

# Exploring the Design Space of Three Criteria Decision Making

Yu Jin , Laura Koesten , and Torsten Möller , University of Vienna, 1090, Vienna, Austria

*This article presents different interface designs of sliders to support decision-making problems with three criteria. We present an exploration of the design space through an iterative development process with eight prototypes and the results of several evaluation studies with visualization experts and nonexperts. Our findings show three candidates for consideration: a standard ternary triangular slider, a novel circular slider, and a standard basic slider displayed three times. All three were considered intuitive and easy to use. The triangular slider is best for exploration with vague user intuition, the circular slider performs best for preference comparisons, and the parallel slider is best for direct preference setting.*

Uncertainty is an inherent part of everyday decision-making, such as when deciding which apartment to rent or which car to buy. Often in such decisions, more than one criterion (factor) is being considered at the same time. The importance of those different factors might only be a vague intuition for the decision-makers. When engaging with the decision task and the available options in more depth, preferences of the factors might change.

In the example of renting an apartment, one might weigh *rent*, *size*, and *location* of the apartment against each other but assign different importance (or weight) to each criterion; a dynamic process which can be difficult and time consuming. These criteria often depend on each other, and the weights assigned can dynamically change during the decision-making process as well as through the exploration of the available options.

In a simple decision-making case, only two factors are considered (say, cost and size). In such scenarios, the weight of the two criteria produces a ranking of the possible options. For instance, the price might have an importance of 75%, while the size (of the apartment) gets the remaining 25% of the weighting.

Hence, the choice of weights can directly produce a ranked list of the available apartments which fit the user's criteria.<sup>1</sup>

For two criteria, a simple slider is sufficient to interact with the entire possible decision space. However, when there are many criteria (a handful to dozens), several complex interfaces have been proposed in the past, such as LineUp<sup>1</sup> and WeightLifter.<sup>2</sup> Yet, such (well-designed but) heavy-handed interfaces might be too complex for simple uncertainty decision spaces of non-expert users.

Returning to renting an apartment, multiple factors, such as size, price, location, amenities, age, heating costs, view, etc., could be considered. Yet, the main criteria are often only a handful or less (such as size, price, location, and maybe the costs of amenities). Hence, the question arises: What would be an intuitive and sufficiently simple interface widget to support decision-making under uncertainty with three criteria? Two criteria are simple, and three criteria have, to the best of our knowledge, not yet been explored.

For such scenarios, a traditional bidirectional slider does not suffice. These scenarios, also standard on online shopping platforms, are commonly supported through filtering interactions. However, this requires people to have specific values, or at least ranges, in mind before selection.

In reality, they may only have preferences for some factors in relation to others; they may have different subjective weighting preferences, which can be complex to articulate; and they may not be able to select specific criteria before exploring the space of possible options and how criteria are related within that space.<sup>3</sup>

---

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see <https://creativecommons.org/licenses/by/4.0/>

This article has supplementary downloadable material available at <https://doi.org/10.1109/MCG.2023.3303672>, provided by the authors.

Digital Object Identifier 10.1109/MCG.2023.3303672

Date of publication 22 August 2023; date of current version 13 September 2023.

It can often be difficult for people to describe the scale of how much more they prefer a certain factor over another (precisely, in a numerical manner).<sup>4</sup> For example, users may prioritize location over the size of an apartment but cannot express this exactly on a numerical scale. For instance, one might attribute great importance to an apartment's location, sacrificing size if necessary, but there comes a size that is considered too small to make the tradeoff worthwhile. This threshold adapts dynamically related to the available options.

Hence, in this article, we designed and compared different interactive tools for three criteria decision-making (3CDM). This is a frequent scenario in everyday decision-making under uncertainty.

## CONTRIBUTION

We explore the design space via eight different high-fidelity prototypes of interactive visualization interfaces, which are evaluated iteratively by visualization experts as well as by nonexperts (people without experience in data visualization). The objective is to support common daily multicriteria decision-making (MCDM) problems. We investigate solutions for a broad set of people, which help to intuitively explore vague preferences by presenting the various widget designs and comparing the usability of each widget while adapting the respective optimum solutions interactively.

Out of these eight prototypes, we find three to be favored under different conditions:

- › The altitude widget [see Figure 3(c)] is best for exploration (i.e., where the user only has a vague intuition of their preferences);
- › The advanced circular widget [see Figure 3(f)] is best for understanding proportion (or part of a whole);
- › The basic three-way widget [see Figure 3(h)] is best for direct, individual input with a precise idea of what the user wants.

## RELATED WORK

### General MCDM Interaction

The community dealing with MCDM has been surveyed by Köksalan et al.<sup>5</sup> and others. While they are mostly concerned with many criteria, computational approaches, and Pareto-front exploration, we are concerned with few (three) tradeoffs and the usability of interfaces.

The visualization community has explored the problem of decisions under multiple criteria mostly from an angle of Visual Parameter Space Exploration.<sup>6</sup> LineUp<sup>1</sup> supports users in defining different

weightings and exploring the options. While LineUp constrains the user to explore a finite set of different weightings, WeightLifter<sup>2</sup> explores the decision-making space comprehensively by allowing all possible (infinite) weightings.

All of these tools have been evaluated for their usability by respective domain experts. However, they remain overwhelming and complicated at first, considering nonexpert scenarios. Hence, these solutions might not be accessible for common decision-making problems with few decision criteria for broader user groups.

### Tradeoff Analysis

There are past works that have specifically targeted tradeoffs or preference settings in the MCDM domain. Hakanen et al.<sup>7</sup> recommend including coordinated multiple views for developing tradeoff analysis, as well as the interaction of Linking-and-Brushing for the development of a multiobjective optimization interactive system to enhance visual analytics.

TOP-slider<sup>8</sup> and Pareto Slider<sup>9</sup> both have a similar design objective. While the Pareto Slider design supports tradeoffs and explores the Pareto optimal solutions (a set of optimal solutions for multiple criteria) in the medical field, TOP-slider targeted a broader audience, including also nonspecialists. Similar to our work, it only focuses on three criteria but on one design, allowing users to adjust their tradeoff preferences directly on three parallel sliders. Compared to the Pareto Slider, it displays a broader range of solutions (e.g., optimum values (Pareto fronts), feasible but nonoptimal values, unfeasible values) differentiated by colors but with almost identical interaction. Importantly, the solutions are updated dynamically based on changes in the sliders on both designs, as recommended by Hakanen et al.<sup>7</sup> when designing multiobjective optimization interactive systems.

A slightly different design to the Pareto Slider, Pareto navigation is also applied in the medical domain for finding a Pareto optimal for IMRT treatment planning.<sup>10</sup> Instead of using parallel sliders for adjusting preferences, the design is based on a star plot. By changing the position of the restrictor (a button that controls the boundaries for planned values) and the selector on each axis, it forms the solution as a space on the plot instead of a single point, where it could be further explored by the user.

A very recent design, VisProm,<sup>11</sup> explores the optimum solutions from a different perspective by using an in-visualization provenance view that focuses on visualizing the history of data. Here the tradeoff problem is addressed by analyzing the high-level needs of experts

and encoding them into low-level provenance objects that can be accessed. It uses scatterplot matrices as the main visualization technique to communicate multi-dimensional data and to display the solution space.

Other forms of preference elicitation techniques can be found in spatial decision making. A study exploring how interactive maps encode geographic data to support spatial exploration and how to could better structure the presentation of tradeoffs was done by Jankowski et al.<sup>12</sup>

### Ternary Plot

By combining multiple criteria, WeightLifter<sup>2</sup> reduced the interface to simple 1-D sliders as well as a Ternary plot. A Ternary plot or ternary diagram is a graph that displays a system of three variables with the constraint that the sum of three variables needs to be one. Hence, the variables are generally represented as percentages. More importantly, it represents a three-component system in a flat (2-D) decision space (as opposed to a 3-D presentation.<sup>13</sup> As we want to avoid 3-D interfaces,<sup>14</sup> the three-way decision problem sits at an intriguing point for special interface design.

## METHODS

We aim for our approach to be simple and usable for nonexpert users. Hence, we focus on the interaction design and visual design space of the 3CDM visualization. In the early prototyping phase, our pilot users were at times unclear about what exact type of three-criteria decision-making interface we were building. In fact, emerging from this phase, we observed three different goals for approaching a 3CDM task:

- › *Preference (our focus):* In this particular scenario, the goal is to determine a preference of multiple criteria and find the optimum result based on the indicated preferences.
- › *Selection:* During a selection task, a user wants to specify a multidimensional value. This means, assuming a three-slider selection widget, they select a particular value on each slider and the combination of the slider values presents a single entity in a 3-D space.
- › *Filtering:* Here, sliders are used to determine “larger than” or “smaller than” ranges on a scale. The combination of intervals on multiple sliders determines a subspace selection in the multidimensional decision space.

We believe that the Selection, as well as the Filtering tasks are relatively well served with a simple

design of three (parallel) value sliders. However, the Preference task is of particular concern since the three values influence each other (accordingly) and is discussed in prior literature.<sup>9,10,15</sup> Incorporating this dependence tends to be harder to conceptualize for users and, therefore, is more complex to design for.

Hence, in this work, we focused on prototyping for preference (tradeoff) tasks. That is, we explored the design space for specifying preferences for three different criteria which should add up to one.

In an initial low-fidelity prototyping phase, we developed 32 prototypes. These prototypes were reviewed and critiqued by three visualization experts via an interview study. Based on these results as well as continued reviewing by two senior visualization experts, we narrowed down our designs to eight high-fidelity prototypes, which were then tested with 20 potential users. Our results are detailed in the “Results” section.

Further, we compared different methods for computing the optimum solution for a tradeoff (preference) task, and based on previous work by Raubal and Rinner,<sup>15</sup> we are using the weighted linear combination (WLC) as our MCDM method. Its simple computation allows great responsiveness for the interaction and is a key factor for perceived usability (see Hakanen et al.,<sup>7</sup> Laurillau et al.,<sup>8</sup> and Monz et al.<sup>10</sup>).

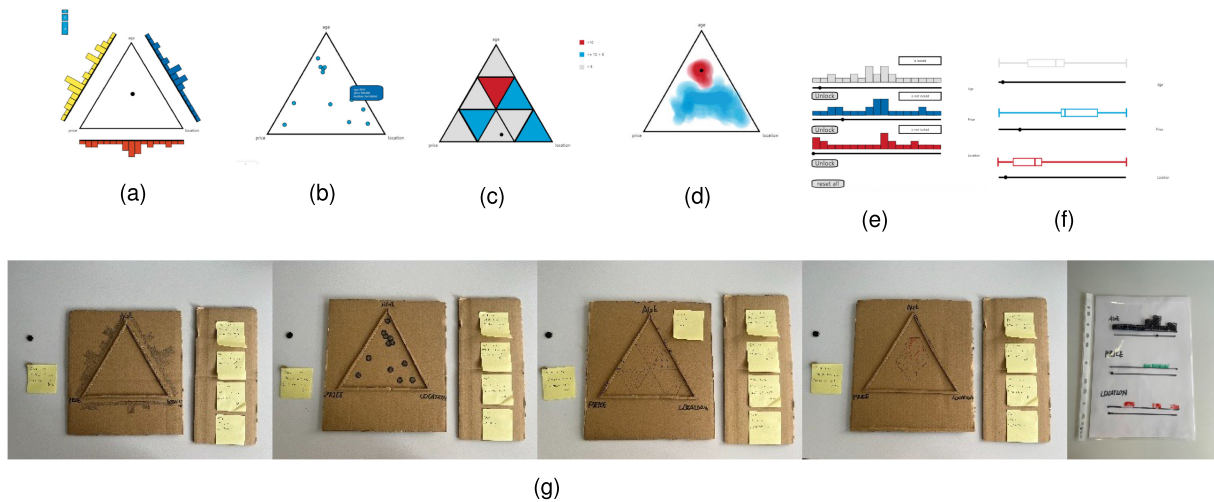
## PROTOTYPING PHASES

The low-fidelity prototyping process encompassed different methods, including sketching, online whiteboard prototypes as well as cardboard prototypes. Those methods aim to enhance the process of collecting expert feedback. The prototypes follow the same Preference task defined in the “Methods” section.

### Low-Fi A: Online Whiteboard and Cardboard Prototypes

Initially, we created six online whiteboard drawings, as shown in Figure 1. The designs were mostly inspired by the ternary plot used by WeightLifter<sup>2</sup> and the interactions supported by the Adobe Color wheel. This led to four variations: a projected histogram style [see Figure 1(a)], a scatterplot style [see Figure 1(b)], as well as two versions of heat maps [Figure 1(c) and (d)]. In addition, we also wanted to contrast the ternary plot to a relatively simple three-slider version of an input widget. Again, using the inspiration of the TOP-slider<sup>8</sup> we designed a histogram-style and a box-plot-style slider [see Figure 1(e)].

After a first round of feedback from one visualization Ph.D. student (E1: researching explanatory interfaces of machine learning algorithms, feedback was sought via a Zoom session) and two senior



**FIGURE 1.** Low fidelity prototypes. (a) Triangular slider with histogram. (b) Triangular slider with scatter plot. (c) Triangular slider with heatmap design 1. (d) Triangular slider with heatmap design 2. (e) Parallel slider with histogram. (f) Parallel slider with box plot. (g) Physical prototypes including plastic foil and cardboard.

researchers (one male and one female, during a Zoom session), we decided not to pursue the box-plot style interface further, as box-plots tend to be inaccessible to a broader audience. We were further inspired by James Dyson's Carboard modeling method,<sup>16</sup> which is quick, low cost, and requires little technical skills from the participants. This allows prototyping at the early conceptual stages of the design process to gather feedback and ideas for alternate design solutions. Hence, a set of cardboard prototypes of the remaining five widget solutions was created, as shown in Figure 1(g). The intention was to discover potential limitations and to further simulate the feeling of the interaction between different slider designs compared to digital prototypes and hand-drawn sketches. Especially for the cardboard, people could interact with the slider by moving a cardboard-cut dot on top. The interaction with the cardboard prototypes did in fact surface limitations of our initial designs, which we iteratively adapted during the design phase. For instance, the limitation of performing a selection task Methods on Ternary widgets (a state where setting 3 criteria to the maximum at the same time could not be reached on this solution).

The five remaining widgets were tested with two additional visualization experts. One (E2) was a Ph.D. student with almost three years of experience in visualization and data analysis. The other expert (E3) was a master's student with two years of experience in data visualization and human-computer interaction.

The next round of interviews with experts E2 and E3 took place in person. The expert interview averaged 30 min and was done individually without communication between

the experts. E1 and E3 suggested guiding lines displayed inside the triangular slider and displaying the current value next to the current position of the slider. Most importantly, one user wanted to move the slider to a position minimizing all values of the three criteria simultaneously.

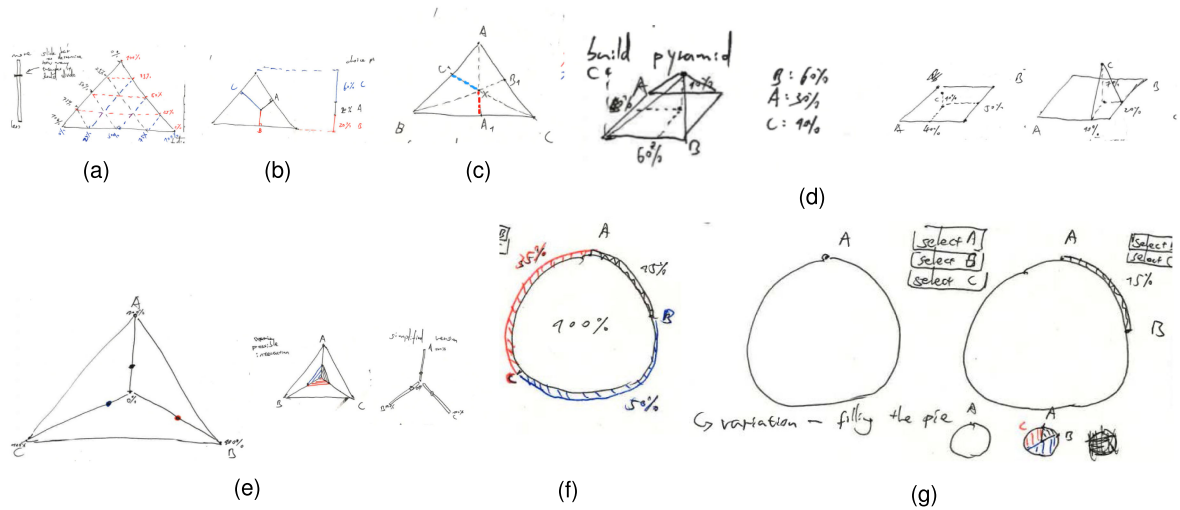
E1 and E2 both suggested that the widget should support further interactions (for example, a zoom-in function) to prevent visual clutter if the data get too dense. E2 and E3 suggested that bivariate color maps are inappropriate for this task.

E2 and E3 both mentioned that the movement inside of a triangular slider for performing data selection varies in sensitivity (i.e., movement in the middle of the triangle does not equal the same movement at the tip of the triangle) a small movement at the tip will generate a substantial amount of change.

This feedback gave us various ideas for a new iteration of low-fidelity prototypes. After clearly articulating the separation of the Selection, Filtering, and Preference tasks, we developed another round of prototypes. Using the feedback, the initial prototypes were refined to six new low-fidelity prototypes in addition to a basic three-way slider (as a baseline). Further, during the development of the high-fidelity prototypes, it became apparent that an essential version of the circular slider was missing [currently seen in Figure 2(g)].

## Low-Fi B: Final Sketches

For the final round of low-fidelity prototypes, we returned to hand-drawn sketches. The motivation was to explore the different design options as much as



**FIGURE 2.** Final low fidelity prototype sketches. (a) Standard ternary widget. (b) Altitude widget. (c) Intersection widget. (d) 3-D Pyramid widget. (e) Propeller widget. (f) Advanced circular widget. (g) Circular widget.

possible, most efficiently and simply. We also found that using hand-drawn sketches for early prototyping would be sufficient to communicate the principal concepts to our participants.

The result of this prototyping stage can be seen in Figure 2. Next, we detail the motivation that went into each of these eight designs.

Figure 2(a) is inspired by the *standard ternary plot*. However, different from a static ternary plot, users have the option to move inside the triangle with the interaction of dragging a dot. Depending on the dot’s location, the percentages of different criteria will change accordingly, and the exact percentages will be displayed in the text next to the slider. This interaction is inspired by Adobe Color Wheel, where the user adjusts the RGB preferences inside the wheel to determine the color. Further to our design, users also have the option to adjust the density of the underlying grid to help them navigate within the triangle. Through a simple binary vertical slider on the left side, users could increase or decrease the grid’s opacity by moving up and down. Here the bottom of the slider signifies transparency (turning OFF the grid completely), and the top presents opaqueness. This design serves two purposes: 1) to help users find a precise location within the decision space (i.e., when they have a specific weight in mind). 2) to enable a fluid exploration (especially when the grid is transparent), i.e., in the case when a user only has a vague intuition. Compared to the similar MCDM tool design of WeightLifter,<sup>2</sup> which starts by adding a slider first inside the Ternary plot, users have the freedom to move arbitrarily inside the triangle.

Figure 2(b), the *altitude plot*, modifies the ternary plot through lines to the edges of the triangle. Similar to Figure 2(a), users could change the location of the dot by dragging. However, the perpendicular lines to the three edges represent three different weights for three criteria. Depending on the dot’s location, the length of the perpendicular line will be changed accordingly. The longer the line is, the higher the percentage or the weight of that criteria. The exact percentages will be displayed on the bar on the right side. This bar further encodes the weights in different categorical colors as well as a bar chart, i.e., by moving the dot, the weight of each criterion represented on the bar will be adjusted accordingly. We hypothesize that looking at the bars might support an understanding of the concept of weights. Similar to the “Overall Trade-off Weight-slider” of WeightLifter (an explicit visual representation of the weight distribution on each criterion), the distribution of weight visualization in our widget is incorporated via the length of each perpendicular line.

Figure 2(c), the *intersection plot* is a small modification of the previous idea, where we measure the extension of the line drawn from each corner of the triangle to the location of the user-controlled dot and extend it until it intersects the opposite edge. Thus A1, B1, and C1 intersect the dotted lines and the edges of the triangle. The length of the lines spanned by A1, B1, and C1 to the (user-controlled) dot will be highlighted through three categorical colors. It then shows the proportion of the entire line from the tips to the intersection, thus representing the weights/percentages of different criteria. Compared to WeightLifter, the

intersection widget completely removed the explicit visualization of the overall weight distribution representation. The user can only perceive the proportion by looking at the different lengths of the colored line on the ternary plot.

Figure 2(d) contains two variations. The main 3-D pyramid design allows users to move the dot on the base plane while the third component would build a pyramid. For a Preference task, if two criteria are fixed, the third one is uniquely determined since they all need to add to 100%. We did acknowledge the complexity of this widget. Hence, we were looking for ways to simplify it and make it more accessible. This motivation led to the next two variations.

In the middle of Figure 2(d) is a variation to the 3-D pyramid widget. It simplifies the design by removing all edges from the pyramid. Hence, we only present the necessary perpendicular dotted line to indicate the percentages or weights of the third criterion. On the left side of Figure 2(d) is another variation to the 3-D pyramid widget. This time, it builds a tetrahedron instead of a pyramid.

Figure 2(e) shows the *propeller widget*, inspired by the MIRA navigator.<sup>10</sup> It contains two variations. The main propeller widget has, in total, three individual sliders with different categorical colors, each pointing from the center toward the corners of the triangle. The lowest value (zero percent) of all three sliders is located at the center, and the highest value (hundred percent) is located at the tips. The concept of only adjusting two criteria and automatically generating the third does not work here. In the scenario where, for example, users select criteria A and B, criteria C would then be automatically generated. As soon as a user wants to adjust criteria C, either criteria A or B would need to be readjusted to limit the total to a hundred percent, which one of the two needs to change is unclear for us to design. To solve this, we let users have the freedom to adjust all three of the sliders without the limitation of the hundred percent total. When the total exceeds a hundred percent, for instance, when A, B, and C are all 50%, in that case, one would simply normalize by the sum.

In the middle is a variation to the propeller slider; it is inspired by the traditional star plot, but has the same concept and interactions as the main design. It links the dot on each slider with different categorical colors, thus forming a star plot. This aims to simplify the comparison between data entries.

To the right, there is another variation to the propeller slider. It simplifies the design by reducing all the edges from the triangle to give it a minimal look.

Figure 2(f) shows the *circular widget*. The idea is to split the whole circle into three parts, where each part

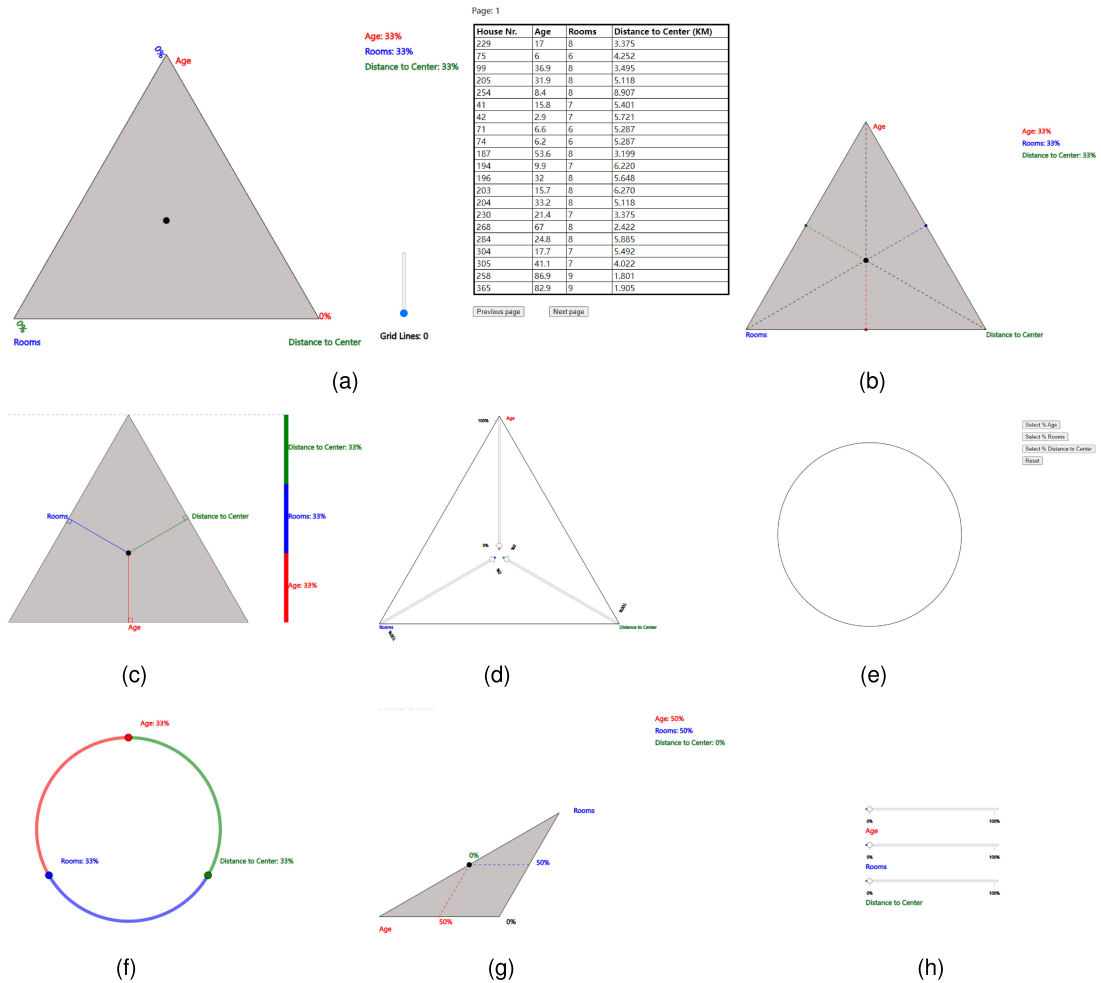
represents one criterion. It shows three dragging dots directly on the circle with different highlighted arc paths in between. Users can drag the dots on the circle to change the percentages freely in any order they like. This allows a quicker adjustment than a triangular slider since the movement space is more limited.

Figure 2(g) shows the *essential circular widget*. The idea came during the development of high-fidelity prototypes. It shows an empty ring without any dots at first. Users can select the criteria in any order using the buttons in the top left corner. The button will be disabled after they are clicked, and a dot for the selected criteria will appear on the ring. Thus, it can be adjusted to a certain percentage. If two of the buttons are being clicked (two criteria have been selected), the last button will be disabled, and the last criteria will be filled on the circle automatically based on the two previously selected criteria. This allows users to prioritize specific criteria by selecting them in order.

## High Fidelity Prototypes

After a discussion with two senior researchers, we decided to develop all seven main designs from Figure 2 as well as the standard basic three-way slider as a baseline. Here, the basic three-way slider consists of individual sliders displayed underneath each other. Each of the individual sliders was named, as shown in Figure 3. Compared to the TOP-slider<sup>8</sup> and Pareto Slider<sup>9</sup> mentioned in the previous section, where a more complex model (Pareto Model) is being used, and the optimum range for each criterion is presented directly on the sliders, our slider design uses a much simpler model. This model does not suggest any optimum range or hint, thus uncoupling the exploration of uncertainty by giving the user full movement (control) over the entire range on each slider. We believe that these prototypes provide a broad enough spectrum of the possible design space of 3CDM widgets. Some of the prototypes were not developed any further. On the one hand, the aim was to keep the user study manageable (keeping the length of the study as short as possible). On the other hand, there were specific reasons for each of the design choices which made them inferior to the others:

- ▶ Version 2 of the 3-D pyramid [see Figure 2(d)]: In the end, the simplification made the 3-D aspect of the widget hard to see. Hence, it was confusing.
- ▶ Version 3 of the 3-D pyramid [see Figure 2(d)]: It was considered to be too identical to the 3-D pyramid widget, not contributing new ideas to the design space.
- ▶ Version 2 of the propeller [see Figure 2(e)]: The star-plot-type interface quickly became cluttered,



**FIGURE 3.** All high fidelity prototypes developed using JavaScript. The ranking table on the right side of the full widget (a) is consistent throughout all prototypes, thus the left figures from (b) to (h) only show the left part of the widget prototype. (a) Standard ternary widget. (b) Intersection widget. (c) Altitude widget. (d) Propeller widget. (e) Circular widget. (f) Advanced circular widget. (g) 3-D pyramid. (h) Basic three-way widget.

considering all possible decision points to be displayed. Hence, we believe this is a drawback that is not easy to overcome.

- › Version 3 of the propeller [see Figure 2(e)]: It was considered too similar to the basic three-way widget. We could not see any advantage to place the sliders in a nonhorizontal way. The horizontal placement would best support the typical reading style of a human user.

Our high-fidelity prototypes were developed using D3.js under the JavaScript programming language. Once the user selected a preference for three criteria, the resulting ranking of the options was displayed as a table, as shown in Figure 3(a).

This table was placed, consistently with all hi-fi prototypes, to the right of the widget but is shown only for the first prototype in Figure 3, due to space constraints.

## EVALUATION

For the evaluation, we used the high-fidelity prototypes from Figure 3. To make the tasks more realistic, we framed a common task: searching to buy an apartment. For our purpose, we used the Boston Housing Dataset to give a more realistic setting. Specifically, we chose the following attributes of the dataset: AGE (proportion of owner-occupied units built prior to 1940), RM (average number of rooms per dwelling), and DIS (weighted distances to five Boston employment centers). We

picked these as we believed that they are somewhat similar to the essential criteria in such a use case (age, rooms, and location). We did not include price since we did not test our prototypes with people from Boston, and the local pricing ranges are very different from those in Boston. To simulate a more realistic house-buying experience, we further modified the data by rounding to the nearest integer to make the options more accessible.

## Experiment Design

A detailed study protocol was designed applying the following four steps:

- 1) Consent form: The user was given a consent form, pointing out the minimal risk of the experiment and that user data will be anonymized while providing contact information in case of further questions.
- 2) Introduction: A short one-minute verbal introduction on the logistics of the study, the motivation of the work as well as the purpose of the prototypes.
- 3) Self-exploration of the interface: Participants were given 10 min to explore all eight different widgets. This gave us the opportunity to observe if the participants could understand how each slider works by themselves without any guidance.
- 4) Study task: The study task was to adjust each interface to the following state:
  - › The percentage of Criterion A has to be more than 50%.
  - › The percentage of Criterion B has to equal the percentage of Criterion C.
  - › The sum of the percentage of Criteria B and C has to be less than 50%.

All three requirements need to be fulfilled to be considered as a correct state.

To get both sufficient quantitative and qualitative feedback for each prototype in one session while having participants focused and interested during the study, we condensed the entire study duration to close to one hour, including one Preference task mentioned above.

To make this task more understandable, we moved it from an abstract setting to the concrete setting of buying a house. We communicated the task to the participants as “Imagine the age of the house is the most important factor for you. Further, imagine that the distance to the city center and the number of rooms are equally important yet less important than the age of the house.”

The same task was then performed on each prototype. While performing the task, participants were asked to speak out loud, verbalizing their thinking process. After finding the correct state (in the participant’s perception), the time consumed for the task was measured. Finally, the user was asked to rate the experience of using the prototype on a scale from 1 to 10. Participants were also asked to discuss each prototype directly after the task and before the testing for the next prototype started. We randomized the sequence of prototypes for each participant with the constraint that at least one different prototype (in terms of concept and design) needs to be in between two similar ones to limit confounding factors (e.g., learning effects).

- 5) Post-task Interview Participants were asked a total of five interview questions and one optional question:
  - a) Which model did you like the most and why?
  - b) Which model did you dislike the most and why?
  - c) What was the feature that was the most interesting to you or would you like to have?
  - d) What was the feature that confuses you the most or which you disliked?
  - e) Which model was the simplest/took the least time for you to understand?
  - f) (optional question) Which model would you see being used broadly in the future?

The questions were asked to elicit participants’ perceptions of the prototypes to enrich the analysis.

## RECRUITMENT

In order to test the high-fidelity prototypes, we ran a pilot with a computer scientist through an online video conferencing platform (Zoom). The pilot went well without any issues and took 50 min. When we ran the full study protocol mentioned in the “Experiment Design” section, we noticed some frequent issues not raised during the previous pilots. Hence, after the first five participants, we made a few changes and separated these five users also in our analysis (see Table 3 in the supplementary material) First, we abandoned the 3-D Pyramid. The feedback for this particular widget was exceptionally negative. None of the participants seemed to recognize that it was designed to be a 3-D widget, hence it created more confusion than planned.

Further, we adjusted the labeling and axis of the propeller widget [see Figure 3(d)], advanced circular widget [see Figure 3(f)], and standard circular widget [see Figure 3(e)], as well as their scale to fit the screen (1920 × 1080 pixels)



better. In addition, the cursor on the basic three-way widget [Figure 3(h)] and propeller changed into a hand pointer instead of an arrow. The A1, B1, and C1 labels of the intersection points were changed into small dots on the intersection widget [see Figure 3(b)]. Comparing the scores of the participants P1 to P5 with the remaining P5 to P20 (Table 3 in the supplementary material) shows an overall improvement in scores (for the later cohort). After these adjustments, the study schedule was reduced slightly in time to 45 min on average.

## Participants

Our study had the purpose of testing usability on the one hand but also to get further feedback on the design space. While some authors have shown that the coverage of usability problems does not significantly increase after only five participants, others have concluded that 10–12 participants are a sensible baseline range.<sup>17</sup> In terms of the design space, we were hoping that others would point us to alternate designs that might be obvious to them but that we missed. After talking to 20 people, no new ideas came up. Hence, we believe that we have reached (theoretical) saturation of design alternatives.

We made use of personal networks and social media groups. Seven sessions were performed in person, and 13 participants participated through an online video conferencing platform (Zoom). All participants were older than 18 years. We had 14 participants aged 18–24 and 6 participants aged 25–40, with an average age of approximately 23. The majority of the participants had just finished their bachelor's degrees. While they have diverse expertise in the field of art, economy, psychology, engineering, language, etc. [as can be seen in the supplementary material (Table 1)], none of the participants have experience with data visualization.

## RESULTS

Both quantitative and qualitative aspects of the evaluation were generally positive for all widgets (except in the case of the 3-D pyramid). Overall, 80% of the participants reported that they are interested and would see themselves incorporating one of the 3CDM widgets for their daily needs.

The Likert scale ratings for the 15 participants can be seen in Figure 4, one histogram for each of the widgets. Further, Figure 4(a) shows 1-D scatterplots for the time each participant took to complete the task for each widget. Here we included the first five participants to demonstrate that the overall time improved, especially for the circular and advanced

circular widgets. Finally, a summary of all results can be seen in the supplementary material (Table 3).

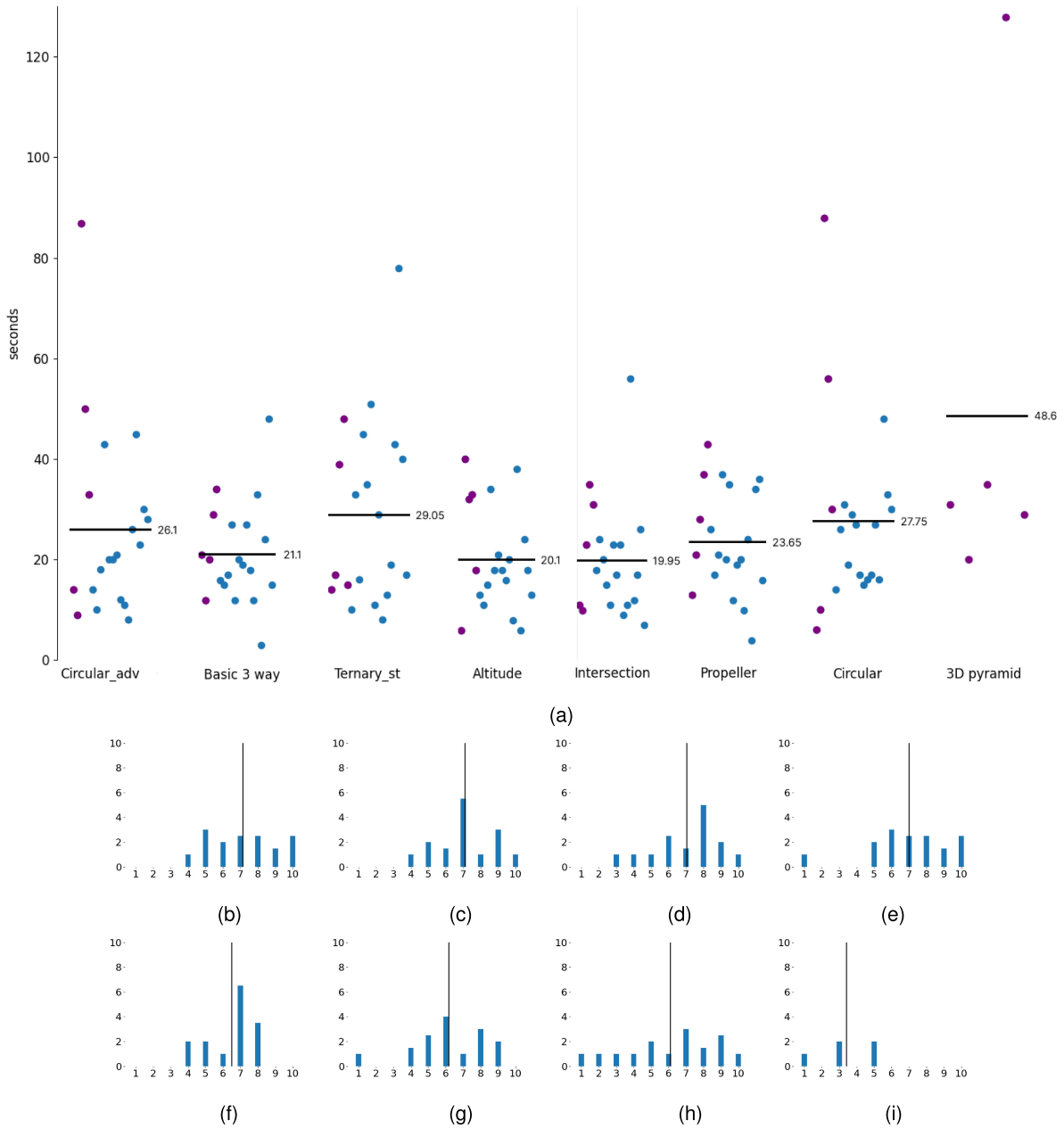
*Errors:* First of all, it is worth noting that there were almost *no errors* made in bringing the widgets into the state as demanded by the test. We conclude that the users understood the task as well as the widgets designed. There were a few outliers, especially among the first five users:

- 1) Propeller widget: P1 accidentally swapped the percentage on criteria A and B, so that criteria A is equal to C and B is more than 50%.
- 2) Standard Circular widget: P3 gave up after not understanding the model and put it in a random state.
- 3) 3-D pyramid: P3 gave up after not understanding the model and put it in a random state. (Similar concerns were raised among the first five users who experienced the Pyramid widget).
- 4) Intersection widget: P12 lost the patience to adjust the spot with a trackpad, even after she explained that she understood the concept and its meaning.

*Scores:* Considering the overall Likert scores, the clear winners would be the advanced circular, the basic three way, and the altitude or standard ternary plot. However, all other designs, although acceptable, only come in at a clear (scoring) distance.

*Timing:* It is perhaps not surprising that the timing for the altitude is best, followed by the intersection, as both of these have only one point to drag (i.e., the interaction is relatively simple). However, people complained about the difficulty of fine-tuning to get the exact percentages as specified [as can be seen in the supplementary material (Table 2)]. In that later sense, the basic-three way version is more precise, yet one has to adjust three sliders instead of one. This tradeoff made it a close third in terms of the average time the users spent solving the task.

*Preferred designs:* The qualitative feedback [as can be seen in the supplementary material (Table 2)] showed that the standard ternary widget, the altitude widget, the advanced circular widget, and the basic three-way slider were the favorite widgets. About 50% of the participants mentioned that they like the preciseness provided by the grid lines in the standard ternary widget and found this feature the most interesting. About 40% of the participants mentioned that they liked the design of the altitude widget, particularly the bar next to it. For participants whose favorite the altitude widget was, the majority mentioned this widget as easy to understand. P20 specifically mentioned: "The right bar further encodes the



**FIGURE 4.** (a) Scatter plot for time spent on each widget. In each plot, the x-axis shows the participants, and the y-axis shows the amount of time (in seconds) they took to complete the task. The black line shows the average time spent on that widget. The first five participants are in violet, and the remaining 15 are in blue. The 3-D pyramid widget was abandoned; therefore, the plot only shows 5 participants on the x-axis. Results show that the Intersection widget has the least time spent. (b)–(i) Likert scale histogram in the order from the highest average rating to the lowest. The black line indicates the average Likert score on the current widget. The x-axis shows the scale from 1–10, the y-axis shows the number of participants that checked that rating. (Note: There are 0.5 on the y-axis because some participants rated in-between the integer scale. For example, P9 rated 8.5 for the propeller. Hence, we counted the participant as 0.5 to both its neighboring values 8 and 9.) (a) Scatter plot for time spent on each widget. (b) Circular advanced. (c) Basic three-way. (d) Ternary standard. (e) Altitude. (f) Intersection. (g) Propeller. (h) Circular. (i) 3-D pyramid.

information of the percentages and can have a better comparison, making it even easier to compare." As for the participants who preferred the advanced circular widget or the basic three-way widget, feedback was similar: they mentioned that the widget for them is the easiest to understand and use. However, the difference is that there are mentions that the advanced circular widget was more fun to play with. P2 also mentioned: "I was not impressed when I first played around with it in the first 10 min. But when I used it for the task, it gave me a good representation of the weighting." As for the basic three-way widget, participants liked the simple movement. P18 mentioned: "it was very simple to operate and a very straight movement rather than moving on an arc. It has also an extra function while others did not."

*Designs with most negative feedback:* On the other hand, the standard circular widget and the propeller are both widgets that received the most negative feedback. For the propeller widget, the feedback was that it seemed nonintuitive and unnecessary to present the widget in a triangular shape.

P6 also mentioned: "the basic three-way would have been enough." As for the standard circular widget, participants found it nonintuitive that the criteria selection has to be selected in an ordered way and that it needs to be reset when trying to adjust or correct the criteria. P5 even mentioned: "If I made a mistake, then I had have to reset the whole thing. I feel that there is a punishment factor to use. I worried about making a mistake while using it." Some participants also articulated stress about wanting to avoid a mistake on the first selection, which meant they would have to restart from the beginning with an empty ring. The ordered selection and resetting features were also voted as the most confusing elements compared to others.

*Other interesting feedback:* We also received interesting general feedback throughout the whole study. Although 55% of the participants found the grid line feature useful in the standard ternary widget, 30% of the participants stated the opposite. They found it better to use without the grid line, as the grid can hinder the usage of that prototype. More interestingly, there were two participants who found it useful but preferred not to have a grid. About 40% mentioned that cursor movement is difficult and imprecise when using the intersection ternary widget, and 30% of the participants commented the same on the altitude slider. That statement did not come up when using the standard ternary widget, although their movements are similar. The reason could be the differences in the visualization and will be discussed further in 8.

A few participants mentioned that the ternary widgets provide a spatial feeling rather than a numerical one.

P7 mentioned: "Easy to use when I have a vague decision, but difficult with specific numbers in mind." She also mentioned that "if you scale down the triangular slider, it would not function well anymore, while if you do that to a circle, it will not affect the usability." This unique and important perspective was not considered during the design stage since the tool could be embedded with other tools at this point. To evaluate whether the statement is true or false, the size perspective needs to be considered and further tested. Three participants mentioned that moving on an arc of the circular widget is not as precise and easy as on a plane like on the triangular widget. P16 said: "it feels faster to find the spot on the circular sliders compared to the triangular sliders." However, the measurement of time-on-task shows that almost all the triangular widgets have faster performance than the circular ones, except the standard ternary widget. Interestingly, participants perceived efficiency differently from their actual performance.

## DISCUSSION

*Uncertainty:* During the design phase of our work, we did not explicitly emphasize uncertainty in each of the widgets prototypes. However, uncertainty is often in close relationship with decision-making,<sup>18,19</sup> as for our widgets, each widget implicitly incorporates uncertainty in decision-making by not supporting the precise placement of the preference setting on each widget. From our observation during the user study, most of the participants would first select the area roughly close enough to their desired preference without looking at the exact preference percentages. They would then spend more time fine-tuning the exact percentage numbers demanded in the study.

*Different mental models:* During our observations across the entire evaluation study, we observed mainly two different general types of thinking and performing among the participants. One tends to only focus on the exact percentages on the right side of the slider, whereas most tend to finish the study task with the exact percentages on each prototype. The other "participant type" tended to look more at the actual slider. On the circular slider, they look at the colors first. If the length of the arc color looks approximately correct, they will then look at the percentages and adjust the numbers to get them exact.

*Complexity of the design:* Some of the participants ( $n=9$ ) had a more technical background. One of them used ternary plots before the experiment. Interestingly, the time he spent on the ternary widgets was longer than some other participants, who did not encounter ternary plots before and whose background

was nontechnical. This suggests the design and the concept of the ternary plot might not require much previous knowledge or technical expertise.

*Controversial design features:* More than half of the participants liked the idea of grid lines on the standard ternary slider, with the attempt to guide them to find the specific state they would like. About 38% of the participants also state that they would like to have this feature on the other ternary widgets. However, the remaining participants state that they do not need or prefer not to have the grid lines on the standard ternary widget. Two participants stated that it is interesting, but they do not need it, as they prefer to have the entire space for exploration. We do not see it as a drawback that some of the widgets invoked playfulness. For example, even though people might not need help from grid lines, it can still be fun to interact with that feature.

*Inaccuracy in visuals:* 40% of the participants mentioned that the movement of the intersection widget is difficult and less precise compared to the other prototypes. A participant mentioned that “a small movement will generate a huge visual effect as well as the percentages,” but in fact, the math and the movement are mostly similar to the two alternative ternary widgets. In other words, a movement on this prototype generates the same amount of changes as the other ternary widgets. However, the reported perception of some participants tells us otherwise. It is interesting to see the different visual interpretations that could cause these differences in perception. As for the Preference task mentioned in the “Methods” section, this interpretation might be misleading as it projects an inaccurate perception and should be avoided. Due to our recruitment via social networks, the people participating in the study were mostly young and university clientele. A broader sampling of the general public might be informative.

*Proportion perceiving in Circular widgets:* During our study, a large number of participants stated that the Circular widget gave them a better sense of the proportion by adjusting the “selector” on the arc, some ( $n=3$ ) participants explicitly stated that circular widgets provide a better feeling of proportion compared to the other widgets. Skau et al.<sup>20</sup> conducted an empirical study on the visual encoding of proportions with different techniques. Their study shows that using arc length as proportion encoding performs more accurately and provides important information for reading values compared to, e.g., area and angle.

## SUMMARY

We considered both quantitative and qualitative measures for all widgets and first eliminated the widgets with the lowest score and negative feedback that left the

remaining widgets: Circular advanced, basic three-way, and ternary standard. The remaining widget designs present three unique design categories (circular sliders, ternary triangular sliders, and parallel standard basic sliders) with distinctive strengths and weaknesses. It is not feasible to clearly declare a winning solution, nor is it possible to combine them into one final prototype.

In summary, we learned that our participants spent more time moving and adjusting within the triangular space. Most of the participants correlated the spacial location to the preference setting rather than purely focusing on the numerical numbers displayed on the axes. Some participants specifically mentioned that the Ternary widget provides more freedom of choice compared to the other widgets, and is good for exploration with only vague intuition. Therefore, it might be better suited for exploration.

However, when using the circular widgets, the basic three-way widget, and the propeller widget, most participants focused on the numbers displayed on the widget. Hence, these might be better suited for specifying a precise preference. However, circular sliders give a better (visual) sense of proportion. As for the basic three-way widget and propeller widget, they both have a more straightforward movement, thus setting the criteria more directly and generally faster than most of the others. From all prototypes, we learned that participants prefer a simpler design with fewer interactions and visual elements.

## CONCLUSION

This study improves the understanding of human interaction with decision-making tools under uncertainty for less professional audiences. Although the results of the study are not showing a clear winner, we were able to narrow the design space to three prototypes: the advanced circular widget, the basic three-way widget, and the altitude widget. All three were intuitive to use. The altitude widget is best for exploration with vague intuition, the advanced circular widget is best for preference comparison, and the basic three-way widget is best for direct preference setting. It was interesting to see each individual perform and interact with different widgets. The positive feedback and discussions further suggest a need for more specific 3CDM widgets. More than half of the participants reported that they would see themselves using one or more versions of this widget for daily uncertainty problems, such as dating, selecting cars, or meal preparation.

## ACKNOWLEDGMENTS

We thank the participants of our studies for their time and contribution.

## REFERENCES

1. S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit, "Lineup: Visual analysis of multi-attribute rankings," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2277–2286, Dec. 2013.
2. S. Pajer, M. Streit, T. Torsney-Weir, F. Spechtenhauser, T. Möller, and H. Piringer, "Weightlifter: Visual weight space exploration for multi-criteria decision making," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 611–620, Jan. 2017.
3. J. W. Payne, J. W. Payne, J. R. Bettman, and E. J. Johnson, *The Adaptive Decision Maker*. New York, NY, USA: Cambridge Univ. Press, 1993.
4. E. Cettolin and A. Riedl, "Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization," *J. Econ. Theory*, vol. 181, pp. 547–585, 2019.
5. M. M. Köksalan, J. Wallenius, and S. Zionts, *Multiple Criteria Decision Making: From Early History to the 21st Century*. Singapore: World Scientific, 2011.
6. M. Sedlmair, C. Heinzl, S. Bruckner, H. Piringer, and T. Möller, "Visual parameter space analysis: A conceptual framework," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2161–2170, Dec. 2014.
7. J. Hakanen, K. Miettinen, and K. Matković, "Task-based visual analytics for interactive multiobjective optimization," *J. Oper. Res. Soc.*, vol. 72, no. 9, pp. 2073–2090, 2021.
8. Y. Laurillau et al., "The top-slider for multi-criteria decision making by non-specialists," in *Proc. 10th Nordic Conf. Hum.-Comput. Interact.*, 2018, pp. 642–653.
9. C. Schumann et al., "Interactive multi-criteria planning for radiofrequency ablation," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, pp. 879–889, 2015.
10. M. Monz, K.-H. Küfer, T. R. Bortfeld, and C. Thieke, "Pareto navigation—algorithmic foundation of interactive multi-criteria IMRT planning," *Phys. Med. Biol.*, vol. 53, no. 4, 2008, Art. no. 985.
11. M. R. Chakhchoukh, N. Boukhelifa, and A. Bezerianos, "Understanding how in-visualization provenance can support trade-off analysis," *IEEE Trans. Vis. Comput. Graphics*, vol. 29, no. 9, pp. 3758–3774, Sep. 2023.
12. P. Jankowski, N. Andrienko, and G. Andrienko, "Map-centred exploratory approach to multiple criteria spatial decision making," *Int. J. Geographical Inf. Sci.*, vol. 15, no. 2, pp. 101–127, 2001.
13. R. J. Howarth, "Sources for a history of the ternary diagram," *Brit. J. Hist. Sci.*, vol. 29, no. 3, pp. 337–356, 1996.
14. T. Munzner, *Visualization Analysis and Design*. Boca Raton, FL, USA: CRC Press, 2014.
15. M. Raubal and C. Rinner, "Multi-criteria decision analysis for location based services," *Geoinformatics*, vol. 36, no. 4, pp. 18–26, 2004.
16. P. Metcalfe and R. Metcalfe, "Design and technology: Diploma course, 2011. [Online]. Available: <https://books.google.at/books?id=IL9cMwEACAAJ>
17. R. Macefield, "How to specify the participant group size for usability studies: A practitioner's guide," *J. Usability Stud.*, vol. 5, no. 1, pp. 34–45, 2009.
18. D. Dubois, M. Grabisch, F. Modave, and H. Prade, "Relating decision under uncertainty and MCDM models," in *Proc. AAAI Fall Symp. Front. Soft Comput. Decis. Syst.*, 1997, pp. 6–15.
19. A. Kangas, J. Kangas, and M. Kurttila, "Uncertainty in multi-criteria decision making," *Decis. Support Forest Manage.*, vol. 16, pp. 55–97, 2008.
20. D. Skau and R. Kosara, "Arcs, angles, or areas: Individual data encodings in pie and donut charts," *Comput. Graphics Forum*, vol. 35, no. 3, 2016, pp. 121–130.

**JIN YU** is a master's student in computer science with the Faculty of Computer Science, University of Vienna, 1090, Vienna, Austria. He is the corresponding author of this article. Contact him at [a11717460@unet.univie.ac.at](mailto:a11717460@unet.univie.ac.at).

**LAURA KOESTEN** is a postdoctoral researcher with the research group for Visualization and Data Analysis, University of Vienna, 1090, Vienna, Austria. Her research interest centers around aspects of human data interaction, such as sense-making with data and visualizations, data discovery and reuse, as well as ethical and collaborative aspects of data-centric work. She received her Ph.D. degree in computer science from the University of Southampton and the Open Data Institute, U.K. Contact her at [laura.koesten@univie.ac.at](mailto:laura.koesten@univie.ac.at).

**TORSTEN MÖLLER** is the head of the Visualization and Data Analysis Research Group, Faculty of Computer Science, University of Vienna, 1090, Vienna, Austria. His research interest includes visualization, computer graphics, and data science. He received his Ph.D. degree in computer and information science from the Ohio State University, USA. He is a member of IEEE. Contact him at [torsten.moeller@univie.ac.at](mailto:torsten.moeller@univie.ac.at).