mutation

Using a dataset from a group of virologists interested in the population dynamics of viral mutation [?], we visualized the variability in viral genomes across 25 populations of the simian immunodeficiency virus (SIV). Next generation sequencing (NGS) allows biologists to collect and analyze large amounts of genomics data. However, these techniques often create data quality issues, as they must align large numbers of potentially ambiguous "reads" of relatively small numbers of base pairs. In the dataset here, it is important not only to identify hotspots in the viral genome (locations with high rates of mutation or variability), but to not be distracted by false positives. The traditional bivariate map (Figure **??**) affords the analysis of only a few, large hot spots. Extending the "confidence fog" metaphor used in viral genome browsers like LayerCake [?], VSUPs (Figure **??**) enable biologists to explore interesting patterns while not being distracted by noise. Smaller scale "hot spots" that are mapped to the same magenta color in the traditional map are visible as bright orange lines. Pink hot spots towards the end of the genome (around reference nucleotide $9,500$) that could be plausibly considered potential regions of interest in the 2D map, are encoded as highly uncertain and of ambiguous value in the VSUPs. In fact, these regions have systematically low coverage; the lack of reliable data for these regions should discourage analysts from giving their apparent pattern equal consideration from the stronger signals earlier in the genome.

EVALUATION

We performed a crowdsourced experiment on Amazon's Mechanical Turk to evaluate the effectiveness of VSUPs for integrating uncertainty and value information in visualizations. This focus on integration meant that we limited our experimental tasks to scenarios where the participants needed to consider both value and uncertainty before making a decision. We gave participants two main tasks:

1. An **identification** task, where we gave participants charts with value and uncertainty information, and asked them to locate specific regions. E.g., "click on the region of the chart with a value of 0.1 and an uncertainty of 0.2."



(a) Traditional Bivariate Map (b) Juxtaposed Univariate Maps (c) Value-Suppressing Uncertainty Palette

Figure 6: The three graph types in our evaluation. Bivariate color ramps ensure that there are orthogonal mappings for each combination of value and uncertainty, but, since color channels are not separable, can afford only a few discrete colors before color categories become unacceptably close, perceptually. Juxtaposed maps, by keeping these channels separate, afford a larger range of colors, but require a visual search task to recover both value and uncertainty in a region. VSUPs have many of the advantages of bivariate maps, but assign more color categories to certain data, at the expense of ambiguity when uncertainty is high.

2. A **prediction** task, inspired by the game Battleship, where we gave participants charts with both *forecast* and *forecast uncertainty* information, and asked the participants to place tokens on the board in order to optimize expected value. E.g., "place your 5 ships on safe locations on the board."

The experiment was a within-subjects design with 1 factor (encoding type) with 3 levels. Figure **??** shows examples of these factor levels. We had 8 replications for each level for the identification task, and 6 for the prediction task, for a total of 24 stimuli for the first task, and 18 for the second. We discuss the two experimental tasks in more detail in the following sections. Experimental materials, including data tables and stimuli generation code, are available at `URL-REMOVED-FOR-REVIEW`.

Identification Task

For the identification task, we gave participants a heatmap, and asked them to click on a region of the heatmap that encodes a particular value,uncertainty pair. To support comparability across conditions, we chose a set of target pairs such that each target mapped to a unique color across all encoding types. This resulted in (4+4+2+1)=11 valid targets. For each trial, these 11 targets were placed randomly in a heatmap; the remaining 13 cells of the heatmap were distractors randomly sampled from this target list.

We selected this task both as a way of assessing the ability of people to successfully encode and decode values using a bivariate map, and to provide training for the prediction task (since we re-used the bivariate maps in that task). We measured performance both in terms of accuracy (did the participant select the correct point) as well as response time.

Prediction Task

For the prediction task, we gave participants the rules of a game like Battleship. Greis et al. [**?**] employ these game-like experimental tasks to assess how different visual designs communicate uncertainty information, which can be abstract or complex, to the general audience. In our task, the participant

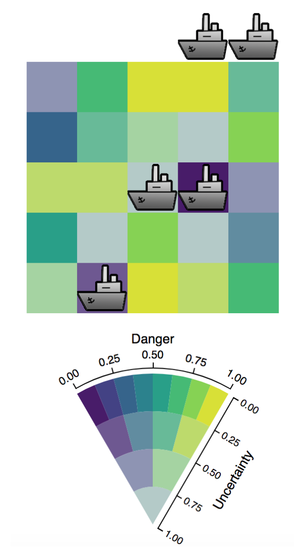


Figure 7: The prediction task. The participant has a list of locations, and ought to place their ships on locations with low probability of attack, and high certainty in this probability.

and a (fictional) adversary have to place tokens representing ships on a spatial grid, with the expectation that certain squares will be hit by missiles. Players have to place all their tokens before continuing. The objective is to minimize the number of your own ships that are hit. In our task, participants were given a map representing the predictions of missile strikes in each location on the grid. The *value* component of the prediction was the likelihood of a missile strike in a particular region. The *uncertainty* component was the confidence in this prediction.

We selected this task in order to promote risk-averse behavior. Tversky & Kahneman [**?**] illustrate that framings in terms of gains or losses produce reliably different outcomes. In particular, there is greater perceived value in avoiding large losses as opposed to striving for a large gain [**?**]. The ideal strategy from a value-maximizing standpoint would be to place tokens on areas with the highest expected value, ignoring the uncertainty information. However, as with roulette and other similar games of chance, the variability in expected value is relevant when considering where to place bets [**?**]. Over the short term, a player may seek out high expected value but risky locations, or, if risk averse, accept lower expected value but lower variability locations. We therefore measured the distribution of both *value* and *uncertainty* of the tokens placed by the participants.

Hypotheses

Across both the identification and prediction tasks, we had several hypotheses, stemming from our belief that VSUPs promote better *integration* between uncertainty and value information, and also encourage *caution* by highlighting the ambiguity or untrustworthiness introduced by uncertain data. In particular:

1. Participants would be **faster and more accurate** when completing the identification task using a VSUP as compared to the other map types. Juxtaposed maps, by adding an additional search task to the lookup task, would have especially poor performance compared to VSUPs.

2. Participants would choose targets with **less uncertainty** in the prediction task when using a VSUP compared to the other maps types. This would result in a tradeoff where they would also choose targets with **less expected value** than the other map types. Again, juxtaposed maps, by increasing the difficulty of retrieving both variables simultaneously, would highlight this difference.

Results

Rather than standard means, Cleveland & McGill [**?**] use 25% trimmed means to report effect sizes in graphical perception tasks. Trimmed means remove the first and last quartiles and then compute the mean from the remaining points. This filtering is important for graphical perception tasks (especially crowdsourced tasks) where error distributions may be long-tailed [**?**]. Trimmed means violate the sampling assumptions required for many null-hypothesis significance tests. Thus, while we report effect sizes in terms of trimmed means, we perform inferential statistical tests on the standard means. Asterisks in our charts denote significant differences, as the

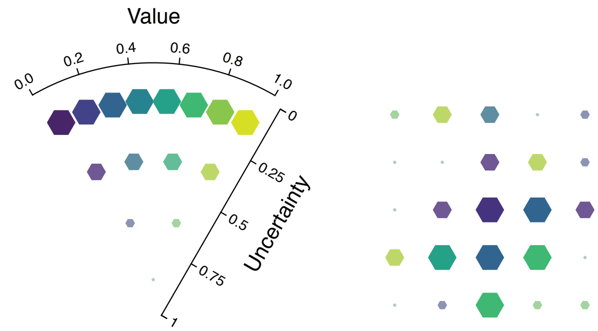


Figure 8: A VSUP where value is mapped to color, and uncertainty is mapped to glyph area. As glyphs become smaller, it becomes more difficult to distinguish their colors [**?**]. VSUPs, by reducing the number of color categories as uncertainty increases, account for this ambiguity. On the right, the VSUP has been used to encode a 5x5 grid of random data.

reported by a post-hoc Tukey's test of Honest Significant Difference (HSD).

Participants

We limited our population to Turkers from the United States, with a prior task approval rate of at least 90%. As the experimental tasks required multi-hue color perception, we presented participants with a set of Ishihara plates [**?**] as a pretest, and excluded participants who either misidentified the values in the plates, or who self-reported as having a color vision deficiency (CVD) in the post-test. Based on piloting, we paid participants $3 dollars for participation, for a target rate of $8/hour. After completion of the main tasks, we solicited demographic information.

We recruited a total of *N* participants for this experiment: *N* female, *N* male, and *N* who declined to state (M_{age}= *N*, SD_{age} = *N*).

Identification Task

Prediction Task

For the prediction task, where there is no dominating strategy, our primary measure was the distribution of guesses, in terms of both their average value, and their average uncertainty. A risky player would choose guesses with high value, regardless of the uncertainty of those points. A more conservative guesser might eschew high-risk, high-reward locations, resulting in a lower average value of guesses, but also lower uncertainty. Our hypothesis is that VSUPs, by communicating uncertainty information at the expense of value information, would promote more conservatives guesses than the other chart types.

DISCUSSION

The results from our initial evaluation of VSUPs are promising, showing that even a general audience can make use of uncertainty information in a reasonable way. The way that uncertainty information is presented can have a measurable impact on decision-making. The VSUP strategy of only "showing" differences when uncertainty is sufficiently low is somewhat analogous to the inferential statistics such as effects tests. VSUPs might therefore be used to promote more caution in judgments, and lead analysts away from spurious signals in data.

Bivariate maps have an inherent limitation in that visual channels are often difficult to attend to separately or orthogonally. Bivariate maps are also complex to interpret. Systems which simultaneously display large amounts of value and uncertainty data to general audiences (such as Pangloss [**?**]) avoid them for precisely these reasons, relying instead on juxtaposed maps. Our results indicate that juxtaposition produces additional problems, forcing users that wish to integrate uncertainty and data to locate an item of interest, and then re-locate that same item in a visually distant map, which may have few visual landmarks in common. This additional step before information fusion is reflected in longer response times, and patterns of decision-making that are less mindful of risk. Given the deleterious effects of latency on data analysis [**?**], designers ought to carefully consider the tradeoffs between expressiveness of visual encoding, and usability.

Limitations and Future Work

In general, VSUPs excel in scenarios where we wish to make firm guarantees about the discriminability of different marks in a chart, but must also represent, in as much detail as possible, two, potentially orthogonal variables of interest. VSUPs have improved data resolution over other bivariate maps by assuming that uncertain values ought to have less visual impact that highly certain values. VSUPs therefore act both as a kind of filtering device as well as a bivariate representation. This assumption holds in scenarios like the virology dataset (Figure **??**), where the expectation is that highly uncertain data is unreliable or should otherwise be downweighted in analysis. If the analyst has a different interpretation of uncertainty, or wishes to quickly and orthogonally analyze the distributions of uncertainty and value, other strategies, such as juxtaposed maps, may be more appropriate. VSUPs are designed for the *integration* of value and uncertainty. Designers should take care in considering when and how this integration is desirable.

Our experiments dealt with cases where both uncertainty and value were represented by color. The perceptual non-separability of color channels is well-known [**?**, **?**], and so the concept of a limited "budget" of distinguishable marks easier to quantify and illustrate. In principle, a VSUP can be created for any combination of visual variables. All that is required is a perceptual model of the interaction between these two variables. Where these models exist, as with the interaction between size and color [**?**], the creation of VSUPs is straightforward. Where these models do not exist, or where the perceptual interaction is too complex to efficiently model, experimental work remains to be done before VSUPs can be considered a feasible design strategy. We are currently experimenting with the creation of VSUPs using these other channels, as in Figure **??**.

Conclusion

Often, uncertainty, data quality, or confidence are considered separate information from the data itself, relegated to tooltips or visually distant supplemental charts. We believe, in contrast, that uncertainty information ought to be directly integrated with charts of data itself. This integration introduces additional complexity in the design and presentation of data. Value-Suppressing Uncertainty Palettes represent one strategy for dealing with this complexity, by assigning marks properties in a way that supports the disambiguation of values in data where uncertainty is low, but suppresses these judgments when uncertainty is high. This decision of how to allocate visual variables promotes patterns of decision-making that make responsible use of uncertainty information, discouraging comparison of values in unreliable regions of the data, and promoting comparison in regions of high certainty.

REFERENCES

1. Jürgen Bernard, Martin Steiger, Sebastian Mittelstädt, Simon Thum, Daniel Keim, and Jörn Kohlhammer. 2015. A survey and task-based quality assessment of static 2D colormaps. In *SPIE/IS&T Electronic Imaging*. International Society for Optics and Photonics, 93970M–93970M.

2. David Borland and Russell M Taylor Ii. 2007. Rainbow color map (still) considered harmful. *IEEE computer graphics and applications* 27, 2 (2007).

3. Nadia Boukhelifa, Anastasia Bezerianos, Tobias Isenberg, and Jean-Daniel Fekete. 2012. Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2769–2778.

4. Nadia Boukhelifa and David John Duke. 2009. Uncertainty visualization: why might it fail?. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 4051–4056.

5. Ken Brodlie, Rodolfo Allendes Osorio, and Adriano Lopes. 2012. A review of uncertainty in data visualization. In *Expanding the frontiers of visual analytics and visualization*. Springer, 81–109.

6. Barbara Buttenfield and M Kate Beard. 1994. Graphical and geographical components of data quality. *Visualization in geographic information systems* (1994), 150–157.

7. William S Cleveland and Robert McGill. 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association* 79, 387 (1984), 531–554.

8. Michael Correll, Adam L Bailey, Alper Sarikaya, David H OâĂŹConnor, and Michael Gleicher. 2015. LayerCake: a tool for the visual comparison of viral deep sequencing data. *Bioinformatics* (2015), btv407.

9. Michael Correll, Subhadip Ghosh, David O'Connor, and Michael Gleicher. 2011. Visualizing virus population variability from next generation sequencing data. In *Biological Data Visualization (BioVis), 2011 IEEE Symposium on*. IEEE, 135–142.

10. M Correll and Michael Gleicher. 2013. Error bars considered harmful. *IEEE Visualization Poster Proceedings. IEEE* (2013).

11. Michael Correll and Jeffrey Heer. 2017. Surprise! Bayesian Weighting for De-Biasing Thematic Maps. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 651–660.
12. Michael W Dobson. 1973. Choropleth maps without class intervals?: a comment. *Geographical Analysis* 5, 4 (1973), 358–360.
13. Richard Dunn. 1989. A dynamic approach to two-variable color mapping. *The American Statistician* 43, 4 (1989), 245–252.
14. W. R. Garner and G. L. Felfoldy. 1970. Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology* 1, 3 (1970), 225–241.
15. Miriam Greis, Passant El Agroudy, Hendrik Schuff, Tonja Machulla, and Albrecht Schmidt. 2016. Decision-Making under Uncertainty: How the Amount of Presented Uncertainty Influences User Behavior. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. ACM, 52.
16. Henning Griethe, Heidrun Schumann, and others. 2006. The visualization of uncertain data: Methods and problems.. In *SimVis*. 143–156.
17. LeGrand H Hardy, Gertrude Rand, and M Catherine Rittler. 1945. Tests for the detection and analysis of color-blindness. I. The Ishihara test: an evaluation. *JOSA* 35, 4 (1945), 268–275.
18. Jeffrey Heer and Michael Bostock. 2010. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 203–212.
19. Jessica Hullman, Paul Resnick, and Eytan Adar. 2015. Hypothetical outcome plots outperform error bars and violin plots for inferences about reliability of variable ordering. *PloS one* 10, 11 (2015), e0142444.
20. Daniel Kahneman and Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society* (1979), 263–291.
21. Christoph Kinkeldey, Alan M MacEachren, Maria Riveiro, and Jochen Schiewe. 2017. Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and recommendations. *Cartography and Geographic Information Science* 44, 1 (2017), 1–21.
22. Christoph Kinkeldey, Alan M MacEachren, and Jochen Schiewe. 2014. How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies. *The Cartographic Journal* 51, 4 (2014), 372–386.
23. Zhicheng Liu and Jeffrey Heer. 2014. The effects of interactive latency on exploratory visual analysis. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 2122–2131.
24. Lydia R Lucchesi and Christopher K Wikle. 2017. Visualizing uncertainty in areal data with bivariate choropleth maps, map pixelation and glyph rotation. *Stat* (2017).
25. Alan M MacEachren. 1992. Visualizing uncertain information. *Cartographic Perspectives* 13 (1992), 10–19.
26. Alan M MacEachren, Robert E Roth, James O'Brien, Bonan Li, Derek Swingley, and Mark Gahegan. 2012. Visual semiotics & uncertainty visualization: An empirical study. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2496–2505.
27. Leonard Mlodinow. 2009. *The drunkard's walk: How randomness rules our lives*. Vintage. 81–104 pages.
28. Dominik Moritz, Danyel Fisher, Bolin Ding, and Chi Wang. 2017. Trust, but Verify: Optimistic Visualizations of Approximate Queries for Exploring Big Data. In *ACM Human Factors in Computing Systems (CHI)*. `http://idl.cs.washington.edu/papers/trust-but-verify`
29. Jean-Claude Muller. 1979. Perception of Continuously Shaded Maps. *Annals of the Association of American Geographers* 69, 2 (1979), 240–249.
30. Shelby L O'Connor, Ericka A Becker, Jason T Weinfurter, Emily N Chin, Melisa L Budde, Emma Gostick, Michael Correll, Michael Gleicher, Austin L Hughes, David A Price, and others. 2012. Conditional CD8+ T cell escape during acute simian immunodeficiency virus infection. *Journal of virology* 86, 1 (2012), 605–609.
31. Bureau of Transportation Statistics. 2016. Airline Delays, Cancellations and Tarmac Times. (2016). `http://www.rita.dot.gov/bts/data_and_statistics/by_mode/airline_and_airports/airline_delay.html`
32. Lace Padilla, P Samuel Quinan, Miriah Meyer, and Sarah H Creem-Regehr. 2017. Evaluating the Impact of Binning 2D Scalar Fields. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 431–440.
33. Maria Riveiro. 2007. Evaluation of uncertainty visualization techniques for information fusion. In *Information Fusion, 2007 10th International Conference on*. IEEE, 1–8.
34. Philip K Robertson and John F O'Callaghan. 1986. The generation of color sequences for univariate and bivariate mapping. *IEEE Computer Graphics and Applications* 6, 2 (1986), 24–32.
35. Bernice E Rogowitz and Alan D Kalvin. 2001. The" Which Blair Project": a quick visual method for evaluating perceptual color maps. In *Visualization, 2001. VIS'01. Proceedings*. IEEE, 183–556.
36. Bernice E Rogowitz, Lloyd A Treinish, Steve Bryson, and others. 1996. How not to lie with visualization. *Computers in Physics* 10, 3 (1996), 268–273.

37. Robert E Roth, Andrew W Woodruff, and Zachary F Johnson. 2010. Value-by-alpha maps: An alternative technique to the cartogram. *The Cartographic Journal* 47, 2 (2010), 130–140.

38. Maureen Stone, Danielle Albers Szafir, and Vidya Setlur. 2014. An engineering model for color difference as a function of size. In *Color and Imaging Conference*, Vol. 2014. Society for Imaging Science and Technology, 253–258.

39. Nassim Nicholas Taleb and Mark Blyth. 2011. The black swan of Cairo: How suppressing volatility makes the world less predictable and more dangerous. *Foreign Affairs* (2011), 33–39.

40. Judi Thomson, Elizabeth Hetzler, Alan MacEachren, Mark Gahegan, and Misha Pavel. 2005. A typology for visualizing uncertainty. In *Electronic Imaging 2005*. International Society for Optics and Photonics, 146–157.

41. Bruce E Trumbo. 1981. A theory for coloring bivariate statistical maps. *The American Statistician* 35, 4 (1981), 220–226.

42. Archie Tse. 2016. Why We Are Doing Fewer Interactives. (2016). https://github.com/archietse/malofiej-2016/blob/master/tse-malofiej-2016-slides.pdf Malofiej.

43. Amos Tversky and Daniel Kahneman. 1985. The framing of decisions and the psychology of choice. In *Environmental Impact Assessment, Technology Assessment, and Risk Analysis*. Springer, 107–129.

44. StÃl'fan van der Walt and Nathaniel Smith. 2015. mpl colormaps. https://bids.github.io/colormap/, (2015).

45. Howard Wainer and Carl M Francolini. 1980. An Empirical Inquiry Concerning Human Understanding of Two-Variable Color Maps. *American Statistician* (1980), 81–93.

46. Colin Ware. 1988. Color sequences for univariate maps: Theory, experiments and principles. *IEEE Computer Graphics and Applications* 8, 5 (1988), 41–49.

47. Colin Ware. 2009. Quantitative texton sequences for legible bivariate maps. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009).

48. Colin Ware. 2012. *Information visualization: perception for design*. Elsevier.