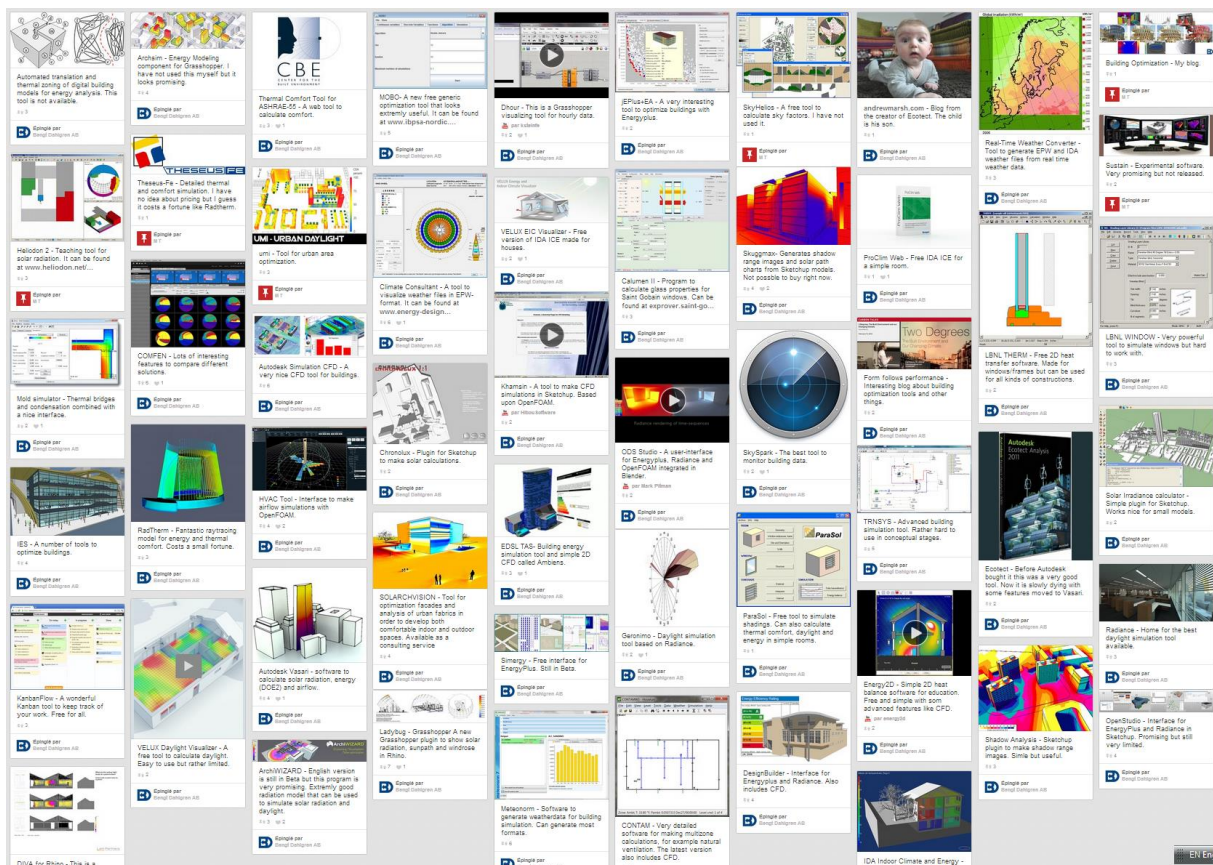


USER-FIT: A framework to evaluate the usability of building performance simulation tools in the architectural design process.

September 2023



Authors

Shady Attia

Maha Attia

Sustainable Building Design Lab, Dept. UEE, Faculty of Applied Sciences, University of Liège, Belgium

Methods Chair, Faculty of Business Administration, Radboud University, the Netherlands

Date de publication : 01 septembre 2023

Report Version: 2.0

Publisher: Sustainable Building Design Lab, Liege University, Belgium.

Citation

Attia, S., Attia, M. (2023) USER-FIT: A framework to evaluate the usability of building performance simulation tools in the architectural design process, Report, ISBN: 9782930909240, Sustainable Building Design Lab, Liege University, Liege, Belgium. DOI: 10.13140/RG.2.2.13584.38400.

© Copyright Sustainable Building Design Lab 2023

Cover Illustration: Building optimization Tools and methods to optimize building according to energy, daylight and thermal comfort, Collection Bengt Dahlgren AB and Max Tillberg

All property rights, including copyright, are vested in I Sustainable Building Design Lab. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the Sustainable Building Design Lab.

Published by Sustainable Building Design Lab. SBD Lab, Building 52, Quartier Polytech 1, Allee de la Decouverte 13A, 4000 Liege, Belgium

Disclaimer Notice: This publication has been compiled with reasonable skill and care. However, neither Sustainable Building Design Lab nor the authors make any representation as to the adequacy or accuracy of the information contained herein or as to its suitability for any particular application and accept no responsibility or liability arising out of the use of this publication. The information contained herein does not supersede the requirements given in any national codes, regulations or usability standards. It should not be regarded as a substitute for the need to obtain specific professional advice for any particular user experience testing. Views, findings and publications in this report do not necessarily represent the views or policies of Liege University.

DOI: 10.13140/RG.2.2.13584.38400.

Table of Content

- 1. Introduction 8
- 2. Literature Review 10
- 3. Research methods..... 14
- 4. Performance Simulation Problems 15
 - 4.1. Decisional Needs and the User 15
 - 4.2. Decision support management 16
 - 4.3. Decision Quality 17
- 5. Definition and stages of usability 20
 - 5.1. Planning 21
 - 5.2. Participants recruitment 21
 - 5.3. Types of usability testing..... 21
 - 5.4. Conducting the usability and validation 22
- 6. Usability Testing Framework for Building Performance Simulation..... 24
 - 6.1. Design Decisional Need 25
 - 6.2. User Need..... 25
 - 6.3. Test Case and Design Tasks..... 26
 - 6.4. Decision Support System Usability Testing..... 26
 - 6.5. Informative Visualization 29
 - 6.6. Decision Support Evaluation..... 29
- 7. Interview Results..... 30
 - 7.1. Validation of the framework 30
 - 7.2. Quotations..... 30
- 8. Discussion..... 30
 - 8.1. Findings and recommendations..... 31
 - 8.2. Strengths and limitations..... 32
 - 8.3. Implications for practice and future research..... 33
- 9. Conclusion 35
- Acknowledgments 36
- Appendix I 37
- Appendix II 38

USER-FIT: A framework to evaluate the usability of building performance simulation tools in the architectural design process.

Abstract There is no consensus on usability testing in studies investigating the integration and decision support of building performance simulation in the architectural design process. This report presents an evaluation and reporting framework for usability testing and user experience investigations of building performance simulation tools in workplace and laboratory settings. A six-step usability-testing framework, USER-Fit, is presented based on the Delphi process involving international experts and a review of standards. A framework lists the steps to perform document and report user experience evaluations during the design workflow and design process. The implementation of the framework will enhance the impact of simulation-based design decision support by ensuring a structured and systematic approach and enabling objective and comparative evaluation of building simulation tools. The framework is the first formal study for documenting and reporting usability studies interventions in the architectural and building design process.

Keywords: user experience, reporting, workflow, design, decision support, building simulation, user friendliness.

1. Introduction

BPS tools are computational decision support tools capable of generating quantitative evidence and insights supporting decision-making about possible performance-based design alternatives and solutions [1]. Thus, the potential of BPS tool integration in architectural and urban design workflow is high for robust decision-making. However, the uptake of those BPS tools in practice is low. A set of conditions impedes architects and urban planners from enabling BPS tools in the design process and relying on computation to yield better decisions under conditions of uncertainty.

Building performance simulation and analysis tools (BPS) have real potential to decrease the environmental impact of buildings and cities and improve energy efficiency and indoor environmental quality through their wide adoption and meaningful use in architectural, engineering, and construction education and practice [1], [2]. This is the rationale behind the American Institute of Architects (AIA), started in 2008 by the Committee on the Environment (COTE), an AIA Knowledge Community, and strengthened by the AIA Commitment to the 2030 challenge to design all new buildings, developments, and major renovations to be carbon-neutral by 2030 [3]. Similarly, the Architects' Council of Europe (ACE) has actively campaigned to open up the sustainability discourse to a more holistic approach, integrating performance simulation in the design process and architectural curricula beyond energy efficiency and across the whole lifecycle of buildings [4], [5]. However, there are huge gaps between the design practice and the potential of BPS to inform design decision-making, primarily due to cognitive, technological, and social challenges [6], [7], [8], [9]. The cognitive challenge mainly concerns usability issues, which did receive significant attention in the BPS community [10], [11], [12], [13], [14]. Unlike many other industries (e.g., aviation, automobile, IT, consumer software, and consumer electronics), where usability is the norm in product design, the practice of usability testing in BPS software and programming interface development has been sporadic, unsystematic casual [7] and shallow [15]. The lack of systematic user experience measurement for BPS tools is partly due to insufficient attention to usability frameworks and methods in performance-based building design and project delivery workflows.

Therefore, there is a need for a BPS-specific usability framework to facilitate the adoption [16] and meaningful use of BPS. A framework that can be used to increase efficiency and productivity, ease of use and learning, increase user retention and satisfaction, decrease human errors, decrease development time and cost, and decrease support and training costs. This research aims to extend the knowledge of usability testing for BPS tools. This paper presents the initial form of a unified framework of BPS usability called USER-FIT. The scope of the research paper is mainly building designers, BPS modelers and developers during early design stages under high uncertainty of design concepts and solutions. Specifically, the study seeks to cover the concept design, developed design and technical design stages based on the RIBA workplan [17]. This qualitative research aims to produce and describe a framework for BPS software usability testing and user experience trials. The paper will serve as a seed for future usability tests. The paper covers usability topics and proposes procedures found in literature and ISO standards - and is limited to developing a theoretical framework without controlled study trials (case study testing). Instead, a triangulation – through the use of multiple data

sources to reach convergence - of the framework and its six working steps has been performed to validate the comprehensiveness and context richness of the study.

This paper provides a valuable contribution to the new body of knowledge on usability testing of BPS used during early design stages, considering existing ISO standards. The detailed criteria, usability testing indicators, and properties describe the critical usability testing measures. This paper includes instructions to apply usability testing to evaluate and measure user experience in a structured and objective way. The suggested framework, USER-FIT, could also be used as a principle for developing BPS-specific guidelines and standards. Contrary to most usability approaches, which concentrate solely on software usability testing [18], we took other adoption and fitness criteria of usability testing into account during the design process. For software developers and architects, this paper sheds light on the importance of software and Graphical User Interfaces (GUI) development and criteria to be embedded during their evaluations. Defining usability testing criteria for the BPS tool can increase the BPS simulation uptake in practice and help the construction sector towards climate change-proof building designs.

2. Literature Review

The International Building Performance Simulation Association (IBPSA) has developed a body of knowledge on user experience since its establishment in the 1980s [19]. Since its establishment, there has been a proliferation of Building Performance Simulation (BPS) tools used worldwide. Until 2017, the U.S. Department of Energy (DOE) maintained an up-to-date listing of 417 BPS tools on the Building Energy Software Tools Directory (BESTD) website, ranging from research software to commercial products with thousands of users (see Figure 01). In 2015, the list was revised to reach 260 tools. In 2018, IBPSA USA started hosting the newly revised list, reaching 214 tools in 2020.

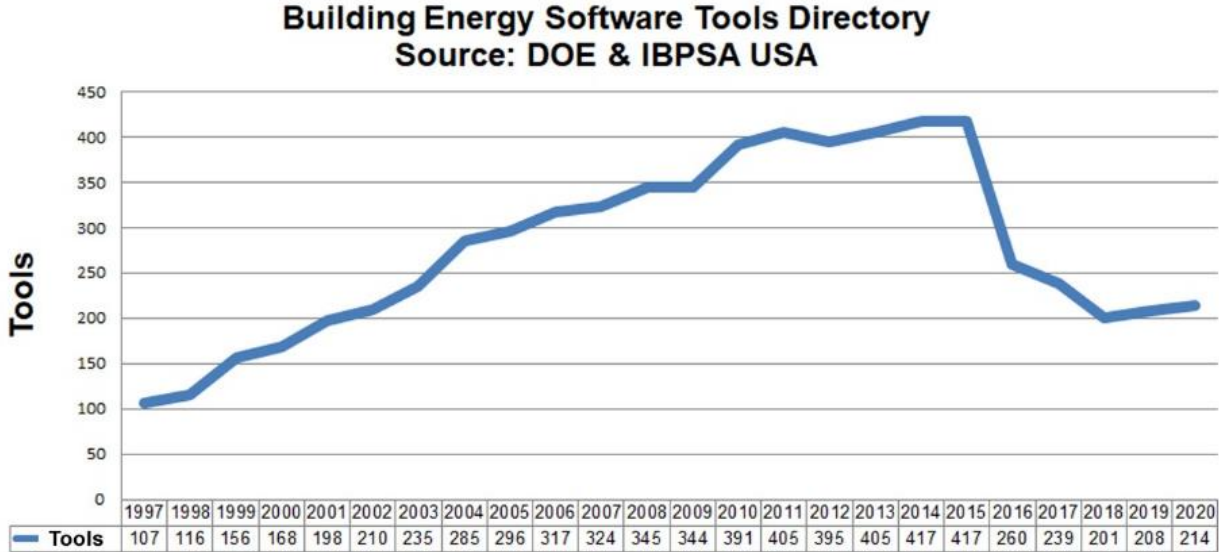


Figure 01, BPS tools developed between 1997- 2020.[20], [21]

Based on Figure 01, there is a general question about the underlying premise of the uptake and usability of BPS tools in the industry. Evidence suggests an increasing use rate and engagement of BPS among building professionals [22]. The penetration of BPS tools used for energy-efficient and green building design has been documented [23] in many studies [24]. The industry has embraced simulation when it comes to things like rating (e.g., LEED, BREEAM, DGNB, NABERS) and ensuring designs meet Energy Performance Regulations. So it is likely that existing tools such as IES, DesignBuilder, OpenStudio, Sefaira and similar are fully adapted for compliance activities. At the same time, the use of BPS for code compliance is facing challenges to move away from a prescription of minimum physical parameters to an increasingly relevant but complex world of performance-based modeling [25], [26]. Also, for compliance checking, far more sophisticated performance criteria are needed than the current practice of simplistic single-number performance targets [27].

Since 2006, there has been a proliferation of parametric simulation tools such as Grasshopper, Ladybug, Honeybee and similar plugins and Application Programming Interfaces (API) for existing simulation engines, such as EnergyPlus and TRNSYS. The plugins and APIs are used to automatically optimize the energy [24], daylight and carbon performance of buildings. However, they are not effectively used, and their uptake remains low in informing green building design[28], [29], [30].

Moreover, the persistent gap between predicted and actual performance, often referred to as the 'Performance Gap,' raises questions about the trustworthiness and effective utilization of BPS in design decision support. There is also the issue of the modelers' experience. BPS tools require engineering knowledge and expertise. Finally, the reason for the low uptake of BPS tools beyond compliance purposes is multifactorial. The challenge of complex building design, a large number of undecided parameters and wide ranges and complicated parameters affect the energy performance of buildings, many of which are out of the scope and expertise of architects and building designers [31]. Unreliable performance prediction is related to many uncertainties, including the multi-criteria nature of the design. This would explain the sharp decline in Figure 01 and explain why we do not see many new tools emerging over the last 10-15 years.

Therefore, this study is focused on the usability of BPS tools and user experience as one of the factors that impede the building simulation in practice. There is a well-established body of knowledge about the fit and usability of building simulation tools. Several studies addressed the effective use of building performance simulation in design and how to measure usability issues and simulation software limitations of the implemented prototypes [32]. The work of Lomas on the empirical validation of building energy simulation programs is one of the earliest comparative studies to evaluate BPS under interfaces [33]. Also, Pati and Augenbroe called for integrating formalized user experience within building design models [34]. Despite the significant progress in advancing building performance simulation applications and the array of tools posted at the directory of the Building Energy Software Tools Directory [35], it is evident that the BPS market uptake in building and city design, beyond compliance, is low [1]. According to the previously identified 'Seven Deadly Sins' of software development by Maver [36] in 1995, in computer-aided architectural design, failure to evaluate BPS tools is a hindering factor for BPS evolution and uptake. As mentioned by Clark [37], "*There is no independent investigation of tool ease of use and applicability to real problems.*" By investigating the user experience problem in the existing literature, we found very important studies that focused on providing credible user feedback from practitioners.

Among those studies that addressed the effective use of BPS in design with Human-Computer Interaction (HCI), several studies were published. One of the earliest studies on building performance simulation usability is the work of Newton et al. [38], who discussed the application of BPS from a user perspective. His work described how designers understand and use the simulation output, explore design options, interact with the software interface, and explore the potential of parametric analysis. Also, Morbizter [39] evaluated the graphical user interface of simulation tools and classified BPS tools based on the levels of ease of use. His work aimed to develop a tool that enables non-simulation experts to create a detailed simulation model and monitor the use of simulation at an early building design stage in an architectural design practice. More formally, Hopfe and Hensen [40] measured the experience of user groups of BPS prototypes, aiming to improve the software design and observe the reactions of simulation users. Also, Mahdavi [41] investigated the building simulation tools and environment usability. He reported that usability features in general and user interfaces in particular lag computational tools. Early simulation tool developers were mostly engineers and physicists, not

experts in HCI (Human Computer Interaction). He confirmed the significant potential for enhancement of user experience.

A series of PhD studies also addressed the usability of BPS tools and conducted several user experience tests. For example, Peterson developed a simulation program for informed decisions of architects during early design [42]. Following Peterson's program, students and researchers conducted usability testing within a master course to assess the operability of the developed tool. Attia [43] developed a tool for designing decision-making zero-energy residential buildings in hot, humid climates. Extensive usability testing took place, investigating the robustness of the final design outcomes and reliability and using building performance simulation tools to inform architects' design decision-making during early design stages.

Furthermore, Negendahl [44] focused on integrating computational collaborative simulation models in the building design phase. Following a case studies approach, he investigated the ease of integrating BPS tools in the design process and collaboration between building designers and simulation modelers. The study targeted improving the speed of multiple and parallel performance evaluations, reducing working hours for the simulation modelers, and informing the design of high-performance buildings. Meex [45] developed early design support for material-related environmental impact assessment of dwellings in Flanders. Usability testing focused on the graphical user interface (GUI), data input and output, and real-time feedback. They suggested three usability testing measures based on the work of Weytiens et al. [15], namely: 1) time of use, 2) adaptability and flexibility, and 3) comparison and feedback loop.

Other studies, like the work of Bazafkan [46], assessed the usability and usefulness of parametric BPS tools in the architectural process. He used some usability testing criteria, including the learning curve, data entry ease and error notification features. Jayapalan [47] conducted a usability evaluation of a building simulation web-based design guide using the ISO 9241 standard. Purup [48] developed a practice-oriented approach using building performance simulation tools to fit the design workflow in the early stages. Purup suggested involving design practice in the BPS tools development without dictating a tool-centric workflow.

According to the examined studies, user experience tests are designed and implemented differently using contradicting measures and criteria. Some limited approaches and frameworks systematically evaluate the usability of BPS tools. The reliance on case studies to test a tool is insufficient and ineffective in addressing workflow and cognitive discrepancies. Designing and implementing BPS tools and Application Programming Interfaces (API) without assessing the users' total experience and tools fitness [49] through usability testing is considered incomplete. BPS methods, tools, and front-end interfaces risk remaining instrumental and academic applications rather than truly influencing architectural design and urban planning practices. As shown in Figure 2, the critical lesson from the BPS usability debate in the last two decades is that more attention should be paid to the user and the design context rather than developing ever more advanced tools [37]. Despite BPS-based decision support being a relatively mature field with more than 40 years [50] of academic

and applied history, it still lacks systematic and rigorous methodologies for assessing the usability of BPS methods and tools, executing usability testing sessions, and managing the knowledge derived.

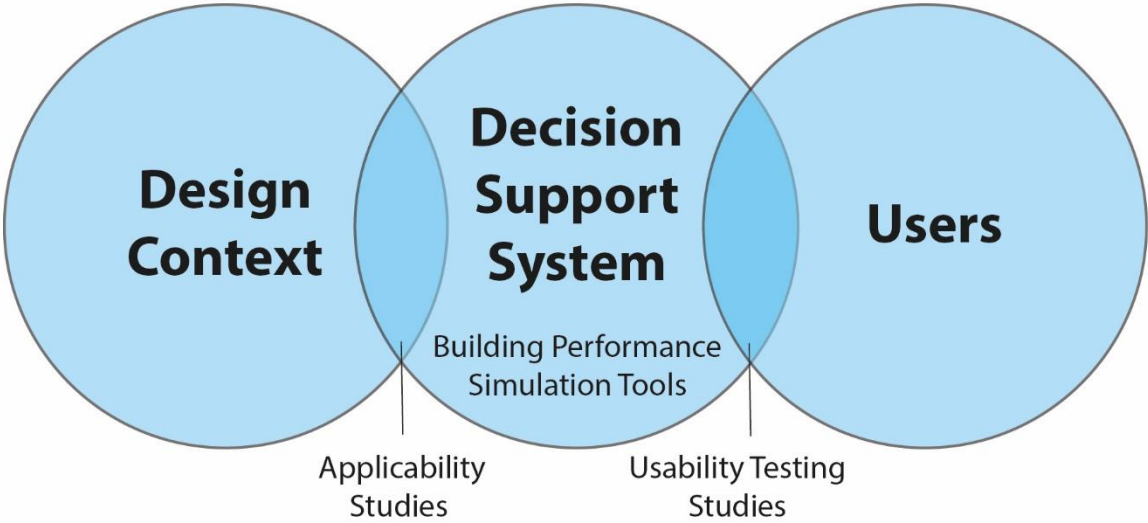


Figure 2, Decision support systems need attention to the user, design, and planning context.

3. Research methods.

This study follows a qualitative research methodology based on a three-step approach to develop and validate the literature-introduced framework. The study is designed as a conceptual article to create new knowledge by building on carefully selected sources of information combined and tested through empirical insights [51]. A detailed report has been published, including the details of the research methodology [52].

First, the systematic literature review focused on recent studies published in the last 20 years and performed usability testing during the development of BPS tools and methods for design decision support during early design. The systematic review protocol was developed with input from systematic review methodologies in the literature. The eligibility criteria and information sources are described in Appendix 1 of the study report [52].

Second, students at Liege University were asked to answer research questions during their attendance at a building performance simulation course. The question was integrated into the course evaluation of a simulation class taught for second-year architectural engineering master's students by the first author and designed based on a previous questionnaire design methodology. Architecture students were asked to rank the most common design activities and where building performance simulation can be helpful during early design stages. The survey question was answered between 2015 and 2023 and was answered by 165 respondents (See Section 4.1).

The third and final research stage involved interviewing practicing architects. Despite the non-empirical nature of conceptual articles [51], researchers in the engineering fields are not used to this type of study without conducting empirical research for validation. Therefore, in-depth interviews took place with BPS tool users [53] to test the developed framework. The criteria for selecting interviewees were that the person should be native Belgian, educated in a Belgian school of architecture and work in a design firm in Belgium. The interviewee had to share a case study of a building simulation exercise, of a project during early design stages, where he/she is using a BPS tool for design decision support. Similar to the approach of Purup, we only interviewed 12 architects [54] in French and Flemish in their offices. The interviewees were recruited through the professional network of the first author of this paper and were conducted randomly depending on their availability. The first author, who has experience with interviews [55], [56], used a semi-structured interview approach. The interview questions were revised by the second author, who is a software engineer specializing in usability testing and can be found in Appendix II of the detailed study report [52].

Moreover, the researchers created memo logs during interviews, allowing learning from the subjects and reflecting on several specific ideas [57]. Memo logs were handwritten during interviews with users. They were useful for improving the quality of the developed framework (USER-FIT) and its inclusiveness. More importantly, the *memo log* was consulted several times during the barrier definition and framework validation stages. The memo logs helped define usability as how a simulation tool is useful, usable, and satisfying for the intended users to inform the decision-making to achieve

sustainable building design. The framework was validated between 2021 and 2023 through interviews that aimed to uncover the point of view of interviewees gradually.

4. Performance Simulation Problems

In this section, we present the summary of the results of the systematic review and memo logs regarding the problems that hinder using BPS in the architectural and engineering practice during the early design stages. A detailed description of the performance simulation problems can be found in the detailed study report [52]. Integrating performance simulation in the early design process is central to the ongoing relevance of the profession of architecture and urban planning within local and global communities. Building performance is a critical event to our shared life on earth. As key leaders in creating the built environment, architects play a pivotal role in determining lifetime building environmental performance. At the same time, various analytical building performance simulation tools for design decision support during early design stages have been developed. In the following section, those barriers are summarized below as three key research gaps:

4.1. Decisional Needs and the User

Based on a sample of 165 early-career architects and building designers, Figure 3 shows the most common way of modeling applied as part of the design activity during early stages involving free-hand drawing, computer-aided design, and building performance simulations. The respondents were asked to indicate the design ways where BPS can play a major role in informing their design. The result confirms the variety of tasks and activities BPS can use during the early design stages.

Ranking of the most common ways of modeling during early design stages

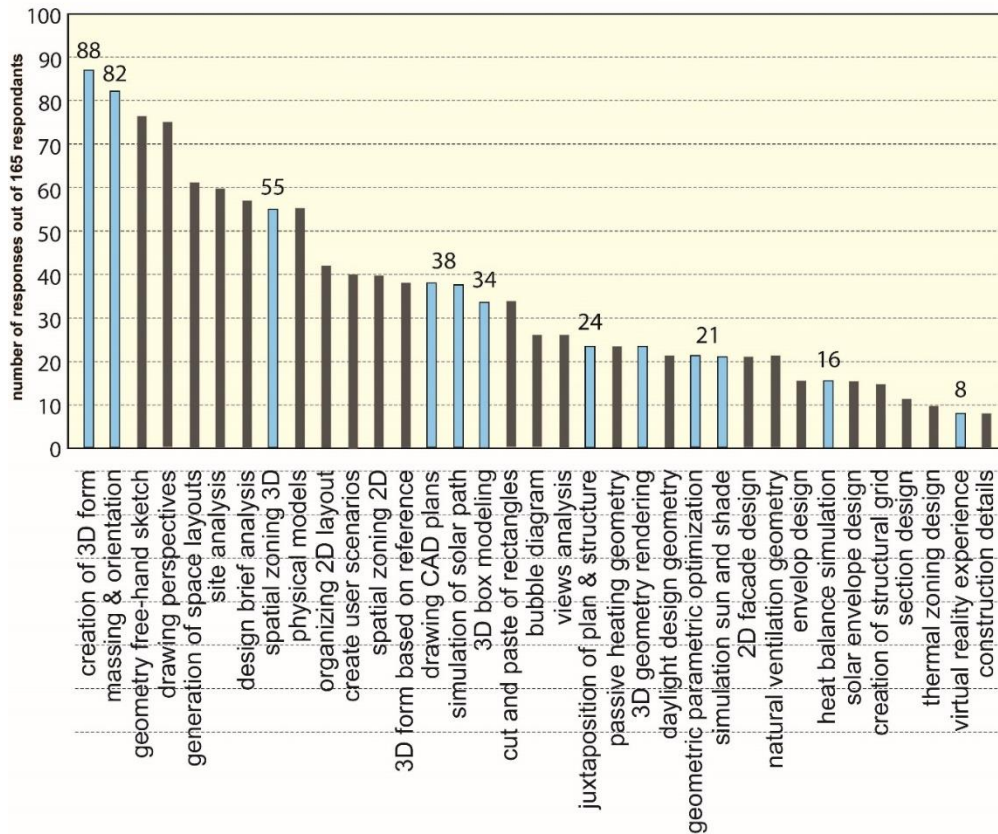


Figure 3 – a sample of architects and building engineers ranking their most common way of modeling during early design stages and indicating (in blue) the potential of BPS use.

In summary, the critical lesson from the integration of building performance simulation in the architectural and planning design process debate in the last decade is that rather than developing ever more advanced tools, more attention should be paid to the user and their way of organizing form and function [58], [59], [60] [61] [62]. Designers suffer from a very low rate of BPS modeling productivity [63] and the interoperability between BIM and BEM [64], [65]. Combining this user perspective with acknowledging the design's multi-collaborative iterative process nature and the 3D geometrical representation [66] logically results in a focus on seamless exchange during the decision support process.

4.2. Decision support management

The clarity of decisions and the ability to guide and communicate is another challenge that hinders the integration of performance simulation in the design process. While trial-and-error processes may be an option to integrate performance simulation in the design process, architects and urban planners usually do not consider it a decision-support option. Managing the design complexity under time constraints requires timely and rich decision support regarding diversity and adaptability. Despite the proliferation of parametric analysis [67] as an advantaged decision support method, the complexity of

the workflow process [62] and the computation time needed to inform the design decision during early design stages is at an early stage of adoption by most BPS tools.

Performance simulation with embedded intelligence is another missing feature in many existing tools. By embedded intelligence, we mean the ability of the BPS tool to reflect on its operational performance by providing advice to users in the form of pre-set building templates, warning or auto-filling features for building code compliance [20]. Comparative approaches to decision support systems are more effective in evaluating a set of well-defined competing options [68]. The ability of many BPS tools to compare different design alternatives, benchmark, or create reference cases cannot be found in most existing BPS tools. To achieve such a comparative decision, the support designer must invest in advanced coding and automation of simulation output post-processing. Remembering earlier design alternatives while running a simulation and visualizing and comparing the improvement between alternative 1 and 2 results is essential for architects and urban planners. On the other hand, multi-objective optimization to explore a nearly optimal solution space or design decision trees (simple and extensive) to explore the influence of geometry, orientation, envelope, structure, material, HVAC, and renewables is hardly found in many tools. Even the ranking of influential parameters [69] or the flexible and adaptive parameterization of sensitivity analysis or uncertainty analysis does not fit users' needs and does not fit the design process [70].

The third issue with BPS decision support systems relates to visualization and the clarity and ability to understand the results and graphs [71]. Very few tools allow interactive visualizations such as parallel, coordinated graphs [72] or interactive Pareto front scatter plots, bar charts, or energy performance rating graphs for compliance with the EPBD or Energy Star requirements. The complexity of post-processing and BPS visualization information overload does not make it easy to understand simulation outcomes or grasp the results [73]. There is a need to standardize the simulation results' functional units and the link between performance information and building model composition [74]. The standardization of visualizations of BPS output can be highly relevant to decision-making in the early stages of the building design process [9].

4.3. Decision Quality

The quality of decision support of performance simulation is based on the robustness of decision-making to yield informed decisions under uncertainty. As a decision support system, BPS can enable practitioners to efficiently explore a space of design alternatives and compare them in terms of performance [75]. In recent decades, many BPS tools have been developed to support architectural design and urban planning in an increasingly fast, user-friendly, and affordable way. Nevertheless, however elegant and sophisticated they may be, these technical developments run the risk of remaining instrumental and academic applications rather than truly influencing architectural design and urban planning practice. Three major obstacles impact the design decision quality and reduce the uptake of BPS in the project delivery process.

Firstly, understanding the architectural design and urban planning context is one of the missing aspects of value-based decision-making. The critical lesson from the BPS debate in the last decade is that rather than developing new advanced tools, more attention should be paid to the design and planning context [70], [76], [77]. Understanding the urban context, such as the solar shading from neighboring buildings or the building occupancy, impacts the building's performance. Unfortunately, in many fast-tracked projects, design teams often need to design for (and policy-makers need to plan for) unknown contexts due to rapid development, which makes simulation outcomes obsolete [78]. Running building performance simulations with a generic context introduces errors compared to the "real" context models and decreases the accuracy and reliability of the results. Therefore, it is critical to inform decisions assuming that the context (physical, occupancy, and climate) matters to reduce the generalized pre-design guidance and increase the decision quality.

Secondly, the variety of key performance indicators (KPI) and the lack of structured approaches to assessing various factors' impact on building performance are reported as a burden in the building simulation literature [6]. Architects and urban planners use indicators to perform decision analysis and scenarios-based modeling to test strategies over myriad plausible paths. Research reports the decision-making stress of designers associated with abundant performance indicators and the lack of decision trees to understand better the associations and correlations that can inform design decision-making [79]. The easy identification and recognition of policy-relevant or standards-compliant scenarios and robust design strategies via building performance tools are part of the quality of the process of decision-making. Therefore, the flexibility of using different performance indicators, functional units, and targets, including prescriptive and performance-based approaches, for compliance or performance thresholds is perplexing.

Moreover, the complexity of building impact assessment during a project's life cycle requires structure and standardization. Using universal and simple performance indicators to evaluate the green house emissions, energy use, and indoor environmental quality can lead to analysis paralysis. Therefore, it is important to standardize the building performance outcomes based on universal functional units and metrics.

Thirdly, more attention should be paid to the decision quality from the user perspective of BPS tools. The decision quality requires well-prepared architects with a good knowledge of building physics [80] and can interpret and use buildings' performance simulation results. Several studies reported the lack of knowledge [81] or low preparedness of many designers [82] and modelers [83] to use BPS tools [84]. However, despite many architects' lack of knowledge and skills in performing simulations [85], little attention has been paid to decision quality from the user perspective. Recognizing the user perspective of BPS tools and acknowledging the interdisciplinary nature of architectural design and urban planning is key to improving the decision quality. Many BPS tools underestimate usability testing and the real complexity of robust design decision-making. Architects and urban planners who use BPS tools face an information overload that results in decision-making stress [86]. Limiting the number of parameters and associating a rank to each parameter reduces the likelihood of incorrect interpretations [87]. Using these limited parameters, modelers could have additional time to acquire

accurate values of influential parameters [88]. Modelers could further quantify the uncertainty of the reduced set of parameters to build confidence in simulation results.

Moreover, the difficulty of using BPS tools is related to the time and investment needed to achieve robust decision-making. Therefore, building performance must lower the decision-making stress and avoid analysis paralysis. Analysis paralysis of design teams is associated with abundance and overshot of information, paralyzing the design outcome.

5. Definition and stages of usability

A usability study is a research method that assesses how easy it is for participants to complete core tasks in a design [89]. Our systematic literature review on usability testing definition and techniques resulted in three ISO standards, namely ISO 9241-210 [90], ISO 25010 [91], and ISO 25066 [92]. As shown in Figure 4 and Table 1, the three standards evaluate user interaction with front-end development or design in creating, testing, and refining the user's side of a website or application. During a usability study researcher, follow participants as they interact with the product. Therefore, users' feedback helps the BPS software design team improve the user experience.

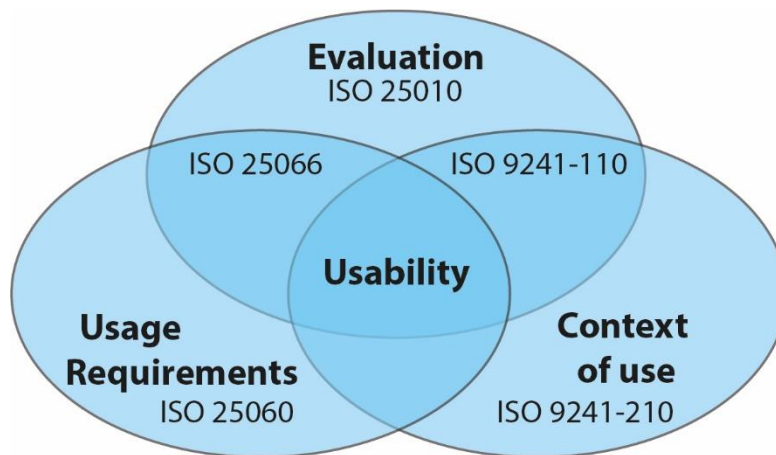


Figure 4 – The key usability testing standards and their domain of operation for software and human computational interactions

Table 1: Summary of relevant ISO standards used as a basis to develop a framework USER-FIT for building performance simulation usability.

Standard	Title	Relevance to USER-FIT
ISO 9241-210: 2019	Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems [93]	<ul style="list-style-type: none"> • Define the approach to centered design for the human operator • Specify requirements for users based on principles of usability
ISO 9241-110: 2020	Ergonomics of human-system interaction — Part 110: Interaction principles [93]	<ul style="list-style-type: none"> • Specify the needs of users in the interaction with the system.
ISO/IEC 25010: 2011	Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models [91]	<ul style="list-style-type: none"> • Specify the features for assessing the quality of a software.
ISO/IEC 25060: 2017	Information technology — User interface component accessibility — Part 25: Guidance on the audio presentation of the text in videos, including captions, subtitles, and other on-screen text [94]	<ul style="list-style-type: none"> • Define terminology information related to the central design of the human • Specify information related to usability
ISO/IEC 25066: 2019	Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Common Industry Format (CIF) for Usability - Evaluation Report [92]	<ul style="list-style-type: none"> • Understand how to document an evaluation of usability.

Usability studies can take place at various points in the product design process. Concept and interface testing are applied at the beginning of the product prototype development with few interactions. A usability study should be conducted when the prototype is fully interactive. This is the

most common time of usability testing because it gives the design team insights on what needs to be revised or added before the product launches. Also, usability testing should be conducted when the application of the software tool is complete. Study participants navigate the software from the opening page to check out as part of any usability. As a researcher, the developer collects feedback while watching the users interact with the prototype. In some usability studies, users get interviewed after interacting with the prototype to get more feedback. The following section describes the key stages of usability testing based on the reviewed ISO standards.

5.1. Planning

One of the key elements of conducting usability testing is to define the study goals and key study questions. By establishing research goals by the developers' team before the usability study begins, designer biases can be avoided. The phrasing of interview questions must be done thoughtfully. Open-ended questions and non-leading questions are recommended. Also, ready-validated tests and metrics are encouraged to be used. Also, the research criteria must be defined beforehand. Based on those elements, a step plan must be established to conduct the study, including the participants' selection process. The planning of the usability testing is the key step for a proper plan to get successful research results and honest feedback.

5.2. Participants recruitment

It is recommended to recruit a handful of participants for the usability study. Five participants are the minimum to start with in a usability study. This sample size is large enough to uncover major user issues but small enough to keep costs and time down. Participants must be found from a representative sample. The small group of testers should represent the key user groups and user groups that are often marginalized. The recruitment can continue beyond five until the research team reaches saturation, and the moderator must understand the feelings and emotions behind their answers and the reason behind each answer. Follow-up questions for more clarity are very encouraged.

5.3. Types of usability testing

Several usability techniques have arisen over the years, with the most popular methods being usability testing and heuristic evaluations. According to Nielsen, '*heuristic evaluation is a process where experts use rules of thumb to evaluate the usability of a user interface independent reviews and reported issues*' [95]. Evaluators use established heuristics (e.g., Nielsen-Molich's) and reveal insights that can help design teams enhance product usability from early in development [96]. However, heuristic evaluations are not recommended for BPS tools because they do not involve the user's opinion in the testing. Therefore, usability testing is the most effective method to evaluate user interface with real tasks through potential users. The problems found with usability testing are true problems in that at least one user encountered each problem.

Among the usability testing methods, there are two types of usability studies: moderated and unmoderated. A moderator guides participants through the study in real-time in a moderated usability study. The moderator aims to help participants interact with the product and collect feedback. On the other hand, an un-moderated usability study does not have a designated moderator. In the unmoderated usability study, participants test the prototypes without human guidance. Usually, the study is recorded on video, and the user experience team reviews the video footage after the study. Moderated and unmoderated usability studies have benefits and limitations depending on the scope and goal of the study. As shown in Table 2, moderated usability testing is more advantageous than unmoderated usability studies; however, both types suffer from subjectivity and bias.

Table 2: Comparison of the advantages and disadvantages of moderated and unmoderated usability studies [97].

	Benefits	Limitations
Moderated usability studies	<ul style="list-style-type: none"> • Guide the participant through the study. • Ask specific questions and follow up in real-time • Rapport building between the moderator and participant 	<ul style="list-style-type: none"> • The moderator could influence or bias the participants • Less flexible • Participants may not identify with the moderator
Unmoderated usability studies	<ul style="list-style-type: none"> • The participant uses the tool in the real world • Participant completes a task on their own time and in their own space • Participants may feel more comfortable giving feedback without others around 	<ul style="list-style-type: none"> • No human guidance if issues arise • No real-time follow-up questions • Little to no control over the environment

Usability testing requires practice and experience so that moderators interact well with participants, which is critical to the success of any usability test and the validity of insights. The book of Joseph, Dumas, Beth, and Loring (2008) presents the golden rules of moderating a usability test [98].

5.4. Conducting the usability and validation

Conducting the usability testing will be mainly based on gathering and collecting data. Major pain points of users must be identified based on specific criteria and usability tests with the product. Chapter 5 presents a comprehensive usability testing framework with several criteria and steps to conduct a usability test for BPS software. After collecting the data based on the proposed criteria and steps, researchers should synthesize the results, looking for the actual meaning in the data. Patterns in quantitative data should be identified, and similar trends in qualitative data from participants' answers to interview questions. The study must be validated before sharing or promoting any insights with developers.

A usability test must be unbiased and neutral regarding the product. As a qualitative method, a usability study is a subjective approach. Therefore, usability researchers must be aware of a series of biases. A bias is a favoring or having a prejudice based on limited information. All humans are susceptible to those biases. The key is to identify those biases and to become aware of them to guard against them. Many biases are likely to come up during the moderation of usability studies, including:

- **Implicit bias:** The collection of attitudes and stereotypes we associate with people without knowledge.

- **Serial position effect:** When given a list of items, people are more likely to remember the first few and the last few, while the items in the middle tend to blur.
- **Friendly Bias:** The tendency of people to agree with those they like to maintain a non-confrontational conversation. Honest feedback is lost.
- **Social Desirability Bias:** People tend to answer questions in a way that others will view favourably.

Identifying the biases and minimizing their effects is part of any usability testing study. Testing and revising the usability testing results allows for more objective evaluations. As a qualitative method, usability studies is a subjective approach that provides many rewarding insights with a shallow risk [99], [100]. But to avoid bias, it is recommended to triangulate the usability tests and feedback forms through various types of tests (see Section 4.3) to validate the findings [101]. Triangulation confirms and validates the quality results using quantitative studies. Using multiple sources and methods can minimize inadequacies in one approach or process.

6. Usability Testing Framework for Building Performance Simulation

The usability testing framework USER-FIT is proposed based on the systematic literature review, memo log, and prolonged engagement. As shown in Figure 5, the framework is developed for usability testing of building performance simulation tools. This framework has divided the process of BPS tools usability testing into six working steps in line with ISO standards [77-79] and collecting systematic data on decision needs, decision support, and decision quality. Each step identifies the responsible persons, activity, input, and output.

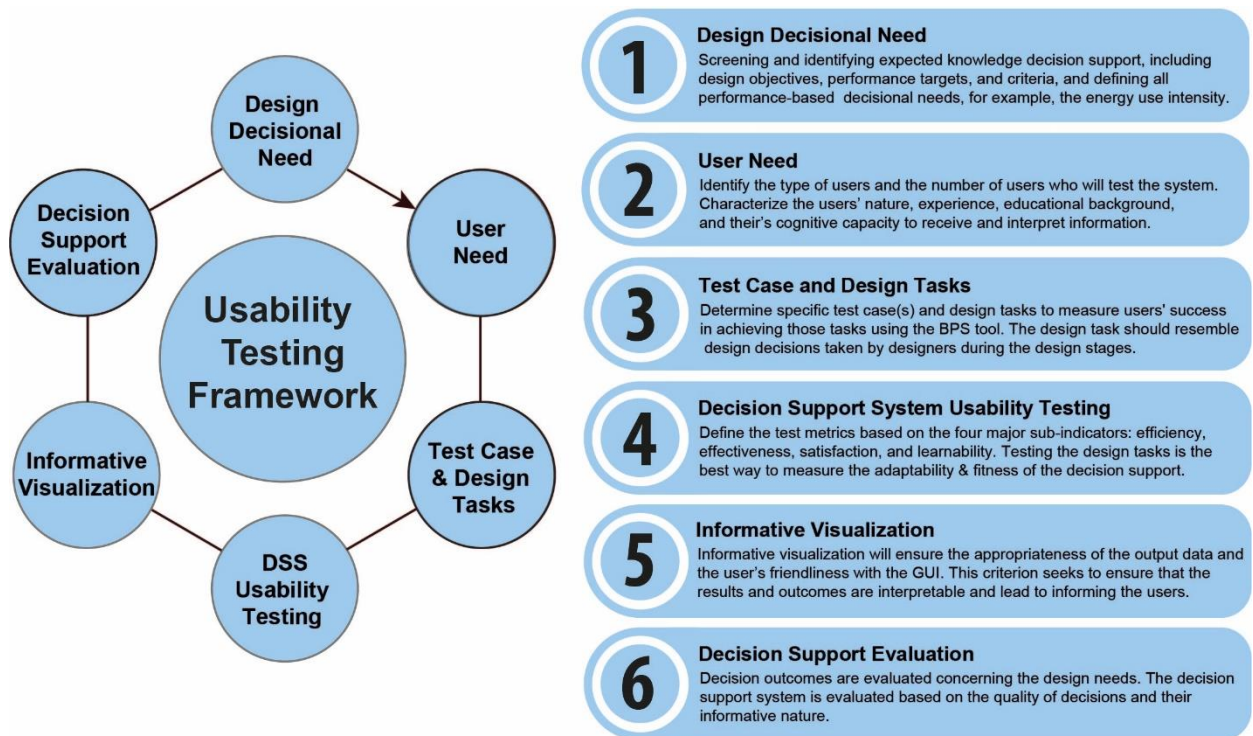


Figure 5, The USER-FIT framework of building performance simulation usability

The first step is to specify the 1) Design Decisional Need to set a goal of developing usability planning and assign actors. The second step is to assess the user's need to define the design decision support needs, design constraints, and relevant parties with the testing process. The third step is to identify test cases and design tasks during the workflow to propose and evaluate the utility of decision support. The fourth step is to study the execution of the usability testing system to support decision-making and set system requirements. The fifth step is to visualize information and test the interaction and interpretation of results. The final sixth step is to evaluate the decision support quality and the design decisions according to the design brief of design decisional need and building performance requirements. The results from usability could be used to revise the tool or software and help detect some errors in interface functionality. USER-FIT framework comprises six-stage usability testing and should be used for communication and documentation. In Sections 5.1-5.6, each stage will be described in detail.

6.1. Design Decisional Need

Assessing the design decision-support needs is key in managing users' expectations and eliciting the key decision determinants. This stage involves screening and identifying the expected knowledge concerning the design intent and decisional conflicts that the BPS should support. The characteristic of the design choice or compliance requirements that need to be made (e.g., geometry, envelope, HVAC systems, renewables), number of design parameters (their range and intervals), design alternatives, performance targets (energy use intensity, greenhouse gas emissions, cost, discomfort) [102], degree of uncertainty, the robustness of outcomes, whether it is irrevocable must be assessed [3]. This stage requires defining benchmarks and key performance indicators, such as Energy Use Intensity EUI or GHG emissions. Therefore, cognizance of the architectural design problem or design situation, alternatives, and outcomes is essential. Evoking the perceived likelihood or probability of design outcomes of each alternative can allow software developers to understand users' decisional needs better. The phase of decision-making in the context of project design work stages (pre-design, schematic design, or design development) must also be determined. The AIA [3] and RIBA [103] defined phases of designing and building a project. Building performance simulation tools' expected role in proactively informing decision-making must be pre-defined. Assessing the design decision determinants can provide software developers and modelers with important cues regarding areas of decisional conflict. Accurate and complete decisional needs assessment sets the BPS tools for effective decision-support interventions. A designer's unique perception of the decision based on their knowledge of parametric variations or design alternatives, possible project outcomes, and the likelihood of their occurrence constitutes an example of one determinant.

6.2. User Need

The second stage of usability testing is to identify the users, their ways of communication, their ability of data interpretation, and the number of users [104] for testing the system or BPS software [105]. It involves assessing the users' requirements for the design and evaluation of the usability of BPS tools. User needs analysis is an important criterion that provides information necessary for the simulation and design decision analysis [106]. User analysis is the process of identifying the types of users, including their needs, ontology, ways of communication, and characteristics of each type of user. For BPS, the type of users includes designers at various levels (e.g., modelers, architects, mechanical engineers, building physicists, and urban planners) and in various specialty areas (e.g., envelope design, HVAC design, acoustic design, lighting design, urban design) and students at various specializations. User characteristics for each type of modeler include knowledge and experience of BPS, education background, knowledge of computers including programming, cognitive capacities and limitations, perceptual variations, age-related skills and personality. [107]. User analysis can help software developers create supportive BPS environments targeted toward specific users and have the proper knowledge and information structure that match the users' literacy level [108]. Several studies investigated the customization of BPS tools based on user needs [9], e.g., architects or engineers, to perform building simulations and inform the design decision [109] [110].

6.3. Test Case and Design Tasks

Determining the test case is identifying the steps of a usability testing operation using a specific design case study. As part of the BPS usability, a set of design tasks must be defined and tested concerning the case study [111]. The set of tasks is then analyzed to identify what makes a task easy or difficult. A critical objective of the test case design tasks analysis is to find out why a task's user interface or simulation result representation (e.g., a bar chart vs. a scatter plot to find the primary energy use of a building over a year) is better than another. Modelers can compare user performance associated with different user interfaces by defining a testing task battery and performing a task analysis for the same design decision implemented in different user interfaces [112]. Moder can perform the comparison in terms of time on task, the number of steps, and mental effort, all metrics of usability efficiency (see 5.4).

Design tasks need to be selected to increase the exposure of architects, engineers, and urban planners to decision support and decrease identified barriers interfering with enhancing designers' knowledge and skills to establish environmental design competencies. Problem-based test case scenarios and specific design tasks, such as determining the nearly optimal window-to-wall ratio for different building orientations or sizing of the HVAC systems, should be developed to build knowledge and skills to use them for performance simulation tools better.

6.4. Decision Support System Usability Testing

The fourth criterion of the USER-FIT framework is to identify the test metrics such as success rate, error rate, efficiency in time on task, user satisfaction, etc. These metrics are selected based on the goal and objective of system testing. The result from usability could be used to revise the system and help detect some errors in system functionality. The usability testing of BPS tools should be user-centered and based on decision-making indicators and metrics. These indicators and metrics should be selected based on the goal, design stage, and decision support system testing objective. In this step, the quality of the BPS tool functionality is assessed to detect some errors and revise the decision support system. Knowledge of designers' backgrounds, contextual design realities, and identified design decisional needs will help determine which design tasks are amenable to decision support. Decision support should be adapted to the designer's characteristics and preferred role in decision-making. Consideration should be given to the designer's experience, specialization, and background.

Four major sub-indicators, efficiency, effectiveness, satisfaction, and learnability should be used for testing. The selection of the decision-support indicators should be guided by implementing the ISO 9241-210 and ISO 25010 requirements.

5.4.1 Efficiency

Efficiency is a useful way to measure the resources expended concerning the accuracy and completeness with which users achieve their goals. This characteristic can be timing the user on a task during the usability testing, comparing users' time with peers, or counting the error rates. The following paragraphs explain the three most common efficiency measures based on ISO 25010 and

[113]. Those three metrics, mouse heat maps and eye-tracking cameras can also be used for more quantitative insights.

- a. *Mean time per task*: The average task completion time of all the users participating in the usability testing becomes a measure of efficiency. Calculating the time on task requires coming up with time-based efficiency. The formula for calculating time-based efficiency is indicated in Equation 1, where the end time is subtracted from the start time [113].

$$\text{Time Based Efficiency} = \frac{\sum_{j=1}^R \sum_{i=1}^N \frac{n_{ij}}{t_{ij}}}{NR} \quad 1$$

Where: N = Number of tasks; R = Number of users; Nij = Result of task i by user j; if the user successfully completes the task, then Nij = 1, if not, then Nij = 0

Tij = The time spent by user j to complete task i. If the task is not completed, time is measured until the moment the user quits the task.

- b. *Overall relative efficiency* is the ratio of the time users take to complete a task to the total time taken by all other users. It can be calculated in Equation 2 [113].

$$\text{Overall Relative Efficiency} = \frac{\sum_{j=1}^R \sum_{i=1}^N n_{ij} t_{ij}}{\sum_{j=1}^R \sum_{i=1}^N t_{ij}} \times 100\% \quad 2$$

- c. *Error rate and frequency of occurrence*: This measure counts the number of errors participants made when attempting a given task. Counting the number of errors made by users can be daunting for many researchers, but this metric provides excellent information about the usability of a system. Learnability considers how easy it is for a user to perform a task the first time they encounter an interface and how many repetitions are required to perform it effectively. To help measure this metric and obtain valuable insights from this metric, a short description needs to be set where details about how to score those errors and the severity of certain errors are given. This metric can show how simple and intuitive the system is.

5.4.2 Effectiveness

Effectiveness is a metric that measures users' task completion and accuracy. The sub-indicator includes wider concepts related to decision-making robustness. Effectiveness can be measured using two main usability metrics: the completion rate, also called the success rate and the accuracy of task results. The following paragraphs explain the two most common effectiveness measures based on ISO 25010 and [113].

- a. *Completion rate*: The percentage of users who successfully completed the tasks. The metric can be collected during any stage of development. Effectiveness can thus be represented as a

percentage by using Equation 3. The completion rate is highly dependent on the context of the task being evaluated.

$$Effectiveness = \frac{\text{number of tasks completed successfully}}{\text{total number of tasks undertaken}} + 100\%$$

- b. *Accuracy of task result*: The accuracy with which users achieve specified goals.

5.4.3 Satisfaction

Satisfaction measures the capacity of the decision support system to satisfy users in a specified context of use. The After Scenario Questionnaire (ASQ) and the System Usability Scale (SUS) provide a quick and reliable way of measuring users' satisfaction. The following paragraphs explain the two most common satisfaction measures based on ISO 9241 and [113].

- a. *Task Level Satisfaction*: After users attempt a task, they answer a questionnaire to measure its difficulty. Typically, ASQ consists of up to five questions. These post-task surveys are usually in the form of Likert-scale ratings and aim to gain insight into the task's difficulty from the participants' perspective.
- b. *Test Level Satisfaction*: Test Level Satisfaction is measured by giving a formalized questionnaire to each test participant at the end of the test session. This estimates their impression of the overall ease of use of the tested system. The SUS consists of a 10-item questionnaire in which respondents have five response options ranging from "strongly agree" to "strongly disagree." Scores should be added, and the total multiplied by 2.5. The mean score of the chosen responses should range between 0 and 100. The higher the score, the more users are satisfied. Any value above 75 is considered a good satisfaction rate.

5.4.4 learnability

Learnability considers how easy it is for a user to perform a task the first time they encounter an interface and how many repetitions are required to perform it effectively. Learnability can be measured using three main usability metrics: first-use learnability, learning curve steepness, and efficiency of the interactions.

- a. *First-use learnability*: How easy is the interface for first-time users to use? This is important for users who plan to use the software or tool only once.
- b. *Learning curve steepness*: How quickly do users improve with repeated interface use? This is important for users who want to feel a sense of progression with the interaction and memorability for getting better at using software or tools.
- c. *Efficiency of the interactions*: How high can users reach productivity once they fully learn how to use the software or tool? This aspect is vital for users who frequently use the software or tool.

The three types of metrics can be measured through a five-scale Likert question or recorded video observation, which timed and recorded the screen to identify hot spots and assume the time taken for

a design task. The study of Baxter and Oatley [114] provides an example of software learning ability measurement.

6.5. Informative Visualization

Informative visualization of the simulation result for a given task performed by a specific type of user allows for extracting meaning. The appropriateness of the simulation data and its representation must be evaluated for any BPS graphical interface or dashboard [115]. Communicating data to BPS users and clients is very difficult. Therefore, it is important to visualize building performance [116]. Different visualizations of a common environmental performance indicator can generate different symbolic meanings, design decision difficulties, and cognitive outcomes. One of the most influential design decision supports means to compare a design alternative or parametric variation of the same building and determine whether it is understandable for the task and the user. An analysis of the informative nature of data visualizations and users' ability to interpret is key. Expert review of graphical usability and the effectiveness of various types of representation is a major step of BPS software evaluations [65]. Communicating data-driven analysis requires high-quality data visualization. This can be only achieved through intensive usability testing of user interface properties and their ability to inform users.

6.6. Decision Support Evaluation

The decision support provided by BPS suits should be evaluated from the perspective of the designers and design teams. Decision outcomes can be evaluated from multiple perspectives. An architect may consider a good decision to be one that best reflects their preferences or causes the least negative consequences. Decision support systems should be able to support high-quality design decisions as one that is informed, consistent, and one in which BPS users express satisfaction with how the decision was arrived at [43]. Moreover, the satisfaction with the design outcome must be evaluated separately from the satisfaction of the decision-making quality discussed in criterion four (see section 5.4). The decision-making process can be evaluated in terms of whether the designer participated and developed more confidence and skills. Skills and capacity building can be assessed by the degree to which designers and design teams can transfer the decision-making skills learned to future decisions.

The decision support evaluation does not only involve users' satisfaction or perception, but it can also involve post-occupancy assessment. The decision outcomes, such as improved environmental building performance, reduced occupancy complaints, reduced operation cost or embodied energy, and appropriate facility management, can also be examined through post-occupancy evaluation [117]. Thus, a designer might have felt empowered and satisfied with the decision support tool, while the design outcome may not have been what they expected. Therefore, it is crucial to determine how valuable was the BPS tools or software in decision-making. The evaluation might include visiting the building after occupancy and reflecting on the energy performance gap [118].

7. Interview Results

The following sections summarize the feedback from interviews conducted with 12 practicing architects. Three females and nine male persons with ages ranging from 35 to 55 years were interviewed. The interviewees were all architects. The interviewees come from six Belgian architectural and engineering firms.

7.1. Validation of the framework

During the early interview stages, the final version of the USER-FIT framework, shown in Figure 5, was not articulated. The questions helped the authors to group and classify the interview responses into six categories. At the end of the interview stage and after analyzing the responses of interviews number seven the architects shared their strong agreement to categorize the usability testing process under six main categories. Saturation was reached through the continuation of the interviewee recruitment for 12 participants. The recruitment processes ended in the summer of 2023, when insights and answers started to replicate and overlap. Data saturation was used as a factor to judge the repetition of answers by experts that indicated similar relevant insights.

7.2. Quotations

The interview analysis and memo logs allowed us to highlight several key responses from interviewees, called quotations. A list of quotations was created and is shown in Table 3. The most relevant and representative quotes to validate the framework are presented. All quotes were addressed by one of the steps of the framework. These verbatim quotations played a key role in clarifying the data and framework steps needed to conduct usability testing for a BPS tool.

Table 3: Some relevant interview quotations

Quotations	Interview #	Step #
'I can only integrate the modeling result during one or two design iterations; the tool has to blend into the workflow of the busy design process.'	2	1
'Time is scarce; the tool should be efficient in terms of time requirements.'	5	2, 4
'I don't care about accuracy, but I care about comparisons. I need to improve my awareness through the BPS tool use and not find absolute performance values.'	9	3
'I need the tool to warn me from what might go wrong with modeling assumptions'	1	4
'The user interface needs to be more intuitive and simple to learn: Short steps, clean UI and easy understanding of results. The tool has to be simple to learn.'	6	4
'The tool is too complex and does not fit in the design process allowing to compare design alternatives. The tool should comparatively report the results through common visual graphs.'	11	5
'I use the tool for insights and relevant findings; I need the tool to help me reflect on the consequences of my design choices.'	7	6
'The simulation results must be accurate to ensure the reliability of results and to help me convince the client.'	4	6

8. Discussion

This paper presented USER-FIT, a usability-testing framework to evaluate building performance simulation tools used to support environmental design decision-making for buildings and cities. The framework has the following working steps: (1) for describing decisional support needs; (2) for defining user needs; (3) for testing the usability of test cases and design tasks; (4) for testing the usability of

the decision support system; (5) for providing informative visualization and (6) evaluation of decision support, once fully developed, for developing BPS Usability guidelines and standards. This section discusses the study findings and positions them regarding the state-of-the-art.

8.1. Findings and recommendations

We approached the usability of BPS in terms of design needs, user needs, informative visualization, and decision support evaluation. Along those criteria, we defined usability around several dimensions: useful, usable, satisfying, and learnable, and listed a set of measures for case study tests and design tasks. The USER-FIT usability-testing framework is coherent and allows the measurement of usability objectively and systematically. Most USER-FIT measures are evidence-based and repeatable based on ISO standards that have been used over thirty years of research in computer-aided design [112], [120]. Unlike most usability approaches, which focus on decision support system usability testing only (Section 5.4), we considered other adoption and fitness criteria of usability testing in the design process. The proposed framework forms a unified approach to measure the usability of BPS in the design process and performance-based design workflows while bringing users to the front.

In summary, the following strategies for further developing and synergizing usability testing are required to facilitate the use of performance simulation tools to enhance the quality of designers' involvement in environmental design decision-making:

1. A capacity-building strategy aimed at informing, influencing, and assisting BPS modelers and other building design professionals to gain competence in providing quality design decision support, learnability, and memorability through pre- and post-licensure education.
2. A regulatory strategy aimed at embedding indicators of quality building design decision-support in legislation and policy for compliance and performance-based design (i.e., accreditations standards, the scope of practice regulation, environmental and sustainability design certification, reflective practice tools);
3. A research strategy designed to evaluate further hypotheses underlying the framework for performance simulation tools in the built environment; interventions for sustainable implementation of building design decision support in routine architectural practice and designers' performance across AEC design firms and with diverse building functions; and
4. An architectural and engineering strategy to balance cost, time, accessibility, and performance (efficiency, effectiveness and satisfaction) of simulation tools, including supporting corrective (error detection) environmental design decision-making.

Facing environmental design decisions can be exceedingly complex for architects, mechanical engineers, and design professionals. Environmental design decisions have evolved from concrete and problem-bound to a constellation of shared, environmentally informed decisions across the AEC design industry continuum. To fully implement environmental design decision-making, architects need to be supported to participate in weighing potential outcomes. BPS software developers are well-positioned to provide the necessary usability testing, but many lack the awareness, knowledge, and

skills. Modelers need opportunities to develop and use new skills to be design-centered and responsive to architects' needs.

Furthermore, design practice environments, regulatory bodies, and educational systems can facilitate the integration of usability testing skills in practice. Promoting policies that endorse architects' decision-support interventions that are practical, visible, accessible, evidence-based, and equitably enforced is essential. Professional accreditation bodies can influence overcoming barriers to usability testing and creating supportive environments, exemplary workflows and software tools. By enhancing modelers' and architects' decision-support skills as individuals and collectively as self-regulating professionals, models can make essential differences in the quality of environmental building design decisions.

8.2. Strengths and limitations

Although this research has primarily been focused on usability testing for architectural performance simulation tools, emerging evidence supports the relevance of the simulation-based design decision support approach for other types of design professionals, such as energy modelers, mechanical engineers, and urban designers. Theoretically, USER-FIT is transferrable across architectural and engineering design environments. Studies conducted in different countries indicate that architects want to participate in making environmental design decisions and an abundance of decisions across AEC design settings that are innately challenging, given the need to weigh the benefits and risks across design alternatives and parametric options [121]. This novel framework will help software developers as well as BPS users. In practice, each design decision can potentially cause uncertainty about the best course of action. Therefore, the framework can be used for tailored coaching to the unique needs of designers and clients to perform BPS. The study successfully triangulated the findings and presented key performance criteria that could be used to redesign BPS products to improve usability.

A usability testing framework like USER-FIT and ISO standards (ISO 25010, ISO 25010 and ISO 9241-210) can be used by independent researchers and modelers of IBPSA to ensure the ease of use of BPS in the architectural and urban design practices. To the best of our knowledge, USER-FIT is the first framework that seeks to be implemented and translated into usability tests by BPS developers in the AEC industry based on Human-Computer Interaction (HCI) principles. Despite the work of Hopfe [40], Mahdavi [122], Bleil De Souza [123] and Petersen [124], [125] many other studies listed in the literature review (Section 2), there has been no study that developed a stepped and systematic approach to evaluate the total user experience of BPS tools users. The paper provides the foundation and essential information to understand usability testing, reporting, and user experience reality in BPS-based and performance-based design. The paper will raise the bar for future usability testing of BPS software for what steps and criteria need to be followed and implemented. The paper provides a simple and first-hand insight into usability tests theory and the theoretical boundaries to develop and perform usability tests by and for users and third-party evaluations.

Although we did not perform a case study or experimental trial, the framework criteria and systematic method can be operationalized for all BPS tools and workflows used during the early design stages. We are aware that the paper would have benefited from a demonstration or experimental trial on the applicability of USER-FIT through user experience testing. However, the embodiment of this work is to present a theoretical framework based on the state of the art. The theoretical framework can be used in the future as a basis for developing customized usability testing.

8.3. Implications for practice and future research

In this study, we presented a consistent and structured framework for usability testing. No matter how dedicated the developers of BPS programs are to providing data-driven simulation results, the ability of BPS tools to explain the building physics phenomena and provide building performance interpretations to the designers remains weak. Most existing BPS tools fail to help users interpret simulation results to understand the buildings' heat balance in the form of design parameters that influence the total heat gains and losses. At the same time, The designer's background and experience are essential determinants of performance-based design. Environmental design problems such as high-embodied carbon materials, large glazing surfaces, undersized HVAC systems, thermal bridges, and discomfort may constrain a building designer's ability to think clearly or participate effectively in decision-making. Involving substitute decision-makers and essential others in decision-making requires architects and designers to develop various decision-support strategies and communication styles. Tailored decision-support for design tasks may include providing focused information, facilitating access to appropriate resources, helping designers re-align project outcome expectations, and assisting in clarification of design preference. In addition, support can involve helping designers strengthen their repertoire of decision-making skills through guidance, coaching, and rehearsal.

Usability testing effectively integrates BPS tools in designer practices during early design stages and throughout the project delivery process [3]. So far, few design firms have developed in-house usability testing protocols and case study testing practices to evaluate the implementation of simulation-based design decision support. Integrating new BPS programs with performance-informed workflows and tools in design firms is a challenge from a business perspective [126]. Specific organizational barriers impacting the implementation of computational design and performance simulation-based design decision support in urban, architectural, and engineering design practice include lack of administrative direction to use building performance simulation to provide design decision support, lack of usability testing of simulation tools during the different stages of projects delivery, limited or no mention of performance simulation-based decision support in the final design, and time pressures that do not integrate simulation results as part of the design process and design reviews iterations, and lack of an interdisciplinary design approach [76].

Future research should focus on the pilot studies and empirical validation of the framework with design professionals. The validation can be achieved during performance-informed design workshops, workflow observation sessions and embedding usability testing in the professional design practice to

reveal users' expectations of BPS tools [127]. Usability testing must be part of any computational or performance-based design approach. At the same time, usability testing, best practice guidelines, and skills are required in design firms to adopt the BPS practice. Usability testing can facilitate integrating the BPS use in the design context to address design problems to design a sustainable built environment. The evaluation of the total user experience of the BPS application during the project delivery process of performance-based designs is vital. It can be done with the help of USER-FIT. To this end, many of them should be used as an organizational-level intervention for testing the usability of BPS in design practice environments.

9. Conclusion

This study conducted a systematic review of recent (2013-2023) Ph.D. studies that performed usability testing during the development of BPS tools and methods for design decision support during the early design stages. Challenges hindering the widespread use of BPS tools among the IBPSA community were highlighted. Next, we proposed a novel usability testing framework called USER-FIT that modelers and designers can adopt in professional practice through simple, applicable evaluation criteria and usability tests within organizational policies and structures that optimize the potential of BPS uptake. The framework is based on six criteria, namely: (1) for describing decisional support needs; (2) for defining user needs; (3) for testing the usability of test cases and design tasks; (4) for testing the usability of the decision support system; (5) for providing informative visualization and (6) evaluation of decision support, once fully developed, for developing BPS Usability guidelines and standards. The study describes the foundation and theoretical background behind the development of the framework. How USER-FIT can be fitted to the project delivery process to improve the usability of BPS in the design process and workflows is presented. This paper states that usability is scientifically and objectively defined through a systematic and unified framework. The paper is a foundation to formalize user experience studies and usability testing for software tools, building performance models, simulation educators and building performance software developers. Directions for future BPS software development and user experience testing were recommended.

Acknowledgments

This research was funded by the Liege University pedagogical support grant for architectural-centered design decision support under the Digital Design Research project. We thank our colleagues, Michael Donn, Steffen Peterson, Ayu Miyamoto, Pil Purup, Andrew Corney and David Goldwasser, for many valuable discussions. We also would like to thank our former and current trainees, Kevin Tabsoba and many others, for conducting various analyses and illustrations cited in this paper.

Appendix I

The systematic review protocol was developed with input from systematic review methodologists found in the literature [128]. The eligibility criteria and information sources are described in Table 4.

Table 4: Summary of inclusion and exclusion criteria of the systematic review protocol

Purpose: Increase the uptake of building performance simulation for design decision support in practice.	
Research Questions: How to evaluate the ease of use of building performance simulation by building designers, modelers and developers?	
Keywords: user experience, usability, testing, trials, modeling, building, design,	Synonyms: building energy modeling, simulation, design process
Data Sources: Web of Science, Scopus, IBPSA conference proceedings, CAAD conference proceedings, cumniAD database, International Standardization Organisation (ISO)	https://clarivate.com/ http://www.scopus.com http://www.ibpsa.org/ http://papers.cumincad.org/ https://www.buildingenergysoftwaretools.com/
Search Strings: “experience” AND “modeling” AND (“user” OR “usability”) (“simulation” OR “evaluation”) AND “design” AND “buildings” AND (“architecture” OR “urban”) AND “systems” AND (“workflow” OR “process”) AND “performance” AND “energy” AND “environment.”	
Inclusion Criteria: Research field: AEC, HCI; Language: English; Publication date: 2010 to 2023; Type of work: scientific articles; Availability: full text; Subject: usability testing, user experience.	Exclusion Criteria: Not belonging to AEC Field; Not in English; Publication before 2010; Not a scientific work; Not related to the subject of interest; Full text unavailable.

Appendix II

Table 5, interview guide translated from French and Flemish based on the methodology used by Attia et al. (2019) [129]

Research Questions	Topic	Usability Concepts
Q1: How do you use BPS tools during the early design stages?	integration S1, S2, S3	design process, workflow
Q2: How do you improve your BPS skills?	user expertise S2	experience, building physics, building engineering
Q3: How do you evaluate the usability of a BPS tool in the design process?	user experience S3, S4, S5	learning curve, total user experience, user interface, speed, steps, simplicity, intuitiveness, performance issues
Q4: How can a BPS tool inform your design decision?	decision support S5, S6	performance requirements, value, quality, advice, assistance

References:

- [1] J. L. Hensen and R. Lamberts, *Building performance simulation for design and operation*. Routledge, 2012.
- [2] K. P. Lam, "Sustainability Performance Simulation Tools for Building Design," *Sustain. Built Environ.*, pp. 589–655, 2020.
- [3] AIA, "Architect's Guide to Building Performance: Integrating performance simulation in the design process," The American Institute of Architects, Washington DC, USA, 2019. [Online]. Available: https://content.aia.org/sites/default/files/2019-06/Energy_Design_Modeling_Guide_v4.pdf
- [4] ACE, "Assessing the Role of Architecture in Building Performance as Defined by the EU Level(s) Scheme," Architects' Council of Europe, Brussels, Belgium, 2019. Accessed: Aug. 07, 2022. [Online]. Available: <https://bit.ly/3Y68w2f>
- [5] ACE, "The architectural profession in Europe 2020," Mirza & Nacey Research Ltd, United Kingdom, 2021. Accessed: Jul. 08, 2022. [Online]. Available: https://www.ace-cae.eu/fileadmin/user_upload/2020ACESECTORSTUDY.pdf
- [6] S. Attia, J. L. Hensen, L. Beltrán, and A. De Herde, "Selection criteria for building performance simulation tools: contrasting architects' and engineers' needs," *J. Build. Perform. Simul.*, vol. 5, no. 3, Art. no. 3, 2012.
- [7] M.-M. Fernandez-Antolin, J.-M. del-Río, F. del Ama Gonzalo, and R.-A. Gonzalez-Lezcano, "The relationship between the use of building performance simulation tools by recent graduate architects and the deficiencies in architectural education," *Energies*, vol. 13, no. 5, Art. no. 5, 2020.
- [8] C. Hopfe, V. I. Soebarto, D. Crawley, and R. Rawal, "Understanding the differences of integrating building performance simulation in the architectural education system," 2017.
- [9] S. Alsaadani and C. B. De Souza, "Of collaboration or condemnation? Exploring the promise and pitfalls of architect-consultant collaborations for building performance simulation," *Energy Res. Soc. Sci.*, vol. 19, pp. 21–36, 2016.
- [10] S. Attia, A. De Herde, E. Gratia, and J. L. Hensen, "Achieving informed decision-making for net zero energy buildings design using building performance simulation tools," in *Building Simulation*, Springer, 2013, pp. 3–21.
- [11] S. Cozza, T. Jusselme, and M. Andersen, "Usability assessment of building performance simulation tools: A pilot study," in *Second International Conference for Sustainable Design of the Built Environment-SDBE 2018*, 2018.
- [12] R. Cetin and A. Mahdavi, *Exploring the availability and usability of web-based building performance simulation tools*. BauSIM 2010-Building Performance Simulation in a changing Environment, 2010.
- [13] J. Hensen, E. Djunaedy, M. Radošević, and A. Yahiaoui, "Building performance simulation for better design: some issues and solutions," in *Proceedings of the 21st Conference on Passive and Low Energy Architecture (PLEA)*, Citeseer, 2004, pp. 1185–1190.
- [14] M. M. Ouf, W. O'Brien, and H. B. Gunay, "Improving occupant-related features in building performance simulation tools," in *Building Simulation*, Springer, 2018, pp. 803–817.
- [15] L. Weytjens, S. Attia, G. Verbeeck, and A. De Herde, "The 'architect-friendliness' of six building performance simulation tools: A comparative study," *Int. J. Sustain. Build. Technol. Urban Dev.*, vol. 2, no. 3, Art. no. 3, 2011.
- [16] H. Son, S. Lee, and C. Kim, "What drives the adoption of building information modeling in design organizations? An empirical investigation of the antecedents affecting architects' behavioral intentions," *Autom. Constr.*, vol. 49, pp. 92–99, 2015.
- [17] RIBA, "RIBA Plan of Work." 2021. Accessed: May 18, 2023. [Online]. Available: <https://www.architecture.com/knowledge-and-resources/resources-landing-page/riba-plan-of-work>
- [18] S.-H. E. Lin and D. J. Gerber, "Designing-in performance: A framework for evolutionary energy performance feedback in early stage design," *Autom. Constr.*, vol. 38, pp. 59–73, 2014.
- [19] S. Oh and J. S. Haberl, "Origins of analysis methods used to design high-performance commercial buildings: Whole-building energy simulation," *Sci. Technol. Built Environ.*, vol. 22, no. 1, pp. 118–137, Jan. 2016, doi: 10.1080/23744731.2015.1063958.
- [20] S. Attia, J. L. Hensen, L. Beltrán, and A. De Herde, "Selection criteria for building performance simulation tools: contrasting architects' and engineers' needs," *J. Build. Perform. Simul.*, vol. 5, no. 3, Art. no. 3, 2012.

- [21] IBPSA USA, "Building Energy Software Tools Directory." Accessed: Dec. 02, 2020. [Online]. Available: <https://www.ibpsa.us/best-directory-list/>
- [22] D. Wang, X. Pang, W. Wang, Z. Qi, Y. Ji, and R. Yin, "Evaluation of the dynamic energy performance gap of green buildings: Case studies in China," in *Building Simulation*, Springer, 2020, pp. 1191–1204. Accessed: Feb. 13, 2024. [Online]. Available: https://idp.springer.com/authorize/casa?redirect_uri=https://link.springer.com/article/10.1007/s12273-020-0653-y&casa_token=QX9Fu3mdfRoAAAAA:mY-7CmRy-cHYISnQ1GG1Chqc6aq_rZahXwZBI5S9pgYbn1dCOwto3uSo4Q4vJnQAqRjwprDdV8xKVay-2A
- [23] T. Han, Q. Huang, A. Zhang, and Q. Zhang, "Simulation-based decision support tools in the early design stages of a green building—A review," *Sustainability*, vol. 10, no. 10, p. 3696, 2018.
- [24] V. J. Gan, I. M. Lo, J. Ma, K. T. Tse, J. C. Cheng, and C. M. Chan, "Simulation optimisation towards energy efficient green buildings: Current status and future trends," *J. Clean. Prod.*, vol. 254, p. 120012, 2020.
- [25] K. Negendahl, "Building performance simulation in the early design stage: An introduction to integrated dynamic models," *Autom. Constr.*, vol. 54, pp. 39–53, 2015.
- [26] G. Augenbroe, "10 The role of simulation in performance-based building," *Build. Perform. Simul. Des. Oper.*, 2019, Accessed: Feb. 13, 2024. [Online]. Available: https://books.google.be/books?hl=en&lr=&id=G4mUDwAAQBAJ&oi=fnd&pg=PT441&dq=the+use+of+BPS+for+code+compliance+is+facing+challenges+to+move+away+from+a+prescription+of+minimum+physical+parameters+to+an+increasingly+relevant+but+complex+world+of+performance-based+modeling&ots=gYtxUK_-q5&sig=YpRkZH_vp4CGzNJQuSAlejQ_Dmk
- [27] M. Donn, "Misuse of Building Performance Simulation," 2023. Accessed: Feb. 15, 2024. [Online]. Available: <https://www.buildingsandcities.org/insights/commentaries/misuse-building-performance-simulation.html>
- [28] K. Papamichael and V. Pal, "Barriers in developing and using simulation-based decision-support software," Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2002. Accessed: Feb. 13, 2024. [Online]. Available: <https://www.osti.gov/biblio/803774>
- [29] M. A. Abdelmegid, V. A. González, M. Poshdar, M. O'Sullivan, C. G. Walker, and F. Ying, "Barriers to adopting simulation modelling in construction industry," *Autom. Constr.*, vol. 111, p. 103046, 2020.
- [30] J. W. Hand, "Removing barriers to the use of simulation in the building design professions," PhD Thesis, University of Strathclyde UK, 1998. Accessed: Feb. 13, 2024. [Online]. Available: https://www.researchgate.net/profile/Jon-Hand/publication/2438187_Removing_Barriers_To_The_Use_Of_Simulation_In_The_Building_Design_Professions/links/0a85e5301c5d944695000000/Removing-Barriers-To-The-Use-Of-Simulation-In-The-Building-Design-Professions.pdf
- [31] R. Rezaee, J. Brown, G. Augenbroe, and J. Haymaker, "The application of inverse approach to the early stage of performance-based building design," in *14th Conference of International Building Performance Simulation Association, Hyderabad, India*, 2015. Accessed: Feb. 13, 2024. [Online]. Available: <https://scholar.google.com/scholar?cluster=17371927719978030828&hl=en&oi=scholar>
- [32] H. Verheij and G. Augenbroe, "Collaborative planning of AEC projects and partnerships," *Autom. Constr.*, vol. 15, no. 4, pp. 428–437, 2006.
- [33] K. J. Lomas, H. Eppel, C. J. Martin, and D. P. Bloomfield, "Empirical validation of building energy simulation programs," *Energy Build.*, vol. 26, no. 3, pp. 253–275, 1997.
- [34] D. Pati and G. Augenbroe, "Integrating formalized user experience within building design models," *Comput.-Aided Civ. Infrastruct. Eng.*, vol. 22, no. 2, pp. 117–132, 2007.
- [35] IBPSA-USA, "Building Energy Software Tools." Accessed: May 18, 2023. [Online]. Available: <https://www.buildingenergysoftwaretools.com/>
- [36] T. W. Maver, "CAAD's seven deadly sins," in *Sixth International Conference on Computer-Aided Architectural Design Futures*, CASA Singapore, 1995, pp. 21–22.
- [37] J. A. Clarke, "Why tools for buildings and cities performance simulation need to evolve," *Build. Cities*, 2020.
- [38] D. Newton, R. James, and D. Bartholomew, "Building energy simulation-a user's perspective," *Energy Build.*, vol. 10, no. 3, pp. 241–247, 1988.
- [39] C. Morbitzer, P. A. Strachan, B. Spires, D. Cafferty, and J. Webster, "Integration of building simulation into the design process of an architectural practice," 2001.

- [40] C. J. Hopfe and J. L. M. Hensen, "Experiences testing enhanced building performance simulation prototypes on potential user group," in *11th IBPSA Building Simulation Conference, 27-30 July*, International Building Performance Simulation Association (IBPSA), 2009, pp. 1976–1982. Accessed: Feb. 13, 2024. [Online]. Available: <https://research.tue.nl/en/publications/experiences-testing-enhanced-building-performance-simulation-prot>
- [41] A. Mahdavi, "The human dimension of building performance simulation," 2011, Accessed: Feb. 13, 2024. [Online]. Available: <https://www.aivc.org/resource/human-dimension-building-performance-simulation>
- [42] S. Petersen, "Simulation-based support for integrated design of new low-energy office buildings," DTU Civil Engineering, Technical University of Denmark, Copenhagen, Denmark, 2011. [Online]. Available: <https://orbit.dtu.dk/en/publications/simulation-based-support-for-integrated-design-of-new-low-energy->
- [43] S. Attia, *A tool for design decision making: zero energy residential buildings in hot humid climates*. Presses univ. de Louvain, 2012.
- [44] K. Negendahl, "Consequence Based Design. An approach for integrating computational collaborative models (Integrated Dynamic Models) in the building design phase," PhD, Technical University of Denmark, Department of Civil Engineering, Copenhagen, Denmark, 2016. [Online]. Available: <https://orbit.dtu.dk/en/publications/consequence-based-design-an-approach-for-integrating-computational>
- [45] E. MEEEX, "Early design support for material related environmental impact assessment of dwellings," PhD, Hasselt University, Hasselt, Belgium, 2018. [Online]. Available: <https://researchportal.be/en/publication/early-design-support-material-related-environmental-impact-assessment-dwellings>
- [46] E. Bazafkan, "Assessment of usability and usefulness of new building performance simulation tools in the architectural design process," PhD Thesis, Wien, 2017. Accessed: Feb. 13, 2024. [Online]. Available: https://www.researchgate.net/profile/Ehsan-Bazafkan/publication/320931455_Assessment_of_Usability_and_Usefulness_of_New_Building_Performance_Simulation_Tools_in_the_Architectural_Design_Process/links/5b58616daca272a2d6677f60/Assessment-of-Usability-and-Usefulness-of-New-Building-Performance-Simulation-Tools-in-the-Architectural-Design-Process.pdf
- [47] R. Jayapalan-Nair, E. Brembilla, C. Hopfe, and J. Mardaljevic, "Usability evaluation of a building simulation web-based design guide on potential user group", Accessed: Feb. 13, 2024. [Online]. Available: https://repository.lboro.ac.uk/articles/conference_contribution/Usability_evaluation_of_a_building_simulation_web-based_design_guide_on_potential_user_group/12621941
- [48] P. B. Purup, "A practice-oriented approach to the development of building performance simulation tools fit for the early stages of building design," 2021.
- [49] S. H. Hamilton, C. A. Pollino, D. S. Stratford, B. Fu, and A. J. Jakeman, "Fit-for-purpose environmental modeling: Targeting the intersection of usability, reliability and feasibility," *Environ. Model. Softw.*, vol. 148, p. 105278, 2022.
- [50] S. Oh and J. S. Haberl, "Origins of analysis methods used to design high-performance commercial buildings: Whole-building energy simulation," *Sci. Technol. Built Environ.*, vol. 22, no. 1, pp. 118–137, 2016.
- [51] E. Jaakkola, "Designing conceptual articles: four approaches," *AMS Rev.*, vol. 10, no. 1–2, pp. 18–26, Jun. 2020, doi: 10.1007/s13162-020-00161-0.
- [52] S. Attia and A. Maha, "USER-FIT: A framework to evaluate the usability of building performance simulation tools in the architectural design process, Report, ISBN: 9782930909240, Sustainable Building Design Lab, . . .," Liege University, Liege, Belgium, 2023. Accessed: Feb. 15, 2024. [Online]. Available: <https://orbi.uliege.be/handle/2268/310019>
- [53] C. Boyce and P. Neale, *Conducting in-depth interviews: A guide for designing and conducting in-depth interviews for evaluation input*, vol. 2. Pathfinder international Watertown, MA, 2006. Accessed: Feb. 14, 2024. [Online]. Available: https://www.academia.edu/download/33661461/m_e_tool_series_indepth_interviews.pdf
- [54] P. B. Purup and S. Petersen, "Requirement analysis for building performance simulation tools conformed to fit design practice," *Autom. Constr.*, vol. 116, p. 103226, 2020.
- [55] S. Attia, R. Lioure, and Q. Declaude, "Future trends and main concepts of adaptive facade systems," *Energy Sci. Eng.*, vol. 8, no. 9, pp. 3255–3272, Sep. 2020, doi: 10.1002/ese3.725.
- [56] Liege, Belgium. *In-depth Interviews*, (2021). Accessed: Feb. 15, 2024. [Online Video]. Available: <https://www.youtube.com/watch?v=z8lc1tUctDI>

- [57] L. M. Connelly, "Trustworthiness in qualitative research," *Medsurg Nurs.*, vol. 25, no. 6, p. 435, 2016.
- [58] S. Attia, E. Gratia, A. De Herde, and J. L. Hensen, "Simulation-based decision support tool for early stages of zero-energy building design," *Energy Build.*, vol. 49, pp. 2–15, 2012.
- [59] D. J. Gerber and S.-H. E. Lin, "Geometric Complexity and Energy Simulation—Evolving Performance Driven Architectural Form," 2013.
- [60] M. Turrin, P. Von Buelow, and R. Stouffs, "Design explorations of performance driven geometry in architectural design using parametric modeling and genetic algorithms," *Adv. Eng. Inform.*, vol. 25, no. 4, pp. 656–675, 2011.
- [61] R. Goldstein and A. Khan, "Simulation-based architectural design," in *Guide to simulation-based disciplines*, Springer, 2017, pp. 167–182.
- [62] J. Natanian, "Beyond Zero Energy Districts: A Holistic Energy and Environmental Quality Evaluation Workflow for Dense Urban Contexts in Hot Climates," PhD Thesis, Technische Universität München, 2021.
- [63] A. Roth *et al.*, "Scripting Frameworks For Enhancing EnergyPlus Modeling Productivity," in *Building Performance Analysis Conference and SimBuild, Chicago IL, September, 2018*, pp. 26–28.
- [64] H. Li *et al.*, "Issues in Bi-Directional Interoperability between BIM and BEM," in *Construction Research Congress 2022*, 2022, pp. 1355–1364.
- [65] H. W. Samuelson, A. Lantz, and C. F. Reinhart, "Non-technical barriers to energy model sharing and reuse," *Build. Environ.*, vol. 54, pp. 71–76, 2012.
- [66] T. Dogan and C. Reinhart, "Automated conversion of architectural massing models into thermal 'shoebox' models," *Proc. BS2013*, 2013.
- [67] K. Feng, W. Lu, and Y. Wang, "Assessing environmental performance in early building design stage: An integrated parametric design and machine learning method," *Sustain. Cities Soc.*, vol. 50, p. 101596, 2019.
- [68] C. J. Hopfe, "Uncertainty and sensitivity analysis in building performance simulation for decision support and design optimization," *PhD Diss Eindh. Univ.*, 2009.
- [69] J. Glazer, "Simulating Buildings While the Design is Still in the Architect's Imagination," GARD Analytics, Inc., Arlington Height, IL (United States), 2020.
- [70] P. B. Purup and S. Petersen, "Requirement analysis for building performance simulation tools conformed to fit design practice," *Autom. Constr.*, vol. 116, p. 103226, 2020.
- [71] S. Korkmaz, J. I. Messner, D. R. Riley, and C. Magent, "High-performance green building design process modeling and integrated use of visualization tools," *J. Archit. Eng.*, vol. 16, no. 1, pp. 37–45, 2010.
- [72] M. Amer, M. Hamdy, T. Wortmann, A. Mustafa, and S. Attia, "Methodology for design decision support of cost-optimal zero-energy lightweight construction," *Energy Build.*, vol. 223, p. 110170, 2020.
- [73] A. Miyamoto, K. Allacker, and F. De Troyer, "Visual tool for sustainable buildings: A design approach with various data visualisation techniques," *J. Build. Eng.*, vol. 56, p. 104741, 2022.
- [74] P. B. Purup and S. Petersen, "Characteristic traits of visualizations for decision-making in the early stages of building design," *J. Build. Perform. Simul.*, vol. 14, no. 4, pp. 403–419, 2021.
- [75] E. Nault, C. Waibel, J. Carmeliet, and M. Andersen, "Development and test application of the UrbanSOLve decision-support prototype for early-stage neighborhood design," *Build. Environ.*, vol. 137, pp. 58–72, 2018.
- [76] S. Alsaadani and C. B. De Souza, "Of collaboration or condemnation? Exploring the promise and pitfalls of architect-consultant collaborations for building performance simulation," *Energy Res. Soc. Sci.*, vol. 19, pp. 21–36, 2016.
- [77] P. Shiel, S. Tarantino, and M. Fischer, "Parametric analysis of design stage building energy performance simulation models," *Energy Build.*, vol. 172, pp. 78–93, 2018.
- [78] A. Nutkiewicz, B. Choi, and R. K. Jain, "Exploring the influence of urban context on building energy retrofit performance: A hybrid simulation and data-driven approach," *Adv. Appl. Energy*, vol. 3, p. 100038, 2021.
- [79] H. Li, Z. Wang, and T. Hong, "Occupant-Centric key performance indicators to inform building design and operations," *J. Build. Perform. Simul.*, vol. 14, no. 6, pp. 814–842, 2021.
- [80] I. Beausoleil-Morrison, "Learning the fundamentals of building performance simulation through an experiential teaching approach," *J. Build. Perform. Simul.*, vol. 12, no. 3, pp. 308–325, 2019.
- [81] P. P. Charles and C. R. Thomas, "Four approaches to teaching with building performance simulation tools in undergraduate architecture and engineering education," *J. Build. Perform. Simul.*, vol. 2, no. 2, pp. 95–114, 2009.

- [82] S. Alsaadani and C. De Souza, "The social component of building performance simulation: Understanding architects," in *Proceedings of the 2012 Building Simulation and Optimization Conference, Loughborough University, Loughborough, UK, September, 2012*, pp. 10–11.
- [83] S. Imam, D. A. Coley, and I. Walker, "The building performance gap: Are modellers literate?," *Build. Serv. Eng. Res. Technol.*, vol. 38, no. 3, pp. 351–375, 2017.
- [84] M.-M. Fernandez-Antolin, J.-M. del-Río, F. del Ama Gonzalo, and R.-A. Gonzalez-Lezcano, "The relationship between the use of building performance simulation tools by recent graduate architects and the deficiencies in architectural education," *Energies*, vol. 13, no. 5, p. 1134, 2020.
- [85] M.-M. Fernandez-Antolin, J. M. del Río, and R.-A. Gonzalez-Lezcano, "Building performance simulation tools as part of architectural design: Breaking the gap through software simulation," *Int. J. Technol. Des. Educ.*, vol. 32, no. 2, pp. 1227–1245, 2022.
- [86] K. Papamichael, "Designers and information overload: a new approach," *Adv. Build. Newsl.*, vol. 1, 1997.
- [87] J. Neale, M. H. Shamsi, E. Mangina, D. Finn, and J. O'Donnell, "Accurate identification of influential building parameters through an integration of global sensitivity and feature selection techniques," *Appl. Energy*, vol. 315, p. 118956, 2022.
- [88] J.-H. Choi, "Investigation of the correlation of building energy use intensity estimated by six building performance simulation tools," *Energy Build.*, vol. 147, pp. 14–26, 2017.
- [89] W. Albert and T. Tullis, *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes, 2013.
- [90] ISO, "ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs): Part 11: Guidance on usability." ISO, 2018.
- [91] ISO/IEC 25010, "System and software quality requirements and evaluation (SQuaRE)," International Organization for Standardization, Geneva, Switzerland, 2011. Accessed: Jan. 01, 2023. [Online]. Available: <https://www.iso.org/standard/52075.html>
- [92] ISO/IEC 25066, "Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Common Industry Format (CIF) for Usability - Evaluation Report," International Organization for Standardization, Geneva, Switzerland, 2016. Accessed: Jan. 01, 2023. [Online]. Available: <https://www.iso.org/standard/63831.html>
- [93] ISO 9241-210, "Ergonomics of Human-System Interaction – Part 210: Human-Centred Design for Interactive Systems," International Organization for Standardization, Geneva, Switzerland, 2010. Accessed: Jan. 01, 2023. [Online]. Available: <https://www.iso.org/standard/52075.html>
- [94] ISO 25060:2010, "Systems and software engineering," International Organization for Standardization, Geneva, Switzerland, 2010. Accessed: Jan. 01, 2023. [Online]. Available: <https://www.iso.org/standard/35786.html>
- [95] J. Nielsen, "How to conduct a heuristic evaluation," *Nielsen Norman Group*, vol. 1, no. 1, p. 8, 1995.
- [96] S. Rosenbaum, "Usability evaluations versus usability testing: When and why?," *IEEE Trans. Prof. Commun.*, vol. 32, no. 4, pp. 210–216, 1989.
- [97] *Usability Testing Tips and Examples, Google UX Design Certificate*, (Sep. 06, 2021). Accessed: Aug. 22, 2022. [Online Video]. Available: <https://www.youtube.com/watch?v=nYCJTea1AUQ>
- [98] J. S. Dumas and B. A. Loring, *Moderating usability tests: Principles and practices for interacting*. Elsevier, 2008.
- [99] L. M. Connelly, "Trustworthiness in qualitative research," *Medsurg Nurs.*, vol. 25, no. 6, Art. no. 6, 2016.
- [100] E. Tenner, "The design of everyday things by Donald Norman," *Technol. Cult.*, vol. 56, no. 3, pp. 785–787, 2015.
- [101] M. M. Byrne, "Evaluating the findings of qualitative research," *AORN J.*, vol. 73, no. 3, Art. no. 3, 2001.
- [102] Z. Afroz, H. B. Gunay, and W. O'Brien, "A review of data collection and analysis requirements for certified green buildings," *Energy Build.*, vol. 226, p. 110367, 2020.
- [103] D. Sinclair, *Guide to Using the RIBA Plan of Work 2013*. Routledge, 2019.
- [104] L. Faulkner, "Beyond the five-user assumption: Benefits of increased sample sizes in usability testing," *Behav. Res. Methods Instrum. Comput.*, vol. 35, pp. 379–383, 2003.
- [105] D. Süsser *et al.*, "Why energy models should integrate social and environmental factors: Assessing user needs, omission impacts, and real-word accuracy in the European Union," *Energy Res. Soc. Sci.*, vol. 92, p. 102775, 2022.

- [106] T. Østergård, R. L. Jensen, and S. E. Maagaard, "Building simulations supporting decision making in early design—A review," *Renew. Sustain. Energy Rev.*, vol. 61, pp. 187–201, 2016.
- [107] T. Østergård, "Proactive building simulations for early design support: Multi-actor decision-making based on Monte Carlo simulations and global sensitivity analysis," PhD, Aalborg, Copenhagen, Denmark, 2017. [Online]. Available: <https://vbn.aau.dk/en/publications/proactive-building-simulations-for-early-design-support-multi-act>
- [108] S. Imam, D. A. Coley, and I. Walker, "The building performance gap: Are modellers literate?," *Build. Serv. Eng. Res. Technol.*, vol. 38, no. 3, Art. no. 3, 2017.
- [109] S. Alsaadani and C. Bleil De Souza, "Performer, consumer or expert? A critical review of building performance simulation training paradigms for building design decision-making," *J. Build. Perform. Simul.*, vol. 12, no. 3, pp. 289–307, 2019.
- [110] S. Alsaadani and C. B. De Souza, "Of collaboration or condemnation? Exploring the promise and pitfalls of architect-consultant collaborations for building performance simulation," *Energy Res. Soc. Sci.*, vol. 19, pp. 21–36, 2016.
- [111] T. Häkkinen, M. Ala-Juusela, and J. Shemeikka, "Usability of energy performance assessment tools for different use purposes with the focus on refurbishment projects," *Energy Build.*, vol. 127, pp. 217–228, 2016.
- [112] R. Zhang and T. Hong, "Modeling of HVAC operational faults in building performance simulation," *Appl. Energy*, vol. 202, pp. 178–188, 2017.
- [113] B. Albert and T. Tullis, *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes, 2013.
- [114] I. Baxter and K. Oatley, "Measuring the learnability of spreadsheets in inexperienced users and those with previous spreadsheet experience," *Behav. Inf. Technol.*, vol. 10, no. 6, pp. 475–490, 1991.
- [115] M. J. Pilgrim, "The application of visualisation techniques to the process of building performance analysis," PhD Thesis, Loughborough University, 2003.
- [116] M. Donn, S. Selkowitz, and B. Bordass, "The building performance sketch," *Build. Res. Inf.*, vol. 40, no. 2, pp. 186–208, 2012.
- [117] Ö. Göçer, Y. Hua, and K. Göçer, "Completing the missing link in building design process: Enhancing post-occupancy evaluation method for effective feedback for building performance," *Build. Environ.*, vol. 89, pp. 14–27, 2015.
- [118] P. De Wilde, "The gap between predicted and measured energy performance of buildings: A framework for investigation," *Autom. Constr.*, vol. 41, pp. 40–49, 2014.
- [119] M. Nisztuk and P. B. Myszowski, "Usability of contemporary tools for the computational design of architectural objects: Review, features evaluation and reflection," *Int. J. Archit. Comput.*, vol. 16, no. 1, pp. 58–84, 2018.
- [120] J. L. Hensen, "Towards more effective use of building performance simulation in design," presented at the In Proc. 7th International Conference on Design & Decision Support Systems in Architecture and Urban Planning, Eindhoven, The Netherlands: TU-Eindhoven, 2004, pp. 2–5.
- [121] J. F. McLennan, *The philosophy of sustainable design: The future of architecture*. Ecotone publishing, 2004.
- [122] E. Bazafkan, U. Pont, and A. Mahdavi, "Usability and usefulness of non-conventional building performance simulation tools in architectural design processes," *Appl. Mech. Mater.*, vol. 887, pp. 219–226, 2019.
- [123] C. Bleil De Souza and S. Tucker, "Thermal simulation software outputs: a framework to produce meaningful information for design decision-making," *J. Build. Perform. Simul.*, vol. 8, no. 2, pp. 57–78, Mar. 2015, doi: 10.1080/19401493.2013.872191.
- [124] P. B. Purup and S. Petersen, "Requirement analysis for building performance simulation tools conformed to fit design practice," *Autom. Constr.*, vol. 116, p. 103226, 2020.
- [125] P. B. Purup and S. Petersen, "Characteristic traits of visualizations for decision-making in the early stages of building design," *J. Build. Perform. Simul.*, vol. 14, no. 4, Art. no. 4, 2021.
- [126] T. Hong, J. Langevin, and K. Sun, "Building simulation: Ten challenges," in *Building Simulation*, Springer, 2018, pp. 871–898.
- [127] S. Cozza, T. Jusselme, and M. Andersen, "Usability assessment of building performance simulation tools: A pilot study," in *Second International Conference for Sustainable Design of the Built Environment-SDBE 2018*, 2018.
- [128] Y. Xiao and M. Watson, "Guidance on conducting a systematic literature review," *J. Plan. Educ. Res.*, vol. 39, no. 1, pp. 93–112, 2019.

[129] Attia, S., Garat, S., & Cools, M. (2019). Development and validation of a survey for well-being and interaction assessment by occupants in office buildings with adaptive facades. *Building and Environment*, 157, 268-276.