Shaping the Future of Architectural Design Tools Through the HCI Paradigm and Collective Human-Machine Intelligence

Conference Paper · April 2025

DOI: 10.1145/3706599.3719720

Yang Chen Lin
National Tsing Hua University

Chien-Hui Su
National Tsing Hua University

Po-Chih Kuo
National Tsing Hua University

Shaping the Future of Architectural Design Tools Through the HCI Paradigm and Collective Human-Machine Intelligence

Yang Chen Lin
yangchenlin@nctu.edu.tw
Institute of Information Systems and Applications
National Tsing Hua University
Hsinchu, Taiwan

Wen-Yen Chung*
quex3201@gmail.com
Institute of Information Systems and Applications
National Tsing Hua University
Hsinchu, Taiwan

Chen-Ying Chien*
chenyingchien1021@gmail.com
Institute of Information Systems and Applications
National Tsing Hua University
Hsinchu, Taiwan

Chien-Hui Su*
fabienne1023@gapp.nthu.edu.tw
Department of Computer Science
National Tsing Hua University
Hsinchu, Taiwan

Po-Chih Kuo
kuopc@cs.nthu.edu.tw
Department of Computer Science
National Tsing Hua University
Hsinchu, Taiwan

ABSTRACT

This paper introduces the "From Design to Dwelling" framework, an HCI-driven approach to bridging architectural design intentions and lived experiences. Central to this framework is the concept of Human-Building Foundation Model (HBFM), which synthesizes multimodal empirical data, including visual, linguistic, and embodied dimensions—to inform AI-aided design tools optimized for human well-being. Our preliminary study analyzed patterns from human annotations, machine-generated visual features, and neural data during first-person navigation of virtual residential interiors across various design styles. Results demonstrate a double dissociation in neural processing: the parahippocampal place area (PPA) correlates with objective visual features, while the retrosplenial complex (RSC) aligns with subjective spatial experiences. While the HBFM is currently conceptual, this work provides an empirical foundation for its future implementation. This work advances both HCI and architectural design by providing empirically grounded methodologies for developing user-centered tools that bridge expert knowledge with quantifiable human experience within architectural spaces.

CCS CONCEPTS

• **Human-centered computing** → *HCI theory, concepts and models*; • **Computing methodologies** → *Cognitive science*; • **Applied computing** → *Computer-aided design*.

KEYWORDS

Human-Building Interaction, AI-aided design, Neuroarchitecture, Multimodal Data Integration, User-centered design

*These authors contributed equally to this research.

CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/2025/04
https://doi.org/10.1145/3706599.3719720

ACM Reference Format:
Yang Chen Lin, Wen-Yen Chung, Chen-Ying Chien, Chien-Hui Su, and Po-Chih Kuo. 2025. Shaping the Future of Architectural Design Tools Through the HCI Paradigm and Collective Human-Machine Intelligence. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3706599.3719720

1 INTRODUCTION

The architectural design process has traditionally relied on the expertise of architects who apply professional knowledge to create built environments. However, there often exists a gap between design intentions and actual user experiences. This fundamental challenge in architecture has led to the emergence of neuroarchitecture, an interdisciplinary field that explores the relationship between built environments and human cognition, behavior, and well-being. By integrating insights from cognitive sciences, psychology, neuroscience, and spatial computing, neuroarchitecture seeks to understand how design shapes human experiences and spatial cognition [17, 20, 37]. Despite its potential to inform evidence-based design practices by examining neural, behavioral, and psychological responses, neuroarchitecture remains challenged in gathering ecologically valid data, capturing the complex interplay between built environments and human behavior, and translating scientific insights into actionable design principles and processes [41].

In this paper, we argue that Human-Computer Interaction (HCI) offers a rich repository of methodological frameworks and theoretical concepts that can advance (neuro)architecture [4, 28, 29, 39] research while simultaneously expanding its own theoretical scope and creating new research opportunities. To fully address how HCI can deal with the challenges of neuroarchitecture, we must expand beyond traditional HCI to encompass the complex interplay between people, technology, space, and society (section 2.1). Interestingly, as we look backward, the evolution of architectural design tools from Computer-Aided Architectural Design (CAAD) to parametric modeling and Building Information Modeling (BIM) has introduced complex algorithmic relationships and collaborative environments that already touch on the types of problems HCI communities seek to address [10, 15, 26, 28, 29]. More recently, the advent of artificial intelligence (AI) systems offering generative capabilities has further influenced architectural practice [40, 42, 43]. Despite these technological advancements, architects often remain disconnected from the lived experiences of users, resulting in designs that fail to fully address real human needs and behaviors.

To address this gap, we introduce "***From Design to Dwelling***" - a research and design framework (section 2.2) that incorporates the concept of the **Human-Building Foundation Model (HBFM)** that captures and integrates both architect and user experiences throughout the architectural design process and everyday dwelling experiences to bridge architectural intentions and user experiences. Using **ecological and multi-modal**—visual, linguistic, and embodied data to inform the development of AI-aided architectural design tools optimized for "well-being" built enviroments. In the current study, we construct a **House Tour Video Dataset** (section 3) with human behavioral, psychophysiological, and subjective annotations, providing a rich resource to examine how individuals interact with and perceive interiors.

Our framework highlights the pivotal role of empirical data in enabling meaningful collaboration between architects and design tools, ultimately optimizing user well-being in architectural outcomes. Additionally, we explore potential challenges associated with data collection, well-being representation, and tool development, drawing on the rich methodologies and theoretical concepts of HCI. Through this approach, our research not only contributes to the advancement of neuroarchitecture but also extends the theoretical and practical boundaries of HCI.

2 SCOPE, FRAMEWORK, AND CHALLENGES

2.1 Fundamental Scope

The trajectory history of HCI is deeply rooted in understanding how humans interact with computational systems, evolving from a focus on individual usability to collective and societal impacts. We propose a paradigmatic shift toward examining the critical intersection of human-space-society relationships, particularly emphasizing the role of architectural experiences in shaping future living spaces and computational design tools. Architectural experiences are inherently multisensory and embodied, arising from the dynamic interplay between individuals, spaces, and technologies. This paper's theoretical scope integrates embodied cognition [19] and grounded cognition theories [6], positing that spatial experiences emerge through dynamic interactions between individuals and their environmental contexts. These experiences are shaped by overlapping sensorimotor patterns and environmental affordances during spatial navigation and interaction [16, 18, 32]. Building upon spatial language theory [30], our scope emphasizes how architectural experiences transcend traditional visual-spatial boundaries to incorporate linguistic and embodied modalities. This integration manifests in professional practice through sophisticated multimodal communication patterns—architects and stakeholders employ hybrid representations combining visual elements with linguistic descriptors to convey design intentions. For example, architects and clients often use a mix of images and language to convey the essence of a design, while potential users articulate spatial "feel" during activities such as house hunting. These multimodal interactions inform design processes that bridge expert knowledge and lived experience. These interactions create a bridge between expert knowledge systems and phenomenological experience, facilitating more nuanced and user-centered design outcomes.

This integrated theoretical scope provides a robust foundation for developing next-generation computational design tools that respond to both individual and collective spatial needs while acknowledging the complex, embodied nature of architectural experience. By synthesizing perspectives from HCI, cognitive science, and architecture, we establish a comprehensive framework for investigating and supporting human-space-society interactions.

2.2 Research and Design Framework: From Design to Dwelling

We present a theoretical study framework (Fig. 1) that centers on three interconnected components. **Human-Building Interaction** [3] connects users' experience-based interactions with built environments.**Human-AI Co-Design** enables architects to collaborate with AI-aided tools that generate design options through knowledge-based processes. These interactions are powered by **Empirical Data from Humans** - visual, linguistic, and embodied data at behavioral and neural levels - which feeds into the HBFM. By integrating experience and knowledge from experts and laypeople, this model facilitates any-to-any generation (like brain-to-video) and enables benchmarking between human and AI designs, creating a comprehensive system for evaluating and improving architectural solutions.

Building on this framework, three main challenges emerge, each offering opportunities for HCI contributions:

(1) **Collecting Ecological and Multimodal Data:** *How can ecological and multi-modal data be effectively collected?* HCI methodologies can facilitate the capture and integration of multimodal data "in the wild" contexts [12, 44], such as logging spatial and non-spatial language from dialogues and diagrams within the design process by experience sampling method or ubiquitous wearable devices combined with Wi-Fi and LiDAR sensors can track user movements and interactions (embodiment data) within physical/virtual architectural spaces [36].
(2) **Defining and Representing Well-Being:** *What exactly constitutes 'well-being' architectural experience, and is there a shared understanding of it among users, AI, and designers?* HCI researchers can utilize mixed methods, such as analyzing user interviews and quantitative data, to explore and reflect on diverse perspectives of well-being, facilitating its integration into design processes [8, 31]. Additionally, visualization research [38] can enable stakeholders to collaboratively interpret well-being data, transforming complex insights into actionable outcomes for more effective and inclusive design practices [23, 24].
(3) **Designing AI-Aided Tools:** *How can AI-aided tools effectively support architects in co-design processes, specifically in terms of human data interpretation and representation within design workflows?* Research-through-design methods, such as formative evalutaion [7] can play a critical role in guiding the development of AI-aided tools.

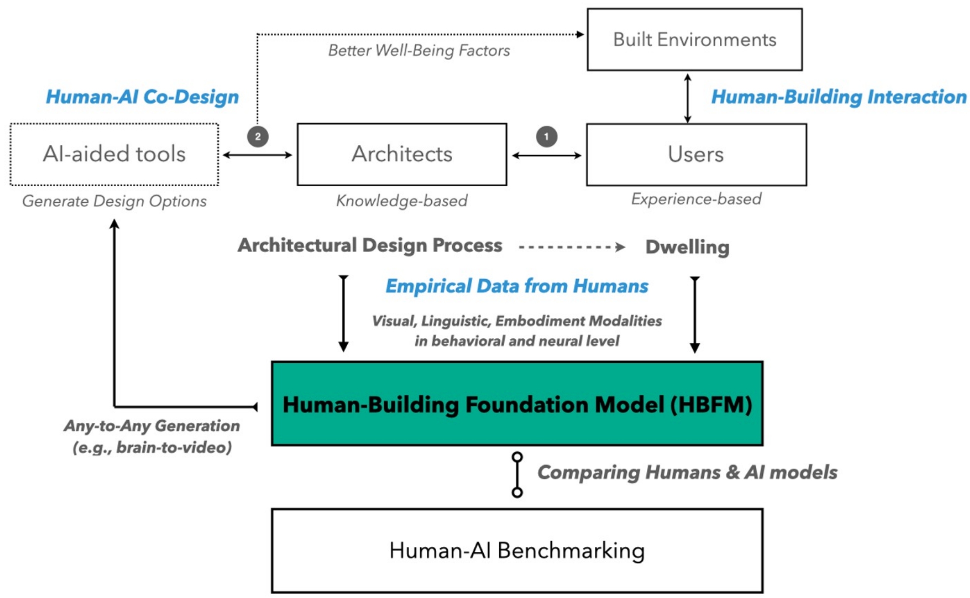


Figure 1: *Our Framework – From Design to Dwelling.* This diagram illustrates the collaborative and collective process involving architects, AI systems, users, and the built environment in shaping the architectural design process to create dwellings that enhance well-being. Architects integrate *knowledge-based* approaches while engaging with users, who provide *experience-based* insights through their interactions with built environments (*Human-Building Interaction*). These interactions, spanning *"from design to dwelling"*, generate *empirical data from humans,* encompassing *visual, linguistic, and embodiment modalities* at both behavioral and neural levels. Central to this framework is the *Human-Building Foundation Model (HBFM),* which is informed by empirical human data to optimize AI-aided design tools for human well-being. This framework also incorporates *Human-AI Benchmarking,* facilitating comparative analyses of human and AI-generated outputs to assess the alignment between human cognition and AI models. The HBFM contributes to the development of *AI-aided architectural design tools,* supporting architects in decision-making with functions such as *"any-to-any" generative capabilities* (e.g., brain-to-video). Label 1: Traditional Design Process – Architects design spaces based on their expertise or vague user feedback. Label 2: AI-Supported Design Process – Architects leverage AI-aided tools to assist in generating and refining diverse design options.

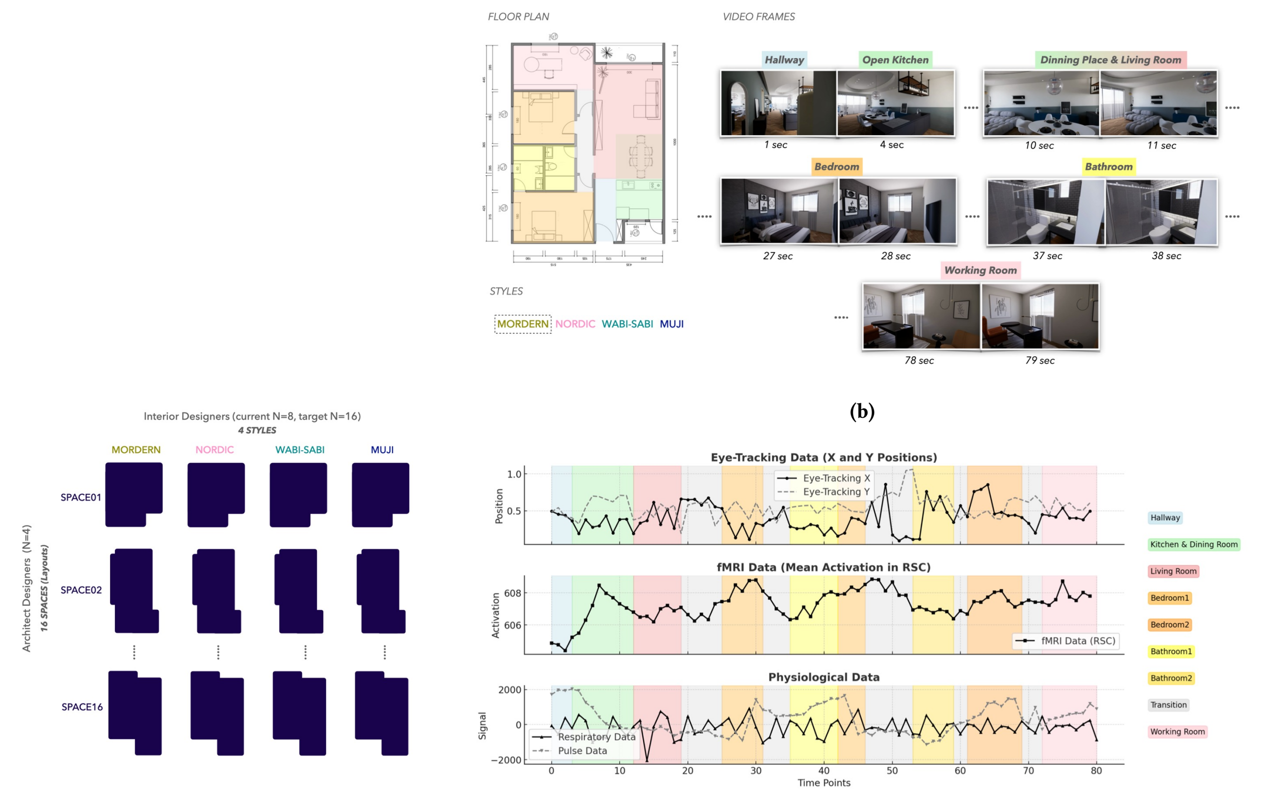
By addressing these challenges, HCI researchers can expand the field's theoretical and practical boundaries, fostering interdisciplinary innovations that redefine how humans and machines collaborate in architectural design.

3 CURRENT STUDY: BUILDING A 3D HOUSE TOUR VIDEO DATASET ALIGNED WITH HUMAN DATA

To operate our framework, we constructed a multimodal dataset capturing human responses to interior architectural spaces. First-person perspective (FPV) walkthrough videos of virtual rendering [2, 25] residential interiors were created in four distinct styles: Modern, Nordic, Wabi-Sabi, and MUJI and were accompanied by complementary human data, including subjective annotations (perception, feelings, description), eye tracking, and fMRI recordings. Two main experiments were conducted using these stimuli: (1) Subjective Annotation Experiment with Eye-Tracking current participants, N=14) and (2) fMRI Experiment (current participants, N=8), to investigate neural and behavioral responses to interior design. The experiments involved participants with normal or corrected-to-normal vision, representing an even gender distribution and a defined age range. Details on participant demographics and inclusion criteria are provided in the Supplementary Material.

3.1 Collecting Interior Design and House Tour Video Creation

3.1.1 Architectural space layouts & interior design styles. The design process involved four licensed architects, each designing four unique spaces, yielding a total of 16 distinct spatial layouts (Fig. 2a). Each space adhered to predefined architectural variables: spatial connectivity, window-to-wall ratio (10–15% or 25–30%), ceiling height (<3.8 m or >3.8 m), and geometric layout (rectangular or elongated townhouse). All spaces included essential residential areas—bedroom, bathroom, kitchen and dining area, living



(a) (c)

Figure 2: (a) The design space comprises 16 distinct spatial layouts (created by 4 architects, each designing 4 unique spaces) combined with 4 different interior styles (Modern, Nordic, Wabi-Sabi, and MUJI) per layout, yielding a total of 64 houses (16 layouts $\times$ 4 styles). (b) Each unique space is presented through an 80-second house tour video, exemplified by a modern-style residential interior following a prescribed circulation path. The sequential progression encompasses distinct functional zones: hallway, kitchen, dining area, living room, bedroom, bathroom, and workspace. (c) Multimodal Data Integration of a House Tour Video from a Participant. Top panel: Eye movement trajectories (X and Y coordinates) recorded during subjective annotation tasks. Middle panel: RSC activation time series from fMRI acquisition. Bottom panel: Concurrent physiological measurements (respiratory patterns and pulse dynamics). Color-coded background regions delineate distinct architectural zones: hallway (blue), kitchen/dining (green), living room (pink), bedrooms (orange), bathrooms (yellow), transitional spaces (gray), and working area (red).

room, study, and hallways with both landscape and utility balconies—within a floor area of 115.7–132.2 m², oriented north-south. These layouts catered to small families while allowing flexibility in lifestyle and marketability.

Interior design was conducted by eight designers (target: 16), each applying one of the four styles to four spaces at a time (Fig. 2a). Modern design emphasized clean lines, functional layouts, and neutral palettes with materials like glass and metal. Nordic design featured simplicity, natural materials, and soft, inviting tones. Wabi-Sabi embraced imperfection and natural textures, favoring earthy tones and an organic flow. MUJI design highlighted minimalism and practicality, with neutral colors and decluttered spaces. Designers provided detailed annotations for furnishings, textures, and materials to ensure aesthetic consistency across the styles. To ensure the basic quality of these designs, each design was reviewed by a professional interior designer.

3.1.2 First-person perspective 3D render house tour video. Using the 3D models, FPV walkthrough videos were rendered in Enscape 3D (version 4.2) following the specified spatial flow. Videos were rendered with default parameters, with the time set to 14:00 and the field of view to 100. These choices were made to ensure consistent lighting conditions and spatial representation across all videos. Based on recommendations from professional interior designers, the time was set to 14:00 to provide balanced daylight illumination, minimizing excessive shadows and strong directional lighting that could affect spatial perception. The 100-degree FOV was selected to offer a wide yet natural perspective, closely mimicking human vision while avoiding distortion. To enhance comparability and realism in the rendered environments, three researchers cross-validated the video paths frame-by-frame, while another three reviewers verified the final output. Each walkthrough video was standardized for experimental use. Videos were 80 seconds long, recorded at 30 frames per second, and prepared in two resolutions: 1920 × 1080 pixels for subjective annotation experiments on a 27-inch monitor and 800 × 600 pixels for fMRI experiments, optimized for MRI-compatible goggles. Consistency in spatial layouts and controlled lighting ensured the stimuli were suitable for investigating responses to architectural aesthetics across modalities. For this paper, we report findings from sixteen videos as a preliminary dataset (Fig. 2b and 2c).

3.2 Collecting Human Data: Experimental Design

3.2.1 Eye tracking, questionnaire evaluation and scene description. The subjective annotation experiment examined participants' visual attention patterns and subjective responses to different interior styles and spaces. To manage experimental duration and participant fatigue, each participant viewed a subset of 8 videos. These comprised two randomly selected spaces, each presented in all four design styles. For each video, eye movements were continuously recorded during viewing using Tobii Pro Spark eye tracker with 60 Hz sampling frequency. After viewing each video, participants provided verbal descriptions of their impressions and completed structured questionnaires about their perceptual knowledge of the space and perceptual evaluations.

The experiment was conducted using PsychoPy software and consisted of three main phases: setup and calibration, a practice session, and the experimental phase. During the setup and calibration phase, the experimental procedures and purposes were explained to each participant, followed by a four-point eye-tracking calibration and validation. If the eye-tracking deviation exceeded 1 degree of visual angle, the calibration process was repeated to ensure measurement accuracy. In the practice session, participants familiarized themselves with the experimental procedure by viewing a randomly selected video stimulus and completing all tasks, including video observation with eye-tracking, verbal description, and perceptual assessment questionnaires (see Supplementary Table 1. and Table 2). The experimental phase involved observing eight video stimuli, each representing two randomly selected spaces presented in four design styles. The presentation sequence was randomized for each participant. During each trial, participants underwent the same procedure as in the practice session: an 80-second video viewing period with concurrent eye-tracking, followed by verbal description and completion of perceptual assessment questionnaires. Questions about light conditions in the perceptual knowledge assessment were adapted from Chinazzo et al. [11], while questions on spatial qualities in the perceptual evaluation were based on Coburn et al. [13]. Accurate timestamps were recorded throughout to ensure precise synchronization of data across all measurements (Fig. 2c).

3.2.2 fMRI experiment. Participants completed a pre-scan questionnaire to provide demographic information, details about their home layout, and preferences for interior design. A brief practice session followed, during which they viewed a 10-second video clip unrelated to the main experimental stimuli. During the scan, participants wore MRI-compatible goggles to view the house tour video stimuli. Functional scans consisted of two runs, each containing eight video stimuli (16 total). Each run was divided into two blocks, with each block featuring four unique spatial layouts within the same interior design style. The blocks were presented in a randomized order across runs, ensuring that each run featured two distinct styles. Within each block, all spaces of the same style were shown consecutively. Each run began with a 10-second dummy scan for signal stabilization. Trials included an instructional screen ("You are about to enter a residential space"), a 2-second fixation cross, an 80-second video, and a 20-second interstimulus interval. Participants were instructed to remain as still as possible throughout the scans, with a break provided between runs to reduce fatigue. For detailed MRI data acquisition, see the Supplementary Material.

3.3 Data Analysis

3.3.1 Video features extraction. We utilized the pre-trained Temporally Sensitive Pretraining (TSP) model [5] to extract spatio-temporal features from our video stimuli. TSP employs a ResNet-based backbone to generate local-level features, which are then aggregated using temporal max-pooling to obtain a global video representation. These extracted features were subsequently used for analysis (see section 3.3.4) to investigate the underlying computational representations of our interior spaces.

3.3.2 Questionnaire evaluation. For the questionnaire data, we analyzed participants' perceptual evaluations in two parts. The first part assessed perceptual knowledge of lighting conditions, including lighting comfort and temperature perception. The second part evaluated perceptual attributes across three main categories: spatial attributes (complexity, organization, naturalness), aesthetic qualities (beauty, personalization, interest, modernity), and emotional responses (comfort, relaxation, vitality). All dimensions in the second part used 7-point scales. To control for individual rating tendencies, responses were normalized within each participant and question dimension, eliminating potential bias from participants' inherent rating patterns. Following normalization, we calculated mean ratings and standard deviations for each dimension to evaluate response consistency across participants.

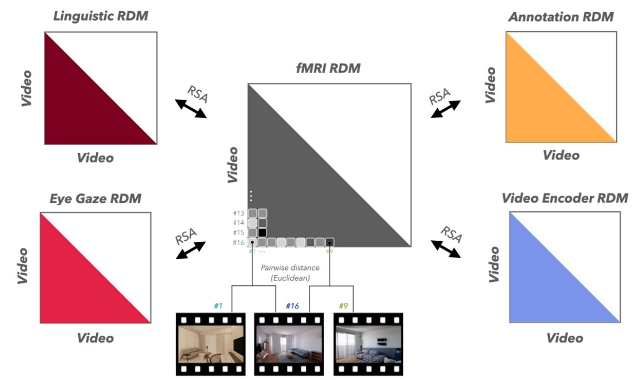
3.3.3 Eye gaze analysis. For the eye movement data, we adopted the predefined metrics [9] to characterize participants' viewing behavior. First, fixations, representing periods of stable gaze position, were identified using a threshold of 200 ms. From these fixations, we computed three primary measures: fixation count (total number of fixations per video), total fixation duration (in seconds), and mean fixation duration (in seconds). Second, we analyzed saccades, the rapid eye movements between fixations, to understand spatial exploration patterns through total and mean saccadic amplitude (in pixels) and saccade count. Third, we examined scanpaths, the complete sequence of fixations and saccades, from which we extracted scanpath length (total distance in pixels), convex hull area (spatial coverage in square pixels), and transition count. Finally, we defined dynamic Areas of Interest (AOIs) based on semantic segmentation of the video content to analyze attention distribution across different spatial regions.

3.3.4 fMRI analysis. The fMRI data were preprocessed using the fMRIPrep pipeline [22]. For detailed pre-processing methods, see the Supplementary Material. For first-level analysis, a General Linear Model (GLM) was implemented using Nilearn's [?] Least Squares All (LSA) approach [35], corporating condition-specific regressors for video stimuli, motion parameters, and polynomial drift terms. This comprehensive preprocessing pipeline ensured optimal signal-to-noise characteristics in the neuroimaging data, establishing a robust foundation for subsequent representational similarity analysis (RSA). We performed RSA using rsatoolbox [1] to explore relationships among fMRI responses, subjective annotations, and video features. fMRI time series data were extracted from four regions of interest (ROIs)—the parahippocampal place area (PPA), retrosplenial complex (RSC), occipital place area (OPA), and early visual cortex (EVC)—using ROI masks from Julian et al. [27]. Representational Dissimilarity Matrices (RDMs) were constructed for each data type (fMRI responses, subjective annotations, and video features) by calculating Euclidean distances between all video stimulus pairs (Fig. 3a). The resulting RDMs were of dimension (16, 16), reflecting the number of video stimuli. To test the statistical significance of correlations between RDMs, we employed a permutation test.

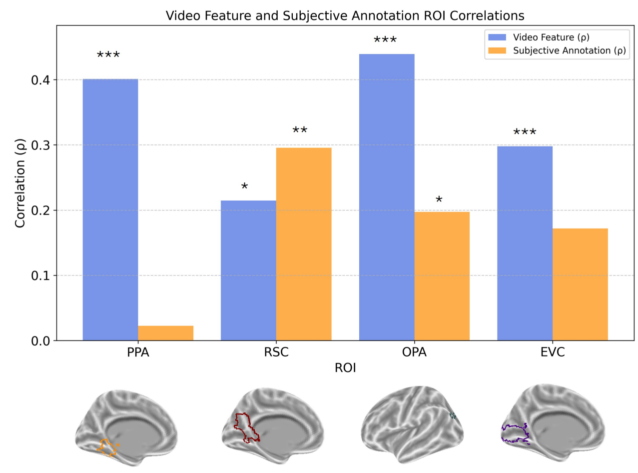
3.4 Preliminary Results and Implications

Our findings reveal complementary insights into how visual attention and neural responses relate to subjective and objective aspects of spatial experience. Eye-tracking data (Supplementary Fig. 1) highlight that reduced visual scanning behavior, such as fewer Fixation Counts and Saccade Counts, correlates with higher complexity ratings, while increased Transition Counts between AOI are associated with diminished relaxation. Broader visual exploration, reflected in Convex Hull Area, aligns with perceptions of warmer lighting. These metrics indicate measurable links between gaze patterns and subjective evaluations. Similarly, fMRI data (Fig. 3b) demonstrate that sensory regions such as the PPA, OPA, and EVC are more strongly correlated with objective, low-level visual features derived from a video encoder model (machine), suggesting a reliance on bottom-up processing. In contrast, the RSC shows stronger alignment with subjective annotations (human), emphasizing its role in integrating higher-level conceptual and contextual information [21, 33]. Together, these results suggest that subjective spatial evaluations arise from an interaction between bottom-up sensory-driven processing and top-down conceptual integration.

Our findings demonstrate the efficacy of the "From Design to Dwelling" framework in advancing interactive systems within HCI research paradigms. The framework's integration of multimodal human data facilitates the development of systems that simultaneously address lower-level sensory processes (mediated through PPA and OPA neural pathways) and higher-level conceptual processes (engaged through RSC mechanisms). This dual-processing approach establishes a theoretical foundation for designing interaction tools that harmonize perceptual engagement with semantic interpretation. The framework's practical applications manifest through two complementary channels: sensory-focused platforms leveraging HBFM's capabilities for high-fidelity environmental simulations (e.g., volumetric rendering, ambient lighting dynamics),



(a)



(b)

Figure 3: (a) Representational Similarity Analysis (RSA) Framework for Multimodal Architectural Experience Data. The current implementation focuses on two primary Representational Dissimilarity Matrices (RDMs): subjective experiential parameters (orange) and video-encoded architectural features (blue), both analyzed through their relationship with the fMRI RDM (gray). Each RDM represents pairwise Euclidean distances between stimulus conditions, enabling quantitative cross-modal comparisons. The framework's architecture anticipates future extensibility to incorporate linguistic response patterns (burgundy) and eye-tracking trajectories (red), establishing a comprehensive analytical pipeline for investigating architectural cognition across multiple representational domains. (b) Comparison of Video Encoder RDM and Annotation RDM with fMRI RDMs. Bar plots show Spearman correlations (ρ) between ROI activity patterns and video features (blue) or subjective annotations (orange) across four scene-selective regions. A double dissociation emerged between PPA and RSC: PPA showed strong correlation with video features ($\rho \approx 0.40$) but minimal with subjective annotations, while RSC displayed the opposite pattern ($\rho \approx 0.30$ for subjective annotations). Brain renders below show anatomical locations of each ROI. (* p<0.05, ** p<0.01, *** p<0.001)

and context-oriented systems utilizing empirical data integration for spatial configuration analysis and comparative architectural taxonomies. Although the HBFM remains a conceptual model, these findings validate its potential as a foundation for developing tools that harmonize sensory engagement with cognitive interpretation.

4 CONCLUDING REMARKS FOR FUTURE RESEARCH

This paper establishes foundational methodologies for developing advanced computational frameworks that integrate behavioral and neurophysiological data into architectural design processes. While HBFM currently exists as a conceptual construct, its potential for informing AI-augmented tools and user-centered design paradigms is implicated by preliminary empirical findings. The proposed implementation of AI-driven systems could facilitate evidence-based design optimization through predictive modeling of human responses to spatial configurations [34], representing a significant advancement in computational design methodologies. By demonstrating HCI's applicability to (neuro)architectural challenges, this work catalyzes research questions among HCI researchers, architects, and cognitive scientists, establishing future investigations at the intersection of human experience, computational design, and built environments.

ACKNOWLEDGMENTS

This research is funded by National Science and Technology Council, Taiwan (NSTC112-2628-E007-013-MY3). We sincerely thank Chunfan Sun, Jihfu Yu, Chia Yin Kuo, Anya Liu, and Yi Shiuan Chiang for their contributions to the architectural and interior design. We also extend our gratitude to our participants for their involvement in this work.

REFERENCES

[1] RSAtoolbox Development Group. [n. d.]. Representational Similarity Analysis 3.0. https://github.com/rsagroup/rsatoolbox
[2] Claudia Ziegler Acemyan and Philip Kortum. 2018. Does the type of presentation medium impact assessments of the built environment? An examination of environmental usability ratings across three modes of presentation. *Journal of Environmental Psychology* 56 (April 2018), 30–35. https://doi.org/10.1016/j.jenvp.2018.02.006
[3] Hamed S Alavi, Elizabeth F Churchill, Mikael Wiberg, Denis Lalanne, Peter Dalsgaard, Ava Fatah gen Schieck, and Yvonne Rogers. 2019. Introduction to human-building interaction (hbi) interfacing hci with architecture and urban design. , 10 pages.
[4] Hamed S. Alavi, Denis Lalanne, Julien Nembrini, Elizabeth Churchill, David Kirk, and Wendy Moncur. 2016. Future of Human-Building Interaction. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, San Jose California USA, 3408–3414. https://doi.org/10.1145/2851581.2856502
[5] Humam Alwassel, Silvio Giancola, and Bernard Ghanem. 2021. TSP: Temporally-Sensitive Pretraining of Video Encoders for Localization Tasks. In *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. IEEE, Montreal, BC, Canada, 3166–3176. https://doi.org/10.1109/ICCVW54120.2021.00356
[6] Lawrence W. Barsalou. 2008. Grounded Cognition. *Annual Review of Psychology* 59, 1 (Jan. 2008), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639
[7] Jesse Josua Benjamin, Joseph Lindley, Elizabeth Edwards, Elisa Rubegni, Tim Korjakow, David Grist, and Rhiannon Sharkey. 2024. Responding to Generative AI Technologies with Research-through-Design: The Ryelands AI Lab as an Exploratory Study. In *Designing Interactive Systems Conference*. ACM, IT University of Copenhagen Denmark, 1823–1841. https://doi.org/10.1145/3643834.3660677
[8] Marit Bentvelzen, Paweł W Woźniak, Pia SF Herbes, Evropi Stefanidi, and Jasmin Niess. 2022. Revisiting reflection in hci: Four design resources for technologies that support reflection. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–27.
[9] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. 2017. Visualization of Eye Tracking Data: A Taxonomy and Survey. *Computer Graphics Forum* 36, 8 (Dec. 2017), 260–284. https://doi.org/10.1111/cgf.13079
[10] Lok Hang Cheung, Juan Carlos Dall'Asta, Giancarlo Di Marco, and Asterios Agkathidis. 2024. Architectural Design Under Pandemic of Today's AI: From Human-Computer Interaction (HCI) to Human-Machine Conversation (HMC). In *Creativity in the Age of Digital Reproduction*, Giancarlo Di Marco, Davide Lombardi, and Mia Tedjosaputro (Eds.). Vol. 343. Springer Nature Singapore, Singapore, 9–17. https://doi.org/10.1007/978-981-97-0621-1_2 Series Title: Lecture Notes in Civil Engineering.
[11] Giorgia Chinazzo, Kynthia Chamilothori, Jan Wienold, and Marilyne Andersen. 2021. Temperature–color interaction: subjective indoor environmental perception and physiological responses in virtual reality. *Human factors* 63 (2021), 474–502. https://doi.org/10.1177/0018720819892383
[12] Patrick Chwalek, Sailin Zhong, Nathan Perry, Tianqi Liu, Clayton Miller, Hamed Seiied Alavi, Denis Lalanne, and Joseph A. Paradiso. 2024. A dataset exploring urban comfort through novel wearables and environmental surveys. *Scientific Data* 11, 1 (Dec. 2024), 1423. https://doi.org/10.1038/s41597-024-04279-9
[13] Alexander Coburn, Oshin Vartanian, Yoed N. Kenett, Marcos Nadal, Franziska Hartung, Gregor Hayn-Leichsenring, Gorka Navarrete, José L. González-Mora, and Anjan Chatterjee. 2020. Psychological and neural responses to architectural interiors. *Cortex* 126 (May 2020), 217–241. https://doi.org/10.1016/j.cortex.2020.01.009
[14]]Nilearn Nilearn contributors. [n. d.]. *nilearn*. https://doi.org/10.5281/zenodo.8397156
[15] Martyn Dade-Robertson. 2013. Architectural User Interfaces: Themes, Trends and Directions in the Evolution of Architectural Design and Human Computer Interaction. *International Journal of Architectural Computing* 11, 1 (March 2013), 1–19. https://doi.org/10.1260/1478-0771.11.1.1
[16] Zakaria Djebbara, Lars Brorson Fich, Laura Petrini, and Klaus Gramann. 2019. Sensorimotor brain dynamics reflect architectural affordances. *Proceedings of the National Academy of Sciences* 116, 29 (July 2019), 14769–14778. https://doi.org/10.1073/pnas.1900648116
[17] Zakaria Djebbara, Ole B. Jensen, Francisco J. Parada, and Klaus Gramann. 2022. Neuroscience and architecture: Modulating behavior through sensorimotor responses to the built environment. *Neuroscience & Biobehavioral Reviews* 138 (July 2022), 104715. https://doi.org/10.1016/j.neubiorev.2022.104715
[18] Zakaria Djebbara, Thomas Parr, and Karl Friston. 2020. Anticipation in architectural experience: a computational neurophenomenology for architecture? https://doi.org/10.48550/ARXIV.2011.03852 Version Number: 1.
[19] Paul Dourish. 2001. *Where the Action Is: The Foundations of Embodied Interaction*. The MIT Press. https://doi.org/10.7551/mitpress/7221.001.0001
[20] John P. Eberhard. 2009. *Brain Landscape: The Coexistence of Neuroscience and Architecture*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195331721.001.0001
[21] Russell A Epstein and Chris I Baker. 2019. Scene perception in the human brain. *Annual review of vision science* 5, 1 (2019), 373–397.
[22] Oscar Esteban, Christopher J. Markiewicz, Ross W. Blair, Craig A. Moodie, A. Ilkay Isik, Asier Erramuzpe, James D. Kent, Mathias Goncalves, Elizabeth DuPre, Madeleine Snyder, Hiroyuki Oya, Satrajit S. Ghosh, Jessey Wright, Joke Durnez, Russell A. Poldrack, and Krzysztof J. Gorgolewski. 2019. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods* 16, 1 (Jan. 2019), 111–116. https://doi.org/10.1038/s41592-018-0235-4
[23] Michal Gath-Morad, Raphaël Baur, and Christoph Hölscher. 2023. The Design-Mind toolkit. Graz, Austria, 51–60. https://doi.org/10.52842/conf.ecaade.2023.1.051
[24] Michal Gath-Morad, Tyler Thrash, Julia Schicker, Christoph Hölscher, Dirk Helbing, and Leonel Enrique Aguilar Melgar. 2021. Visibility matters during wayfinding in the vertical. *Scientific Reports* 11, 1 (Sept. 2021), 18980. https://doi.org/10.1038/s41598-021-98439-1
[25] Lara Gregorians, Pablo Fernández Velasco, Fiona Zisch, and Hugo J. Spiers. 2022. Architectural experience: Clarifying its central components and their relation to core affect with a set of first-person-view videos. *Journal of Environmental Psychology* 82 (Aug. 2022), 101841. https://doi.org/10.1016/j.jenvp.2022.101841
[26] Harri Haapasalo. 2000. *Creative computer aided architectural design: an internal approach to the design process*. Oulun Yliopisto, Oulu. OCLC: 890194490.
[27] J.B. Julian, Evelina Fedorenko, Jason Webster, and Nancy Kanwisher. 2012. An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage* 60, 4 (May 2012), 2357–2364. https://doi.org/10.1016/j.neuroimage.2012.02.055
[28] David Kirsh. 2019. Do Architects and Designers Think about Interactivity Differently? *ACM Transactions on Computer-Human Interaction* 26, 2 (April 2019), 1–43. https://doi.org/10.1145/3301425
[29] Jakub Krukar, Ruth Conroy Dalton, and Christoph Hölscher. 2016. Applying HCI Methods and Concepts to Architectural Design (Or Why Architects Could Use HCI Even If They Don't Know It). In *Architecture and Interaction*, Nicholas S. Dalton, Holger Schnädelbach, Mikael Wiberg, and Tasos Varoudis (Eds.). Springer International Publishing, Cham, 17–35. https://doi.org/10.1007/978-3-319-30028-

3_2 Series Title: Human–Computer Interaction Series.

[30] Barbara Landau and Ray Jackendoff. 1993. "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences* 16, 2 (June 1993), 217–238. https://doi.org/10.1017/S0140525X00029733

[31] Bokyung Lee, Michael Lee, Pan Zhang, Alexander Tessier, Daniel Saakes, and Azam Khan. 2021. Socio-Spatial Comfort: Using Vision-based Analysis to Inform User-Centred Human-Building Interactions. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (Jan. 2021), 1–33. https://doi.org/10.1145/3432937

[32] Isabella Pasqualini, Maria Laura Blefari, Tej Tadi, Andrea Serino, and Olaf Blanke. 2018. The Architectonic Experience of Body and Space in Augmented Interiors. *Frontiers in Psychology* 9 (April 2018), 375. https://doi.org/10.3389/fpsyg.2018.00375

[33] Michael Peer, Iva K Brunec, Nora S Newcombe, and Russell A Epstein. 2021. Structuring knowledge with cognitive maps and cognitive graphs. *Trends in cognitive sciences* 25, 1 (2021), 37–54.

[34] Kai-Florian Richter, Ben Weber, Brett Bojduj, and Sven Bertel. 2010. Supporting the designer's and the user's perspectives in computer-aided architectural design. *Advanced Engineering Informatics* 24, 2 (2010), 180–187.

[35] Jesse Rissman, Adam Gazzaley, and Mark D'Esposito. 2004. Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage* 23, 2 (Oct. 2004), 752–763. https://doi.org/10.1016/j.neuroimage.2004.06.035

[36] Andrew J. Sonta, Perry E. Simmons, and Rishee K. Jain. 2018. Understanding building occupant activities at scale: An integrated knowledge-based and data-driven approach. *Advanced Engineering Informatics* 37 (Aug. 2018), 1–13. https://doi.org/10.1016/j.aei.2018.04.009

[37] Esther M. Sternberg and Matthew A. Wilson. 2006. Neuroscience and Architecture: Seeking Common Ground. *Cell* 127, 2 (Oct. 2006), 239–242. https://doi.org/10.1016/j.cell.2006.10.012

[38] Melanie Tory and Torsten Moller. 2004. Human factors in visualization research. *IEEE transactions on visualization and computer graphics* 10, 1 (2004), 72–84.

[39] Himanshu Verma, Hamed S. Alavi, and Denis Lalanne. 2017. Studying Space Use: Bringing HCI Tools to Architectural Projects. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 3856–3866. https://doi.org/10.1145/3025453.3026055

[40] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y. Lim. 2019. Designing Theory-Driven User-Centric Explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–15. https://doi.org/10.1145/3290605.3300831

[41] Sheng Wang, Guilherme Sanches de Oliveira, Zakaria Djebbara, and Klaus Gramann. 2022. The embodiment of architectural experience: A methodological perspective on neuro-architecture. *Frontiers in human neuroscience* 16 (2022), 833528.

[42] Shun-Yu Wang, Wei-Chung Su, Serena Chen, Ching-Yi Tsai, Marta Misztal, Katherine M. Cheng, Alwena Lin, Yu Chen, and Mike Y. Chen. 2024. RoomDreaming: Generative-AI Approach to Facilitating Iterative, Preliminary Interior Design Exploration. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–20. https://doi.org/10.1145/3613904.3642901

[43] Yunkai Xu and TianTian Yu. 2022. Visual Performance of Psychological Factors in Interior Design Under the Background of Artificial Intelligence. *Frontiers in Psychology* 13 (July 2022), 941196. https://doi.org/10.3389/fpsyg.2022.941196

[44] Cansu Çetin Er and Oguzhan Özcan. 2024. Learning from Users: Everyday Playful Interactions to Support Architectural Spatial Changes. *Proceedings of the ACM on Human-Computer Interaction* 8, CHI PLAY (Oct. 2024), 1–25. https://doi.org/10.1145/3677085