Design space visualization for efficiency in knowledge discovery leading to an informed decision

Audrey Abi Akle, Stéphanie Minel, Bernard Yannou

Abi Akle A., Minel S., Yannou B., 2017. Design space visualization for efficiency in knowledge discovery leading to an informed decision, In 21st International Conference on Engineering Design (ICED), August 21-25, Vancouver, Canada.

DESIGN SPACE VISUALIZATION FOR EFFICIENCY IN KNOWLEDGE DISCOVERY LEADING TO AN INFORMED DECISION

Audrey Abi Akle(a,b), Stéphanie Minel(a), Bernard Yannou(b)

(a) ESTIA, Bidart, France
(b) Laboratoire Genie Industriel, CentraleSupélec, Université Paris-Saclay

ABSTRACT

Design space exploration refers to the systematic activity of discovery and evaluation of the elements in the design space in order to identify optimal solution by reducing the design space toward an area of performance. Designers sample thousand design points iteratively, explore the design space, gain knowledge about the problem and make design decision. According to the literature, Design Space Exploration results in a decision of quality called informed decision which is supported by information visualization. Indeed, the representation of design points is seen as primordial to gain understanding of the problem and make an informed decision. Thereby, in our work, we try to identify which graph is the most suited to the discovery phase and allows designers to make an informed decision. We designed a web platform with four design problem and carried out an experiment with 42 participants. It results one graph more suited to make a decision of quality and to gain the most understanding: the Scatter Plot Matrix.

KEY WORDS

Visualization, Computer aided design (CAD), Decision making, Design by shopping

1 INTRODUCTION

There is a paradigm where designers shop for the best solution. It is called Design by Shopping and was coined by Balling (Balling, 1999). Indeed, Balling noted that the traditional optimization-based design process to “formulate the design problem, obtain analysis models and execute an optimization algorithm” leaves designers unsatisfied. Designers like consumers want to “shop” to gain an insight into trades, feasible and impractical solutions, and to learn about their alternatives before making decisions. Design by Shopping, firstly, allows designers to explore the design space, and, secondly, optimize and choose an optimal solution from a set of possible designs, and then develop realistic expectations with regard to what is possible. One embodiment of this paradigm is the Design Space Exploration (Simpson et al., 2008). With Design Space Exploration (DSE), designers sample thousand (and more) design points iteratively, explore the design space that is a multidimensional set of data, gain insights and knowledge about the problem and make design decision. In DSE the design decision is performed following the discovery and evaluation of the elements in the design space in order to identify optimal solution by reducing the design space toward an area of performance. Based on the work of (Miller et al., 2013) we identify that exploring the design space consists of three main phases: (1) Discovery: acquire knowledge and understanding of the problem, (2) Narrowing: active pursuit of a design by eliminating sets, exploring limits, highlighting preferences, etc. and (3) Selection: check satisfaction (see Figure 1). Considering the knowledge discovery in DSE, we find one point that challenges us. Indeed, some authors refer to informed decision making (Sulllivan et al., 2001), (Mavris et al., 2010), (Chandrasegaran et al., 2013). From our literature review dealing with this decision type (see section 2); we particularly identify a need in information visualization for decision-makers.

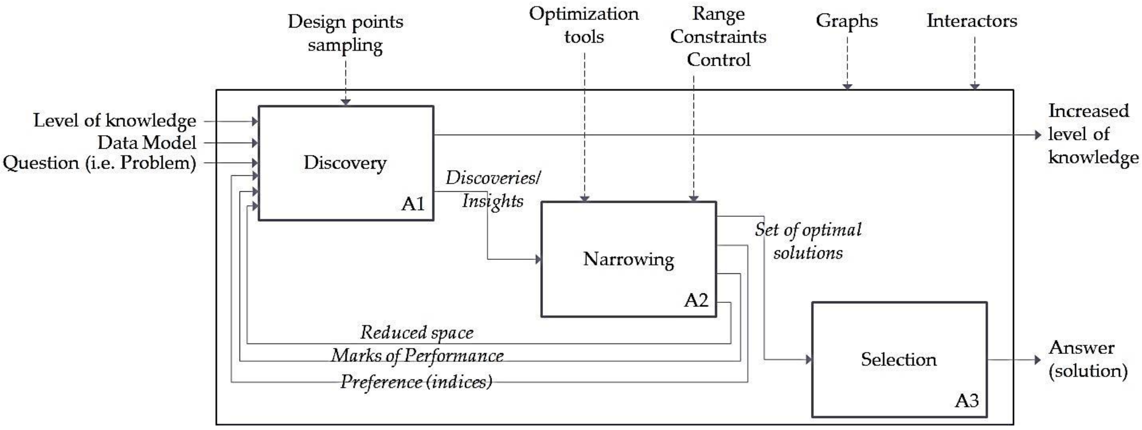

Figure 1. Design Space Exploration Process

Graphs are useful for quickly visualizing the feasible solutions as opposed to impractical solutions, as well as those violating engineering constraints or client requirements. Three different situations can be represented with more or less data: representing the single vector of design parameters featuring the product solution, the single vector of solution performances for feasible solutions or two sets of design parameters and corresponding performances for feasible. There are already many works in visual design and it has already been shown that fast graphical design interfaces impact user performance in terms of design efficiency, design effectiveness and the design search process (Ligetti *et al*., 2003).

We identify several tools for exploring the design space. For example, there are the ARL Trade Space Visualizer (Stump et al., 2004), the VIDEO tool (Kollat & Reed, 2007), the LIVE tool (Yan et al., 2012) and the Rave tool (Daskilewicz & German, 2012). Thus we find in these tools, various graphs: Scatter Plot Matrix, 2D or 3D Scatter Plot, Parallel Coordinate Plot, Bars Chart and Treemap. In addition, it has already been shown in a simplified framework that Parallel Coordinate Plot is the most suitable graph to selection in Design by Shopping (Abi Akle *et al*., 2015, 2016). It therefore appears both that knowledge discovery phase for insights gain and understanding of the design problem is a key element of the Design by Shopping. On the other hand, we observe that information visualization is an indispensable element to the practice of DSE. Thereby, our research is motivated by a question: **What graph allows designers to be effective in the discovery phase and results in an informed decision?**

We have thus identified three graphs useful for representing multidimensional sets of data (>3 variables) and with an unlimited number of design points: Simple Scatter Plot (SSP), Scatter Plot Matrix (SPM) and Parallel Coordinate Plot (PCP). We carried out experiments with 42 participants and designed a web platform with four design problem to solve. The platform allows to generate an unlimited number of design points (random or Pareto sampling), to reduce the design space with a range constraints controller, to visualize preferences, etc. in order to mimic the design activity. We identified a graph more suited to the discovery phase and to an informed decision making in design space exploration: the Scatter Plot Matrix (SPM).

2 INFORMED DECISION

A notion used in several disciplines seems to be of importance to be in a situation to make a decision of quality, this is an informed decision.

In monitoring & supervision field, Ireson states that "The management of this mass of information is crucial in aiding the decision-making process, ensuring, as far as possible, that the responders have full situational awareness to make informed decisions" (Ireson, 2009). With the same idea, Bass (2000) and Riveiro et al. (2008) indicate the need for "situational awareness" for the formulation of an informed decision. Bass adds the need to "fusing data into information and knowledge, so network operators can make informed decisions" (Bass, 2000). In business / marketing field, Lurie and Mason (2007) suggest managing a large data set and the use of visualization tools could lead toward an informed decision. Glaser and Tolman (2008) link the informed decision to the process of analysing large amounts of data, "tracking" of performance and detecting patterns and trends. Information systems field informs us that "the making of informed decisions requires the application of a variety of knowledge to information" (Wiederhold, 1992). In the building field, making informed design

decision needs to manage a large amount of information on the detailed design options and properties and to operate simulations of their performance. For them, the designer needs a large design space and an overview of parameter changes consequences to gain a deep understanding of the performance and so make an informed design decision (Petersen & Svendsen, 2010). Russell et al. (2009) consider that "Visual analytics, the science of analytical reasoning facilitated by interactive visual interfaces, has the potential to improve the construction management process through the enhanced understanding of project status and reasons for it, better informed decision making" (Russell et al., 2009). In information visualization field, Keim et al. (2006) indicate that for an informed decision, "it is indispensable to include humans in the data analysis process to combine flexibility, creativity, and background knowledge with the enormous storage capacity and the computational power of today's computers". Later, Keim et al. (2008a, 2008b) state that "visual analytics" is the system to make an informed decision. It guarantees the full support of the user in navigating and analysing the data, memorizing insights and making informed decisions. Also, we find that "a tight coupling between cognition, interaction and visual analytics is necessary to enable the user to make informed decisions" (Meyer et al., 2010).
Finally, we find the term informed decision in the design engineering field. Wood et al. (1992), in preliminary design, model and manipulate such uncertainties in a computer-assisted environment, under the hypothesis that doing so will allow the designer to make faster and more-informed decisions. Sullivan et al. (2001) show, in particular, to make informed decisions about the choice of design rules and clustering of design parameters, designers needed to know how changes in the environment would affect them. Chandrasegaran et al. (2013) indicate that "an effective computer support tool that helps the designer make better-informed decisions requires efficient knowledge representation schemes". Mavris et al. (2010) argue that the integration of "visual analytics" in the design process provides designers the ability to gain knowledge and insights needed with the justified means of making an informed decision. The visualization seems essential to facilitate the generation of hypotheses and the formulation of an informed decision. They point out the data, knowledge, and insight necessary for the formulation of informed decisions is generated throughout the design process (Mavris et al., 2010).
From the literature, we identify eight themes that appear to contribute to the formulation of an informed decision (see Table 1). Furthermore, we identify four of these themes that are most widely used by the authors and seem essential in the definition of an informed decision: "Knowledge and insights gain", "Visualization", "Analysis and treatment" et "Manipulation and management".

Table 1. Themes that appear to contribute to the formulation of an informed decision

Themes	Contributors
Situational awareness	(Ireson, 2009), (Bass, 2000)
Knowledge and insights gain	(Sullivan *et al*., 2001), (Keim *et al*., 2006, 2008a, 2008b), (Glaser & Tolman, 2008), (Petersen & Svendsen, 2010), (Mavris et al., 2010)
Visualization	(Chandrasegaran et al., 2013), (Lurie & Mason, 2007), (Bass, 2000), (Mavris et al., 2010), (Keim et al., 2008b), (Russell et al., 2009), (Meyer et al., 2010), (Riveiro et al., 2008)
Manipulation and management	(Ireson, 2009), (Chandrasegaran et al., 2013), (Lurie & Mason, 2007), (Wood et al., 1992), (Petersen & Svendsen, 2010), (Mavris et al., 2010), (Riveiro et al., 2008), (Meyer et al., 2010)
Analysis and treatment	(Keim et al., 2006, 2008a, 2008b), (Russell et al., 2009), (Meyer et al., 2010), (Riveiro et al., 2008), (Mavris et al., 2010)
Understanding	(Petersen & Svendsen, 2010), (Russell et al., 2009), (Mavris et al., 2010)
Human and cognition	(Keim et al., 2006), (Wood et al., 1992), (Mavris et al., 2010), (Meyer et al., 2010)
Transformation	(Chandrasegaran et al., 2013), (Bass, 2000), (Keim et al., 2008b)

We know that, in engineering design, once the design has been formalized, a necessary design task is to make a selection from amongst candidate designs or parametric values (Otto & Antonsson, 1993). The main challenge lies in resolving the inherent trade-offs that exist between the overall system and subsystems, and between conflicting and competing objectives (Abi Akle et al., 2015, 2016). Thus we define an informed decision in Design by Shopping as the selection of a design point, among several others that will achieve optimal benefits and minimum inconvenience, following an iterative and interactive treatment and analysis process in which designers are gaining understanding, knowledge and insights with visualization and manipulation of large sets and / or data model.

3 GRAPHS ADAPTED

In our context of design space exploration in order to make an informed decision several graphs (design space representations) are available to us. So, we are in a case-representation of multidimensional sets of data with an unlimited number of alternatives (design points). Based on the work of Miettinen (2014), Wegman (1990) and Keim (2000) amongst others about graph characteristics, we identify the Simple Scatter Plot (SSP), the Scatter Plot Matrix (SPM) and the Parallel Coordinate Plot (PCP) (see Figure 2). Thus, we propose comparing these tree graphs under the ability to make an informed decision in design space exploration.
Scatterplot is a conventional method to visualize the relationship between two variables. As a visual structure, Scatterplot uses the position to encode the values of two variables and their relationships. This is a projection of the data (represented by the design points) in a 2D space. Siirtola (2007) considers the Scatterplot is useful for easily detect non-linear patterns and positive or negative correlations between variables. The Scatterplots are Cartesian representations and therefore have a long history with, for example, the learning of linear algebra at school. Due to this training, it results in a strong development of intuitions about the appearance of this type of representation (Wegman, 1990).

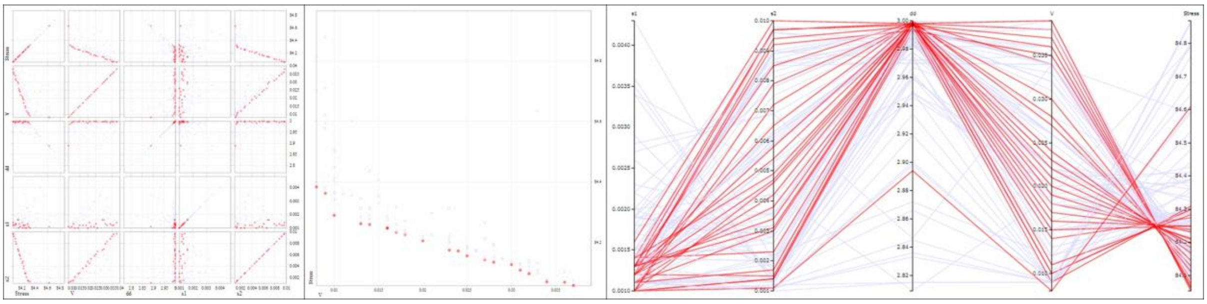

Figure 2. Illustration of the SPM, the SSP and the PCP

The Scatterplot Matrix is a collection of Scatterplots Simples (x-y) ordered by pairs. This representation provides an overview of data. The size of the matrix depends on the number of visualized variables. The Scatterplots are duplicated in the matrix relative to the diagonal. It is possible to use colour glyphs, shape, or add a size marker to add a supplementary variable. Ware (2004) highlights that, although the colour used with Scatterplots allows the identification of clusters and patterns, interpretation could be difficult.
Parallel Coordinates Plot is defined as a graph displaying multiple criteria without drastically increasing the complexity of the display (Inselberg, 2009). In PCP, variable values are displayed on separate axes laid out in parallel. The design points (or alternatives) are depicted as profile lines that connect points on the respective axes. According to Gettinger et al. (2013), this representation can be readily interpreted and provides a good overview. Furthermore, patterns such as positive, negative and non-trivial (multiple) correlations, may quickly be identified at a glance.

4 EXPERIMENTAL DESIGN

To answer our research question "What graph allows designers to be effective in the discovery phase and result in an informed decision?" we conducted a controlled experiment that adopted a between subject approach. Each participant performed the experiment on one graph and resolved 4 design optimization problems. They are classic problems of mechanical design from literature: two-member Truss design (Truss), gear train design (Gear), multiple-disk clutch brake design (Disk) and pressure vessel design (Vessel). The first three problems come from the work of Deb and Srinivasan (2006) and the fourth problem comes from (Canbaz, 2013). To realize the experimentation we have developed a web "platform" (http://these.aaa.alwaysdata.net/expe2/) where problems and the three graphs are available. This platform allows, among other the generation of design points (random or non-dominated Pareto solutions sampling), glyph colouring according to the designer preferences and reducing the design space.

4.1 Procedure

The experiment is divided into two main phases: (1) training part and (2) test part. Moreover, we incorporate a type of milestone during the session: multiple choice forms. We use three forms, one at the beginning, another between the training part and the test part and one at the end of the session.
The experimentation is sized to be limited to a two-hour session. The training part (1) is used to upgrade the level of knowledge of participants. It is divided in three steps: a crash course, a "getting started" step with a tutorial to resolve the Truss problem and a practice phase with the platform guide where participants resolve the Gear problem. The test part (2) is the phase where the graphs are tested and we realize our measurements. The tests are performed on two design problems consecutively without help supports: the Vessel problem followed by the Disk problem. The instruction for both problems is to solve the bi-objective optimization problem using the method of Design Space Exploration to select an optimal solution with a justified means. The optimization problems are bi-objectives that is to say there are two antagonistic performance variables to maximize or minimize. For each of the two problems, the participants have a 10-minute time limit and after each problem, a questionnaire is given to the participant to gather confidence and to know the information acquired during exploration, which enabled him to make his/her decision (selecting a solution).

4.2 Design problems

As already mentioned, we use four bi-objective design problems for our experiment. The first two are used during the training part and a description is available in (Deb and Srinivasan, 2006). The first problem for the test is the Vessel problem described as follow: It is a design problem of a cylindrical thin walled pressure vessel with hemispherical ends. There are three design variables (R, T, L) and two performances (W, V). The objectives are minimizing W and maximizing V with controlling R, T and L while satisfying constraints C1 to C7. Vessel problem nomenclature and constants are: W is Weight of the pressure vessel in lbs, V is Volume of the pressure vessel, R is Radius, T is Thickness of the pressure wall, L is Length of the cylinder, P is Pressure inside the cylinder, UTS is Ultimate tensile strength of the vessel material and equals 35 klb, d is density of the vessel material and equals 0.283 and Circ is Circumferential stress (see Figure 3).
The second problem for the test is the Disk problem described as follow: In this problem, a multiple clutch brake needs to be designed. Two conflicting objectives are considered: minimization of mass (M) of the brake system in kg and minimization of stopping time (S) in seconds. There are five decision variables Ri, Ro, t, F, and Z (see Figure 3), where Ri is the inner radius in mm, Ro is the outer radius in mm, t is the thickness of discs in mm, F is the actuating force in N and Z is the number of discs (or friction surfaces). All five variables are considered discrete and their allowable values are given below: 60 < Ri < 80; 90 < Ro < 110; 1 < t < 3; 600 < F < 1000; 2 < Z < 10.
All performance formulas, constraints and bounds of the two problems are available on the web platform.

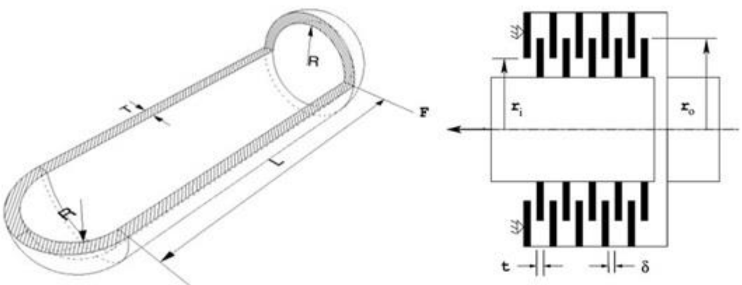

Figure 3. Illustration of the Vessel problem (on the left) and the Disk problem (on the right)

4.3 Measurements

In our work, variables were either measured during the test with an eye-tracking system the Tobii X2 or collected through the questionnaires. We use three types of measurements: those controlled during the test and those collected with questionnaires.

The controlled measurements are used to verify that there are no differences between the three groups of participants (one group per graph) concerning: the knowledge level from the three Multiple-Choice-Questionnaires (MCQs) and the "performance" of the final solution selected (is it a Pareto solution?).

Then, we have three measurements realized during the test:

- The number of discoveries / insights realized by the participant knowing that for each problem, there are seven discoveries in order to make a fully informed decision (see **Erreur ! Source du renvoi introuvable.**). The collect of the number of discoveries was performed by analysis of the video screen in conjunction with eye-tracking record.
- We also measure the time passed before the participant realizes the first discovery (in seconds). This time is not dependent on elapsed time to understand the problem because the timer is triggered only when the participant has read the description of the problem and s/he makes the first design points sampling. This second measurement is for us a first clue to the speed of discovery with the three graphs.
- The third measurement is the average time to complete a discovery (in seconds) i.e. the total time to perform all the discoveries divided by the number of discoveries realized. This measurement is for us a second clue to the speed of discovery with the three graphs.

Table 2. List of the discoveries for the two problems of the test

Problem	Discoveries from a global point of view	Discoveries from a local point of view (Pareto solutions)
Vessel	Positive correlation (trend) between R & V	Positive correlation between W & L
	Positive correlation (trend) between R & T	Positive correlation between V & L
	Positive correlation (trend) between T & V	The solutions tend to T = 4
		The solutions tend to R = 36
Disk	Negative correlation (trend) between S & Z	Positive correlation (trend) between Ri & Ro
		A transition point for Ro = 100
		A transition point for M = 1
		The solutions tend to t = 1
		The solutions tend to F = 1000
		Positive correlation between M & Z

Finally we use a post-problem questionnaire to know if participants can "justify" their decision based on the information acquired during exploration (correlation, trend, transition point, etc.). Note that for the analysis of this indicator, we only check if the participants justify their decision or not (nominal qualitative variable). We cannot analyse the amount of information that participants use to justify their response because the amount depends on the participant. Similarly, the white forms are not considered (because lack of response does not mean that participants do not know how to justify their decision).

5 RESULTS

We have a sample of 42 subjects and it is a between approach (3 groups). So we have three samples of N = 14 subjects in each group / graph. For the analysis of data, we apply different statistical tests and we consider a significance level α=10%.

5.1 Differences and similarities between the three groups of participants

First of all, we analyse the data obtained with MCQs (Multiple-Choice-Questionnaires) to verify if we can distinguish designer profiles within testers and / or if there is a difference between groups (i.e. three groups - three graphs). To detect designer profiles, we use the Friedman test (and the Wilcoxon Signed-Rank test if a post-hoc analysis is required) because MCQs answers give us qualitative ordinal variables and the groups are paired (i.e. within-approach). To detect difference between groups, we use Kruskal-Wallis test because in this case the analysis is a between-approach.

Table 3. Results of the Friedman and Wilcoxon tests for the intra-graphs analysis

		SSP	PCP	SPM
Friedman test		csqr=23.89, df=2 and p<0.0001	csqr=26.14, df=2 and p<0.0001	csqr=21.09, df=2 and p<0.0001
Wilcoxon Signed-Rank Test	MCQ1 vs. MCQ2	W=-105, Z=-3.28 and p=0.0005.	W=-105, Z=-3.28 and p=0.0005.	W=-107.5, Z=-3.36 and p=0.0004
	MCQ2 vs. MCQ3	W=-76, Z=-2.64 and p=0.0041.	W=-103, Z=-3.22 and p=0.0006	W=-53, Z=-1.83 and p=0.0336

The results for the three graphs, presented in Table 3, allow us to conclude that all participants have a novice profile at the start of the test and they all gained knowledge (in design space exploration method): MCQ1<MCQ2<MCQ3. The average scores of MCQs are depicted in Figure 4.
The results presented in **Erreur ! Source du renvoi introuvable.** show us that there is no significant difference between the three graphs for the three MCQs answers. Thus, there is no difference of knowledge between the groups.

Table 4. Results of Kruskal-Wallis tests for the inter-graphs analysis

		MCQ1	MCQ2	MCQ3
Kruskal-Wallis Test		H=0.61 p=0.7371	H=1.55 p=0.4607	H=4.2 p=0.12
Rank	SPM	23.1	22	16.8
	PCP	19.5	18.4	26.3
	SSP	21.9	24.1	21.5

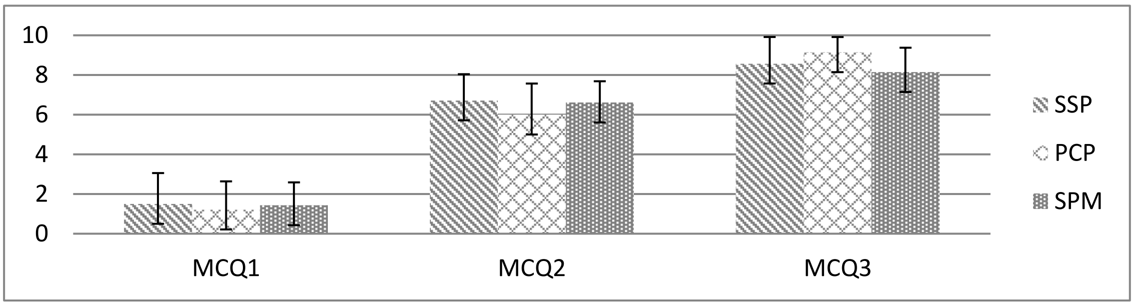

Figure 4. Average scores of MCQs and standard deviation

Then, we analyse the solutions selected by participants. For this indicator we analyse the data from the two problems together. Thus, our study has two variables: the group (SPM, PCP and SSP) and the "Pareto" (yes or no). We obtain (from the descriptive statistics analysis) that over 20% of our results (half) had an expected frequency less than 5 (i.e. 95% for the non-dominated) that is why we realize the Chi2 test with Yates correction to address our statistical assumptions: $X^2 = 1.647$, df = 2 and p = 0.44. We should infer that there is no difference between the three groups.
The first part of the analysis allows us to conclude that there is no difference in knowledge between the three groups and they have almost all selected a Pareto solution. We can thus free ourselves to divide the groups for the following analyses.

5.2 Measurements during test: discovery variables

Unfortunately, we have different sample sizes because all subjects were unable to perform the tests in their entirety. This is due to bugs in the platform. All discovery variables measured are quantitative. Therefore we apply the ANOVA-between statistical test and pairwise t-tests post-hoc analysis if necessary (i.e. if ANOVA is significant). We considered the following statistical hypotheses: H0: there is no difference between the graphs for the discovery phase and H1: there is a difference.
For the Vessel problem, we obtain for the SSP (n =12) an average of 2.66 discoveries, 209.6 seconds elapsed before the first discovery and 198.4 seconds mean time per discovery, for PCP (n =12) 3.25 discoveries, 115.8 seconds elapsed before the first discovery and 131.1 seconds mean time per discovery and for SPM (n =13) 6.3 discoveries, 14.8 seconds elapsed before the first discovery and 65.1 seconds mean time per discovery. ANOVAs give significant results: $F(2,34)= 24.27$ and $p<0.0001$ for the number of discoveries, $F(2,33)= 8.65$ and $p= 0.000954$ for the time before the first discovery and $F(2,33)= 6.45$ and $p= 0.004321$ for the meantime per discovery.

Regarding the Disk problem, we obtain for the SSP (n =13) an average of 2.92 discoveries, 175.2 seconds elapsed before the first discovery and 129.1 seconds mean time per discovery, for PCP (n =14) 3 discoveries, 107.2 seconds elapsed before the first discovery and 93.3 seconds mean time per discovery and for SPM (n =13) 5.4 discoveries, 52.4 seconds elapsed before the first discovery and 75.3 seconds mean time per discovery. ANOVAs give significant results: $F(2,37) = 17.01$ and $p<0.0001$ for the number of discoveries, $F(2,37)= 9.17$ and $p= 0.000583$ for the time before the first discovery and $F(2,37)= 5.71$ and $p= 0.006899$ for the meantime per discovery.
So we perform a post-hoc analysis for the two problems (see Table 5). The averages for each variable are depicted in Figure 5.

Table 5. Results of the t-test for the three discovery variables for the 2 problems

Problem	T-test pairwise comparison	Number of discoveries	Time before the first discovery	Mean time to complete a discovery
Vessel	SSP vs PCP	t(22)=-0.92 p=0.18	t(21)=-1.57 p= 0.066	t(21)= 1.44 p= 0.0823
	SSP vs SPM	t(23)=6.62 p<0.0001	t(22)= 3.69 p= 0.0006	t(22)= 3.19 p= 0.002
	PCP vs SPM	t(22)=5.86 p<0.0001	t(23)= 4.55 p<0.0001	t(23)= 3.52 p= 0.0009
Disk	SSP vs PCP	t(25)=0.16 p= 0.437	t(25)= 2.39 p= 0.0123	t(25)= 1.94 p= 0.0319
	SSP vs SPM	t(24)= 4.99 p <.0001	t(24)= 4.44 p <.0001	t(24)= 3.24 p = 0.0017
	PCP vs SPM	t(25)= 5.08 p <.0001	t(25)= 1.89 p = 0.0352	t(25)= 1.44 p = 0.081

The results of the t-test pairwise comparison indicate that there is a significant difference between SPM and SSP; SPM and PCP for the number of discoveries and between SSP, PCP and SPM for the other two indicators. SPM is significantly different from the other two and gets the best scores.

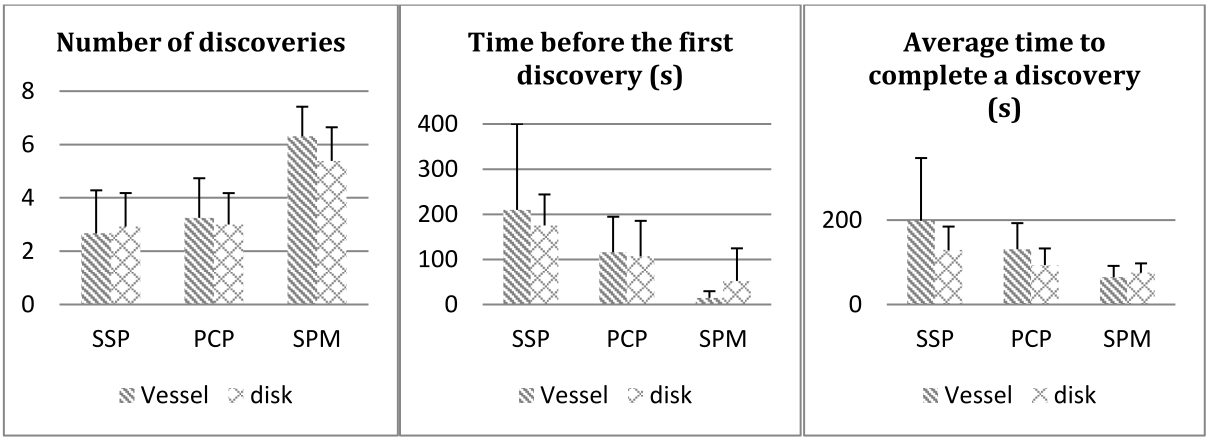

Figure 5. Average and standard deviation for the three discovery variables

Statistical analysis performed on three discovery variables allows us to conclude that the Scatter Plot Matrix (SPM) is the most relevant graph for the discovery phase. We should also indicate that the Simple Scatter Plot (SSP) is the graph getting the worst results for the discovery phase.

5.3 Justification of the decision

For this indicator we analyse the data from the two problems together. Our study has two qualitative variables: the group (SPM, PCP and SSP) and the answer to question (decision justified or not). We apply a test of Chi2: $X^2 = 4.747$, df = 2 and p = 0.093. The result is significant. We then operate the Khi2 tests in pairs:

- SSP vs PCP : X2=0.18, df=1 and p=0.671
- SSP vs SPM : X2=2.99, df=1 and p=0.084
- PCP vs SPM : X2=4.447, df=1 and p=0.035

There is a significant difference between the SPM and SSP graphs; and SPM and PCP. Decisions are justified with SPM. We conclude that SPM is the graph with which the participants made the most informed decisions.

Abi Akle A., Minel S., Yannou B., 2017. Design space visualization for efficiency in knowledge discovery leading to an informed decision, In 21st International Conference on Engineering Design (ICED), August 21-25, Vancouver, Canada.

6 DISCUSSION AND CONCLUSION

We have presented an optimization strategy to focus on informed decision as competiveness leverage. Hence, designers must iteratively build the design based on experimentation. They have to choose with a large, and often cumbersome, set of alternatives. Our research helps in better understanding the performance impacts of the data presentation on the system. We identify the Simple Scatter Plot (SSP), the Scatter Plot Matrix (SPM) and the Parallel Coordinate Plot (PCP) as graphical optimization. Thanks to our experiment we aimed to draw out clear recommendations regarding the choice of a graph for the discovery phase in Design Space Exploration to make an informed decision. We have shown how to select an optimal solution in a set of feasible solutions defined by their design and performance value vectors. From the results of our tests, design space exploration is improved while using SPM graph to present data. Designer seems to be more confident and made informed decisions depending on the graphical interface proposed. SPM is most appropriate for the discovery phase because this phase involves understanding the problem by observing interactions between the variables and SPM is the graph that most reveals these types of interactions (such as clusters, correlations, etc.) (Keim, 2000). Moreover, we believe that it inspires designers because it uses Cartesian representations (i.e. facilitates interpretation) and gives an overview of the dataset and therefore of the problem. Our results are consistent, however there is still a need to test our models and build on this work to improve the accuracy of our results, especially focusing with the impacts of graphical visualization depending on the designer's expertise.

REFERENCES

Abi Akle, A., Minel, S., & Yannou, B. (2015). Graphical support adapted to designers for the selection of an optimal solution in design by shopping. In DS 80-6 Proceedings of the 20th International Conference on Engineering Design (ICED 15) Vol 6: Design Methods and Tools-Part 2 Milan, Italy, 27-30.07. 15.

Abi Akle, A., Minel, S., & Yannou, B. (2016). Information visualization for selection in Design by Shopping. Research in Engineering Design, 1-19, doi: 10.1007/s00163-016-0235-2

Balling, R. (1999, May). Design by shopping: A new paradigm?. In Proceedings of the Third World Congress of structural and multidisciplinary optimization (WCSMO-3) (Vol. 1, pp. 295-297).

Bass, T. (2000). Intrusion detection systems and multisensor data fusion. Communications of the ACM, 43(4), 99-105.

Canbaz, B. (2013). Preventing and resolving design conflicts for a collaborative convergence in distributed set-based design (Doctoral dissertation, Ecole Centrale Paris).

Chandrasegaran, S. K., Ramani, K., Sriram, R. D., Horváth, I., Bernard, A., Harik, R. F., & Gao, W. (2013). The evolution, challenges, and future of knowledge representation in product design systems. Computer-Aided Design, 45(2), 204-228.

Daskilewicz, M. J., & German, B. J. (2012). Rave: A computational framework to facilitate research in design decision support. Journal of computing and information science in engineering, 12(2), 021005.

Deb, K., & Srinivasan, A. (2006, July). Innovization: Innovating design principles through optimization. In Proceedings of the 8th annual conference on Genetic and evolutionary computation (pp. 1629-1636). ACM.

Gettinger, J., Kiesling, E., Stummer, C., & Vetschera, R. (2013). A comparison of representations for discrete multi-criteria decision problems. Decision support systems, 54(2), 976-985.

Glaser, S. D., & Tolman, A. (2008). Sense of sensing: from data to informed decisions for the built environment. Journal of infrastructure systems, 14(1), 4-14.

Inselberg, A. (2009). Parallel Coordinates: VISUAL Multidimensional Geometry and its Applications.

Ireson, N. (2009, June). Local community situational awareness during an emergency. In Digital Ecosystems and Technologies, 2009. DEST'09. 3rd IEEE International Conference on (pp. 49-54). IEEE.

Keim, D. A. (2000). Designing pixel-oriented visualization techniques: Theory and applications. Visualization and Computer Graphics, IEEE Transactions on, 6(1), 59-78.

Keim, D. A., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006, July). Challenges in visual data analysis. In Information Visualization, 2006. IV 2006. Tenth International Conference on (pp. 9-16). IEEE.

Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008b). Visual analytics: Scope and challenges (pp. 76-90). Springer Berlin Heidelberg.

Keim, D., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J., & Melançon, G. (2008a). Visual analytics: Definition, process, and challenges (pp. 154-175). Springer Berlin Heidelberg.

Kollat, J. B., & Reed, P. (2007). A framework for visually interactive decision-making and design using evolutionary multi-objective optimization (VIDEO). Environmental Modelling & Software, 22(12), 1691-1704.

Ligetti, C., Simpson, T. W., Frecker, M., Barton, R. R., & Stump, G. (2003). Assessing the impact of graphical design interfaces on design efficiency and effectiveness. Journal of Computing and Information Science in Engineering, 3(2), 144-154.

Lurie, N. H., & Mason, C. H. (2007). Visual representation: Implications for decision making. Journal of Marketing, 71(1), 160-177.

Mavris, D. N., Pinon, O. J., & Fullmer Jr, D. (2010). Systems design and modeling: A visual analytics approach. In Proceedings of the 27th International Congress of the Aeronautical Sciences (ICAS), Nice, France.

Meyer, J., Thomas, J., Diehl, S., Fisher, B. D., Keim, D. A., Laidlaw, D., & Ynnerman, A. (2010). From Visualization to Visually Enabled Reasoning. Scientific visualization: Advanced concepts, 1, 227-245.

Miettinen, K. (2014). Survey of methods to visualize alternatives in multiple criteria decision making problems. OR Spectrum, 36(1), 3-37.

Miller, S. W., Simpson, T. W., Yukish, M. A., Bennett, L. A., Lego, S. E., & Stump, G. M. (2013, August). Preference Construction, Sequential Decision Making, and Trade Space Exploration. In ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (pp. V03AT03A014-V03AT03A014). American Society of Mechanical Engineers.

Otto, K. N., & Antonsson, E. K. (1993). "The method of imprecision compared to utility theory for design selection problems", Design Theory and Methodology–DTM'93, 167-173.

Petersen, S., & Svendsen, S. (2010). Method and simulation program informed decisions in the early stages of building design. Energy and buildings, 42(7), 1113-1119.

Riveiro, M., Falkman, G., & Ziemke, T. (2008, July). Visual analytics for the detection of anomalous maritime behavior. In Information Visualisation, 2008. IV'08. 12th International Conference (pp. 273-279). IEEE.

Russell, A. D., Chiu, C. Y., & Korde, T. (2009). Visual representation of construction management data. Automation in Construction, 18(8), 1045-1062.

Simpson, T. W., Spencer, D. B., Yukish, M. A., & Stump, G. (2008, September). Visual steering commands and test problems to support research in trade space exploration. In 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (pp. 10-12).

Siirtola, H. (2007). Interactive visualization of multidimensional data. Tampereen yliopisto.

Stump, G., Yukish, M., Martin, J. D., & Simpson, T. W. (2004, August). The ARL trade space visualizer: An engineering decision-making tool. In 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (Vol. 30).

Sullivan, K. J., Griswold, W. G., Cai, Y., & Hallen, B. (2001, September). The structure and value of modularity in software design. In ACM SIGSOFT Software Engineering Notes (Vol. 26, No. 5, pp. 99-108). ACM.

Ware, C. (2004). Information Visualization: Perception for Design (Interactive Technologies). Morgan Kaufmann 2nd edition.

Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. Journal of the American Statistical Association, 85(411), 664-675

Wiederhold, G. (1992). Mediators in the architecture of future information systems. Computer, 25(3), 38-49.

Wood, K. L., Otto, K. N., & Antonsson, E. K. (1992). Engineering design calculations with fuzzy parameters. Fuzzy Sets and Systems, 52(1), 1-20.

Yan, X., Qiao, M., Li, J., Simpson, T. W., Stump, G. M., & Zhang, X. (2012, January). A Work-Centered Visual Analytics Model to Support Engineering Design with Interactive Visualization and Data-Mining. In System Science (HICSS), 2012 45th Hawaii International Conference on (pp. 1845-1854). IEEE.