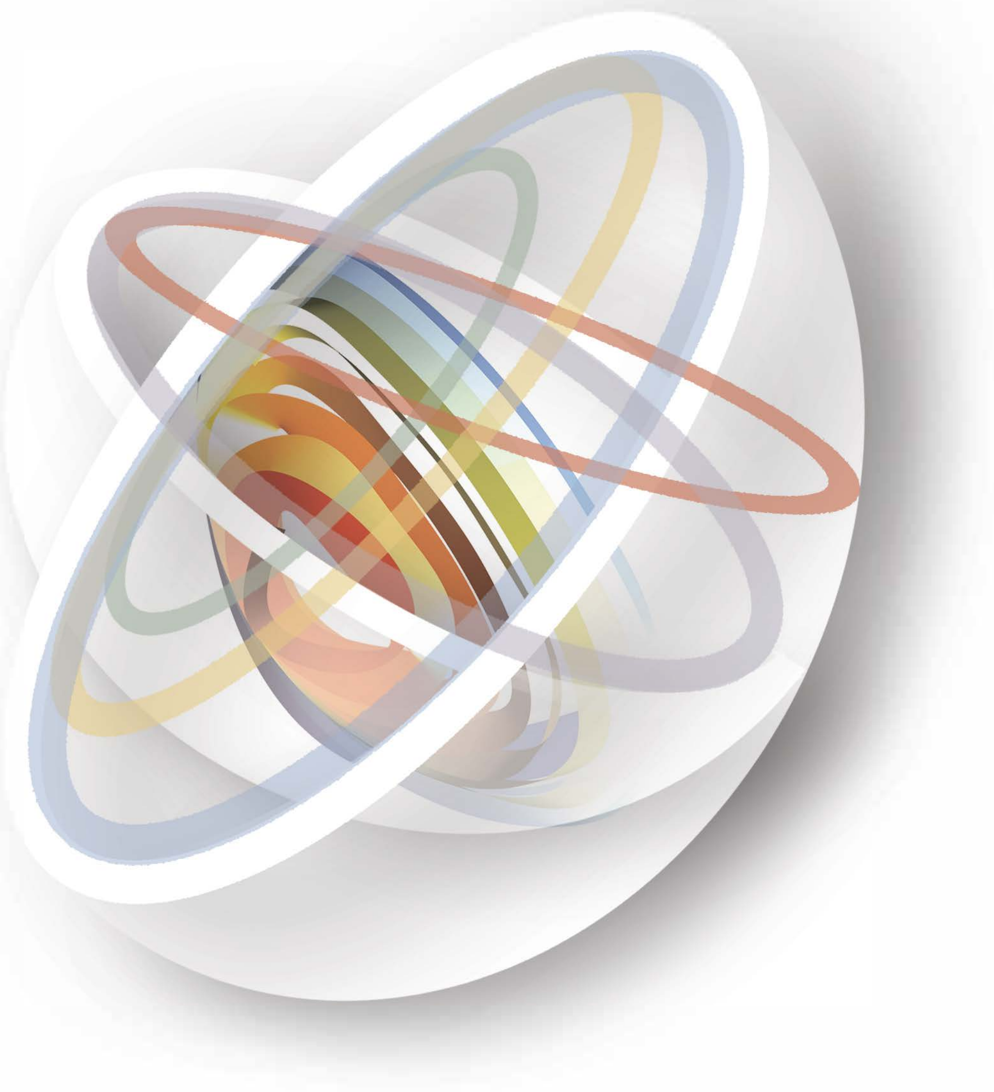
International Conference for

Sustainable Design of the Built Environment SDBE 2018

Proceedings



Editors

Heba Elsharkawy
Sahar Zahiri
Jack Clough

Usability assessment of building performance simulation tools: a pilot study

Stefano Cozza[1], Raphaël Tuor[2], Andreas Sonderegger[3], Thomas Jusselme[1], Denis Lalanne[2] and Marilyne Andersen[1,4]

[1] Building 2050 Research Group, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland.
[2] Human-IST Institute, University of Fribourg, Switzerland.
[3] EPFL+ECAL-Lab, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland.
[4] Laboratory of Integrated Performance in Design (LIPID), School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland.

Abstract: Due to climate change, the built environment is facing increasingly strict environmental targets. Thus, architects are challenged to design evermore high-performing buildings, a task for which they can no longer depend solely on their experience and intuition. Building performance simulation (BPS) tools have become central in this context to support the design process.
Yet, several studies show that such tools are still not widespread among practitioners at early design stages. Despite significant efforts made to deliver more “architect-friendly” tools, a gap remains between the expected use and the reality, highlighting the need to adapt the design-approach when developing such tools. A user-centred design approach seems promising for increasing the usability and acceptance of BPS tools, and should be fine-tuned through multiple iterations between BPS developers and potential users via usability assessments. However, as usability assessment has its origins in the domain of human-machine interaction, no methodology has been proposed yet specifically for BPS tools. This paper is the result of a first interdisciplinary pilot study, describing and evaluating a usability assessment method for a new BPS tool that supports the low carbon building design process. Usability, the reliability of the tool and its usefulness are amongst the dimensions that have been assessed with a selected population of future users. Moreover, recommendations and guidelines for the reproducibility of the test are provided. The study shows that both, the quantitative and qualitative results gathered through a usability assessment are insightful to develop a BPS tool that is efficient, satisfactory, pleasant to use and widely adopted by designers.

Keywords: UX, Usability, Building Performance Simulation, User Test

Introduction

It has been estimated that, in Europe, built environment accounts for more than 40% of total CO_2 emissions and 45% of the total final energy demand (Sutherland et al., 2015). Thus, energy efficiency in buildings has become a key energy policy in order to reach the environmental targets settled for 2050 (EU Parliament, 2014). To meet these ambitious objectives, architects and engineers are increasingly challenged and can hardly depend solely on their experience and intuition. Consequently, building performance simulation (BPS) tools have a great potential in supporting the design process (Hensen, 2004; Negendahl, 2015).

The use of BPS has significantly grown in the last 10 years, mostly due to the use of environmental performance rating tools to obtain building certification. An important portion of assessments in LEED and BREEAM in Europe, or MINERGIE in Switzerland, are based on the building’s energy use, predicted by using BPS. However, these tools are supposed to assist architects during the whole design process, especially in the early design phase, in which important parameters affecting the building performance are addressed.

Despite significant efforts made to deliver more “architect-friendly” tools, several studies show that such tools are still not widespread among practitioners and most of the time are used only to obtain a certification (Attia et al., 2012; Bleil de Souza et al., 2015; Soebarto et al., 2015). The same authors identified the poor usability of these tools as the main reason of the gap between their expected use and the reality. Most architects who use BPS tools in design practice are very concerned with the usability of interfaces, generally criticizing the ease of use of graphical user interfaces (GUI) deployed in this domain (Hensen, 2004; Reinhart et al., 2006; Shi et al., 2013).

Assessment of usability is an important exercise in the domain of human-machine interaction. Usability assessment, generally conducted throughout the whole development process, requires experts, developers and real user test (UT) to assess usability of a prototype (Lewis, 2006). While the method is well established and widely used in systems design, no methodology has been proposed and evaluated so far addressing particular requirements of BPS tools. This paper proposes a complete usability assessment method for BPS tools, and the results of its first application on a small sample of users.

State of the Art

Several authors have hence criticized poor usability as a common and constant negative feature of existing tools. Six BPS tools were compared in Weytjens et al., (2011) showing that no single tool is entirely satisfying architects' needs, mainly due to poor communication and visualization of the results. Farzaneh et al., (2015) examined three different tools, and found that most often preference was given to the one with the best graphical representation of outputs and inputs. Both the above results are confirmed in an extensive comparison of the ten major BPS tools on the market conducted by Attia et al., (2009). Findings show that architects need a tool that provides a comprehensible graphical representation and a simple navigation.

Comparable conclusions have been drawn in various studies based on interviews and web-based surveys aiming at the assessment of problems in the daily use of BPS by architects (Hopfe et al., 2006; Reinhart et al., 2006; Soebarto et al., 2015). Results indicate that users look for better tools with more user-friendly interfaces and with a shorter learning curve (Mahdavi et al., 2003).

Despite the obvious importance of usability in the process of designing software and interfaces, findings of such studies indicate that developers of BPS tools have not entirely addressed this issue so far. Usability design signifies a constant improvement of the interaction of users with a system, by obtaining information about performance (e.g. efficiency and effectiveness of use), the likes, dislikes, needs, emotions and understandings of real users interacting with a system in a real operational environment (ISO 9241-210:2010).

Efforts have been conducted in different fields, but only a few publications relate usability engineering to BPS software (Holzinger, 2005; Hopfe et al., 2009). Previous work has focused on assessing the usability of a BPS tool through a heuristic evaluation demonstrating the feasibility of this expert-based evaluation method (Struck et al., 2010). However, Folmer et al., (2002) conclude their comprehensive review that none of the various usability engineering methods alone (i.e. testing, inspection or inquiry) have the capacity to support BPS software architecture.

Although the results of this literature review clearly highlight the importance of usability engineering for the development of BPS tools, the review showed also that usability seems to remain a problem with existing tools used in practice. One reason for this could be that in the development of such tools, aspects of human-centred usability design are not applied thoroughly. While only few studies have described a complete usability methodology to assess usability of BPS tools, there is a great potential for further research to improve these tools in terms of usability.

Usability and human-centred systems design

To ensure the development of tools that are easy, satisfying and fun to use, it is important to guarantee that the needs and limitations of the user are taken into account throughout the whole development process (Rubin et al., 2008). This "user-centred" approach embodies three main principles of design: a) early focus on users and tasks, b) empirical measurement, and c) iterative design (Gould et al., 1985).

This implies that designers should bear the end user in mind throughout the whole design process. For that purpose, the authors propose to include the user very early into the design process. To do so, users should use prototypes of the product to carry out real tasks while their performance and reactions to the product are observed, recorded, and analysed.

With regard to the scope of the usability evaluation, two different approaches can be distinguished: formative and summative usability testing. Formative testing is conducted throughout the whole development phase in an iterative design process with the goal to gather qualitative information about weaknesses and operation problems of a product. Summative testing on the other hand aims to collect quantitative data about the accomplishment of task goals. It is often conducted at the end of specific phases in the product development process or at the end of the development process (Lewis, 2006; Rubin et al., 2008).

However, usability testing is not the only technique applied in product development practice that allows evaluating the usability of a product. Other popular methods are, for example, cognitive walkthroughs, heuristic evaluations, checklists or interviews, and focus groups (Jordan, 1998; Kuniavsky, 2003; Nielsen, 2003). Each of those methods feature some advantages and disadvantages. Nevertheless, usability tests and focus groups are the only methods that consider representative end users and provide empirical data as requested in the user-centred design principles (Gould et al., 1985). Therefore, usability testing and focus groups are among the most important and widely applied methods in usability practice.

The case study, ELSA

In order to realize efficient buildings with a low environmental impact, the analysis of the energy performance must be integrated since the early design stage. Designers are then facing the problem of quantifying and foreseeing the CO_2 emissions of a building when a lot of major architectural and technical solutions (under the environmental point of view) are still undefined. This leads to a time mismatch between the moment a building is designed and major decisions are taken, and the moment when a complete knowledge of the consequences of a design choice is gained.

We proposed in a previous study a methodology to address this different detail resolutions in energy simulation (Jusselme et al., 2016). This method has then been put into practice in the research framework of the smart living lab (Jusselme et al., 2017), developing a first functional prototype called ELSA, which stands for Exploration tool for Sustainable Architecture (elsa.epfl.ch). This tool couples parametric energy simulation with sensitivity analysis (SA), in order to provide useful information about the connections between different design parameters and their influence on the performance of the building. In order to perform a SA, it was necessary to assess a large number of design alternatives. This led to the creation of a large database, with thousands of alternatives. To explore the database, several innovative data visualization techniques have been tested in order to find the most suitable one to stimulate the user to investigate the design space (Jusselme et al., 2017).

It was indeed crucial to evaluate the usability of the different GUI, unveiling strengths and weaknesses of each. Accordingly, we performed a UT in order to receive constructive feedback for further development of the tool. Two different GUI were compared. The first one, in Figure 1, uses the Parallel Coordinates (PC) visualization technique, which allows to interact with a large number of parameters and to explore relationships among these.

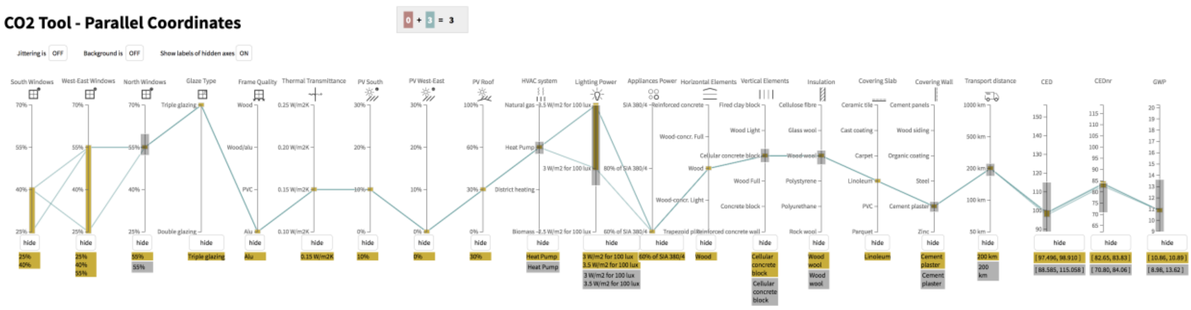


Figure 1. Parallel Coordinates

This visualization method is expected to empower the designer with data analysis capabilities that are important in the early stages of building design. Each vertical graduated axis represents one parameter, with each value being displayed as one tick. A design alternative is represented by a polyline crossing all axes at the corresponding parameter values. With the use of brushing methods, the user can specify a range of values on each dimension. This reduces the number of design alternatives displayed and unclutters the overview. More details on the use of Parallel Coordinates to display building energy performance data can be found in Jusselme et al., (2017). The second GUI, depicted in Figure 2, is called Stacked Coordinates (SC). The same has been specifically designed for this case study by the EPFL+ECAL-Lab Design research Center (Koller and Florentine, 2016).

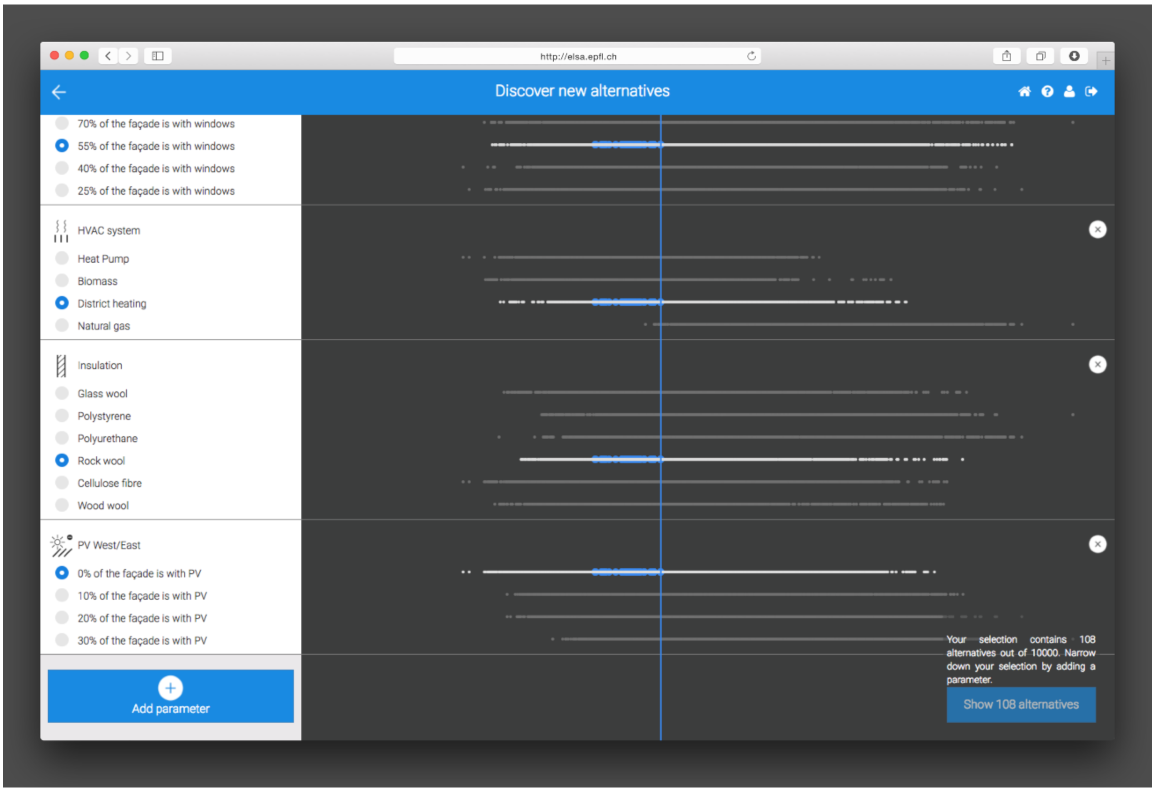


Figure 2. Stacked Coordinates

Using SC, users can take one design parameter after another and assess the impact of each of them. Not only how they affect building performance, but also how they reduce the number of solutions for all the next choices still to come. The interface suggests the most influential design parameters, guiding the users through the desiderata options.

User test objectives

The study was intended as a first and unique pilot work to unveil the complexity of running a usability study for BPS. Moreover, the UT was meant to improve the usability of the ELSA tool under development with the results of our summative and formative tests. We defined the key dimensions to assess through the UT:

- User satisfaction and perceived usability.
- User learning.
- Affect and emotions.

Along these dimensions, we wanted to compare the usability and usefulness of two BPS tools in the early design phase. We wanted to verify the three following hypotheses:

1. Whether the use of SC or PC increases the feeling of control of the architect towards design-oriented tasks.
2. Whether using SC or PC increases (and to what extent) the knowledge of the user regarding environmental performance.
3. Whether the SC receives higher usability ratings than PC in the user evaluation thanks to its decision-tool oriented features including suggestions of the most impactful parameters and the constructivist approach, which help to reduce cognitive load and hence increase user performance and satisfaction.

Study design and procedure

We designed our study as a "within-group" experiment with Master-students in Architecture. The independent variable was the type of visualization tool used: SC or PC, meaning that each participant used both visualization tools to complete the same set of tasks. Participants who used SC in the first session had to use PC in the second session, and conversely. We conducted the study over two different sessions, three weeks apart. Each session was divided into 10 steps. We chose to use both summative and formative approaches already mentioned in the previous section. The 10 steps were as follows:

1 - Assessment of affective state – baseline measure

Emotions are an important component of the user experience. They can strongly affect how a user perceives and appreciates an interface (Agarwal et al., 2009). Thus, affective state was measured based on the Self-Assessment Manikin model (SAM; Bradley et al., 1994) twice during each testing session, before (baseline measure) and after interaction with the system (step 6). On three single items, SAM assesses pleasure (positive – negative), arousal (excited – calm) and dominance (high – low) on a nine-point scale. This instrument provides a quick, reliable and non-linguistic way to assess the emotions of the participant related to their response to the tool.

2 - Introduction class

The performance of the user in building design tasks using BPS software is influenced by his knowledge regarding the design of energy-efficient buildings. Since one of the goals of this study was to assess the gain in knowledge thanks to the two tools proposed, this part consisted in giving some information to the user about the context of use of the tool, while giving minimal information about BPS. We explained the goal of the user test and what was expected of the participants. We gave background information about the ELSA project, followed by an explanation of how an energy assessment tool can be used in order to design energy-efficient buildings. We were careful not to go into details on environmental concepts, to avoid bias of knowledge tests' results (steps 3 and 8). For the sake of reducing effects linked to social comparison, we informed participants that we would evaluate the performance of the two groups rather than their individual performance (Buunk *et al.*, 1990).

3 - Knowledge test I

In order to assess the usability of BPS, a summative approach seems essential. This test allows to determine whether the user gains knowledge by using any BPS tool. *Learning* was measured via a knowledge test. This test consisted in 10 multiple-choice questions (e.g. "Which window type is best regarding CO_2 emissions?") assessing participants' knowledge regarding environmental awareness and the design of energy-efficient buildings before an interaction with the tool took place. A similar test was conducted a second time after using the tool in order to assess the learning outcome of tool usage. The second time we randomized the order of questions to avoid any bias related to the order, and to minimize priming effects (Tourangeau et al., 2000) (step 8).

4 - Tutorial

At this step we divided the class into two groups. The group was determined by the ID randomly assigned to each participant at the beginning of the session. Each group was assigned to one of the two visualization tools, and we explained its functioning. Participants used the tool on their own computer (Mac and PC laptops) in order to reduce the extraneous cognitive load as much as possible (Chandler et al., 1991). Their working environment was thus as familiar as possible: they used their own computer, devices and operating system. We set up a video-capture tool on the computer screen of each participant. Screen recordings allowed to measure the time needed to accomplish each task.

5 - Tasks

Again, the summative approach was used in this step. This step allows measuring the performance of the user in a set of tasks that are inherent to the design of energy efficient buildings. It consisted of six specific tasks. Participants were asked to perform different practical activities using only the tool. The tasks were classified into four types:

- Exploration: to find out the relations between different design parameters and the environmental performance of the buildings. Discover and understand how different design choices affect the energy consumption.
- Configuration: to introduce and evaluate their own project. Participants were asked to input the project they had in mind to realize their building and estimate its environmental performance.
- Improvement: to reduce the total environmental impact of their project. Once that a first assessment of the project was done, users were requested to change some design choices to improve the overall performance.
- Frequency: to create a project following some specific constrains. In this case the freedom of the users was limited to few parameters to quickly reach the desired result.

6 - Emotional Test II

After all tasks were completed, we asked participants to rate their affective state using the SAM a second time, to measure the emotional response after the use of the tool.

7 - UX Questionnaire

For this step, we chose to use another summative approach and we asked the participants to rate the *perceived usability* of the tool using the widely accepted System Usability Scale (SUS; Brooke, 1996). This instrument can be considered an industry standard and offers several advantages: it can be used on small sample sizes, it is reliable and it can effectively differentiate between usable and unusable systems (Bangor et al., 2008). Questions in this test include: “How much did you trust in the information?”, “How much did you like the tool?”, “Would you use it on a regular basis?”.

8 - Knowledge test II

9 - Break

10 - Focus Group Discussion

The end of the session seems to be the ideal time to run a formative test: users have been using the tool for an extended period of time, and they identified weaknesses and strengths of the BPS tool. In our experiment, we gathered each group in a separate room and asked for their feedback regarding the tool and its usefulness in the early design stages. More specifically, with the scope of a formative usability evaluation (Lewis, 2006), this discussion was intended to help understanding which features were considered useful and which not, and how the interface could be improved in order to suit the needs of architects. The questions we asked were all the following:

- How would you use this tool in your future practical work?
- What was useful, what was not?
- What do you think about the design of the interface? What could be improved?
- Do you think that this tool might improve the communication with stakeholders in future project?
- What makes you trust the tool and the information it provides?

Results

In this section, we present the results of the pilot study, conducted over the two sessions (I and II). Ten subjects took part in Session I, and eight took part in Session II. The whole study was designed as a within-group experiment, therefore we removed the answers of two subjects in the first session since they did not attend the second session, for a total of eight participants. Due to the small sample size typical in usability tests, most statistical comparisons did not show significant differences. Therefore, effect sizes and confidence intervals are reported in addition.

System Usability Scale (SUS) Questionnaire

Analysis of the data on perceived usability (c.f. Figure 3) indicated that both systems do not differ with regard to their obtained ratings on the SUS scale ($t(7) = 0.08$, $p > 0.05$) with a Cohen's d (using pooled variance) of 0.04, representing a very small effect. Although no difference in the usability-evaluation was observed, confidence interval for the usability-ratings of the PC tool was considerably larger than for SC.

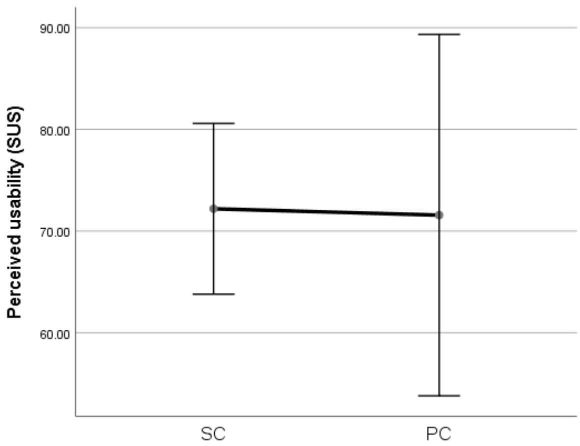


Figure 3. Average Usability score for each question. Error bars represent 95% Confidence Interval.

The total usability score for the two visualizations was very similar as the Stacked Coordinates earned 72.2 (*SD* = 10.0) while the Parallel Coordinates earned 71.6 (*SD* = 21.3). According to established norms (Bangor et al., 2008), both scores can be considered "excellent".

Emotional reaction

Participants had to report their experienced emotions before and after using the tool. Figure 4 displays the change scores (score after interaction – baseline measure) for each of the three SAM measures. Results indicate that measures of pleasure were very similar for both tools. A paired t-test indicates no significant differences ($t(7) = 0.48$, $p > 0.05$, $d_{pooled} = 0.25$, small effect). Similar results were obtained for the measure of arousal ($t(7) = 0.67$, $p > 0.05$, $d_{pooled} = 0.31$, small effect), and dominance ($t(7) = 1.13$, $p > 0.05$, $d_{pooled} = 0.50$, medium sized effect). Interestingly, the slightly higher dominance ratings for the PC tool show a smaller confidence interval than the dominance-ratings for the SC tool.

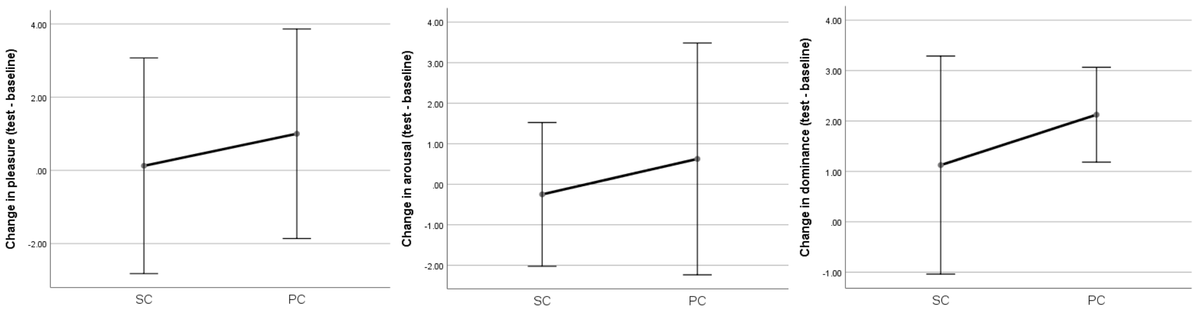


Figure 4. Average scores of SAM items for Parallel and Stacked Coordinates, before and after use, by type of emotion. Error bars show 95% confidence intervals.

Performance indicators

The main function of the tools is to support architects to improve environmental performance of their designs. One task required participants to input their own project and then improve it by using the BPS tool. A comparison of the change in energy consumption after use of the tools indicates that improvements in energy consumption were higher when participants used PC compared to SC (Figure 5). This effect was significant ($t(7) = 2.36$, $p < 0.05$) with a Cohen's *d* (using pooled variance) of 0.945, representing a large effect.

Additionally, in a configuration task, participants were asked to filter out different design parameters and find one design alternative complying with the given environmental indicator threshold. With SC, participants needed lesser parameters compared to PC (see Figure 5), indicating that SC requires significantly less steps to reach a given target ($t(7) = 2.87$, $p < 0.05$) with a Cohen's d (using pooled variance) of 0.961, representing a large effect.

In the frequency task, participants had to input a set of parameter values, and then answer with the exact amount of alternatives left. SC users gave 50% of correct answers, and PC users gave 12.5% of correct answers, non-parametric analysis (Wilcoxon signed-rank test due to violation of normality-assumption) however indicates that this difference is not significant ($T+ = 12.0$, T- = 3, $z = 1.342$, $p > 0.05$, $r = 0.335$, representing a medium effect).

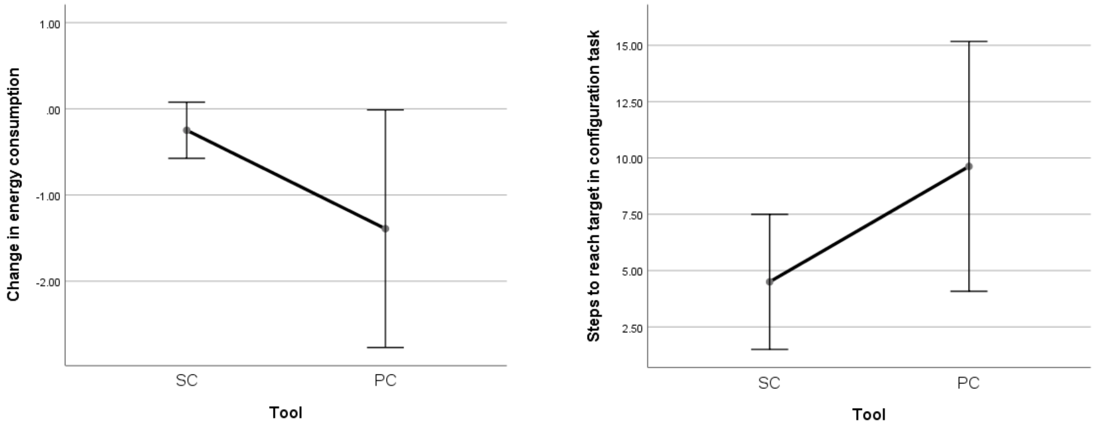


Figure 5.Performance indicators after use of SC or PC tool. Error bars show 95% confidence intervals.

Knowledge test

We computed the total score for all the questions of the knowledge test. The mean values and 95% CIs are shown in Figure 6. Overall, the scores increased for both tools, meaning that participants gained knowledge while using them. A repeated ANOVA measure indicated that this effect was statistically significant ($F\ (1,7) = 7.0$, $p > 0.05$, $\eta^2_p = 0.5$). The comparison between the two tools did not reach significance level ($F\ (1,7) = 2.3$, $p > 0.05$), neither did the interaction of tool x time ($F\ (1,7) < 1$, $\eta^2_p = 0.045$). Interestingly, CI for PC increases after interaction with the tool compared to SC which decreased.

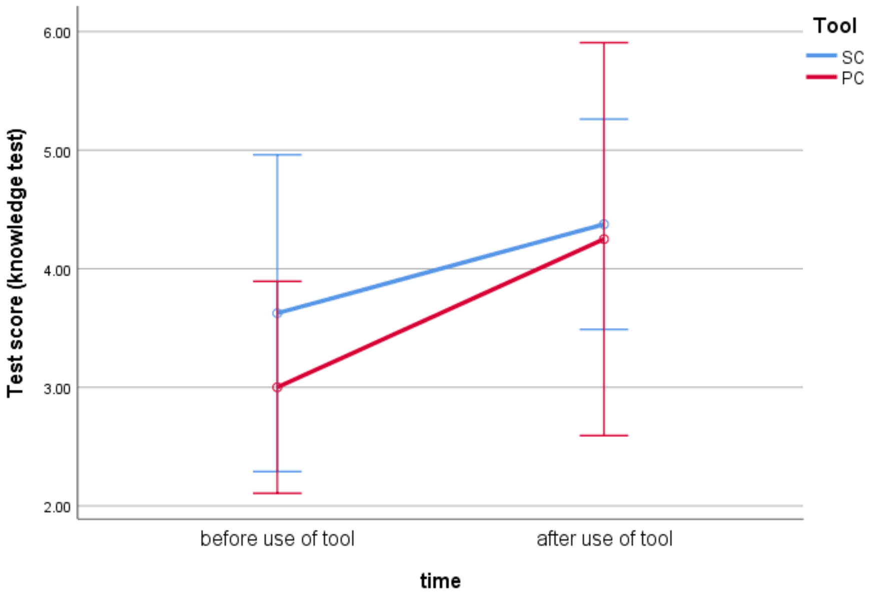


Figure 6. Final score in the knowledge test. Error bars show 95% confidence intervals

Focus Group

The group discussion was very useful to gather some qualitative results about the whole experience and the two GUI. In this moment the users felt free to share their impressions and we could really understand the strengths of both visualization methods.

The main strengths of the Parallel Coordinates can be summarized through the following statements: "*it is easy to track a single project from the left to the right across each polyline*", "*it is useful to see how, at the beginning of the design process, just changing a parameter the building performance can vary a lot*", and "*it seems very professional because it uses a scientific way to represent the data*". On the other side the principal weak point of this visualization is that "*at first sight it is overwhelming, looks too complicated*".

Instead, in the Stacked Coordinates, proposing parameters through a sensitivity analysis was really appreciated: "*It is nice to have the two ways of describing the project (as you want and per sensitivity), because it shows us that we should think differently regarding the parameter definition into the design process*". In the case of the SC the novelty in the design has been turned into a down factor, as "*if you go to a client with it they will think it is more commercial and they will not trust the results so much*".

Discussion

The results of this pilot study have enhanced our understanding of usability in building performance simulation tools. Moreover, the work showed the great potential of a usability assessment for increasing the acceptance of BPS tools. The combination of the quantitative results gathered through a summative approach, and the qualitative ones obtained through a formative approach, have brought benefits in the development of ELSA.

The summative approach was positively used to assess the user learning, user satisfaction, and perceived usability of the tools. The quantitative results to evaluate these dimensions were obtained through the knowledge test, the performance indicators, the SAM, and the SUS. These methods represent an innovative application of the human-centred systems design approach applied to BPS tools and offer considerable insights regarding the two GUI. The Stacked Coordinates provided a better understanding of all the features of the parameters and the process, as showed by the results of the configuration tasks. Moreover, SC implied a lower cognitive load, providing more guidance to reach the objectives, as confirmed by the frequency tasks. Parallel Coordinates instead gave more interactivity and flexibility, as the whole set of alternatives is visible from the beginning. The result of the improvement task showed a higher improvement of the environmental performance of the project using PC.

Alternatively, the formative approach was used to better understand the wishes and the requirements of the users, based on which it has been possible to continue the development of ELSA. The group discussion, combined with the summative results, lead us to think that coupling both SC and PC strengths into a single tool could empower the user. Indeed, PC helps to give an overview on the dataset and the correlations between dimensions, while SC helps the user to choose the best parameters and reduces cognitive load with its constructivist "building blocks" approach. Once the prototype of the new tool is available, it will have to be tested again with this UT, in order to compare the results with the previous versions and observe the improvements achieved through this usability assessment.

Conclusion

The literature review revealed that despite usability being a crucial aspect in the process of designing software and GUI, so far no usability assessment method has been proposed specifically for BPS tools. This paper aimed to fill this gap, proposing a complete User Test to assess the usability of the new developed energy tool, ELSA. According to the user-centred design approach the usability of a product should be tested very early in the development process. Therefore, two functional prototypes were provided to the users, based on two visualizations methods: Stacked and Parallel Coordinates.

The objective of the UT was to verify three hypotheses regarding the usability and usefulness of two BPS tools in the early design phase. The first one, stating that the use of SC and PC increases the feeling of control of the architect towards design-oriented tasks, was partially confirmed by the results of the SAM test. Both visualizations indeed increase the dominance of the user after a period of use. The second hypothesis was instead fully validated by the knowledge test. The user gains expertise on the design of energy-efficient buildings by using both PC and SC, as the results of the knowledge test increased for both visualizations. Finally, the third hypothesis was verified using the SUS test. A higher usability rating was found for the SC, due to its decision-tool oriented features. However both SC and PC achieved high scores in the System Usability Scale.

Based on those findings, ELSA will be developed further combining the key features of the two proposed GUI, with the aim to create a new prototype which can be tested with users again. This iterative process of user testing and further development is crucial in the user-centred approach and should lead finally to a functional product that corresponds with the requirements of the user. The essential role of usability tests in such a user-centred design approach found in the literature is as well confirmed for the case of BPS and indicates the importance of this evaluation method for the development of usable tools.

Acknowledgment

The work presented in this paper has been funded by the State of Fribourg (message du Conseil d'Etat au Grand Conseil 2014-DEE-22). We thank Andreas Koller for his valuable contribution to this work.

References

Agarwal, A. and Meyer, A. (2009) Beyond usability: evaluating emotional response as an integral part of the user experience. in *CHI '09 on Human Factors in Computing Systems*.

Attia, S. *et al.* (2009) Architect Friendly: a Comparison of Ten Different Building Performance Simulation Tools. in *Eleventh International IBPSA Conference*, pp. 1–8.

Attia, S. *et al.* (2012) Selection criteria for building performance simulation tools: Contrasting architects' and engineers' needs. *Journal of Building Performance Simulation*, 5(3), pp. 155–169.

Bangor, A., Kortum, P. T. and Miller, J. T. (2008) An Empirical Evaluation of the System Usability Scale. *Intl. journal of Human–Computer Interaction*, 24(6), pp. 574–594.

Bleil de Souza, C. and Tucker, S. (2015) Thermal simulation software outputs: a framework to produce meaningful information for design decision-making. *Journal of Building Performance Simulation*, 8(2), pp. 57–78.

Bradley, M. M. and Lang, P. J. (1994) Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), pp. 49–59.

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

Buunk, B. P. *et al.* (1990) The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology*, 59(6), pp. 1238–1249.

Chandler, P. and Sweller, J. (1991) Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, 8(4), pp. 293–332.

EU Parliament (2014) Energy Efficiency and its contribution to energy security and the 2030 framework for climate and energy policy. pp. 1–8.

Farzaneh, A., Monfet, D. and Forgues, D. (2015) Usability and information management of energy simulation inputs: a comparison between 3 tools. in *14th Conference of International Building Performance Simulation Association*.

Folmer, E. and Bosch, J. (2002) Architecting for usability: A survey. *The Journal of Systems and Software*, 70(1–2), pp. 61–78.

Gould, J. D. and Lewis, C. (1985) Designing for usability: key principles and what designers think. in *Communications of the ACM*, pp. 50–53.

Hensen, J. L. M. (2004) Towards more effective use of building performance simulation in design. in *Proceedings of the 7th international Conference on Design & decision support systems in architecture and urban planning*, pp. 2–5.

Holzinger, A. (2005) Usability engineering methods for software developers. *Communications of the ACM*.

Hopfe, C. *et al.* (2006) Computational Simulation Tools For Building Services Design – Professional's Practice and Wishes. in *17th International Air-conditioning and Ventilation Conference*.

Hopfe, C. and Hensen, J. (2009) Experiences testing enhanced building performance. in *Eleventh International IBPSA Conference*, pp. 1976–1982.

ISO 9241-210:2010 (2010) Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems.

Jordan, P. W. (1998) *An Introduction To Usability*.

Jusselme, T. *et al.* (2016) Towards a Pre-Design Method for Low Carbon Architectural Strategies The case of the smart living building in Fribourg, Switzerland. in *PLEA 2016 Los Angeles - 32th International Conference on Passive and Low Energy Architecture.*

Jusselme, T., Brambilla, A., *et al.* (2017) *Smart living building research program - Executive summary*.

Jusselme, T., Tuor, R., *et al.* (2017) Visualization techniques for heterogeneous and multidimensional simulated building performance data sets. in *International Conference for Sustainable Design of the Built Environment- SDBE London 2017*.

Koller, A. and Florentine, J. (2016) *Report on User Perception, Functionality and Visual Strategies*.

Kuniavsky, M. (2003) *Observing the User Experience: A Practitioner's Guide to User Research*.

Lewis, J. R. (2006) Usability Testing. in *Handbook of Human Factors and Ergonomics*, pp. 1275–1316.

Mahdavi, A. *et al.* (2003) An inquiry into the building performance simulation tools usage by architects in Austria. in *Eighth International IBPSA Conference*, pp. 777–784.

Negendahl, K. (2015) Building performance simulation in the early design stage: An introduction to integrated dynamic models. *Automation in Construction*, 54, pp. 39–53.

Nielsen, J. (2003) Usability 101: Introduction to Usability, pp. 1–13.

Reinhart, C. and Fitz, A. (2006) Findings from a survey on the current use of daylight simulations in building design. *Energy and Buildings*, 38(7), pp. 824–835.

Rubin, J. and Chisnell, D. (2008) *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*.

Shi, X. and Yang, W. (2013) Performance-driven architectural design and optimization technique from a perspective of architects. *Automation in Construction*, 32, pp. 125–135.

Soebarto, V. *et al.* (2015) Capturing the Views of Architects About Building Performance Simulation To Be Used During Design Processes. in *Conference Proceedings Building Simulation 2015*, pp. 1480–1487.

Struck, C., Hensen, J. and Plokker, W. (2010) Usability evaluation of a prototype design tool for uncertainty propagation and sensitivity analysis. in *Proceedings of the 3rd German/Austria IBPSA Conference, 22-24 Sept. 2010 Technical University of Wien*, pp. 22–24.

Sutherland, G. *et al.* (2015) *Implementing the Energy Performance of Building Directive.*

Tourangeau, R., Rips, L. J. and Rasinski, K. (2000) *The Psychology of Survey Response*, *Cambridge University Press*.

Weytjens, L. *et al.* (2011) The "Architect-friendliness" Of Six Building Performance Simulation Tools: A Comparative Study. *International Journal of Sustainable Building Technology and Urban Development*, 2(3).

De Wilde, P. and Prickett, D. (2009) Preconditions for the use of simulation in M&E engineering. in *Proceedings 11th International IBPSA*, pp. 414–419.

ISBN: *978-1-9997971-9-5*
Project website: *www.newton-sdbe.uk*
Twitter: *@Newton_SDBE*