USER-BASED EVALUATION OF AN INTERACTIVE EXPERT SYSTEM FOR FULL-YEAR DAYLIGHTING DESIGN SUPPORT

Jaime M.L. Gagne¹, Marilyne Andersen¹,²

¹Building Technology Program, Department of Architecture, Massachusetts Institute of Technology (MIT), Cambridge MA, USA
²Interdisciplinary Laboratory of Performance-Integrated Design (LIPID), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

contact email: marilyne.andersen@epfl.ch

ABSTRACT

This paper presents the results of an original design process-oriented user study conducted on an interactive expert system specifically developed for full year, climate-based daylighting design support. The aim of the study was to determine how well its decision-making algorithm would work when independent human interactions and decisions were included into the process through a two-step process, first based on design heuristics then using the expert system. The results of this evaluation demonstrate that the expert system is generally successful as a performance-driven design tool and as a method for influencing and educating designers in ways that they can improve the daylighting performance of their designs. It also demonstrates the relevance of the proposed user study to validate design support tools that account for the unpredictability of the inherently ill-defined architectural design process itself.

INTRODUCTION

Designing spaces that are able to balance illumination, glare and solar gains over a whole year is a real challenge, yet a problem faced every day by building envelope designers. The main challenge resides in the reconciliation of the many factors influencing how daylight and sunlight each interact with the built environment and in the great variations they show in intensity and distribution depending on location, weather and time. To assist architectural designers in this search, a full year daylighting simulation method was developed, called Lightsolve, meant to be used early on in the design process when façade and space details have not yet been defined. It takes a new perspective on daylighting analysis, focusing on the variation of daylight performance over the day and the year by combining temporal performance visualization with spatial renderings (Andersen et al., 2008; Kleindienst & Andersen, 2012), and includes an expert system to support a guided search process (Gagne et al., 2011).

One of the underlying principles in terms of how daylighting performance is evaluated in Lightsolve is, on one hand, to make it specific to the user’s own performance objectives and to his or her chosen areas of interest within the considered design project, and on the other hand to combine a synthetic perspective of full-year data with a visual impression of what the space looks like over time. Examples of interactive search methods that would accept user input and grant him or her a large degree of control include human-guided search algorithms (Anderson et al., 1999), as well as knowledge-based (Fazio et al. 1989) or expert systems (Luger, 2004), in which human expert knowledge about a specific domain is encoded in an algorithm or computer system (Zemler et al., 2007; Kalay, 2004; Mashood et al., 2007). In the daylighting domain, such a system would function as a virtual consultant, guiding the designer towards design modifications that improve overall daylighting performance. Knowledge-based systems have already been successfully implemented for artificial lighting (Guo et al. 1993; Jung et al., 2003) and in simplified scenarios for daylighting (Paule et al., 2011; Ochoa & Capeluto, 2009). The expert system considered in the present paper differs from these previous efforts in that it allows a comprehensive understanding of daylighting and offers user interactivity regarding design choices. Details about this approach have been published in (Gagne et al., 2011) and will only be briefly summarized here.

The Lightsolve expert system is intended for early design stage exploration and performance-based decision support while respecting the role of the architect and his or her design intent. The expert system allows designers to create a 3d model of their own design and to input project-specific performance goals for illumination and glare within the space. It uses a fuzzy rule-based logic (Siler & Buckley, 2005) in combination with a database of pre-computed simulation data (Gagne & Andersen, 2011) in order to create a list of potential design changes that could improve the daylighting performance of the initial design. The system has been implemented in Google SketchUp as a part of the Lightsolve project. A user interface has also been developed which displays the current performance of the design and the list of suggested design changes to the user (Figure 1). In a previous paper (Gagne et al., 2011), we were able to demonstrate that the expert system is able to propose design recommendations that improved building performance when automatically accepted in the absence of user interaction. The aim of this paper is to assess the system for use by human designers.
The participants were also selected so as to represent a variety of backgrounds in daylighting. They were asked to rate their experience level in working with daylighting using one of 4 categories: experienced, intermediate, novice, or none. The group was fairly evenly split, with four (self-ranked) experienced daylighters, three intermediates, and five novices. No participant chose “no daylighting experience”.

**Study Procedure**

The user study was conducted as a series of brief segments, which included three design sessions and two questionnaires. The total amount of time used for each participant was approximately two hours. The design problem will be described in detail in the next section. The same design problem was solved by each participant three times: first by hand, then using the expert system, and finally by hand again. During all sessions, participants were allowed access to pencils, blank paper, a calculator, and a stereographic sun course diagram for Boston, MA.

The sessions of the study were organized as follows:

- Introductory Questionnaire (basic information on design and daylighting background)
- General Tutorial (including task description and brief explanation on daylighting metrics)
- Design Session #1 (design problem to complete by hand from an initial massing model and daylighting performance goals)
- Design Session #2 (same design problem using only the expert system)
- Design Session #3 (final version of design)
- Final Questionnaire (satisfaction with the final design, experience of using the expert system; use the tool in a real design context).

In Design Session #1, the participant was asked to design two façades on the massing model and attempt to meet the daylighting performance goals as well as satisfy him- or herself as a designer. An example of a previous façade design was provided, along with the performance of that example design (Figure 2a). During this design session, the participant was asked to sketch his or her design by hand and to draw the final design on a template sheet.

In Design Session #2, the participant began with the same example model and was allowed to choose to accept or decline design changes suggested by the expert system, to choose the magnitude of the design change, and to return to previous design iterations. The participant was also allowed to explore designs which resulted in decreased performance if desired. During this session, the participant was not allowed to change the design by hand or in SketchUp.

In Design Session #3, the participant was told that this design would be considered “final”. He or she

**METHODOLOGY**

To evaluate the performance of the expert system when incorporated into the design process, a user study was conducted during which designers were asked to interact with the system and solve a daylighting design problem as a three step process: first, using their own intuition, next, using the expert system, and finally, again using their own intuition.

This format was designed to determine whether the use of the expert system was able to improve the daylighting performance of each participant’s second and final designs when compared to his or her initial design.

Several important results were expected from this study. The first is an assessment of the ability of the expert system to find designs with improved daylighting performance when a human user is allowed to interact with it in an independent way. The second is an evaluation of the expert system process as a method for improving a designer’s intuition about daylighting and influencing him or her to consider design elements which result in good daylighting performance. The third focuses on user satisfaction and the acceptance of the expert system by designers.

**Profile of Participants**

Twelve participants were included in the user study. Because the goal of the study was to evaluate the expert system for use by designers, only those with at least one prior degree in architecture and with at least one year of experience working for an architectural design or architectural consulting firm were allowed to participate.

The median work experience in a design firm was 3.8 years (minimum 1 year, maximum 8 years). Of the twelve participants, four had previously completed a bachelor’s level degree in architecture and eight had completed a master’s level degree in architecture. Additionally, most participants were in the process of completing a second or third degree in architecture or a related field at the time of the study.

The user study was conducted as a series of brief segments, which included three design sessions and two questionnaires. The total amount of time used for each participant was approximately two hours. The design problem will be described in detail in the next section. The same design problem was solved by each participant three times: first by hand, then using the expert system, and finally by hand again. During all sessions, participants were allowed access to pencils, blank paper, a calculator, and a stereographic sun course diagram for Boston, MA.

The sessions of the study were organized as follows:

- Introductory Questionnaire (basic information on design and daylighting background)
- General Tutorial (including task description and brief explanation on daylighting metrics)
- Design Session #1 (design problem to complete by hand from an initial massing model and daylighting performance goals)
- Design Session #2 (same design problem using only the expert system)
- Design Session #3 (final version of design)
- Final Questionnaire (satisfaction with the final design, experience of using the expert system, use the tool in a real design context).

In Design Session #1, the participant was asked to design two façades on the massing model and attempt to meet the daylighting performance goals as well as satisfy him- or herself as a designer. An example of a previous façade design was provided, along with the performance of that example design (Figure 2a). During this design session, the participant was asked to sketch his or her design by hand and to draw the final design on a template sheet.

In Design Session #2, the participant began with the same example model and was allowed to choose to accept or decline design changes suggested by the expert system, to choose the magnitude of the design change, and to return to previous design iterations. The participant was also allowed to explore designs which resulted in decreased performance if desired. During this session, the participant was not allowed to change the design by hand or in SketchUp.

In Design Session #3, the participant was told that this design would be considered “final”. He or she
could revisit either or both of the designs produced during the first two sessions or completely start over.

The purpose of the three design session format was to determine if the process of using the expert system was able to positively influence each participant's final design. To create an initial design during the first design session, participants relied primarily on their own intuition and understanding of daylighting. Participants were not told how well these initial designs performed based on the daylighting goals. During the second session, participants all worked with the same starting design, which in many cases was quite different from their own initial design. During this session, however, participants were able to view the performance of the model after they applied various changes to it.

One hypothesis of this study was that if participants chose design changes that resulted in improved performance during the expert system session, they might elect to apply some of those design changes to their designs during the third session. A corollary of this hypothesis was that if participants made these design changes to their own initial designs, the performance of those designs should improve. This improved performance would indicate that participants were able to learn something about working with daylighting by using the expert system, and that the process of using the expert system, even for a seemingly unrelated design but for the same space characteristics and goals, could influence participants to incorporate certain design elements into their own designs.

**Design Problem**

The participants were asked to work through a conceptual design for the façade of a school library wing in Boston, MA which should use natural light instead of artificial light as much as possible. This design problem was developed to be of medium level difficulty so that designers who were experienced with daylighting concepts could create a very good solution using only their intuition. The problem was meant to be challenging for those designers who were not experienced with daylighting, but not so difficult as to discourage them.

Participants were informed that they were taking over the project from a colleague who had already started working on the design. They were required to keep the original massing model that their colleague had designed (footprint, wall heights, and interior walls). However, they were allowed to change the façade elements as necessary to meet the daylighting goals. They were allowed to choose the size and placement of windows, the types of glass used, and the types, size, and placement of shading devices. The two façades that were considered were those oriented towards South and East.

The library space has three main areas: a double-height main study area, which should receive abundant light; a smaller study area that overlooks the main study area, which should receive an adequate amount of light; a rare book room, in which light must be carefully controlled. In each area, an illuminance sensor plane was modeled at work plane height. The original design and the location of sensors within the space are shown in Figure 2. Participants were told that, based on the client’s description of the space, their firm had decided that the specific daylighting goals they should work towards were:

- **Main Study Space:** min 500 lux is desired, down to 400 lux is acceptable; no max.
- **Small Study Space:** min 200 lux is desired, down to 0 lux is acceptable; max 800 lux is desired, up to 1000 lux is acceptable.
- **Rare Book Room:** no min; max 200 lux is desired, up to 400 lux is acceptable.

![Figure 2 Library massing model with (a) South and East example façades and (b) three considered areas](image-url)

Participants were told that if the illuminance on the entire area of a sensor plane falls within the desired range during all daylit times of the year, the performance of that sensor would be 100%. Participants were also provided the average performances of each sensor in the example design:

- **Main Study Space:** 80%
- **Small Study Space:** 65%
- **Rare Book Room:** 70%
- **Average of All Spaces:** 72%

Finally, the participants were informed that the client had requested a certain aesthetic which must be maintained. The following design rules were given:
• Windows must be rectangular or square.
• Glass may be transparent or translucent.
• Glass may not be tinted with color.
• Shading devices must be opaque, and must be vertical or horizontal.
• Both vertical and horizontal shading devices may be used on the same window.
• No advanced systems, such as light-redirecting or responsive systems, may be used.
• It is up to the designer to determine if a uniform façade aesthetic should be maintained.
• The design should achieve the daylighting goals and also satisfy the designer.

Procedure for Modeling Designs
For this study, it was necessary to determine the performance of models from all three design sessions, including those which the participants completed by hand.

Following each participant's study session, the authors constructed SketchUp models of the participant's initial and final designs and calculated the performance on all sensor planes using the LightSolve Viewer (LSV), that was specifically developed as an interactive daylight rendering engine (Cutler et al., 2008).

The dimensions and locations of windows and shading devices on each façade were modeled based on the template sheet drawings. An example template sheet and the corresponding SketchUp model are shown in Figures 3 and 4, respectively.

Figure 3 Example template sheet with façades drawn

For glazing types, participants were allowed to check one option in each of two categories, “View” and “Amount of Light Let In” as described below:
• View: transparent (all specular transmittance), translucent/frosty (equal specular and diffuse transmittance), opalescent (diffuse only).
• Amount of Light Let In: most (e.g. single-glazed clear, \( \tau = 80\% \)), intermediate (e.g. double-glazed low-e, \( \tau = 60\% \)), least (e.g. neutral tint, \( \tau = 40\% \)).

If a participant selected “translucent/frosty” and “intermediate” for instance, the glazing would be modeled as 30% specular transmissivity and 30% diffuse transmissivity.

Figure 4 SketchUp model based on façade drawings

RESULTS
Performance of Designs
To determine the performance of each model produced during the three design sessions based on LSV renderings, the performance was considered to be the percentage of the total area of each sensor plane that the illuminance was calculated to fall within the desired goal range, averaged over the whole year. The performance of each sensor plane was averaged into a single value, which represented the total performance of each design over the whole year. A design which met all goals would be one for which this average value would equal 100%.

Performances for designs produced during each of the three design sessions are shown for all twelve participants in Figure 5. These results have been ordered from least successful to most successful in terms of average whole-year performance across all three illuminance goals for the first design session. Throughout the study, each participant also had access to an example design and its performance. For reference, the performance of this example design is indicated in Figure 5 as a dashed line.

Figure 5 Performance of all produced designs (example design performance shown as dashed line)

During the first session, participants were asked to complete the design problem by hand, using only their intuition. The mean performance of the twelve designs produced during the first session was 73.9%,
which is similar to the performance of the example design shown to each participant (71.9%). Six participants produced designs that performed at least 10% above the example, three participants produced designs that performed similarly to the example, and three participants produced designs that performed well below the example.

During the second design session, participants were asked to use the expert system for a fixed amount of time, starting with the example model with a performance of 71.9%. In general, most participants were able to make four design iterations during the allotted time (one participant was able to make five iterations and one was only able to make three iterations). The mean performance of the designs produced during this session was 87.6%, and the performances of all twelve final designs were higher than that of the example model, which indicates that every participant was able to improve the performance of the starting model by using the expert system. Additionally, eleven out of twelve designs were improved by 10% or more; the design which saw the smallest improvement was created by the participant who was only able to complete three design iterations during the session.

The result that every participant was able to find a better performing design than the starting model in session #2 is important because it demonstrates that the expert system can improve the performance of designs even when the participants’ unique sets of design choices were introduced into the process. Although participants only had a short amount of time to interact with the system, and although participants were not specifically told to choose design changes which improved performance, the expert system was able to find good solutions nevertheless.

It is also interesting to note that each participant made a unique set of design decisions and that while the final performances of some models were similar, no two participants ended up with the same final design. One result that is evident from Figure 4 is that the set of performances of the designs found by the expert system was more uniform than those designed strictly by the participants during the first session. Using the expert system, those participants who struggled during the first session were able to find designs which performed similarly to the designs generated by those who were successful in the first session. Additionally, 9 out of 12 of the second session designs outperformed the designs created by the same participant during the first session.

During the final design session, participants were asked to revisit the same problem for a third time and to draw their final design by hand, again using only their intuition. Participants were not restricted and were allowed to draw inspiration from either or both of the first two design sessions. They could also completely start over if desired. It was the hope of the authors that during the final session participants would combine their initial design with elements from the expert system design to create a better performing final design.

The mean performance of the final twelve designs was 82.9%, which was 9.0% higher than the mean performance of the initial set of designs. Additionally, nine out of twelve designers were able to produce final designs which performed the same or better than their initial designs. This improvement was particularly evident for those designers who had the least performing designs from the first session. These results are interesting because participants completed both the first and third sessions by hand, with no performance feedback. During the second session, participants did not work with their own initial design, but instead with an example design that may have had little in common aesthetically with their own initial design. This result indicates that the process of using the expert system, even with a generic example instead of their own design, was able to improve the intuition of some of the designers about ways in which performance could be increased.

As far as the comparison of results between participants with varying levels of daylighting expertise goes, one initial hypothesis was that participants with little to no previous daylighting experience would benefit more from the process of using the expert system than participants who had more substantial previous experience working with daylighting. The results of this study instead indicate that those who benefited the most from the process of using the expert system were those who produced the least successful initial designs, and that these participants did not necessarily consider themselves inexperienced at working with daylighting.

Based on the performance of their initial designs, the participants can be divided into three general groups: Group A consists of three participants whose initial designs performed worse (more than 20% lower) than the example design, Group B consists of three participants whose initial designs performed about the same as the example (within 3.5%), and Group C consists of six participants whose initial designs performed better than the example by 10% or higher. The performances of each of the three groups are shown in Figure 6.

![Figure 6 Mean performance for groups A, B, and C](image)
Figure 6 shows clearly that the greatest benefits from using the expert system occurred for those in Group A, the participants whose initial designs performed the least successfully. These results are intuitive as there was more room for improvement if participants began with a lower performing design than if they began with a high performing design. However, these results also demonstrate that the expert system allowed those participants who produced the weakest initial designs to ultimately produce designs whose performance approached those developed by the more successful designers. While the difference in mean performance between Groups A and C was close to 40% for the initial designs, the process of using the expert system reduced this difference to only 10% for the final designs.

Additionally, the results show that the mean performance of the Group A final designs was about the same as the mean performance of the Group B initial designs, and likewise, the mean performance of the Group B final designs was about the same as the mean performance of the Group C initial designs. Such results indicate that the process of working with the expert system between the first and third design sessions effectively allowed participants in the lower two groups to “move up” one group.

It is interesting to note that Groups A and B each consisted of one self-rated novice, one intermediate and one experienced daylighter. Group C consisted of three novices, one intermediate, and 2 experienced daylighters. For this particular study, the self-rated experience level of each participant had little to no correlation with his/her initial design’ performance.

**Qualitative Results**

In addition to quantitative results based on design performance, the user study produced qualitative results, based primarily on the participants’ responses on the final questionnaire and observed behavior of participants during the study. This section includes these qualitative results as well as a brief analysis of the influence of the expert system on the aesthetics of each participant's final design.

To assess how the expert system might influence design, the first 2 questions on the final questionnaire asked the participants how they felt about their final design as compared to their first design, when they considered performance and aesthetics. These two questions were designed to help determine how the process of using the expert system during the second design session influenced each participant's final design. The participants’ responses to these questions are shown in Figure 7a. Each participant was also asked directly if the process of using the expert system influenced his or her final design (Figure 7b).

From Figure 7b, one can note that eleven out of the twelve participants felt that their final designs performed better or the same as their initial designs, but that the aesthetics of their final designs were the same or worse than their initial designs. These responses indicate that in many cases, the expert system may have influenced the designers to sacrifice aesthetics in some way for performance, despite the fact that they were told during each design session that they were supposed to try to meet the daylighting goals and to satisfy themselves as a designer.

A visual examination of the designs produced by the participants over the course of the study also makes apparent the influence of the expert system on the final designs. In many cases, the final design is aesthetically derivative of both the participant’s initial design and the design produced during the expert system session. Two example sets of such designs are shown in Figure 8. In both sets, the participant combined elements from his or her original with elements from the expert system design to create a final design.

In terms of educational value, it was shown using quantitative data (see above) that the process of using
the expert system helped many of the participants improve their designs, particularly those whose initial design did not perform successfully. Participants were also asked two questions on the final questionnaire about the educational value of using the expert system: first, whether they thought that they learned something new which helped them approach the specific design problem, and second, whether they thought that they learned something new about daylighting in general. The responses to these questions are shown in Figure 9.

The majority of participants responded that they learned a “small amount” about both the specific design problem and about daylighting in general, a positive result. That they learned a “small” rather than a “large” amount may be partially due to having only a limited amount of time (40 min) to work with the tool, or to not having offered a more moderate option (like “fair amount” e.g.). It is interesting to note that although two participants claimed that they did not learn anything new about the design problem by using the expert system, these 2 participants were those who saw the highest improvement between their initial and final designs (30.2% and 34.4%).

In terms of assessing how the expert system can be used to support design, participants were asked two questions on the final questionnaire to help determine whether they were satisfied with the expert system as a design tool: first, whether they would consider using the expert system again for a studio project, and second, whether they would consider using the expert system again for a professional design project. The participants' responses to these questions are shown in Figure 10.

For both studio and professional projects, about two-thirds of the participants responded that they would consider using the expert system. Given the limited amount of time and control that participants had in using the expert system during the study, this is a positive result. There were no participants who responded that they would “definitely not” use the tool for a future project, and only two participants replied that they would “probably not” use the tool for a professional project. In general, the responses to these questions indicated mostly positive reception of the expert system.

CONCLUSION

This paper presented an evaluation of the expert system as a user-interactive method for performance-driven exploration and as a design tool, based on the results of a user study. During the user study, twelve designers were asked to solve a design problem with multiple daylighting goals, first using their own intuition, and second, using the expert system. The designers were then asked to solve the design problem a third time, again using their own intuition. This study procedure was developed to discover if the expert system had positively influenced the performance and aesthetics of the final designs, as compared to the initial designs. The study participants were also asked to fill out a questionnaire which allowed them to assess their own designs and their experience using the expert system.

The results of the user study were generally positive and indicate that many of the major goals of the expert system as a user-interactive tool were met. One important result was that every participant was able to find a design with improved performance during his or her session with the expert system. While the (Gagne, Andersen & Norford, 2011) paper verified that the expert system could successfully work towards improved designs in the absence of a human user (similarly to traditional optimization like Genetic Algorithms e.g., as in Gagne & Andersen, 2012), this paper demonstrated that the expert system algorithm is also successful when human input is included in the process.

Another important result of the user study was that many of the participants were positively influenced by the process of using the expert system. Most participants also seem to have learned something about the specific design problem which allowed them to intuitively develop better performing designs after they had interacted with the expert system. These results were supported by both the data and by participant response. A final important result is that the majority of the participants responded that they would use the expert system again for a studio or professional design project.

One possible limitation of the user study is the short amount of time that each participant was able to spend designing and interacting with the expert system.
system. Because the sessions were restricted to a maximum of two hours, the designers may not have been able to respond as creatively as they may have been with more time. In many design situations, the designer is allowed to assess and redesign many times before selecting a final design. Therefore, some of the behavior observed during the study may not have been indicative of how participants would have acted in a less formal and less time-constrained environment. Additionally, the number of participants involved in the study was relatively small. Nevertheless, the study was able to provide a glimpse at the nature of the human design process and how human designers might respond to a tool such as the expert system.

An unexpected outcome of the study was that no two participants used the expert system in exactly the same way, i.e. all participants made different sets of decisions. One consequence of this behavior was that the final design found by each participant was unique. This result was somewhat surprising because the authors assumed that some of the novice users might use the expert system as an optimization method rather than as a design tool by choosing only the first design suggestion given at each step and by always accepting the best performing option. Instead, however, it was found that all participants had strong opinions of their own about which design changes to try and about how their final design should look. This type of behavior supports the idea that many designers would not readily accept a design solution generated by a “black box” algorithm. Instead, the highly interactive nature of the expert system allowed each participant to remain actively involved in the expert system design process by retaining control over design decisions. It is the belief of the authors that the participants’ mostly positive receptions to the expert system were due to the interactive nature of the expert system tool.

ACKNOWLEDGEMENT

The authors were supported by the Massachusetts Institute of Technology and would like to acknowledge additional support from the Ecole Polytechnique Fédérale de Lausanne (Prof. M. Andersen) and the Martin Family Society for Sustainability (Dr. J.M.L. Gagne). They would like to thank Prof. Les Norford for his insightful advice for this study.

REFERENCES


Zeiler W., Savanovic P., Quan jel E. 2007. Design decision support for the conceptual phase of the design process, In Proc. IASDR, Hong Kong.


