Emerging Methodology to Inform Design Evaluation: Mind the Perception

MADLEN SIMON
University of Maryland

MING HU
University of Maryland

Keywords: design evaluation, virtual reality, brainwave response

The predominant tools for evaluation of built environment features have been user response surveys and expert panel scoring, applied to actual environments or to visual representations of urban environments in drawing or model form (Nasar 1994, Ewing and Handy 2009, Adkins et al 2012). Here we propose to test the validity of combining electroencephalography (EEG) and virtual reality (VR) to overcome the problem of confounding variables in real environments or their representations and to elicit actual user responses in real time. This research combines a neuroscientific technology with an emerging design technology to record electrophysiological brain activity of participants in a well-controlled three-dimensional virtual audiovisual environment. Experimental subjects are immersed in three different virtual urban settings while wearing EEG equipment. A device called Emotive EPOC Insight, a low-cost mobile EEG recorder, will be employed to monitor the brain activities. The aim of this research project is to develop and test a methodology using data-driven approach, rather than user-reported, responses for evaluation of built environment design features.

INTRODUCTION

What would it mean if data management were at the core of our discipline? Data gives insights into users' responses to environments. The more data, the richer the insights, promoting a better fit of environment to people. How can the design process be informed by data on human response to space and place? Stakeholder preference of one alternative over another plays an important role in design process, especially in dealing with multi-objective design problems in which designers juggle competing objectives. Current tools for design evaluation are surveys, scorecards, and verbal comments. The goal of this research project is to develop, test, and validate a data-driven approach for design decision-making. Such a framework would facilitate participation and action by multiple decision-makers and stakeholders, offering insights into the architectural design process. This paper presents an experiment combining an immersive virtual environment (VR) and electroencephalogram (EEG) as a promising tool to evaluate alternative options during the early design stage of a project. More precisely, the objective is to (a) develop a data-driven approach for design evaluation and (b) understand the correlation between end users' preference and emotional state. To our knowledge, this is the first time that the combination of immersive virtual reality technology and brainwave response monitoring has been proposed to study the design validation method in architecture.

VIRTUAL REALITY AND EEG USAGE IN DESIGN AND RESEARCH OF THE BUILT ENVIRONMENT

The concept of mixed reality, which includes both virtual reality (VR) and augmented vitality (AV) was first coined by Milgram and Kishino (1994). VR is a commonly known technology that can add the dimensions of immersion and interactivity to three-dimensional, computer-generated models, offering an experience that does not exist in the conventional form of representation (Burdea and Coiffet 1994, Stouffs et al. 2013). VR offers the possibility to experience sensations and movement in a simulated environment of the proposed design solutions. This experience is often difficult to fully grasp from two-dimensional rendering or even a three-dimensional model. VR provides the opportunity for end users or clients to experience space and building prior to construction. Consequently, the application of VR in the architecture, engineering, and construction (AEC) field is wide, from design to construction process simulation and communication in a collaborative decision-making environment (Milovanovic et al. 2017).

Relatively low-cost mobile electroencephalography (EEG) data acquisition equipment, combined with virtual reality technologies, enables neuroscientific measurement outside the clinical context. An EEG device measures the electrical signals produced by the post-synaptic activity of large neuronal assemblies firing in coherence from locations in the higher layers of the cortex (Marvos et al. 2016, Buzsaki 2006). Signals detected allow researchers and designers to gain new insights into ways that people perceive and respond to the built environment and its design features. In the 1980s, researchers in the design and built environment fields began to experiment. Ulrich (1981) examined EEG and heart rate data to study perceptions of urban and rural environments. Later, Kim et al. (2010) used functional magnetic resonance imaging (fMRI) to assess brain activity patterns, concluding that subjects' emotional responses to natural and urban environments were different. Aspinall et al. (2015) used a mobile EEG headset, Emotiv EPOC, to study responses of participants who took a 25-minute walk through three areas of Edinburg, Scotland. Results indicated a high-dimensional

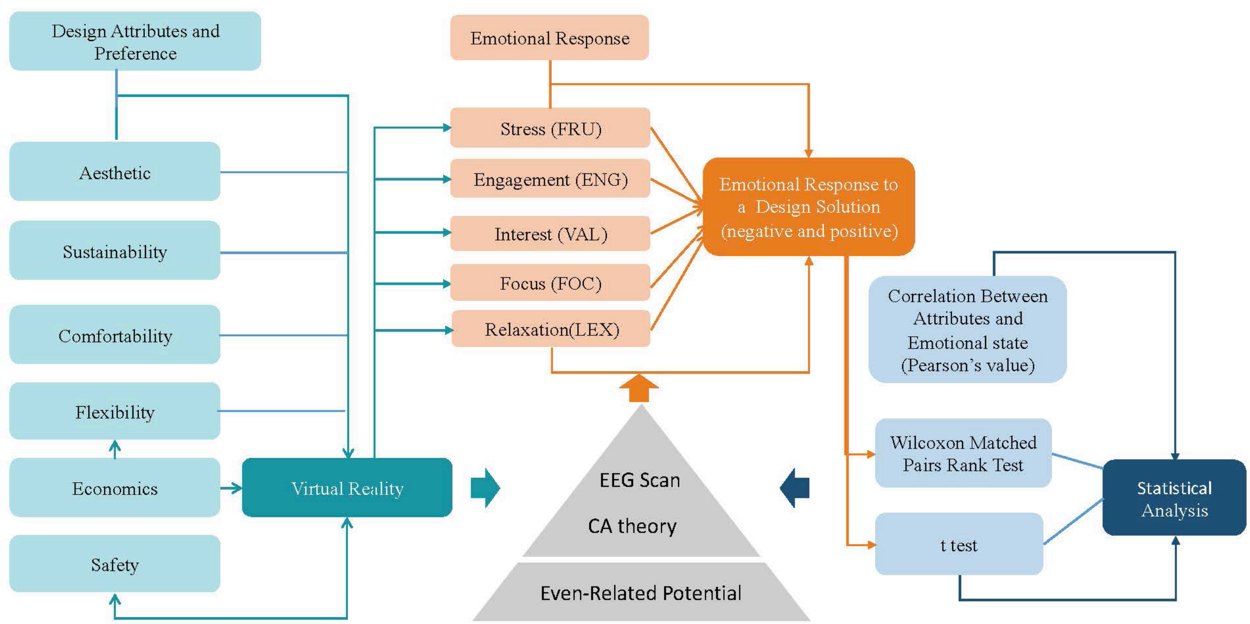

Figure 2.. Research methodology

correlation between low frustration, engagement, arousal and higher mediation when moving into the green space zone.

METHODOLOGY: SAMPLE, EXPERIMENT, AND SETUP

Theoretic base

This research method is based on the event-related potential (ERP) neuroscience approach and cognition architecture (CA) theory. ERP studies brain activity in response to visual stimuli (Picton et al. 2000). More specifically, it compares the neural signal (brainwave) during different conditions to determine whether and how the brain responds to different stimuli. CA was proposed by Herbert Simon, a pioneer in artificial intelligence, in 1960. CA has since then been explained, expanded, and further developed by researchers, mainly from the psychology and computer science domains. It has also been applied to research on design thinking. Leading researcher William Mitchell (1990) explained the basic trial-and-error structure of design process and stated that different types of computational devices may be used to generate proposals, test them, and apply control strategies. CA theory can be translated into a method of scoring and providing principles to demonstrate how specific physical/virtual environments can influence our mental state (Hollander and Foster 2016). Our hypothesis is that a design proposal based on the participants' preference could result in better evaluation scores and stimulate a more positive mental/emotional state. To this extent, the emotional state could be used as supplementary evidence to evaluate design proposals. In order to validate our hypothesis, we designed a case study in which participants experienced two alternative design solutions and we captured their responses to determine whether participants might experience a higher positive emotional state, such as interest and engagement, and less negative emotion (stress) as a result of being immersed in their preferred environment (the environment was constructed in VR). It is useful to illustrate the methodology framework in a diagram (refer to figure 1). The goal was to discover whether mental state corresponds with ERP and CA theory.

Case design and tools

This research combines a neuroscientific technology (EEG) with an emerging design technology (VR) to compare emotional stage levels (interpreted from the EEG raw signal) of participants in a well-controlled, three-dimensional virtual environment. A virtual reality set (Samsung Odyssey) and EEG equipment (Emotiv EPOC Insight) were used to collect design evaluation data from six males and two females (aged 18–60). All test participants at University of Maryland were either architecture school faculty members, architects, or architecture and engineering students. The design task was to create a layout best-suited to the preference and needs of an architecture department. The test layout occupied a single floor of an L-shaped building on a university campus. The space included lecture rooms, conference rooms, offices for professors and staff, computer labs, and other spaces (e.g., restrooms, stairs, and elevators). The design alternatives were proposed independently by two architectural designers. Test participants acted as clients who would be end-users of the building. Before the experiment began, the clients and designers agreed that six attributes (flexibility, economics, comfort, safety, sustainability, and aesthetics) were important to consider when designing the floor plan layout, although no specific interpretations for the various attributes were given at that point. Although we

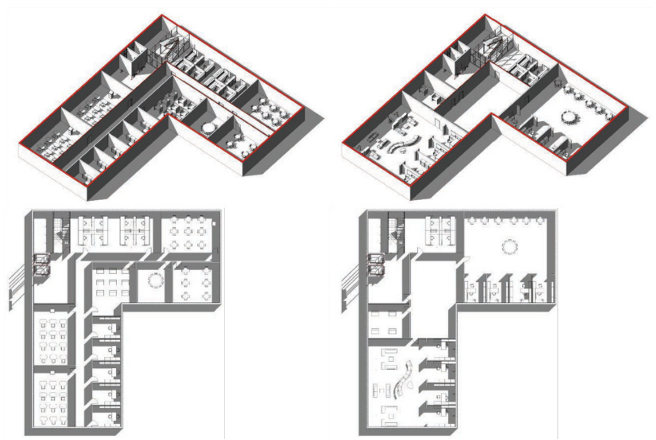

Figure 2. Two design options (Autodesk Revit Model)

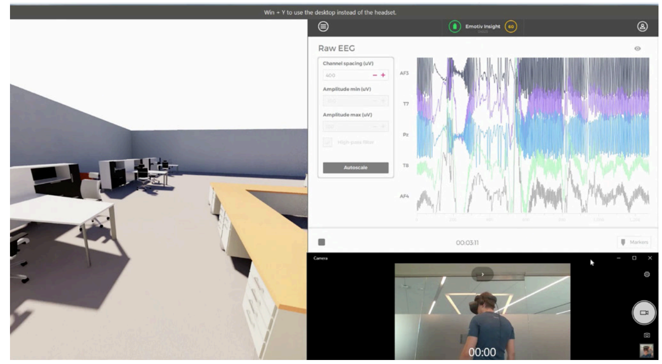

Figure 3 VR and EEG Testing

only chose six attributes, we acknowledge that other attributes would be considered in real practice, such as the initial cost, long-term profit, and ease of maintenance and repair. The above attributes were used to test the hypothesis, and the same framework can be applied to other design attributes. The case design proceeds in two stages.

In Stage I, we determined the primary factors that the design alternative would focus on based on the potential clients' preferences. In order to identify thesw clients' and end users' preferences, the research team conducted two surveys (Zhuang et al. 2017). Survey one identified the different weight sets for three primary players—architects, engineers, and members of the public who lack professional training—in the early design stage. We determined different weight sets regarding the six primary design attributes for these three groups. Preliminary results revealed that architects paid the greatest attention to comfortability (functionality), engineers cared most about safety, and the public's greatest concern was flexibility. Based on the different weighting sets, a predication model was created to predict each group's preference of certain projects. Survey two was conducted using an existing online database of architectural projects to verify the predication model. Ten specific projects were selected from an online architectural project library (Archdaily). Based on the results from the two surveys, we determined that comfortability, flexibility, and safety weighed the most in determining the clients' preference for a certain design proposal.

In Stage II, two design options were proposed: OPT A and OPT B. OPT A was designed without knowledge of the primary weighting factors (comfortability, flexibility, and safety) whereas OPT B was designed with all the information. We built two virtual models reflecting the two design options (refer to figure 2) and asked test participants to experience the two design options in an immersive virtual environment (VR). To test the proposed framework, this project has focused on the interior design, the exterior wall features were excluded, such as windows and door. The floor and ceiling were set as the generic types and the same in both models (design options). The furniture types, colors and texture are consistent in the both models as well, the difference and contrast were mainly focused on the floor plan layout.

We then used an EEG device to record their brainwaves and emotional response to those two different environments. Afterwards, each participant completed the survey, scoring the quality of design—on a scale of 1 to 7—on comfortability, safety, flexibility, aesthetic value, sustainability, and potential cost. Finally, they provided a preference score for both design options.

Virtual reality (VR) model setup

In general, there are three types of VR technologies: (1) direct VR into 3D modeling software, (2) VR with a game engine, and (3) a 360-degree panorama picture (Burdea 1994). In this research project, direct VR integration technology was applied due to the speed and quality of the work. Firstly, a 3D model was created with Autodesk Revit based on the initial design. Then, a plug-in tool called ENSCAPE was used to translate 3D Revit to a virtual environment. This is the only program that can be directly integrated into multiple 3D modeling software, which creates high requirements for the hardware (computer); however, it allows for instantaneous design changes in a virtual environment. Two models were set up and transferred to VR (see figure 3). potential cost. Finally, they provided a preference score for both design options.

Mobile electroencephalograph (mobile EEG): Brainwave recording and measurements

In this research, the Emotiv INSIGHT EEG headset was chosen based on its relatively low cost and the accompanying software, which can aggregate raw data. The headset was fitted on the underside of the VR headset (refer to figure 7). After being fitted with the device, each participant was asked to focus on a blank background for 10 seconds, attentive to his or her own breathing, which allowed us to detect the baseline emotional state of the participants. Afterward, we immersed them into two different design options, with each design option taking approximately 15-20 minutes to experience. The device consisted of five sensors positioned on the wearer's scalp according to the international 10–20 system: the antero-front (AF3, AF4), parietal (Pz) and temporal sites (T7, T8). Brainwaves were measured through those five channels in terms of amplitude (10–100 microvolts) and frequency (1-80Hz) at 128 samples per second per channel (Aspinall et al. 2015, Emotiv website). The four main brainwaves/bands measured and recorded were beta, alpha, theta, and gamma. The beta wave is associated with engaged brain activities, such as learning, working, and speaking. The alpha wave, in contrast, represents non-arousal brain activity; a person who is sitting down and resting is often in the alpha state. The theta wave represents the state of free flow of ideas as well as less engagement in the current physical state. For instance, an individual who often runs outdoors has a state of mental relaxation and is prone to a flow of ideas. Theta waves, known as "suggestive waves," also appear during daydreaming, which is considered a positive mental state since these waves suggest an open mindset. They also imply deep emotional connections to others or objects; in this case, the built environment. Furthermore, theta waves have the benefit of improving creativity and intuition (Emotiv). The gamma wave is a more recent discovery and is involved in processing highly complex tasks with healthy cognitive functions.

After the raw EEG data was collected from each participant, the signals were analyzed with the software Emotiv Pro (developed by Emotiv) and categorized into one of five emotional states: engagement, focus, interest, stress, or relaxation. Engagement (ENG) typifies a mixture of attention and concentration, with high scores indicating higher productivity. It measures the level of immersion in the moment and contrasts with boredom. Engagement is characterized by an increased physiological arousal as well as beta waves and attenuated alpha waves. The greater the attention, focus, and workload, the greater the output score reported by the detection. Focus (FOC) is a measure of the depth of attention as well as the frequency that attention switches between tasks. In contrast, interest (VAL) provides a measure of affinity to tasks, with low scores indicating aversion and mid-range scores indicating neither like nor dislike. Stress (FRU) is indicative of several outcomes: low to moderate levels of stress can improve productivity whereas a higher level tends to be destructive and can have long-term consequences for health and wellbeing (Ramirez and Vamvakousis 2012, Hollander and Foster 2016, Sena et al. 2016). Finally, relaxation (LEX) is a measure of ability to switch off and recover from intense concentration. Each of the five measurements was given a number between 0 and 1. Scores provided by Emotiv were useful indicators for our pilot study; however, we recognize the limitations in the validity and reliability of this device. Following each participant's immersion in one of the VR designs, Emotiv Pro provided temporally based scores for each of the five emotional states, which were later used in the statistical analysis.

Subjective evaluation

After immersion in the two VRs, an eight-item questionnaire measuring the six design attributes of the building was given to each participant. Each attribute was rated on a seven-point Likert scale. The six attributes represented the criteria defined in the case design stage: aesthetics, safety, flexibility, comfortability, sustainability, and economics (potential cost). At the end, test subjects were also asked to provide a preference score ranking one design option over the other.

DATA ANALYSES AND FINDINGS

Questionnaires and EEG results

Altogether, six participants provided 56 scores for six attributes plus an overall preference score for each of the two design options. Design option B (OPT B) received a higher score than design option A (OPT A) in all six design attributes. In OPT B, comfortability, aesthetics, and flexibility were rated as the top three attributes, aligning with the fact that OPT B was created based on knowledge of the clients' preferences. The much higher overall preference score of OPT B validates the point that giving more consideration to the clients' preferences could result in a much higher satisfaction rate. One interesting finding was the fact that aesthetics was rated as one of the top three attributes by participants who favored the preferable option (OPT B) while aesthetics was not identified as an important design attribute by any of the three groups (architects, engineers, and members of the public) through the survey. This might be explained by the innate nature of human beings as visual thinkers who do not consciously or proactively acknowledge the important role of aesthetic value in our decision-making, particularly in design context. Scientists agree that humans possess five basic senses: smell, hearing, touch, taste, and vision. However, the human brain expressly prioritizes just one sense: vision (Hollander and Foster 2016). Furthermore, Kandel (2012) stated that about half of the sensory information reaching the brain is visual. The reason why none of the groups listed aesthetics as their primary design attribute needs to be understood and potentially represents the next research focus.

Next, we examined whether higher evaluation scores were correlated with a positive emotional state of the participants. Each participant had an approximate 20-minute recording, which generated more than 1,400 data points. Overall, OPT A stimulated more engagement, interest, and attention while OPT B generated greater relaxation and stress, with stress levels in the moderate range, which could indicate higher productivity (Emotiv). Unlike the scores from the questionnaires, which indicate a clear preference for OPT B from all participants, EEG data indicated that participants had varied responses—including negative, positive, and neutral—toward the two design options. In order to further understand the varied responses, we examined the data from individual participants and conducted a statistical analysis to determine whether there were significant differences in participants' emotional responses to the two design options. The five emotional states were used as a proxy measurement of participants' preference for design options. The following section explains the findings.

TWO STATISTICAL ANALYSES

Statistical analysis: Wilcoxon signed- rank test

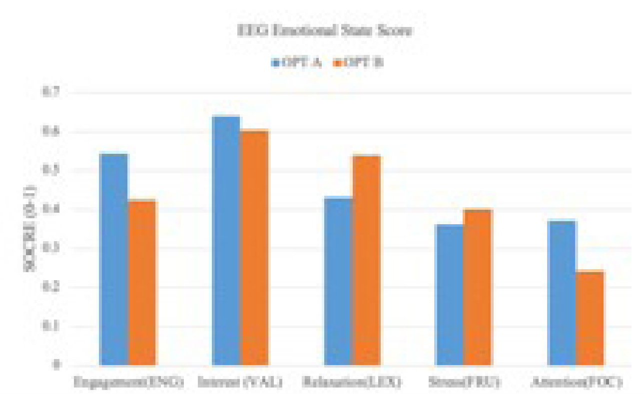

Figure 4. Emotional state score from the EEG recording

The EEG data did not directly lead to the overall evaluation of design options, and the emotional state could not be directly translated as negative or positive toward the design solutions. Therefore, instead of looking at the representation of the individuals' emotional state, the authors examined whether there was a significant difference in how participants responded to the different design options. A null analysis was appropriate for verifying the hypothesis. The Wilcoxon signed-rank test is commonly used to test for a difference in a paired observation, and a sign test is often used to test the null hypothesis.

Table 1. Wilcoxon matched pairs signed-rank tests for responses to OPT A and OPT B

	OPT A	OPT B	Rank	α = 0.05
Response 1	0.718	0.227	12	
Response 2	0.703	0.699	1	
Response 3	0.497	0.648	-9	
Response 4	0.355	0.370	-3	
Response 5	0.325	0.054	11	
Response 6	0.533	0.471	5	
Response 7	0.564	0.631	-6	
Response 8	0.406	0.667	-10	
Response 9	0.372	0.467	-8	
Res 10	0.354	0.262	7	
Res 11	0.517	0.5	4	
			64	Postive Sum
			-63	Negative Sum
			64	Test statistic (W)

The analysis considers one null hypothesis: H01: There is no significant difference between the participants' negative and positive emotional responses to OPT A and OPT B. The alternative hypothesis is: Ha1: There is a significant difference between the participants' negative and positive emotional responses to OPT A and OPT B.

Descriptive results: Wilcoxon signed- rank test

For null hypothesis H01, among the five different emotional responses to the two design options, three response types were higher for OPT A while the other two were lower for OPT A (figure 11). The Wilcoxon test score (W), 64, was higher than the critical value used for a two-tier test of 52. Based on these results, we could not reject null hypothesis H01. Instead, we should reject the alternative hypothesis, Ha1. In conclusion, there is no emotional state difference between the participants' positive and negative responses to the two design options (refer to table 1).

The rejection of the null hypothesis suggests that the overall positive or negative emotional state does not directly affect or correlate with how participants answered the design evaluation. Depending on the importance of design attributes, the preferred design solution might stimulate a negative emotion, and the less preferred design might stimulate a more positive emotion, such as interest and engagement. Participants clearly preferred OPT B overall, which received higher scores; however, their emotional responses did not show clear negative or positive direction.

CONCLUSION AND DISCUSSION

The research project has developed and tested a data-driven approach for design validation. Such a framework would facilitate participation and action by multiple decision-makers and stakeholders as well as lend insights into any design process marked by the characteristics of an architectural design process. Based on the available data from this experimental study, there is no cohesive conclusion about whether a positive emotional state (brain activity) can be correlated with a higher scoring design evaluation. Likewise, a negative emotional state does not automatically result in negative design evaluations. The limitation of this research that could be improved in the future work are: (1) Interpretation of the brainwave is mainly conducted through a predetermined algorithm created and managed by the company that made the device. The mechanism of translating brainwaves into emotional scores is unknown. Future research should consider applying a more transparent approach by using third-party software. (2) The small sample size did not enable us to run a multivariate statistical analysis. For future research, larger samples and datasets are needed for a full in-depth analysis. Additional experiments and data are needed for further studies, and virtual environment design could be more defined with higher graphic quality. This model has great potential to open new avenues for inquiry of how technology-based tools can be leveraged to influence mainstream design choices that incorporate clients' and end users' preferences. Furthermore, the use of an EEG device allows us to enter the research arena of how brainwaves respond to different design solutions.

ENDNOTES

1. Kim, Gwang-Won, Gwang-Woo Jeong, Tae-Hoon Kim, Han-Su Baek, Seok-Kyun Oh, Heoung-Keun Kang, Sam-Gyu Lee, Yoon Soo Kim, and Jin-Kyu Song. "Functional neuroanatomy associated with natural and urban scenic views in the human brain: 3.0 T functional MR imaging." Korean Journal of Radiology 11, no. 5 (2010): 507-513.
2. Kandel, Eric R. The age of insight: The quest to understand the unconscious in art, mind, and brain, from Vienna 1900 to the present. Random House Incorporated, 2012.
3. Milgram, Paul, and Fumio Kishino. "A taxonomy of mixed reality visual displays." IEICE TRANSACTIONS on Information and Systems 77, no. 12 (1994): 1321-1329.
4. Milovanovic, Julie, Guillaume Moreau, Daniel Siret, and Francis Miguet. "Virtual and Augmented Reality in Architectural Design and Education: An Immersive Multimodal Platform to Support Architectural Pedagogy." (2017).
5. Mitchell, William J. The logic of architecture: Design, computation, and cognition. MIT press, 1990.
6. Stanney, Kay M., Dylan D. Schmorrow, Matthew Johnston, Sven Fuchs, David Jones, Kelly S. Hale, Ali Ahmad, and Peter Young. "Augmented cognition: An overview." Reviews of human factors and ergonomics 5, no. 1 (2009): 195-224.
7. Sena, Pasquale, Matteo d'Amore, Maria Antonella Brandimonte, Rolando Squitieri, and Anita Fiorentino. "Experimental framework for simulators to study driver cognitive distraction: brake reaction time in different levels of arousal." Transportation research procedia 14 (2016): 4410-4419.
8. Royal Institute of British Architects. Client & architect: Developing the essential relationship. https://books.google.com/books?id=G3UDjwEACAAJ, 2015
9. Rittel, Horst WJ, and Melvin M. Webber. "Dilemmas in a general theory of planning." Policy sciences 4, no. 2 (1973): 155-169.
10. Ramirez, Rafael, and Zacharias Vamvakousis. "Detecting emotion from EEG signals using the emotive epoc device." In International Conference on Brain Informatics, pp. 175-184. Springer, Berlin, Heidelberg, 2012.
11. In International Conference on Brain Informatics (pp. 175-184). Springer, Berlin, Heidelberg.
12. Stouffs, R., P. Janssen, S. Roudavski, and B. Tunçer. "What is happening to virtual and augmented reality applied to architecture." In Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2013), vol. 1, p. 10. 2013.
13. Simon, Herbert A. The sciences of the artificial. MIT press, 2019.
14. Ulrich, Roger S. "Natural versus urban scenes: Some psychophysiological effects." Environment and behavior 13, no. 5 (1981): 523-556.
15. Zhuang, Jun, Ming Hu, and Fatemeh Mousapour. "Value-driven design process: A systematic decision-making framework considering different attribute preferences from multiple stakeholders." Journal of Solar Energy Engineering 139, no. 2 (2017): 021001.