Validation of a questionnaire for assessing perceptions of lighting characteristics in daylit spaces

CORALIE CAUWERTS[1], MAGALI BODART[1]

[1]Architecture & Climat, Université catholique de Louvain (UCL), Louvain-la-Neuve, Belgium

ABSTRACT: This study is realized in the frame of a work aiming at evaluating the potential of several presentation modes whose images displayed on 2D, 3D and HDR monitors for studying the perceived visual appearance of daylit spaces. The paper discusses two metrological qualities (sensitivity and validity) of a questionnaire developed in this context. This questionnaire is composed of rating scales, multiple choice questions, and questions based on blank sketches. A first group of forty-three subjects visited actual rooms and responded to the questionnaire. A second group of forty-two participants responded to the same questionnaire without visiting the actual rooms but only on the basis of blank sketches. Results show that multiple choice questions are complementary to rating scales. The comparison of perceptions experienced by the participants having visited the actual rooms to responses given by the control group shows that luminous stimuli influence the way the participants rated the rooms. At last, perceptions experienced by the participants visiting the rooms are in accordance with the objective analysis of the rooms' luminous conditions realized on the basis of luminance and illuminance measurements.
Keywords: daylighting, perceptions, questionnaire, validation

INTRODUCTION

In the field of artificial lighting, Veitch et al. proposed in 1996 to define lighting quality as "the degree to which the luminous environment supports the following requirements of the people who use the space: visual and post-visual performances, social interaction and communication, mood state, health and safety as well as aesthetic judgments (assessments of the appearance of the space or the lighting)" [1].

Even if daylight is variable in intensity and color and that daylit spaces generally offer a view toward the outside, daylight is first and foremost a lighting source. The definition, developed in the frame of artificial lighting, seems to be adapted to daylighting. To develop tools to help the architect to design high quality luminous environments, this definition should nevertheless be revisited in integrating the architectural design process. As explained in [2], high quality luminous ambiance will be achieved in satisfying the three main components of architecture (*venustas*-beauty, *utilitas*-utility, *firmitas*-durability) set out by Vitruvius. Fig. 1 illustrates this process, in a sustainable development context.

Currently, in the field of daylighting, most of performance indicators are developed to inform about utility and durability dimensions (e.g. daylight factor (DF), daylight glare probability (DGP), useful daylight illuminance (UDI), daylight autonomy (DA), and so on). Few tools are available for assessing the beauty dimension. However, the appearance of the lighting and the space as well as the creation of emotions are probably the main driving force of the designer. And, in order to evaluate this aesthetic dimension, architects often visualize simulation images on their computer screen.

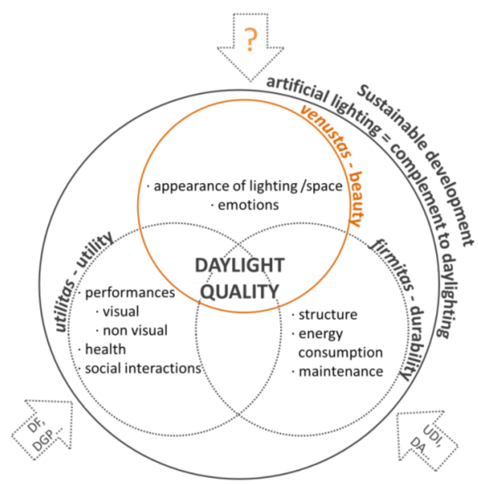

Figure 1: Design process of high quality daylit spaces, in a sustainable development context.

In lighting research too, images are often used to study how the luminous environment influences people perceptions.

REPRESENTATION OF THE REAL WORLD

In comparison with other representation modes of the real world such as mock-ups or scale models, images offer some non-negligible advantages. Indeed, images and more particularly virtual renderings make possible to fix all variables except the one being studied.

Moreover, images are less expensive than mock-ups or scale models. At last, in the field of daylighting, images are a way to overcome the uncontrollable variability of light and to ensure that all the participants visualize the scene under the same luminous conditions.

Since the seventies and the first researches carried out by Flynn on the perception of lit environment [3], imaging technologies have been strongly developed to approach some characteristics of the human vision. Currently, images can be presented in various formats and on various display devices. However, since the work realized by Hendrick et al. to determine if slides reproduce perceptions experienced in the real world [4], few validation works have been done.

OBJECTIVES

This study is realized in the frame of a work aiming at evaluating the potential of virtual renderings as well as the interest of several image formats and display devices for evaluating visual appearance of daylit scenes. In order to test the hypothesis that perceptions experienced in the real world are reproduced using images, perceptions experienced by people in actual rooms are collected using a questionnaire and compared to those experienced by people visualizing the same scenes displayed in various ways on several kinds of monitors.

This paper presents the first phase of the work which consists in validating the questionnaire developed for studying perceptions of the visual appearance of daylit spaces. Two questions are addressed: *Does the lighting questionnaire measure what it is supposed to measure that is, perceptions of the visual appearance of daylit spaces?* And, *is the questionnaire sensitive enough to highlight differences between rooms?*

PARTICIPANTS AND PROCEDURE

To respond to these questions, a first group of 43 participants visited actual rooms. Perceptions of the appearance of the lighting and the space were collected using a questionnaire.

The visit of the real rooms took place in Louvain-la-Neuve, Belgium (50.66° North, 4.56° East) on March 9th 2012 around solar noon, in a period of time as short as possible to minimize daylight variations linked to the sun position as well as weather variations. As it was not possible to visit the room in a single and large group, participants visited the rooms by groups of five or seven. The first group of participants started the visit at 11.00 am while the last group finished at 14.20 pm. On the average, the visit took 45 minutes by subgroup of participants. A unique order of visit was fixed to ensure that all the participants visit the rooms in luminous conditions as similar as possible. The bias highlighted in [6] and linked to the order in which stimuli are presented to the participants is assumed.

Before the visit of the real rooms, participants received instructions and printed questionnaire booklets. Unclear vocabulary was defined. As the rooms were located in several buildings, each group of participants was lead from room to room by a guide. Participants were asked to walk across the corridor and stop themselves at the level of the mark indicated on the ground. They should look straight ahead as they crossed the corridor, and they should not turn around. They were asked to respond to the questionnaire after having immersed themselves in the room during 30 seconds.

To evaluate the potential bias of the response instrument and based on the work of Danford and Willems [7], a second group of 42 participants (control group) responded to the same questionnaire without receiving luminous stimulus. This second group presents characteristics similar than the real world group in terms of native language (French speaking), age (participants are between 18 and 26 years old), educational background (university students) and gender (40% men and 60% women).

ASSESSED ROOMS

Four rooms were assessed. They were located in three distinct but close buildings and were chosen for several reasons. First, the four rooms share a same function: they are all corridors. Corridors were chosen for the aesthetic dimension probably perceived as dominant regarding dimension of usage.

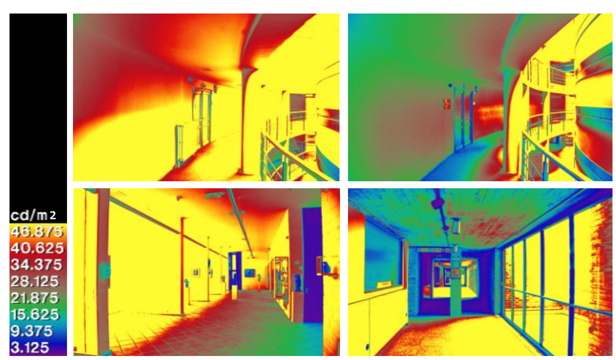

Figure 2: The four assessed corridors. False color luminance maps (cd/m²).

Moreover, corridors present the advantages that participants, during the experiment, feel as much as possible in a real context of use. The four rooms were also selected to maximize as much as possible visual appearance differences. A particular attention was given on some practical reasons like accessibility, cleanness, and calm during the visit.

RESPONSE INSTRUMENT

The questionnaire, presented in Table 1, is composed of rating scales developed around the following dimensions, based on the work of Bülow-Hübe [5]: light level, coloration, distribution, directivity, glare and contrast. Additional questions on pleasantness and enclosedness have also been asked but are not discussed in the present paper.

Table 1: Lighting questionnaire

Descriptive scales (6-grade scales)
Corridor is: dim-bright (light level)
You are: in the dark-in the light (light level)
Corridor is: neutral-colorful (coloration)
Corridor is visually: cold-warm (coloration)
Light is : neutral- colorful (coloration)
Contrast in the corridor is: high-low (contrast)
Distribution of light is: varied-monotonous (distribution)
Shadows are: sharp-blurry (directivity)
Textures are: sharp-blurry (directivity)
Corridor is: comfortable-glaring (glare)
You are disturbed by glare from a window: little-much (glare)
You are disturbed by glare from a surface: little-much (glare)
Scales of appreciation (5-grade scales)
You would prefer the corridor to be: more-less bright
You would prefer the corridor to be: more-less colorful
You would prefer the corridor with a contrast: higher-lower

In addition to the questions presented in Table 1, a series of non-conventional questions, based on blank sketches, were developed with the aim to link objectives maps (luminance-based maps) with subjective maps (participant sketches). The aim of these questions was to evaluate the ability of the observer to detect some zones in his visual field, as well as his ability to discriminate surfaces for brightness, roughness and uniformity, in a built environment so, in context.

Participants were first invited to circle on blank sketches (see Fig. 3) attractive areas and materials perceived as emphasized by light, and to color brightest and dimmest areas, respectively in red and blue.

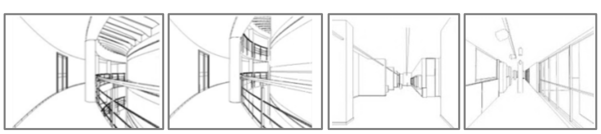

Figure 3: Blank sketches

Participants were also asked, in each room, to compare two walls, on 5-point rating scales, for brightness, uniformity and roughness, and to classify the brightness of three punctual zones on a continuous scale (see walls #1 and #2 and punctual zones a, b and c in Fig. 4).

Finally, at the end of the visit, after having rated the four corridors, participants were asked to respond to a series of multiple choice questions (MCQ). They should determine which room, among the four visited, they perceived as the brightest, the least bright, the most colorful, the least colorful, the one presenting the highest contrast and the one presenting the lowest contrast.

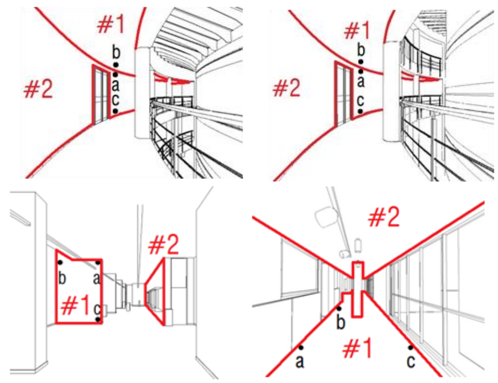

Figure 4: Sketches indicating walls and punctual zones to be compared in each room.

METHOD

In order to check the sensitivity of the rating scales, a first analysis was carried out on the responses collected in the real world. This analysis aimed at determining if each rating scale makes possible to distinguish the four rooms.

Then, to evaluate the validity of the questionnaire, additional analyses were done. First, rating scales were compared to responses given to multiple choice questions. Then, real-world group responses were compared to control group responses to check that luminous stimuli really influence the way participants respond to the questions. At last, subjective responses given by the participants of the real-world group were compared to physical measurements done in the rooms the day of the experiment.

RESULTS

Statistical analyses were performed using R software [8]. For reasons of clarity, statistical results are summarized in this paper. Detailed results are available from the authors on request.

Sensitivity of the rating scales

Fig. 5 presents mean score given by the participants of the real-world group at each rating scale, in each room. Sensitivity of the rating scales was tested in performing ANOVAs and multiple comparison tests on these responses. P-values for multiple test comparison were adjusted using the Bonferroni correction. Rooms sharing a same letter (a, b, c or d) in Fig. 5 do not differ significantly. According to the results of the statistical tests, five questions (marked with an asterisk in Fig. 5) do not present a high level of sensitivity as they did not make possible to create distinct groups of rooms presenting significant differences according to statistical tests.

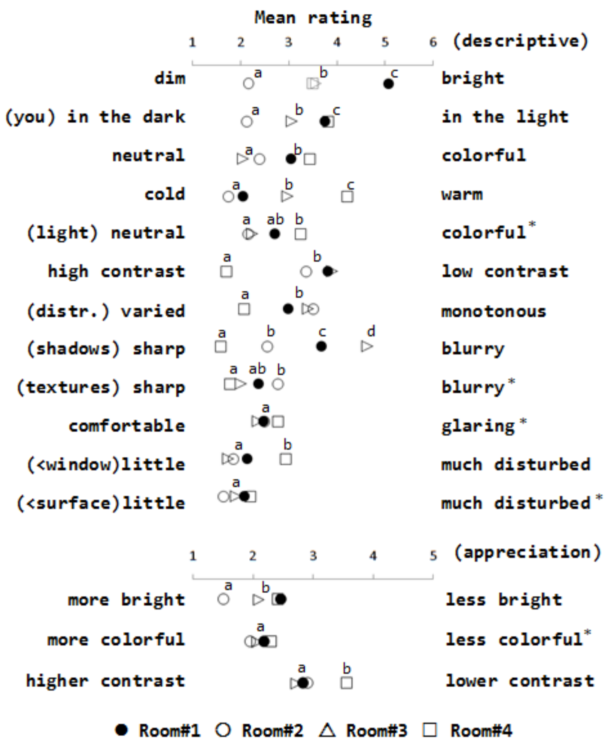

Figure 5: Rating scales. Mean scores (real world group). Letters (a, b, c and d) indicate groups of rooms resulting from the statistical tests. Asterisk (*) indicates questions less discriminating.

It was also observed that responses to the three appreciation scales are less discriminating than the corresponding descriptive ones as less groups of rooms are created: more-less bright (2 groups) vs. dim-bright (3 groups); more-less colorful (1 group) vs. neutral-colorful (2 groups); high-low contrast (2 groups) vs. higher-lower contrast (2 groups).

Validity – Rating scales vs. MCQ

In order to check the consistency of the responses, ratings scales were compared to multiple choice questions (MCQ).

Chi-squared tests and multiple comparison tests were performed to determine the significance of the results to the multiple choice questions. P-values for multiple test comparison were adjusted using the Bonferroni correction. Table 2 presents the frequency and statistical significance resulting from the multiple comparison tests performed on MCQ.

Table 2: Frequency and statistical significance resulting from the multiple comparison tests performed on MCQ.

MCQ	Room			
	#1	#2	#3	#4
the brightest	**31*****	0***	2**	10
the least bright	0***	**23*****	10	10
the most colorful	9	0***	3*	**31*****
the least colorful	5	13	**23*****	2**
the highest contrast	3*	0***	2**	**38*****
the lowest contrast	11	9	**22*****	0***

Significance : * = p< 0.05; ** = p< 0.01; *** = p< 0.001

The first rooms is perceived as the brightest (X-squared= 50.86; p-value=3.97E-12). The second room is perceived as the dimmest (X-squared=18.61; p-value=6.41E-05). The third room is perceived as the less colorful (X-squared=9.50; p-value=6.41E-05) and the one presenting the lowest contrast (X-squared=14; p-value=7.31E-04). At last, the fourth room is perceived as the most colorful (X-squared=7.45; p-value= 3.97E-12) and those presenting the highest contrast (X-squared=9.50; p-value=3.29E-21).

These results are consistent with those obtained using rating scales. However, MCQ make possible the distinction between some rooms that the rating scales do not. For example, while the rating scale related to the colorfulness does not make possible to distinguish either the two lowest colorful rooms (Room#3 and Room #2) or the two most colorful rooms (Room#4 and Room#1), the multiple choice questions make possible to determine that the most colorful is Room#4 and that the less colorful is Room#3.

Validity – Real world group vs. Control group

In order to check that the luminous stimulus really influences the way participants respond to the questionnaire, responses given by the participants having visited the actual rooms were compared to those of the participants having responded to the questionnaire without having visited the rooms.

Fig. 6 presents the differences between real-world group and control group responses to the rating scales, for the fourth room. As shown in Fig. 6, scores deviate more from the mean of the scales when participants receive a luminous stimulus (real-world group) than when they do not receive this stimulus (control group). Two-way ANOVAs (presentation mode * room) and multiple comparison tests were performed on data collected using rating scales to determine if, as expected, differences between real world group and control group are significant. According to statistical results, all the rating scales present either a presentation mode effect or an interaction effect indicating that the effect of presentation mode (real world or control) varies according to the room.

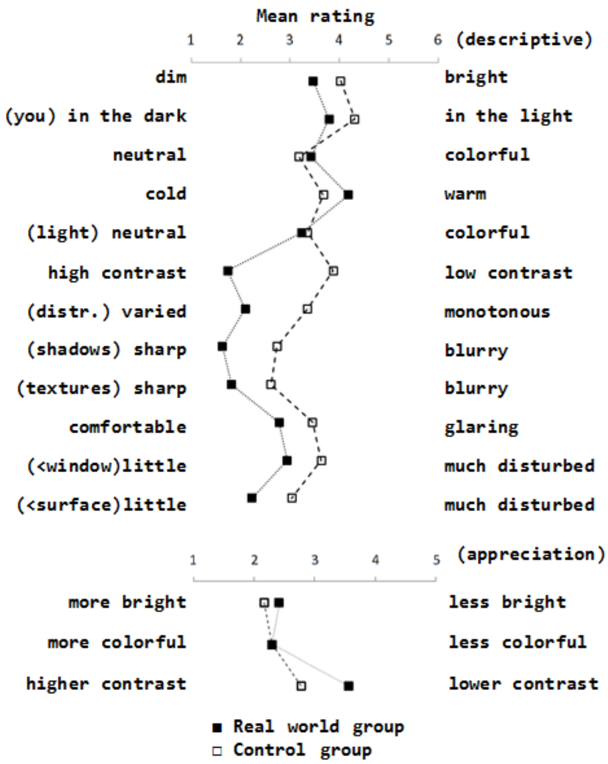

Figure 6: Rating scales. Mean scores (real world group vs. control group). Room #4.

Fig. 7 and Fig. 8 illustrate responses of the participants to the first type of non-conventional questions. A visual comparison between real-world group sketches and control group sketches indicates that participants are clearly influenced by the luminous stimuli.

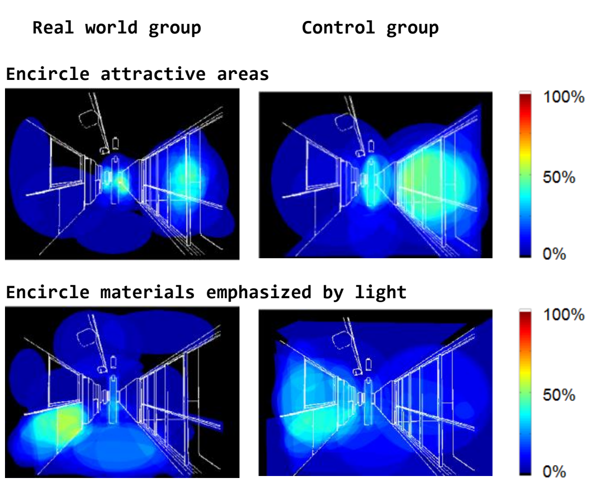

Figure 7: Percentage of participants (real world vs. control group) who circle the area of interest. Room #4.

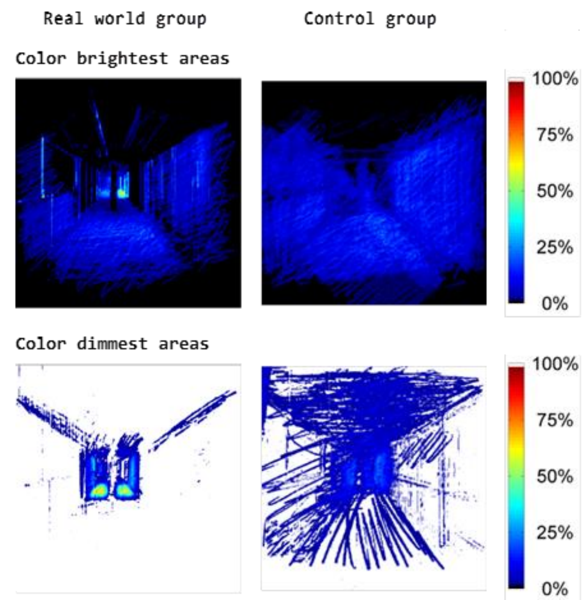

Figure 8: Percentage of participants (real world vs. control group) who color the brightest and dimmest areas. Room #4.

Fig. 9 illustrates the paired-comparison of walls for brightness, uniformity and roughness. Two-way ANOVAs (presentation mode * room) and multiple comparison tests were performed on these data to determine if differences observed between real-world group and control group are significant. Only the third room presents differences for brightness (○). The four rooms present a significant effect for uniformity (△). Concerning roughness (□), three rooms (Rooms #2, #3 and #4) present a significant effect.

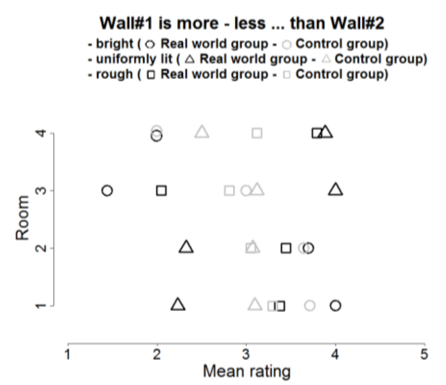

Figure 9: Comparison of walls in terms of brightness, uniformity and roughness.

At last, Fig. 10 presents the classification of the three punctual zones a, b and c in terms of brightness. In three rooms (Rooms #1, #2 and #4), the three points are ordered similarly by the real-world group than by the control group.

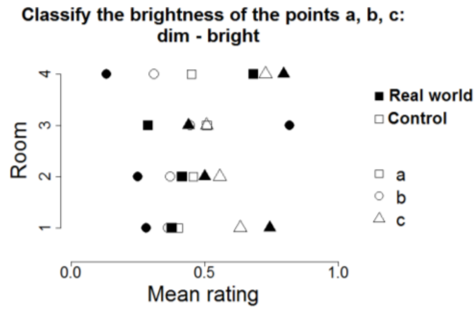

Figure 10: Classification of a, b and c in terms of brightness.

Comparison of walls or punctual zones for brightness suggests that participants of the control group have the ability to analyze where daylight is coming from and to guess how it is distributed. Indeed, in the three rooms where they can locate the windows (the first, the second and the fourth room), the points a, b and c are correctly ordered.

Validity – Subjective vs. objective assessment

On the basis of physical measurements and HDR pictures taken during the visit of the actual rooms, some indicators of performance (maps of luminances and contrast, vertical illuminance at eye level, CIE L*a*b* coordinates, vertical to horizontal ratio…) were calculated to objectively evaluate the visited rooms. This objective evaluation did not show contradiction with subjective scores. Moreover, the low DGP values calculated based on HDR pictures and vertical illuminance measured at eye level explains why rating scales related to glare do not make possible to differentiate the rooms: given that the rooms do not present risks of glare, the sensitivity of these questions cannot be evaluated.

DISCUSSION

The analysis of the sensitivity of the questions revealed that scales of appreciation are less discriminating than descriptive ones. Moreover, questions linked to the perception of the color of light and to the perception of the textures did not make possible to differentiate the four rooms. To help the participants to discriminate the rooms, an identical poster of fruits as well as a sculpture could be placed in the rooms. Rating scales linked to glare did not make possible to differentiate rooms, which are perceived as not glaring. The absence of glare in the four rooms was confirmed by DGP calculation.

The comparison between rating scales and multiple choice questions revealed that there is no contradiction between rating scales and MCQ. However, the two types of questions are complementary as they do not bring the same information: while rating scales give information regarding intensity, the multiple choice question makes possible the identification of the extrema.

At last, contrary to the results obtained by Danford and Willems in [8] where no difference was observed between control groups and groups receiving a luminous stimulus, the comparison of real-world group and control group responses suggests that the presence of the luminous stimuli influenced the way participants respond to the lighting questionnaire. However, it was observed that participants have the ability, on the basis of sketches, to evaluate where daylight comes from and to guess how it is distributed.

CONCLUSION AND FURTHER WORK

Results showed that multiple choice questions are complementary to rating scales and that the luminous stimuli influenced the way participants rated the rooms. At last, perceptions experienced by the participants in the actual rooms are in accordance with the objective analysis of the rooms' luminous conditions realized on the basis of luminance and illuminance measurements.

The work will continue in comparing perceptions experienced in the real world with those experienced by other people visualizing virtual renderings of the rooms and photographs realized the day of the visit by the first group of participants.

ACKNOWLEDGEMENTS

Coralie Cauwerts is a research fellow of the Fonds de la Recherche Scientifique – FNRS. The authors thank students who have participated in the experiments.

REFERENCES

1. Veitch, J. and Newsham, G. (1996). Determinants of lighting quality II: Research and recommendations, 104th Annual Convention of the American Psychological Association. Toronto, Canada, August.
2. Chaabouni, S. (2011). Voir, savoir, concevoir: une méthode d'assistance à la conception d'ambiances lumineuses par l'utilisation d'images références.
3. Flynn, J., Spencer, T., Martyniuk, O. and Hendrick, C. (1973). Interim study of procedures for investigating the effect of light on impression and behavior. *Journal of the Illuminating Engineering Society*, 3: p. 87-94.
4. Hendrick, C., Martyniuk, O., Spencer, T. and Flynn, J. (1977). Procedures for investigating the effect of light on impression. Simulation of a real space by slides. *Environment and Behavior*, 9: p. 491-510.
5. Bülow-Hübe, H. (1995). Subjective reactions to daylight in rooms: Effect of using low-emittance coatings on windows. *Lighting Research and Technology*, 27: p. 37-44.
6. Fotios, S., Houser, K. & Cheal, C. (2008). Counterbalancing needed to avoid bias in side-by-side brightness matching tasks. *Leukos*, 4: p. 207-223.
7. Danford, S. and Willems, E. (1975). Subjective responses to architectural displays: A question of validity. *Environment and Behavior*, 7: p. 486-516.
8. R Development Core Team 2010. R 2.11.1.