



HAL
open science

The TOP-Slider for Multi-criteria Decision Making by Non-Specialists

Yann Laurillau, van Bao Nguyen, Joëlle Coutaz, Gaëlle Calvary, Nadine
Mandran, Fatoumata Camara, Raffaella Balzarini

► **To cite this version:**

Yann Laurillau, van Bao Nguyen, Joëlle Coutaz, Gaëlle Calvary, Nadine Mandran, et al.. The TOP-Slider for Multi-criteria Decision Making by Non-Specialists. Proceedings of the 10th Nordic Conference on Human-Computer Interaction (NordiCHI '18), Sep 2018, Oslo, Norway. pp.642-653, 10.1145/3240167.3240185 . hal-03223755

HAL Id: hal-03223755

<https://hal.science/hal-03223755>

Submitted on 11 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The TOP-Slider for Multi-criteria Decision Making by Non-Specialists

Yann Laurillau, Van-Bao Nguyen, Joëlle Coutaz, Gaëlle Calvary, Raffaella Balzarini
Nadine Mandran, Fatoumata Camara

Université Grenoble Alpes, CNRS, LIG, France
name.surname@univ-grenoble-alpes.fr

INRIA, LIG, France
name.surname@inria.fr

ABSTRACT

The TOP-Slider revisits the classical slider widget to help non-experts to find optimal compromises between tightly coupled, possibly conflicting, criteria. The TOP-Slider is designed to make the tradeoff between criteria both observable and intelligible while hiding the complexity of an underlying optimization algorithm. As a case study, a TOP-Slider has been assessed in terms of perceived affordance, intelligibility and incentive, when used in conjunction with an e-coach energy management system to interactively explore Pareto optimal solutions and to choose the best compromise between comfort and financial cost. A preliminary semi-controlled empirical study with 16 participants using 3 interdependent criteria shows that the TOP-Slider can serve as a sound basis to support non-expert users in making informed compromises.

Author Keywords

Tightly coupled sliders, multidimensional visualization, multi-criteria decision making, Pareto optimality, energy management, e-coach, optimization problem.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

INTRODUCTION

Multi-criteria decision making consists of searching for optimal compromises between multiple, possibly conflicting criteria. Decision makers must compare several alternatives to select the most preferred solution. Methods, algorithms, and interactive tools have been developed in the field of engineering design to support domain experts in making informed decisions [14]. By contrast, we target layman users, as making compromises is a rather frequent and difficult task in everyday life. For example, to obtain a bank loan, one must deal with multiple criteria to obtain the best compromise that minimizes the cost of the loan and its

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

NordiCHI'18, September 29-October 3, 2018, Oslo, Norway

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-6437-9/18/09...\$15.00

<https://doi.org/10.1145/3240167.3240185>

duration while maximizing the amount of the loan.

Targeting non-domain experts raises the following research questions: How to present the alternatives of a solution space in an effective way? How to support the exploration of the solution space? How to hide the complexity of the optimization problem while facilitating the understanding of the mutual influence between the criteria? How far the underlying mathematical model should be made explicit and observable without losing precision and rigorosity? To address these research questions, we propose the Trade-Off-Pareto-Slider widget or, in short, the TOP-Slider¹.

The design of the TOP-Slider is grounded on the “conceptual integration” or “blending” cognitive theory [21] as we associate a familiar widget, the slider, with an optimization problem. In addition, sliders being widely used for interactive visualization, it is reasonable to hypothesize that the TOP-Slider supports the incentive to explore the solution space through a “what-if” process.

Solution spaces covered by the TOP-Slider are Pareto optimal sets, as they are widely used for multi-criteria optimization problems. A Pareto optimal set is the set of Pareto optimal solutions for a given optimization problem. This set is called Pareto front, or Pareto frontier. A solution is considered Pareto optimal if no other solution exists that is better for a criterion without being worse for some other criteria [3, 26]. We also say that it is not “dominated” by any other solution. Without the specification of an additional subjective preference, all Pareto optimal solutions are mathematically equally good. The TOP-Slider is intended to support layman users in finding and selecting the solution that satisfies their preferences.

In the following, we present previous work on sliders developed for supporting multi-criteria tasks, then we justify and describe the design of the TOP-Slider, followed by a detailed presentation of the method used to assess our initial design as well as the results of the experiment. Before concluding with a summary of our contribution, we discuss these results in light of our requirements, the limitations of the experiment, and related work.

SLIDERS AND MULTI-CRITERIA TASKS

Since its introduction in graphical user interfaces (GUI), the slider widget has been extended in various ways to support

¹<https://pareto-sliders.firebaseio.com/> (improved version)

multi-criteria tasks. A multi-criteria task is a task that involves the simultaneous consideration for multiple criteria to reach a particular goal. These tasks include data filtering and exploration of multidimensional datasets, and choosing a preferred solution among a fixed set of alternatives as in multi-criteria decision making.

Sliders for Data Filtering and Exploration

The classical GUI slider has been extended in a number of ways to facilitate the exploration of multidimensional datasets, using additional cursors, combining the slider with another visualization technique (such as a histogram), exploiting color-coding and brushing, or binding several sliders in a tightly manner.

In their seminal work on visual information seeking (VIS), Ahlberg and Shneiderman [2] proposed the QuerySlider widget to specify a sub-interval of values by the way of two cursors, as well as the AlphaSlider [1] to specify an index in a list of alphanumeric data (Figure 1a). They are connected such that one change in a slider may affect the others. As shown in Figure 1b, the Dynamic Query Slider [11] extends QuerySlider by displaying a histogram in the sliding area to express the data distribution for the associated attribute in that interval.

Eick [5] proposes four versions of the slider widget to support data visualization. For instance, the sliding area may embed a color scale, show data distribution, or allow turning on or off regions of interest based on colored (on) or grey (off) areas.

Scented widgets augment the familiar GUI widgets (e.g. check buttons, lists, hierarchical lists) with visual information scent cues as “*appropriate visual navigation cues can support users by guiding and refining their exploration*” [23]. In particular, the slider is augmented with a bar chart that encodes two variables with visual cues (i.e. height of a bar and color opacity). The authors have identified seven scent encodings and associated guidelines.

Greis et al. [6] (Figure 1c) go one step further with the

investigation of entering uncertain data: the slider bar displays a color gradient to visualize an underlying probability distribution function that reflects the level of uncertainty.

As shown in Figure 1e, SketchSliders [20] are flexible sliders, sketched by the user, aimed at exploring multidimensional datasets on large wall displays. The SketchSliders are flexible enough to support multiple branches for a more precise exploration.

Multiple sliders may work in a tightly coupled manner where a change in one slider may have an impact on the others. Their interdependency may, or may not, be expressed explicitly. In the FilmFinder application [2], setting a value on one slider impacts the others but the relationships are not explicitly visible. Influence Explorer on the other hand [22] uses color-coding (i.e. “color linking”) to make the relationships observable, and when one selection has been made on a slider, the histograms of the other sliders are updated automatically. Based on crosstets, Perin et al. [17] promote a new interaction technique to manipulate multiple sliders simultaneously by crossing gestures: a single crossing gesture suffices to select and modify all the sliders simultaneously.

Sliders for Multi-criteria UIs

The slider metaphor has also served as a basis for the design of multi-criteria user interfaces (UIs), most notably EZChooser, Sliding Rods, and Influence Explorer. Some of them, such as EZChooser are intended for the general audience while others such as Influence Explorer have been designed for domain experts.

EZChooser [25] allows users to select an item (for example, a car) from a database, based on the specification of multiple attributes to form a query. Horizontal bargrams are used in parallel to represent these attributes, one bargram per attribute. A bargram is similar to a histogram slider. As shown in Figure 1f, within a bargram, bars are represented by a set of contiguous and horizontal clickable buttons whose size is proportional to the associated bar. Clicking a

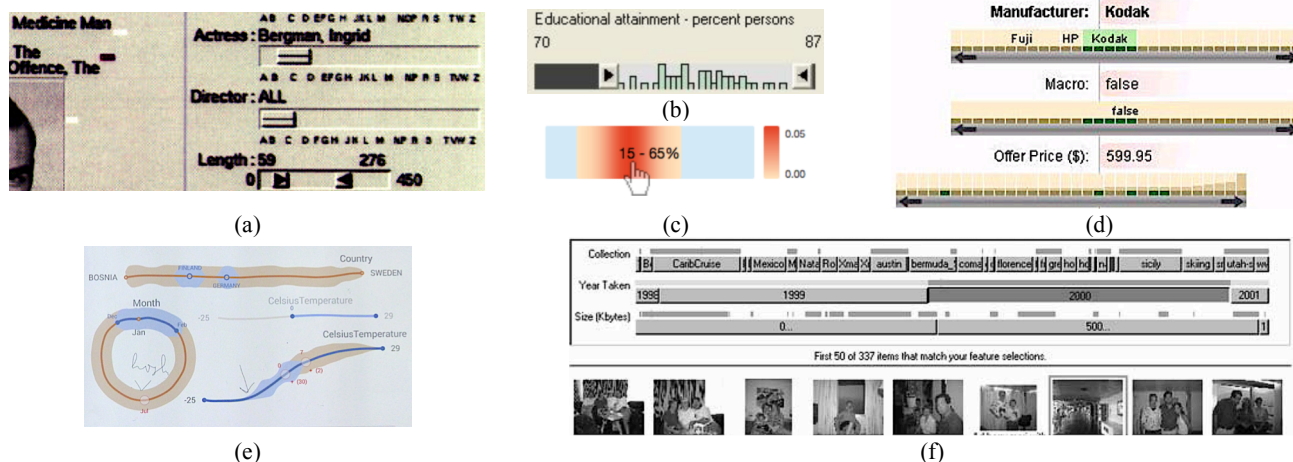


Figure 1. (a) FilmFinder [2] (b) Dynamic Query Slider widget [11], (c) Uncertainty slider [6], (d) Sliding Rods widget [10], (e) SketchSliders [20], (f) EZChooser [25].

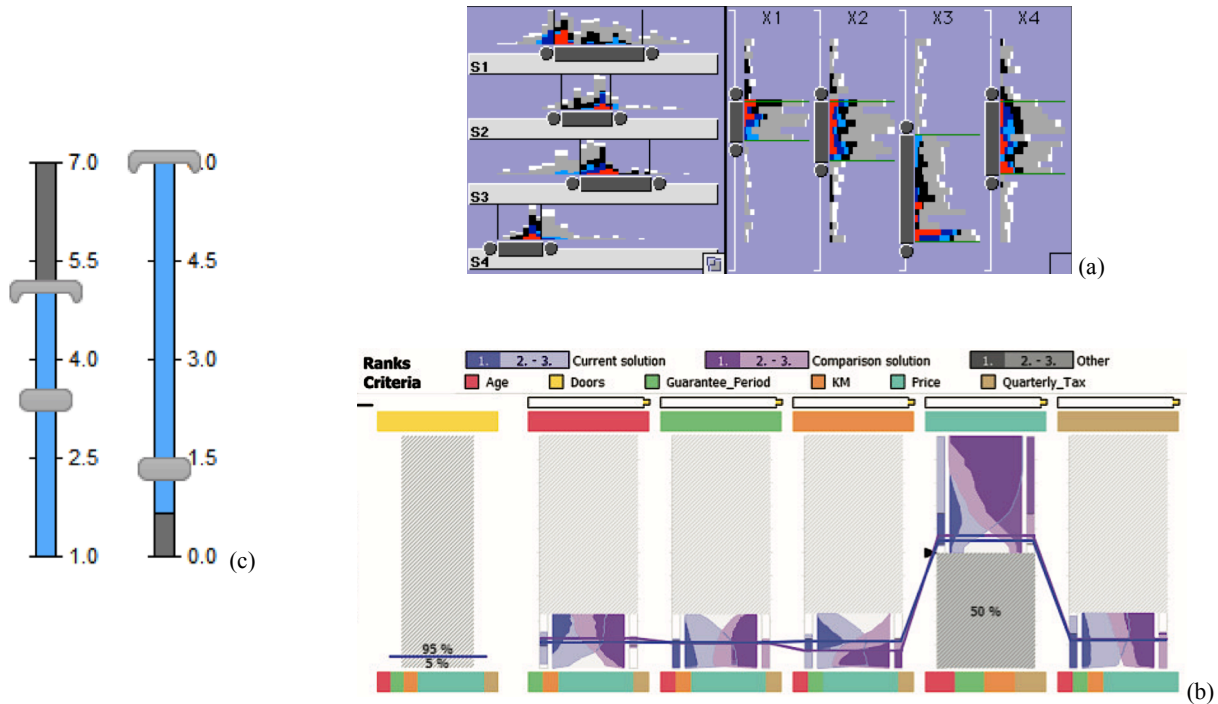


Figure 2. (a) Influence Explorer [22], (b) WeightLifter [16], (c) Pareto Sliders for surgery [18].

bar allows users to restrict their selection. Similar to EZChooser, SGViewer [19] improves filter coordination among bargram widgets.

To help users to “understand global information and their relationships”, Lanning et al. [10] propose the Sliding Rods for multidimensional data exploration and querying. In their MultiNav tool, each Sliding Rods widget, organized in parallel rows, is associated with an attribute of the data space. As shown in Figure 1d, the sliders are horizontally moveable in order to keep the focused item in the center of the screen.

Influence Explorer [22] aims at exploring relationships between parameters within a multidimensional dataset. As shown in Figure 2a, users can specify attribute ranges using sliders, coupled to a histogram. Links between the attributes rely on a “color inking” representation embedded in the histogram where red, for example, indicates a compatible range of values between parameters. Influence Explorer has been used in industrial contexts by domain experts.

Sliders for Multi-criteria optimization UIs

Multi-criteria optimization UIs, a subclass of multi-criteria UIs, is an important field of research concerned with multi-criteria decision making [23] as well as with visualization tools. In multi-criteria optimization, there is rarely a unique optimal solution and the choice of a solution depends on the user’s preferences [16]. The tools dedicated to multi-criteria optimization typically rely on bar charts, scatter plot matrices, petal diagrams, tabular visualizations, glyphs such as the Chernoff’s faces, parallel coordinates, spider diagrams, and multiple ways of combining these basic techniques [14].

The Parallel Coordinates technique is a projection of N-dimensional datasets onto 2-dimensional datasets [8]. Each attribute is represented by a parallel axis. A related approach is the Self Organizing Maps technique [9] where a N-dimensional dataset is mapped onto a lower dimensional dataset (e.g. color hexagonal neurons). In the spider diagram, each axis represents a criterion and the solution space is visualized as a spider web. This technique facilitates the comparison between two solutions as it highlights the strength and weaknesses for each criterion.

Rendering the solution space and the criteria of an optimization problem is necessary, but not sufficient for users to make informed decisions. In some situations, the decision maker needs to specify the relative importance of criteria so that the system can rank the solutions according to these preferences. This capacity is provided by WeightLifter [16] where users are able to set weights between the criteria. As shown in Figure 2b, similar to parallel coordinates, a column represents a criterion, one color per criterion (e.g. light green for car price). It represents the weight space constraints.

Closest to our work, the Pareto Slider was designed to support the exploration of Pareto optimal solutions for planning medical radiofrequency ablation [18]. As shown in Figure 2c, each criterion is represented by a slider that includes two types of cursors: the usual cursor and a restrictor. The lower bound of a slider is the best value for the criterion whereas the upper bound denotes the worst value. By moving the restrictor, the user can limit the range of acceptable values for the criterion. A “blue color-coding” denotes the range of acceptable values whereas eliminated

ranges are colored in grey. Consequently, a cursor cannot be moved outside its blue range. In addition, the restrictor impacts the other sliders as it may reduce their blue range of acceptable values. Then, moving the cursor of one slider moves the cursor of the other sliders automatically so that the new position of the cursors corresponds to a Pareto optimal solution. The strategy used for choosing the new Pareto optimal solution among the possible ones is decided by the designer, not by the end-user.

All of these interactive visualization techniques are intended for domain-experts in technical application domains (e.g. engineering, medicine). As a result, they are found to be difficult to understand, especially by novice users [16].

THE TOP-SLIDER

As a case study, the TOP-Slider is used in conjunction with an e-coach energy management system to interactively explore the Pareto front computed by the e-coach and to select a satisfying compromise between thermal comfort, air quality, and financial cost. Based on the preferred solution selected by the end-user, the e-coach is then able to suggest optimal actions [7], such as opening/closing doors and windows, to optimize energy consumption. Before describing the TOP-Slider *per se*, we need to illustrate what “exploring a Pareto front” means.

Exploring a Pareto Front

As a simple illustration, the Pareto front of Figure 3 shows the set of optimal solutions for two conflicting criteria, thermal comfort and financial cost. The front delimits the frontier between the set of feasible but not optimal solutions (the grey zone in Figure 3) from the set of inaccessible solutions (the red zone). The yellow zone corresponds to unwanted ranges of values for the criteria. For example, temperature is not comfortable when below 16°C. The shape of a Pareto front is generally similar to a convex or concave monotonic function [12].

As there is no unique solution to an optimization problem, it is up to the user to explore the Pareto front to find the preferred tradeoff between the criteria. For instance, from the non-optimal solution A of Figure 3, several optimal solutions may be reached: B1 as a significant reduced

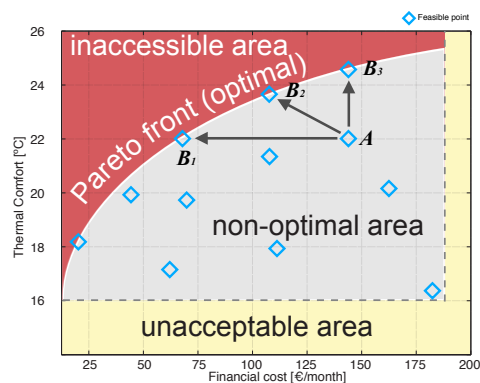


Figure 3. Example of a Pareto for two objective functions: cost and comfort.

financial cost at constant thermal comfort, B3 as a significant increase in thermal comfort at constant financial cost, or B2 as a small reduction of the financial cost for a small increase of thermal comfort.

DESIGN OF THE TOP-SLIDER

Like the Pareto Slider designed for planning surgery described above, the TOP-Slider is composed of a set of parallel sliders, each of which is associated to a criterion of the optimization problem (see Figure 4 as an illustration with three criteria: financial cost, thermal comfort, and air quality). As with standard sliders, the position of the cursor on a slider denotes the currently selected value for that criterion.

The sliders are augmented with tightly coupled visual features to support decision making through a “what-if” process. These features represent (1) the interdependence between the criteria, (2) the impact of the modification of the value of one criterion on the others, (3) and for each criterion, the current ranges within which its values may fall, each range resulting from the Pareto front calculated for the current optimization problem.

Color encoding is used to discriminate the three sorts of “Pareto ranges”: white denotes a range of values situated on the Pareto front; grey, a range of feasible but non-optimal values, and red for ranges of unfeasible values. For example, in Figure 4, the choice for financial cost (66€) and thermal comfort (22.4°C) is optimal whereas air quality is impossible to satisfy (620 ppm).

Inspired by Parallel Coordinates technique [8], interdependence between the criteria is made explicit with pairs of dashed white lines where a pair pops up when a cursor is selected and links this cursor to the boundaries of the optimal range of the other sliders. For example, in Figure 4, the cursor of the financial cost is currently selected: two pairs of lines have appeared to show the impact of the current choice on the range of the optimal values for the other two criteria. In addition, as the cursor of a slider is moved, the ranges of the other criteria are updated according to the underlying Pareto model. Tight coupling between cursor movements and Pareto ranges makes explicit the impact of a selection of one criterion

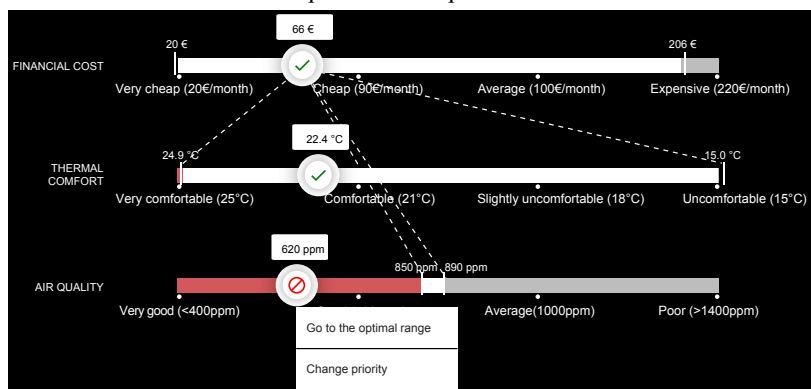


Figure 4. TOP-Slider (initial design) with three criteria: financial cost, thermal comfort, and air quality.

value on the remaining criteria. In order to allow users to explore tradeoffs that are not necessarily Pareto optimal, moving one cursor does not move the other cursors.

Three icons are used for additional recommendations: a warning sign (\triangle) to recommend the user to choose an optimal value; a forbidden sign (\emptyset) to indicate that the current selected value is impossible to reach; a check mark (\checkmark) when the cursor lies within an optimal range. For the two first cases, a pop-up menu is accessible to automatically move the cursor at the closest value of the optimal range as well as to change the priority between the criteria, that is, their subjective relative importance.

EXPERIMENTAL EVALUATION

In their exploratory study of three elementary multidimensional visualizations for their ability to support decisions (Parallel Coordinates, Scatterplot Matrix, and Tabular Visualization), Dimira et al. observe that “evaluating visualizations for their ability to support decisions is difficult, and there is a lack of methodological guidance in the information visualization literature on how to do so” [4]. The problem is two-fold: (1) Objective measures are not enough to capture the quality of a decision, given that “finding a good compromise” is by essence, subjective. Subjective measures such as self-reported satisfaction are useful, but may be unreliable. (2) There is a lack of clear references for identifying an appropriate baseline for comparative assessment. As a result, we have proceeded according to a multistep, iterative and prospective approach.

As a first step of the development process of the TOP-Slider, we have performed a preliminary intra-evaluation. In this section, we describe the design of this early assessment involving 16 users to verify that the graphical and interactional design of the TOP-Slider satisfies the following requirements: (R1) affords interactive exploration; (R2) makes intelligible the interdependence of the criteria; (R3) motivates users to find a suitable compromise between the criteria.

Method

Participants and Apparatus

We recruited 16 subjects by email and word of mouth (12 men and 4 women). Ages ranged from 17 to 71 of which 6 over 40, with an average of ~ 38 . 10 were computer scientists (1) and students (9), and 6 were family members (of which 4 retired healthy persons), but none of them was involved in the project. All of them were confident in using computers and tablets. The subjects signed consent, and were not paid for their participation in the experiment.

Most of the participants were familiar with digital sliders and/or with physical sliders as in cars to control air stream. No participant has ever used tightly coupled sliders. 7 participants have concerns with energy consumption and financial cost and 3 of the retired participants use a

technical solution to manage their own consumption at home (e.g. programming heating periods).

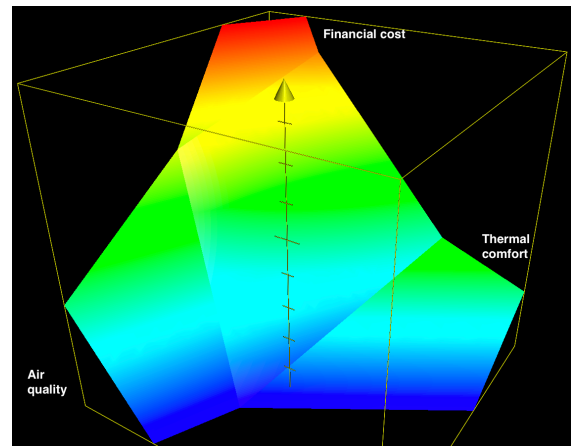


Figure 5. The 3D Pareto model used in the experiment (the color gradient highlights the z-axis).

The 16 participants performed their tasks with an iPad 2² (9,7") running the user interface of the TOP-Slider shown in Figure 4. The TOP-Slider is implemented in JavaScript as the client of a web application using the Polymer programming toolkit. Users' actions are time-stamped and logged on a secure server. The TOP-Slider is model agnostic. In this experiment, the Pareto front is computed based on a piecewise multivariate function (hyperplanes) represented in Figure 5. For the e-coach application, it is implemented to support the exploration of a mathematical model of a Pareto front obtained by interpolating a set of measures that characterize the actual physical conditions of the habitat.

Task domain and Scenario-based Tasks

Energy consumption, which has become a major societal challenge both at the political and individual levels, was selected as the task domain. For context of use, we chose a smart home augmented with an e-coach that provides end-users with suggestions for actions (e.g. opening/closing doors and windows) based on their preferred compromise between thermal comfort, air quality and financial cost.

Using the TOP-Slider shown in Figure 4, all participants were asked to perform the same task – to find a preferred compromise for three scenarios of increasing difficulty: easy, medium, difficult. Each scenario was designed to cover an area of the Pareto Front.

Procedure and Data Collection

The experiment was organized into four successive phases to collect both quantitative and qualitative data: (1) Introduction, (2) training, (3) task execution, (4) semi-structured interview and SUS-based questionnaire (System Usability Scale).

² Our project partners defined tablets as the target device.

Introduction. The experimenter presented the 4 steps of the experiment as well as the context for the task domain – a smart home augmented with an e-coach, and a tablet to interact with the system. No details were provided about the user interface, nor about the underlying Pareto model. Participants were made aware that the session was audio-recorded and that interaction with the device was captured.

Training (phase #2). This phase was meant to get users familiar with the TOP-Slider through a free and decontextualized exploration. No objective was specified. Participants were instructed to think aloud as they were interacting with the device. The goal of this phase was to elicit the perceived affordance of the TOP-Slider and make sure that the participants had discovered the visual elements of the TOP-Slider. No time limit was imposed. The participants were told to start phase #3 when they felt confident.

Scenario-based task execution (phase #3). The goal of this phase was to observe how manipulation and visual UI elements helped users to find a suitable and satisfying combination for each interdependent criterion. The 16 participants were instructed to think aloud as they were executing the task “express your preference” for the three following scenarios:

Easy: The requirements of the scenario map to criteria values that all belong to an optimal range (optimal compromise) “*It is winter. You are financially well off but you are not wasteful. At home, you prefer to live with a reasonable temperature (norms recommend 21°C), and a standard air quality suits you (European norms recommend a level of particulate matter between 400 and 600 particulates per m³ of air). How do you communicate your preferences to the system?*”

Medium: Although participants are implicitly asked to find a combination on the Pareto front, the requirements target solutions in the non-optimal but feasible area. “*You are concerned with air quality. As this is important for you, you open windows and doors every morning to renew the air inside your home, even in winter. You also like feeling very comfortable with regard to temperature (you do not care if the temperature is above the norm – 21°C). However, you control your energy consumption to avoid a heavy bill. How can you express these preferences to the system?*” In this scenario, financial cost is constraining (i.e. “to avoid a heavy bill”). Participants have to decide how to lower their requirements in terms of thermal comfort and/or of air quality.

Difficult. In this case, the requirements target solutions in the non-feasible area. “*In your daily living, you save money as much as possible. It is winter. You are back home after a cold day. You feel sick with flu. You want to be warm, with very good air quality because of the flu.*” This scenario, which stipulates “*best comfort and air quality at lowest*

financial cost”, requires the participants to find a compromise between the three criteria.

Semi-structured interview and SUS (phase #4). The interview was guided by a questionnaire covering the following categories:

- **Affordance.** (1) Do you see the visual elements while moving a cursor? (2) What do you understand about the animation of visual changes?
- **Intelligibility.** (1) How did you proceed to make a choice? (2) How did you manage to find a suitable/appropriate choice? (3) Does the ability to freely move the cursors motivate you to test different combinations? And to find the best combination for you?
- **Utility.** Generally speaking, what do you think of the use of a slider to manage your own energy consumption?
- **Energy management.** How do you manage energy consumption at home?
- **Use of sliders in general.** Have you ever used sliders in your daily life? Can you give examples?

To conclude this phase and the experiment, we submitted a non-modified standard SUS questionnaire with a discrete scale ranging from 1 (strongly disagree) to 5 (strongly agree).

RESULTS

Raw Data

All participants, but one, filled in the SUS questionnaire. A total of ~336 minutes of audio recording has been transcribed into quotes annotated according to the following coding scheme: affordance and manipulation, intelligibility, utility for energy management, and experiment-related issues (such as understanding the scenarios and/or the context of use). For each topic and participant, we have counted positive and negative quotes. The result is the extraction of 291 quotes of which 5 are out of context and thus discarded; 42 quotes were related to energy management and the use of physical and/or digital sliders in daily life. Table 1 shows the resulting distribution of the quotes. We applied a binomial test on the positive quotes over the number of participants/category with a 95% confidence interval.

The average number of actions per scenario is the following: Easy: 8.33 actions; Medium: 16.93 actions; Difficult: 20.06 actions. The average interaction duration per scenario is: Easy: 2 mn 1 sec; Medium: 2 mn 56 sec; Difficult: 2 mn 21 sec. We have no clear explanation why the mean duration for completing the difficult scenario is less than that for the Medium Scenario – although they performed more actions for the difficult task. A possible hypothesis is that subjects abandoned the difficult task more quickly. Testing in real world conditions with an e-coach running would provide more insights.

Overall Feedback

Thirteen participants referred to the utility of the TOP-Slider (13 quotes) and the majority of them (11/13, $p=0.02$) clearly found it usable: “It is clear, intelligible” (P2); “It is very good. I like very much having lines between the slider widgets with the slices. And to see what is possible” (P4); “It is intuitive. It is very well designed although it took me some time to understand” (P7); “It allows me to make a choice depending on the different constituents” (P10). This is consistent with the results of the SUS questionnaires with an average score of 77.2/100. In detail, except for Question #1 (“I think that I would like to use this system frequently”), the score is very good (between 80/100 and 90/100) indicating that the TOP-Slider is easy to use and to learn. As for Question #1, we hypothesize that the result is rather a consequence of the application domain than that of a major usability issue, as in real life, 3 participants asserted that they were not concerned with energy consumption at home.

Category	Total # of quotes	# of part.	# of positive quotes	# of negative	p-value (95%)
Useful	13	13	11	2	0.02
Satisfaction	24	13	11	2	0.02
Easy to use	14	11	9	2	<0.01
Links	27	14	14	0	<0.01
Red zone	47	14	14	0	<0.01
White zone		14	14	0	<0.01
Grey zone		9	7	2	0.18
Check icon	32	13	13	0	<0.01
Forbid. icon		8	9	0	<0.01
Warning icon		13	10	3	0.09
Coupling	61	16	15	1	<0.01
Compromise	42	14	11	3	0.05
Priority	37	16	16	0	<0.01
Motivation	17	15	13	2	<0.01

Table 1. Categories of quotes and their distribution. For the quotes that cover multiple topics, one quote for each topic (one per row) is counted. # of part. column represents the number of participants having spoken about the category. Among this number, we have counted the number of positive (# of positive) and negative (# of negative) quotes related to satisfaction and/or intelligibility depending on the category. Only one positive or negative quote is counted in case of multiple quotes for the same participant and topic. The p-value column is the result of a binomial test with a 95% confidence interval (# of positive vs # of negative quotes).

With regard to finding compromises, thirteen participants made comments about their satisfaction (24 quotes). Most of the participants (11/13, $p=0.02$) were satisfied: “Yes, that’s a good one [i.e. combination]” (P5); “Here, I’m good [...]. It is the best compromise” (P8); “If I play [with the

cursors] [...], yes, it is a good one” (P13). Two participants (P6, P7) were not fully satisfied: “There is no ideal solution, it is embarrassing” (P6); “I had only one choice [...] but it is nice.” (P7).

Eleven participants mentioned the usability of the TOP-Slider (14 quotes). Nine participants (9/11, $p<0.01$) found the widget easy to use: “It indicates me how to modify” (P2); “It is fluid” (P3); “We can see very well [how it works]” (P10); “It is clear, intuitive” (P12); “I quickly understood [...]” (P14). Two participants (P6, P11) had difficulties: “The intervals worried me at the beginning” (P6); “The white interval is very narrow” (P11). For the latter, this is due to the finger-based interaction already mentioned above.

Perceived affordance

Pairs of lines

Fourteen participants referred to the dashed lines and their meaning (27 quotes). All participants (14/14, $p<0.01$) understood that the slider widgets are tightly coupled: “Visually, it is nice, there are funnels and the sliders show in which area we can go and how it influences” (P3); “we can see easily the links” (P6); “The lines are very clear. When you move a cursor, I see the lines moving. It is not visually disturbing” (P7); “There are multiple dashed lines and lines between the circles [cursors], we clearly see they are linked” (P10); “I saw lines of correlation” (P12).

Among these participants, twelve participants noticed the tiny vertical lines coupled to, and terminating the dashed lines. They (12/12, $p<0.01$) understood that they represent and delineate an interval of recommended values: “It moves simultaneously, and we see the boundaries at the same time” (P1); “There are lines that delimit the intervals depending on the [financial] cost” (P2); “There is an area of possible values with boundaries that are displayed [...]” (P5); “You have limits, boundaries, it helps to optimize” (P14).

Color-coding of the Pareto ranges

Fourteen participants spoke about the colored intervals and their meaning (47 quotes). All the participants (14/14, $p<0.01$) understood that a red interval indicates an incompatible range of values: “When we are in an impossible zone [...]” (P3); “I saw with the red zones that I could not optimize [...]” (P5); “Red, we can’t go there, it is impossible” (P10). Similarly, all participants found white intervals obvious: “The white zone is what is possible” (P4); “The white areas, it is when the three variables are compatible” (P8); “White is acceptable” (P10). Only nine participants referred to grey intervals and seven did not pay attention to them. Seven participants out of nine understood its meaning: “The grey zone is an overcharge compared to what we can get” (P4); “Here [grey interval], I think this is a useless overcharge” (P8); “The grey part on the right, it is the value we should not go beyond” (P9); “The grey [interval], after the white [interval], it is not acceptable”

(P10). Two participants (P7, P16) had difficulties with the grey intervals: “*I have not well understood the grey zone*” (P7); “*The red, it is not good, the two others are better*” (P16).

Icons

Thirteen users mentioned icons and their meaning (32 quotes). All the participants (13/13, $p < 0.01$) found the ‘check’ symbol obvious, clearly indicating that the cursor was in the optimal interval.

Some participants did not pay attention to the ‘forbidden’ symbol. As this symbol is complementary to the red zone, it did not harm the interaction significantly. However, one participant (P3) did not understand its meaning.

Thirteen participants paid attention to the ‘warning’ symbol. Ten participants (10/13, $p = 0.09$) understood that this symbol indicates a non-recommended area: “*The orange icon tells me ‘warning’*” (P2); “*The /\ tells me that I’m over the limit*” (P7); “*/\ if I’m not in a valid area*” (P11); “*/\? It means warning?*” (P13). However, three participants (P3, P8, P16) had troubles in identifying its meaning: “*The orange thing [/\ symbol], I don’t know what it is*” (P8); “*There is the /\ symbol. I do not understand, I don’t care*” (P16).

Intelligibility

Interdependent sliders

All the participants spoke about their comprehension of the coupled sliders (61 quotes). A wide majority of participants (15/16, $p < 0.01$) inferred without obvious difficulty that the criteria, each represented by one slider widget, are interdependent and have mutual influence: “*A good thing, when I move one, the others react*” (P1); “*Each factor influences the others [...]*” (P3); “*These are coupled variables*” (P5); “*I need to make a choice to make it compatible between them*” (P8); “*If I lower it [air quality], the financial cost will increase*” (P9); “*If I raise it [thermal comfort], the financial cost will augment*” (P14). Although P16 understood the coupling (“*Temperature influences air quality and how much it will cost me*”), this participant had difficulties in understanding how criteria are mutually influenced: “*It doesn’t work [...] I have finished [scenario #1] but I have not understood why*”. This may explain the relatively low SUS score for participant P16.

Need for a compromise, finding a preferred compromise

Fourteen participants referred to their need for finding a compromise and/or a suitable/optimal choice (42 quotes). The majority of the participants (11/14, $p = 0.05$) inferred without obvious difficulties that:

(1) They had to find a compromise between the criteria to stay on the Pareto front (optimality): “*[we can] sacrifice air quality if we don’t have the budget for*” (P3); “*I choose to make a compromise among the three [criteria]*” (P6); “*If I set [air quality] to 400 ppm, I have to pay more. I will set [air quality] to 500 [ppm] with the lowest cost*” (P7); “*I can*

make a sacrifice on temperature” (P11); “*I have to come to a compromise*” (P16).

(2) They could iterate to find a more appropriate combination (i.e. optimization): “*As we can see all the possibilities, it incites me to seek for different combinations*” (P4); “*I tried multiple combinations, we are free [to move the cursors] and I don’t want to miss one*” (P10); “*It incites me to search for [the best] ratio among solutions*” (P11); “*If I have the budget, how much air quality can I gain [...]?*” (P12); “*You have limits, you have boundaries, it helps to optimize*” (P14).

However, three participants (P1, P7 P15) had difficulties to find a compromise and/or to infer that they had to make a compromise: “*It is very difficult to obtain what we want*” (P1); “*It is very difficult to save a lot of money. I will keep the maximum [value] but I still have the exclamation point [/\ sign]*” (P7); “*Why can’t I have everything I want?*” (P15).

Priority-based manipulation

All the participants (16/16, $p < 0.01$) made explicit the method they used to set the cursors (37 quotes). Two strategies were considered: (1) to choose a first-class criterion (i.e. cursor) and to set the others consequently; (2) to order the criteria by priority, starting with the criteria with top priority: “*The price is the most important criteria. I depend on it primarily*” (P1, P4); “*I focused on air quality and it guided me for the others*” (P2, P14); “*I created a hierarchy of my priorities*” (P7); “*My priority was the temperature*” (P12).

Motivation for testing multiple combinations

Fifteen participants spoke about how the system motivates them to test several solutions. Most of the participants (13/15, $p < 0.01$) positively stated that they were motivated to search for a better combination of criteria. However, two participants (P1, P5) out of fifteen were not motivated: “*I stayed focused on the first criteria [financial cost]*” (P1); “*No [about searching for another solution], I have just narrowed down my limits*” (P5).

DISCUSSION

Overall, the results of this first step experiment are consistent with our requirements in terms of affordance (R1), intelligibility (R2), and incentive to explore (R3). Additionally, the participants have elaborated strategies for finding compromises, setting priorities between the criteria. We close this discussion with the description of the design improvements suggested by the experimental results.

R1: Effective Affordance

Participants felt confident with classical slider widgets as it is widely used in interactive software. A large majority of the participants appreciated the graphical design of the TOP-Slider, as well as the graphical animation (e.g. colored intervals). They found the graphical design clear without overloading the individual slider widgets. In terms of affordance, most of the added graphical elements added to

the classical slider widgets are sufficiently meaningful: dashed lines to make concrete the actual coupling between the criteria; colored intervals to make explicit the Pareto ranges. The ‘check’ and ‘forbidden’ icons were found meaningful but half of the participants did not pay attention to the ‘warning’ icon or had difficulties to understand its meaning. Correlatively, similar issues were raised with the grey intervals for the non-optimal intervals of values. In addition, the ‘go to the optimal range’ contextual menu entry was rarely used.

R2: Intelligibility

The participants discovered tightly coupled sliders for the first time. Although they found it unusual and quite disturbing during the training phase (phase #2), most of the participants inferred and understood the underlying logic/model of the TOP-Slider. In particular, they understood (1) the interdependency of the variables represented by each individual slider, (2) that each individual slider includes three types of intervals of values (optimal, non-optimal overcharge, impossible); (3) the need for making tradeoffs in order to set all the cursors in a white interval. In addition, most of the participants were satisfied with their solutions. However, two participants (P15, P16) showed real difficulties. As mentioned above, we hypothesize that this issue is a consequence of the application domain and of the lack of concern for energy consumption. As well, two participants (P1, P7) found it hard to set a compromise. For P7, based on her/his statements, we hypothesize that the difficulty is related to the ‘warning’ icon. For P1, the difficulty is not significant as P1 achieved the three scenarios very quickly with very few actions (9 actions for scenario 3 in less than a minute).

R3: Incentive to Explore

The capacity to manipulate the TOP-Slider has motivated most of the participants to improve a selected combination for either a slightly better one or a very different but more satisfactory one. In addition, the statements recorded show that these participants adopted a ‘what-if’ approach. At the opposite, three participants stayed stuck on one criterion (financial cost for P1; thermal comfort for P13 and P16). Once again, based on their statements, we hypothesize that this is due to the application domain and the personal situation: P1 is a student with limited financial resources whereas P13 and P16 are not concerned with energy consumption.

Priority-based interaction

The participants clearly adopted a priority-based strategy to find a suitable compromise. In the version used for the experiment, the TOP-Slider supports priority setting, but it does so implicitly: a criterion gets priority through the contextual menu or when the associated cursor is moved. This design solution, which was somewhat disturbing for some participants (P13), was then revised. This is discussed next.

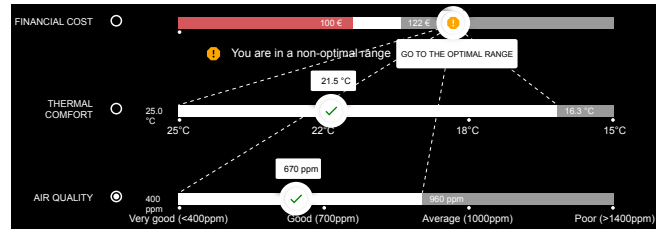


Figure 6. Improved version of the TOP-Slider.
<https://pareto-sliders.firebaseio.com/>

Design improvement

Based on the experimental results and issues raised by some participants, the TOP-Slider has been revised in the following ways, as shown in Figure 6:

(1) To improve priority-based strategies, menu-based priority setting between criteria is now replaced with explicit interaction: A clickable radio-button is associated with each individual slider (left side). Whereas in the previous version of the TOP-Slider the system sets the priority dynamically to the criterion whose cursor is selected, users are now explicitly able to opportunistically decide which criterion becomes key, and which criterion is currently key.

(2) To improve affordance and intelligibility for situations where the cursor for a criterion lies outside of an optimal range, an additional textual label coupled with the appropriate icon are now displayed below the appropriate interval to explain the situation. In addition, a corrective action is proposed as a speed-up button. When clicked, the cursor is automatically moved to the nearest optimal value.

LIMITATIONS

Our study has limitations that will be addressed in future work. In particular, the scalability of the TOP-Slider to support additional criteria has not been addressed. One approach is to draw on mathematical methods, such as dimension fixing, that address the problem of visualizing Pareto solutions for more than three objectives [3, 14]. Similarly to Monz et al. [15], a locking mechanism could be used to identify the two dependent criteria and thus reduce the dimension to three objectives: a primary objective and two secondary dependent objectives.

Our recruitment strategy (word of mouth, email, and relatives) imposed restrictions on the representativeness of the participants. Although representativeness is an important factor for sound assessment and generalization, we believe that the scope of the participants was sufficient for an early phase of the development process. We have been able to identify key spots for improvement (see above) as well as to confirm the *raison d'être* of tightly coupled sliders as a basis for supporting multi-criteria decision making.

A second evaluation experiment has been conducted with ~200 participants with the improved version of the TOP-Slider using the Surgery Pareto Slider presented above as a

baseline for comparison. Most notably, the Pareto slider allows users to restrict the domain of values for any criteria via restrictors, and moving one cursor enforces the other cursors to move automatically to an optimal solution. A preliminary analysis of the quantitative data collected during the experiment shows that users reach their objective more quickly with the TOP-Slider than with the Pareto Slider while exploring more solutions with less hesitation.

RELATED WORK

With the notable exception of the Pareto Slider designed for planning surgery, to the best of our knowledge, no slider widget has been designed to explore Pareto fronts.

Most revisited sliders focus on data base query and data base exploration [1, 2, 10, 11, 19, 20, 25]. The Pareto slider approach (i.e. the “Surgery Pareto slider” and our TOP-Slider) complements these interaction techniques by addressing decision tasks explicitly with some similarities with Influence Explorer [22]. As in Influence Explorer, the sliders of the Pareto approach are tightly dependent, and they both use color-coding to differentiate parameter ranges.

Influence Explorer [22] and WeightLifter [16] both target experts in engineering design, while the Surgery Pareto Slider targets surgeons. By contrast, we target layman users that need to find a suitable compromise without needing to know the underlying optimization mathematical model to perform their decision task.

Interestingly, WeightLifter allows users to specify relative weights between the criteria. From this perspective, our TOP-Slider as well as the “Surgery Pareto slider” are less powerful, as users can specify only one criterion as the most important.

Cursor coupling is a notable design distinction between the “Surgery Pareto slider” and the TOP-Slider. In the “surgery” slider, moving one cursor to specify a new value for a criterion enforces the other cursors to move automatically to a position that corresponds to a Pareto optimal solution. This solution is selected by a “black box” algorithm, not by the end-user. Therefore, the system may choose an undesired optimal solution. Consequently, users cannot choose non-optimal solutions. As the surgery system has been evaluated with two surgeons only, we cannot assert whether this lack of freedom is a strong restriction on end-users’ objectives.

By contrast, with TOP-Slider, moving one cursor does not move the other cursors. Instead, it shows its impact on all the value ranges in a strongly-coupled manner. The other cursors may then find themselves in a grey area (denoting a non optimal but reachable value) or in a red area (corresponding to a non-reachable value), or in a white Pareto optimal range. As discussed above, the TOP-Slider proposes corrective actions for the red and grey areas. End-users are not forced to accept the recommendations for the grey areas. After all, in some situations, users may not be

looking for optimality, but look for “good enough” optimality provided it corresponds to their needs – typically being ready to pay extra money for significant additional comfort and air quality because one has the flu.

As recommended by Matejka et al. from their study on slider decorations for rating tasks [13], the TOP-Slider uses dynamic feedback to improve speed and precision: each slider reports the value associated to the cursor current position. According to same study, tick marks along the sliders should be avoided as they introduce bias into the user’s choice. While using two labels at the ends of sliders may reduce bias, we have adopted multiple tick marks along each slider as additional information about the domain concepts. In addition, given the context of our target domain, influencing the user to choose values that favor energy savings is *a priori* appropriate.

CONCLUSION AND TAKE AWAY MESSAGE

In this paper, we have motivated and described the design of the TOP-Slider, a novel combination of tightly coupled sliders that allow non-domain experts to find optimal compromises between interdependent, possibly conflicting, criteria. The TOP-Slider (1) makes observable the solution space through a set of dynamic colored bands (i.e. optimal, sub-optimal, and non-possible value ranges), (2) supports the exploration of the solution space using sliders, a familiar interaction technique, with dynamic labeling attached to the cursors, (3) facilitates the understanding of the mutual influence between the criteria of the optimization problem using two complementary dynamic visual features: the pop up pairs of linking lines coupled with the dynamic update of the colored bands, (4) hides the complexity of the underlying mathematical model, while providing users with explanations using, when appropriate, dynamic textual labels and icons, as well as suggestions for corrective actions as speed-up buttons.

In the next step of the development process, a systemic approach to assessment will be used with the integration of the TOP-Slider with the actions plans recommended by the e-coach (e.g. opening and closing doors and windows). This action plans proposed by the system satisfies the tradeoff specified by the inhabitant using the TOP-Slider.

ACKNOWLEDGMENTS

This work has been partially sponsored by the French Agence Nationale de la Recherche (ANR), project INVOLVED reference ANR-14-CE22-0020, as well as by “Program Investissement d’Avenir” project reference ANR-11-EQPX-0002, Amiqua4Home, and by Eco-SESA, a “Cross Disciplinary Program” project of the IDEX of Université Grenoble-Alpes, France, ANR project ANR-15-IDEX-02.

REFERENCES

1. Christopher Ahlberg and Ben Shneiderman. 1994. The Alphaslider: A Compact and Rapid Selector. In *Proceedings of the SIGCHI Conference on Human*

- Factors in Computing Systems* (CHI '94), 365–371. DOI:<https://doi.org/10.1145/191666.191790>
2. Christopher Ahlberg and Ben Shneiderman. 1994. Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '94), 313–317. DOI:<https://doi.org/10.1145/191666.191775>
 3. Gautam Agrawal, Christina Bloebaum, Kemper Lewis, Kevin Chugh, Chen-Hung Huang, and Sumeet Parashar. 2004. Intuitive Visualization of Pareto Frontier for Multiobjective Optimization in n-Dimensional Performance Space. In *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference. American Institute of Aeronautics and Astronautics*. DOI:<https://doi.org/10.2514/6.2004-4434>
 4. Evanthia Dimara, Anastasia Bezerianos, and Pierre Dragicevic. 2018. Conceptual and Methodological Issues in Evaluating Multidimensional Visualizations for Decision Support. *IEEE Transactions on Visualization and Computer Graphics*. 24, 1 (January 2018), 749–759. DOI:<https://doi.org/10.1109/TVCG.2017.2745138>
 5. Stephen G. Eick. 1994. Data Visualization Sliders. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology* (UIST '94), 119–120. DOI:<https://doi.org/10.1145/192426.192472>
 6. Miriam Greis, Hendrik Schuff, Marius Kleiner, Niels Henze, and Albrecht Schmidt. 2017. Input Controls for Entering Uncertain Data: Probability Distribution Sliders. In *Proceedings of ACM on Human-Computer Interaction*. 1, EICS (June 2017), 3:1–3:17. DOI:<https://doi.org/10.1145/3095805>
 7. H el ene Haller, Van Bao Nguyen, Gilles Debizet, Yann Laurillau, Jo elle Coutaz, and Ga ell Calvary. 2017. Energy consumption in smarthome: Persuasive interaction respecting user’s values. In *Proceedings of the 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, 804–809. <https://doi.org/10.1109/IDAACS.2017.8095199>
 8. Alfred Inselberg and Bernard Dimsdale. 1987. Parallel Coordinates for Visualizing Multi-Dimensional Geometry. In *Computer Graphics 1987*. Springer, Tokyo, 25–44. DOI:https://doi.org/10.1007/978-4-431-68057-4_3
 9. Teuvo Kohonen, Manfred R. Schroeder, and Thomas S. Huang (Eds.). 2001. *Self-Organizing Maps* (3rd ed.). Springer-Verlag New York, Inc., Secaucus, NJ, USA. DOI:<https://doi.org/10.1007/978-3-642-97966-8>
 10. Tom Lanning, Kent Wittenburg, Michael Heinrichs, Christina Fyock, and Glenn Li. 2000. Multidimensional Information Visualization Through Sliding Rods. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (AVI '00), 173–180. DOI:<https://doi.org/10.1145/345513.345306>
 11. Qing Li and Stephen C. North. 2003. Empirical comparison of dynamic query sliders and brushing histograms. In *IEEE Symposium on Information Visualization 2003* (IEEE Cat. No.03TH8714), 147–153. DOI:<https://doi.org/10.1109/INFVIS.2003.1249020>
 12. Rui Ma, Kai Li, Xuan Li, and Zeyu Qin. 2015. An economic and low-carbon day-ahead Pareto-optimal scheduling for wind farm integrated power systems with demand response. *Journal of Modern Power Systems and Clean Energy*. 3, 3 (September 2015), 393–401. DOI:<https://doi.org/10.1007/s40565-014-0094-7>
 13. Justin Matejka, Michael Glueck, Tovi Grossman, and George Fitzmaurice. 2016. The effect of visual appearance on the performance of continuous sliders and visual analogue scales. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI 2016), 5421–5432. DOI:<http://dx.doi.org/10.1145/2858036.2858063>
 14. Kaisa Miettinen. 2014. Survey of methods to visualize alternatives in multiple criteria decision making problems. *OR Spectrum*. 36, 1 (January 2014), 3–37. DOI:<https://doi.org/10.1007/s00291-012-0297-0>
 15. Michael Monz, Karl-Heinz K ufer, Thomas R. Bortfeld, and Christian Thieke. 2008. Pareto navigation—algorithmic foundation of interactive multi-criteria IMRT planning. *Physics in Medicine and Biology*. 53, 4 (February 2008), 985–998. DOI:<https://doi.org/10.1088/0031-9155/53/4/011>
 16. Stephan Pajer, Marc Streit, Thomas Torsney-Weir, Florian Spechtenhauser, Torsten M oller, and Harald Piringer. 2017. WeightLifter: Visual Weight Space Exploration for Multi-Criteria Decision Making. *IEEE Transactions on Visualization and Computer Graphics*. 23, 1 (January 2017), 611–620. DOI:<https://doi.org/10.1109/TVCG.2016.2598589>
 17. Charles Perin, Pierre Dragicevic, and Jean-Daniel Fekete. 2015. Crossets: manipulating multiple sliders by crossing. In *Proceedings of Graphics Interface 2015* (GI 2015), 233–240. DOI:<https://doi.org/10.20380/GI2015.30>
 18. Christian Schuman, Christian Rieder, Sabrina Haase, Katrin Teichert, Philipp S us, Peter Isfort, Philipp Bruners, and Tobias Preusser. 2015. Interactive multi-criteria planning for radiofrequency ablation. *International Journal of Computer Assisted Radiology and Surgery* (CARS), 10(6), 879–889. DOI:<https://doi.org/10.1007/s11548-015-1201-6>
 19. Mark Sifer. 2006. Filter co-ordinations for exploring multi-dimensional data. *Journal of Visual Languages & Computing*. 17, 2 (2006), 107–125.

- DOI:<https://doi.org/https://doi.org/10.1016/j.jvlc.2005.10.005>
20. Theophanis Tsandilas, Anastasia Bezerianos, and Thibaut Jacob. 2015. SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, 3255–3264. DOI:<https://doi.org/10.1145/2702123.2702129>
 21. Mark Tunner and Gilles Fauconnier. 1995. Conceptual Integration and Formal Expression. *Metaphor and Symbolic Activity*. 10, 3 (1995), 183–204. DOI:https://doi.org/10.1207/s15327868ms1003_3
 22. Lisa Tweedie, Robert Spence, Huw Dawkes, and Hus Su. 1996. Externalising Abstract Mathematical Models. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '96)*, 406–412. DOI:<https://doi.org/10.1145/238386.238587>
 23. Mark Velasquez and Patrick T. Hester. 2013. An analysis of multi-criteria decision making methods. *International Journal of Operations Research*. 10, 2 (2013), 56–66.
 24. Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. 2007. Scented Widgets: Improving Navigation Cues with Embedded Visualizations. *IEEE Transactions on Visualization and Computer Graphics*. 13, 6 (November 2007), 1129–1136. DOI:<https://doi.org/10.1109/TVCG.2007.70589>
 25. Kent Wittenburg, Tom Lanning, Michael Heinrichs, and Michael Stanton. 2001. Parallel Bargrams for Consumer-based Information Exploration and Choice. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology (UIST '01)*, 51–60. DOI:<https://doi.org/10.1145/502348.502357>
 26. Victor M. Zavala. 2012. Real-time resolution of conflicting objectives in building energy management: an utopia-tracking approach. In *Proceedings of the 5th national conference of IBPSA-USA*, 1–6.