sustainability

MDPI

Article

Visual-Aesthetic Convergence and Performative Gaps Between Sustainability and Biomimicry in AI-Generated Interiors

Turgut Kalay [1,*] and Ayşenur Kandemir [2]

[1] Department of Interior Architecture, Kütahya Dumlupınar University, Kütahya 43100, Turkey
[2] Department of Interior Architecture, İstanbul Nişantaşı University, Istanbul 34398, Turkey; aysenur.kandemir@nisantasi.edu.tr
* Correspondence: turgut.kalay@dpu.edu.tr

Abstract

Text-to-image generative artificial intelligence tools are increasingly used in interior architecture and design education, yet it remains unclear whether biomimicry and sustainability retain their distinct design intentions when visualized through such tools. This study comparatively examines biomimicry (B), sustainability (S), and combined (B + S) conditions in AI-generated interior images from 22 student projects in a Biomimicry and Space Design course, using descriptive content analysis. Visual-aesthetic similarity (VA1–VA5, Gower), biomimetic representation (B1–B4), sustainability strategies (S1–S7), and integration level (IL0–IL3) were coded as separate layers, with prompt (P-) and visual-output (V-) levels analyzed independently. The B and S outputs showed high visual-aesthetic similarity (mean Gower 0.82; 81.8% of projects scoring ≥0.80); however, biomimetic representation was strong at formal and structural levels but was verified in no project at the functional-mechanism level (V-B4). Sustainability prompt–output correspondence was 26.3% overall, 66.7% for visually assessable strategies, and 0% for strategies requiring specification or performance evidence. Combined outputs remained additive or coordinated (IL1–IL2), and none reached synergistic integration. In this course-based dataset, nature-oriented visual features were represented more readily than functional and performance-related dimensions; visual similarity is not equivalent to environmental performance, suggesting that prompt literacy and critical evaluation deserve attention in design education.

Keywords: sustainable interior design; design education; biomimicry; generative artificial intelligence; text-to-image generation; prompt–output correspondence; visual-aesthetic homogenization

Check for updates

Academic Editor: Nikos A. Salingaros

Received: 8 October 2026
Revised: 2 November 2026
Accepted: 4 November 2026
Published: 6 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Generative AI in Design and Design Education

Rapid advances in generative artificial intelligence technologies are transforming visual thinking and representation processes in architecture and design, particularly through models that generate visual content from textual inputs. By enabling designers to rapidly produce alternative visual proposals from textual descriptions, these models are reshaping ideation processes in the early stages of design [1]. With the growing use of tools such as DALL-E, ChatGPT, Midjourney, and Stable Diffusion, generative AI has increasingly been used to produce design proposals in the fields of interior architecture and industrial design [2]. Although current text-to-image models can interpret various design components, such as material selection, furniture layout, and lighting, through textual inputs, this capability also raises questions about how design intent is interpreted by the model and to what

extent textual design knowledge is transferred to the visual output [3]. Beyond its potential to support efficiency, creativity, and design quality in architectural design, AI should also be developed in ways that can contribute to more sustainable, equitable, and resilient design solutions and should be addressed through interdisciplinary collaboration [4].

The growing use of generative AI in design also requires existing pedagogical approaches in architecture and design education to be reconsidered. Digital and AI-oriented technologies are transforming how design is taught, conceptualized, and practiced, prompting renewed discussion of traditional methods in architectural education [5]. Research on the use of generative AI in design education shows that, alongside its potential to support creative processes, pedagogical and ethical issues, including the quality of production in creative disciplines, also need to be considered [6]. Similarly, although generative AI offers new opportunities for learning and production in educational settings, limitations such as the generation of inaccurate information need to be critically addressed [7]. Therefore, the use of generative AI in design education can be understood not merely as a technical process for producing visual outputs, but as a new form of design interaction that requires designers to communicate their intentions to the model and critically evaluate the resulting outputs. Indeed, prompts that express design intent clearly and in a structured manner have been shown to play an important role in reducing discrepancies between designers' expectations and generated images [8]. Within this framework, a prompt is more than a textual command that initiates visual generation; it serves as a guiding tool that determines the level of detail and priorities through which design knowledge is communicated to the AI system.

1.2. Biomimicry and Sustainability in Interior Design

Biomimicry is defined as an interdisciplinary approach based on the deliberate transfer of biological models, principles, and mechanisms into design processes to address contemporary technical and ecological challenges [9]. In this study, biomimicry refers to this design approach as a whole, whereas the adjective biomimetic describes design features, references, or forms of transfer derived from it, such as biomimetic representation or biomimetic transfer [9]. By enabling knowledge transfer between biology and the disciplines of design, engineering, and architecture, this approach aims to adapt principles and processes observed in nature to design problems [10]. Benyus, who has played an important role in the development of biomimicry thinking, argued that nature should be regarded not merely as a source of formal inspiration but as a model from which to learn in developing more sustainable solutions [11,12]. The literature shows that biological knowledge can be transferred into design at formal, structural, mechanical, and functional levels through biomimetic approaches [13]. This distinction clarifies the difference between superficial formal resemblance and the deeper, functional translation of biological principles into design. Biomimicry practices may range from reductionist to holistic approaches; however, the imitation of specific organism characteristics or biological processes does not necessarily guarantee environmental sustainability [14].

Sustainable interior design is a multidimensional design approach that addresses the reduction in environmental impacts together with the promotion of user health and comfort. Growing environmental responsibility in interior design has led to a reconsideration of design decisions, from material selection to spatial solutions, in order to create healthier and more environmentally responsive environments [15]. Because interiors are components of the built environment that interact directly with users, design decisions made in these spaces can generate environmental impacts throughout their life cycle, requiring interior architects to take these impacts into account [16]. Sustainable design strategies aim to improve the environmental performance of buildings and their components while

reducing energy consumption and adverse impacts on the natural environment. Given the direct and indirect relationships between interior spaces and the environment, making the environmental consequences of design decisions visible is particularly important in interior architecture education [17]. In sustainable architecture and interior design, the efficient use of energy, water, and material resources; the use of renewable energy sources; the promotion of natural daylight; and the provision of healthy indoor environmental conditions are among the main areas of application [18]. Within this context, material and resource use constitutes an important design concern; reducing material use, limiting the consumption of natural resources, considering recycling opportunities, and reducing waste with high environmental impacts are among the key strategies of sustainable design [19]. During the use and operation phases, energy-efficient lighting and plumbing systems, renewable energy applications, and technologies for monitoring resource consumption can contribute to improving environmental performance [20]. Sustainable interior design can therefore be understood as a holistic approach in which the relationships among materials, energy, technology, user requirements, and environmental conditions are considered together [21]. In interior architecture education, developing pedagogical structures that enable students to apply sustainability principles at different stages of the design process is important for transforming sustainable design knowledge into creative design decisions [22].

Biomimicry can contribute to the development of sustainable solutions by transferring principles of adaptability, efficiency, and resource use observed in natural systems into design. However, the transfer of a biological model or mechanism into design does not in itself guarantee environmental sustainability; sustainable outcomes depend on the level at which biological characteristics are adapted and the purpose they serve within the design process [23]. In this context, the sustainability potential of biomimetic design can be strengthened through more comprehensive applications that move beyond formal resemblance and address ecosystem- and process-level principles [24]. This relationship becomes particularly evident when biomimicry is considered together with circular design principles; the reuse of materials, reduction in waste, and development of systems compatible with natural cycles can strengthen the functional connection between biomimicry and sustainability [25]. In design education, combining biomimicry with environmentally responsible material approaches has also been shown to contribute to students' knowledge and design outcomes related to sustainable interior design [26]. Biomimicry and sustainability should therefore be regarded as closely related approaches that nevertheless require distinct design criteria and dimensions of evaluation.

1.3. Nature-Oriented Aesthetics and Generative AI

Sustainability in the built environment is not limited to measurable environmental performance; environmentally responsive design approaches can also develop a distinct aesthetic vocabulary, a phenomenon discussed within the framework of "sustainability aesthetics." This aesthetic framework includes visible features such as visual connection with the environment, natural materials, natural forms, color, the integration of ecosystems into design, and natural light, suggesting that certain aesthetic characteristics may overlap across biophilic design, sustainability aesthetics, and other nature-oriented approaches [27]. Because this shared vocabulary coincides to a large extent with that of biophilic design, which pursues a different aim from both biomimicry and sustainability, the distinction between biophilia and biomimicry is discussed in Section 4.1. Text-to-image generative AI systems have been reported to systematically emphasize certain aesthetic features and modes of representation over others, while insufficiently specified prompts may tend toward a limited set of default visual patterns [28]. Recent discussions in architecture similarly suggest that generative AI tools trained on comparable datasets may converge toward similar

visual vocabularies and may contribute to formal and aesthetic homogenization in the absence of critical design intervention [29]. However, different generative AI platforms have also been shown to exhibit distinct interpretive tendencies in terms of prompt adherence, creative expansion, and aesthetic interpretation; visual convergence should therefore not be regarded as an invariant or unidirectional outcome across all models [30]. Taken together, these studies raise an empirical research question: whether biomimicry and sustainability, despite their distinct conceptual and performative foundations, retain their differentiated design intentions when visualized through generative AI, or instead converge toward a shared visual-aesthetic repertoire.

1.4. Previous Research on Generative AI in Nature-Based and Sustainable Design

Recent studies on the use of generative AI in nature-based and sustainable design processes have followed several research directions. In biomimetic design, researchers have examined how generative AI interprets biological references at formal, functional, and structural levels through prompts structured at different levels of detail [31]. Similarly, designs generated using the same biomimicry-based prompts across different generative AI tools have been compared in terms of aesthetic value, functional potential, manufacturability, and sustainability, showing that the tool used can influence visual and design interpretation [32]. Studies on prompt structure have also systematically examined how prompt categorization, level of detail, and mode of expression affect the resulting images [2]. In sustainable design, integrating real-world material and carbon data into generative AI processes has been shown to guide users' prompts and design decisions toward more environmentally responsible objectives, while research in interior architecture has highlighted the potential of AI to support material selection, energy use, and sustainable design decisions [33,34]. Studies in design education have also examined generative-AI-assisted studio processes that incorporate biomimetic or nature-inspired design approaches, focusing on their role in students' conceptual thinking, creative production, and design development [35,36].

These strands of research do not point in the same direction. Studies emphasizing the potential of generative AI report that it can support material, energy, and other sustainable design decisions [33,34] and explore its contribution to conceptual thinking and creative production in biomimetic and nature-inspired studios [35,36], and reviews of AI-driven biomimetic optimization present AI as a means of translating biological patterns into resource-efficient architectural solutions [37]. Studies that examine the generated images themselves report more limited results: text-to-image models reproduce the geometric analogies of biological forms without conveying the underlying form–function relationship [31], favor recurring aesthetic features and default visual patterns [28], and may contribute to formal and aesthetic homogenization [29], while the outputs obtained from comparable prompts differ across tools [30,32]. The literature therefore offers no consistent answer to whether generative AI supports or undermines nature-based and sustainable design intentions.

We interpret this inconsistency as largely a consequence of how the two concepts and the outputs are evaluated. First, both terms are used to denote either a visible design language or a functional and performance-related principle: biomimicry ranges from formal resemblance to the transfer of biological mechanisms [13,14], and sustainability from an aesthetic vocabulary to measurable environmental performance [27]. Second, studies differ in the level at which they evaluate generative AI; positive assessments typically concern ideation, creativity, or design potential, whereas critical assessments concern what the generated images actually depict. Third, the term generative AI covers both performance-driven computational design, in which environmental criteria can enter the generation process, and text-to-image models, in which they enter only through the wording of the

prompt [33,37]. When these levels are conflated, the visual plausibility of nature-oriented images can be read as evidence of conceptual or environmental fidelity. Our working interpretation is therefore that text-to-image models readily reproduce the visual cues shared by biomimicry and sustainability but do not differentiate their functional and performance-related intentions, which would explain why the literature reports both potential and homogenization. Examining this interpretation requires separating what is requested in the prompt, what is visually represented, and what can be verified from an image, and comparing biomimicry and sustainability within the same dataset under separate and combined conditions; to our knowledge, such a comparison has not been reported.

1.5. Aim, Research Questions, and Contributions

In this context, the present study comparatively examines AI-generated interior images produced under biomimicry (B), sustainability (S), and combined biomimicry + sustainability (B + S) conditions in 22 student projects developed in the Biomimicry and Space Design course. The study aims to determine the extent to which biomimicry and sustainability, despite their distinct conceptual foundations, exhibit similar visual-aesthetic profiles; the forms of representation through which biomimetic references appear in visual outputs; the extent to which sustainability strategies explicitly specified in prompts can be visually verified in the generated images; and the levels of integration achieved when the two approaches are requested together. AI-generated images are not treated as direct evidence of environmental performance; rather, the analysis is limited to visually verifiable features and the correspondence between design intentions expressed in the prompts and their representation in the outputs. Accordingly, the study addresses the following research questions:

RQ1. To what extent do biomimicry (B) and sustainability (S) outputs exhibit similarity in terms of shared visual-aesthetic features?

RQ2. Which forms of biomimetic representation are most prominent in biomimicry-focused outputs?

RQ3. To what extent can the design strategies explicitly specified in sustainability prompts be visually verified in the generated outputs?

RQ4. What levels of integration are observed in outputs in which biomimicry and sustainability are requested together (B + S)?

This study examines biomimicry and sustainability within the same design education dataset under both separate and combined conditions. Rather than offering a further assessment of the potential of generative AI for either approach, the study addresses the inconsistency outlined above by examining whether the two approaches remain distinguishable at the level of visual representation, while their functional and performance-related intentions are assessed separately. The study makes three contributions. Empirically, it compares, within a single design-education dataset, how biomimicry and sustainability are represented when they are requested separately and together. Methodologically, it separates the prompt (P-) and visual-output (V-) layers, thereby distinguishing between a strategy being requested in the prompt and its verifiable representation in the image, and it uses a transparent coding framework that keeps visual-aesthetic similarity conceptually distinct from biomimetic and sustainability performance. For design education, it clarifies which aspects of biomimetic and sustainable design intentions AI-generated images convey and which require other forms of evidence. In doing so, the study seeks to contribute empirical evidence to current discussions on the risk of formal and aesthetic homogenization in generative AI outputs [29,30] and to identify issues that deserve attention in sustainable interior design education.

2. Materials and Methods

2.1. Research Design and Data Set

This study uses a descriptive research design based on content analysis to comparatively examine how biomimicry and sustainability are represented when considered separately and together in a generative-AI-assisted visual production process. The dataset comprises 22 student projects developed within the Biomimicry and Space Design course. For each of the 22 projects, three complete output groups were compared: B outputs prioritizing biomimicry, S outputs prioritizing sustainability, and B + S outputs in which both approaches were requested together. Here, biomimicry denotes the transfer of biological forms, principles, or mechanisms into design, and sustainability denotes design strategies that reduce environmental impacts and support user health and comfort (Section 1.2). The visual outputs were generated using different generative AI tools, as documented in Section S2 of the Supplementary Materials. Because the study did not employ a controlled single-model design, platform identity was not treated as an independent analytical variable.

The units of analysis are the project-level prompts and the final visual outputs generated from these prompts. All 22 projects were included in the visual comparisons. Prompts and final images were available for all three conditions in all 22 projects, as documented in the students' final submissions; no prompt was reconstructed, and no output was substituted. Three non-perspective outputs, a material detail collage and a material palette board in P03 and an isometric plan render in P17, were not coded, in accordance with the unit-of-analysis rule of the codebook (Supplementary Materials, Section S1). A project-by-condition inventory is provided in Table S6 of the Supplementary Materials. Although P11 and P20 included general guidance related to sustainability, none of the strategies represented by codes S1–S7 were explicitly requested. Therefore, project-level percentages for these projects were reported as N/A rather than assigned a value of zero.

Visual-aesthetic similarity, biomimetic representation, the visibility of sustainability strategies, and the level of integration in the combined condition were treated as separate analytical dimensions. This distinction was made specifically to prevent a space that appears natural or green from automatically being interpreted as evidence of performance-based sustainability.

2.2. Course Context and Image-Generation Procedure

The dataset derives from the Biomimicry and Space Design course, a fourth-year elective offered in the eighth semester of the undergraduate interior architecture program; the outputs analyzed were produced in the spring semester of the 2025–2026 academic year. The study was designed retrospectively, after the course had been completed and the assignments had been submitted; the course itself was conducted as a regular educational activity rather than as a research intervention. The course began with a theoretical phase in which students were introduced to the concepts of biomimicry and sustainability and researched their applications in interior design. Floor plans were then identified through group research, and each student developed an individual project. In most projects, the selected floor plan served as the spatial basis for the B, S, and B + S conditions; in P11 and P16, condition-specific floor plans were used, and no floor plan was documented for P04, P12, P15, and P18. The projects collectively addressed different functional programs, such as libraries, cafés, and offices.

Image generation followed a fixed sequence. Each student first wrote biomimicry prompts and generated the B outputs, and then wrote sustainability prompts for the same project and generated the S outputs. The choice of generative AI tool and the number of generation attempts were left to the students; the tools used in each project are documented in Section S2 of the Supplementary Materials. The projects generated with ChatGPT used GPT-5 (OpenAI, San Francisco, CA, USA), those generated with Gemini used Gemini 3 Pro (Google, Mountain View, CA, USA), and P15 used Nano Banana Pro v7.1.7 (Google, Mountain View, CA, USA);

generation settings were not documented in the students' submissions, and the images were not post-processed (for example, retouched or upscaled) before being placed on the presentation boards. During the in-class review of the B and S outputs, students observed that the two sets appeared visually similar. Before the combined stage, students were therefore guided to specify sustainability more explicitly through decisions on materials, lighting, ventilation, furniture, and surface finishes, after which they generated the B + S outputs. The S outputs analyzed in this study are the initial sustainability outputs generated before this guidance; the guidance thus affected only the B + S condition.

The analysis was based on the students' final submissions. For each condition, the images analyzed were those that the students themselves selected and included in their final submissions, together with the prompts submitted with them. The authors did not select among alternative generations, and intermediate generations were not part of the dataset. Student selection may nevertheless have favored images that the students considered visually successful, which is inherent to a course-based dataset. Visual inputs were not standardized across projects. In P05, the B and S outputs were generated by transforming the same existing interior image while preserving its layout and perspective, and in P20, the S outputs were derived from the B outputs (Supplementary Materials, Table S1). The combined condition was the least standardized: in P16, P19, and P20, the B + S outputs were produced by merging or combining the previously generated B and S outputs; in P05 and P07, they were generated from the same existing interior image or from an uploaded interior image; and in P18, the B + S prompts were identical to the B prompts (Supplementary Materials, Table S1). These projects were retained because they reflect the procedures that students actually followed in the course.

Because the B and S outputs of each project were produced by the same student, within the same project, and with freely chosen tools, the B-S comparison is a within-project comparison of student-generated outputs under naturalistic course conditions rather than a controlled experiment. The similarity between the two conditions may therefore arise from the generative models, the content of the students' related prompts, the shared project concept and floor plan, shared or sequential visual inputs, or the students' selection of final images, and the present design does not allow these sources to be separated. At the same time, since the S outputs were generated before the in-class guidance described above, the B-S similarity cannot be attributed to that guidance. Accordingly, the term convergence is used in this study to describe the similarity observed within this particular set of images rather than as a general property of generative AI outputs.

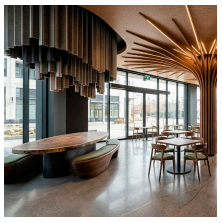
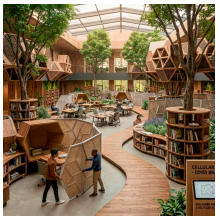
The use of the completed course outputs for research was reviewed by the Science and Engineering Sciences Scientific Research and Publication Ethics Board of Kütahya Dumlupınar University, which issued an exemption for this retrospective analysis (decision date: 28 September 2026; decision no.: 2026/09). Written informed consent was obtained individually from each student whose work is included, after the final course grades had been announced; students were informed that participation was voluntary and that declining would have no effect on their grades. Because grading had been completed before the students were invited to participate, their decisions could not influence their course assessment. The consent form covered the publication of the students' prompts and AI-generated images in open access format, and the projects are identified only by anonymized codes (P01–P22).

2.3. Visual Materials

To provide readers with a direct overview of the visual scope of the dataset, one representative Biomimicry (B) output from each project is presented in Table 1. These images do not substitute for the analytical coding; they were selected only to illustrate the general character of the material. Aesthetic superiority was not used as a selection criterion.

The images were not selected or evaluated according to biophilic principles; their visual features were coded only with the indicators described in Section 2.4. Instead, priority was given to a clearly readable interior perspective that represented the spatial design language of each project under the B condition. The complete sets of visual outputs for the B, S, and B + S conditions are provided in Section S2 of the Supplementary Materials.

Table 1. Representative biomimicry (B) visual outputs from projects P01–P22.

Project Code	P01	P02	P03	P04
Image				
Project Code	P05	P06	P07	P08
Image				
Project Code	P09	P10	P11	P12
Image				
Project Code	P13	P14	P15	P16
Image				
Project Code	P17	P18	P19	P20
Image				
Project Code	P21	P22		
Image				

2.4. Coding Framework

Coding was conducted across four complementary layers: visual-aesthetic feature codes (VA1–VA5), biomimetic representation codes (B1–B4), sustainability strategy codes (S1–S7), and integration levels in the combined outputs (IL0–IL3). The VA, B, and S codes were recorded as binary variables. 1 = present and verifiable in the relevant data layer, and 0 = not verified. For the prompt codes, a value of 0 means that a feature was not explicitly specified; for the visual codes, it means that a feature was not visually verified, which may reflect either its observable absence or the impossibility of determining it from an image (Section 2.5). Prompts and visual outputs were treated as separate data layers; prompt-based codes were prefixed with P-, while visual-output-based codes were prefixed with V-. This distinction separated a conceptual strategy being specified in the prompt from its verifiable representation in the visual output. All codes were applied consistently across all 22 cases using definitions and decision rules established before the final coding cycle.

The first coding layer comprised the VA1–VA5 indicators, which were developed to assess the extent to which the Biomimicry (B) and Sustainability (S) conditions exhibited similarity in their visual-aesthetic language (Table 2). These indicators were not intended to measure sustainability performance or the depth of biomimetic representation; rather, they were used to compare visible aesthetic features recurring across the two conditions. Each of these features also corresponds to an attribute of biophilic design, namely natural materials, natural colors, plants, natural light and views, and naturalistic shapes and forms. The indicators were derived from the sustainability-aesthetics framework [27] and were not designed as a biophilic assessment, but this correspondence is taken into account in the interpretation of the results (Section 4.1). Accordingly, VA5 was kept distinct from the sustainability strategy code S2. For VA5, it was sufficient for a space to appear bright due to large windows, a visual connection with the outdoor environment, or a clearly visible daylight effect. In contrast, V-S2 required daylight to be represented as a design strategy through an identifiable feature such as a skylight, atrium, solar-shading device, or a comparable element.

Table 2. Visual-aesthetic codes (VA1–VA5).

Code	Criterion	Definition
VA1	Organic/Curved form	Presence of pronounced curved, flowing, or organic geometries in the space or furniture
VA2	Natural-looking material	Visual predominance of surfaces and textures evoking natural materials such as wood, stone, earth, bamboo, or rattan
VA3	Nature-inspired color palette	Predominance of a natural or earthy palette, including earth tones, wood tones, beige, neutrals, and greens
VA4	Planting/Biophilic component	Visibility of plants, green walls, integrated landscaping, or similar biophilic elements in interior or exterior spaces
VA5	Daylight/Outdoor connection	Visually identifiable large glazed surfaces, visual connections with the outdoor environment, or a pronounced natural lighting effect

Biomimetic representation was assessed using four complementary indicators, ranging from the visible formal expression of a biological reference to its translation into a spatial, structural, or functional design principle (Table 3). Codes B1–B4 were treated as representational features that could coexist within the same project rather than as mutually exclusive levels. Prompts and visual outputs were coded separately as P-B1–P-B4 and V-B1–V-B4,

respectively. However, because the final prompt–output correspondence analysis focused on the sustainability dimension, no separate percentage-based correspondence rate was calculated for biomimicry. Instead, the frequencies of the P-B and V-B codes are reported together with the number of projects in which a requested feature was visually verified, and the requests for a biological function (P-B4) are examined project by project (Section 3.2; Supplementary Materials, Table S5).

Table 3. Biomimetic representation codes (B1–B4).

Code	Criterion	Definition
B1	Identifiable biological reference	A specific biological reference, such as a tree, shell, honeycomb, vein network, or mycelium, is explicitly named in the prompt (P-B1) or identifiable from the image alone (V-B1)
B2	Morphological/Pattern transfer	Transfer of biological form, geometry, texture, or pattern into the design language
B3	Spatial/Structural transfer	Translation of a biological reference into a larger-scale structural or spatial configuration, such as columns, ceilings, circulation, shells, partitions, or spatial organization
B4	Visually identifiable functional/environmental mechanism	A biological function related to ventilation, shading, water management, thermal behavior, or similar processes is identifiable as an operational mechanism in the image, rather than merely described in the text

In applying these codes, the biological reference had to be identifiable from the image itself for V-B1; a reference named only in the prompt justified P-B1 = 1 but not V-B1 = 1. Because both layers of a project were coded by the same coder, the prompts were known when the visual codes were assigned, so the independence of the V- layer rests on this evidence rule rather than on blinding. Knowledge of the prompts is more likely to have encouraged than discouraged the assignment of 1, so this limitation would tend to raise rather than lower the visual frequencies; the low rates of visual verification reported for V-B4 and for the sustainability strategies are therefore unlikely to have been underestimated for this reason. B3 and B4 were distinguished by what had to be visible. B3 concerns configuration: the biological reference organizes a primary structural or spatial element, such as supports, the ceiling or roof structure, partitions, or circulation, rather than only surfaces or furnishings, and no function needs to be depicted. B4 additionally requires the image to show how a feature derived from a biological function produces an environmental effect, such as ventilation, shading, water collection, or thermal regulation, in a way that cannot be attributed merely to the presence of the element; a canopy that casts shade or an opening that admits air does not by itself meet this condition, and neither does kinetic or responsive behavior that cannot be observed in a still image. For example, tree-like branching columns supporting the ceiling would be coded as B3 but not B4, and a ventilation function stated in the prompt but not legible in the image would be coded as P-B4 = 1 and V-B4 = 0. The complete decision rules and further examples are given in the codebook (Supplementary Materials, Section S1).

Sustainability strategies were examined using codes S1–S7 to distinguish between strategies explicitly specified in the prompts and their verifiable representation in the visual outputs (Table 4). The same code set was applied to prompts as P-S1–P-S7 and to visual outputs as V-S1–V-S7. The specification of a strategy in a prompt was not considered sufficient to assign a value of 1 to the corresponding visual code. For example, features such as FSC certification, recycled content, or low-carbon materials can be specified textually but cannot be verified from a render alone. Therefore, they were not coded as present

under the corresponding V-S code unless identifiable visual evidence was available in the output. This distinction enabled the textual articulation of sustainability and its visually verifiable representation to be assessed separately. Table 4 also gives, for each strategy, the visual evidence required for V-S = 1 and the type of evidence on which its verification depends (Section 2.5); together with the correspondence rates in Equations (2)–(4), these criteria constitute the indicator set used to assess the visual verifiability of sustainability strategies in this study. No relative weights were assigned to the strategies. Because the images provide no information on environmental performance, weighting the strategies by their environmental importance would imply a performance assessment that the visual evidence cannot support; equal weighting instead gives each requested strategy the same influence on the correspondence rates. What this indicator set can and cannot establish about actual environmental performance is discussed in Sections 4.3 and 4.7.

Table 4. Sustainability codes (S1–S7).

Code	Strategy	Definition	Visual Evidence for V-S = 1	Type of Evidence
S1	Material/Resource efficiency	Explicit and visually verifiable representation of material efficiency, reduced resource use, or resource efficiency	A material- or resource-efficiency system identifiable in the image; certification, recycled content, or carbon-related properties cannot be inferred from appearance	Specification or performance evidence
S2	Daylighting strategy	Presence of an identifiable design feature, such as a skylight, atrium, dedicated daylight opening, solar shading system, or overhead structure that filters daylight	Skylight, atrium, light well, solar-shading device, or daylight-filtering roof or façade system; a large window alone is insufficient	Visually assessable
S3	Passive ventilation/Climate control	Visual representation of natural ventilation, the stack effect, cross-ventilation, passive cooling, or a similar climatic mechanism	Louvers, chimneys, ventilation openings, double-skin systems, or another passive climate system	Partially assessable
S4	Renewable energy	Visibility of solar panels/PV, wind turbines, or another renewable energy system	Photovoltaic panels, turbines, or another energy-generation system	Visually assessable
S5	Water management	Representation of rainwater harvesting, water storage, reuse, a visible water cycle, or similar systems	Cistern, collection channel, or visible rainwater or greywater system	Visually assessable

Table 4. *Cont.*

Code	Strategy	Definition	Visual Evidence for V-S = 1	Type of Evidence
S6	Circularity/ Reuse/Recycling	Visibility of identifiable elements associated with recycling, reuse, disassembly, or a circular system	Recycling station, reuse system, or reused element whose previous function remains recognizable	Partially assessable
S7	Indoor environmental quality	Explicit representation of indoor physical environmental performance, such as acoustics, air quality, low VOC levels, or thermal comfort, through a visually verifiable indicator	Acoustic or indoor environmental quality system identifiable in the image; plants or a calm atmosphere alone are insufficient	Partially assessable

Note. Each strategy was coded as a binary variable, separately in the Sustainability-only prompt (P-S) and in the final images (V-S), according to the visual evidence criteria; the complete decision rules are given in the codebook (Supplementary Materials, Section S1). All strategies were weighted equally. The correspondence rates (Equations (2)–(4)) express the proportion of requested strategies that were visually verified; they do not measure environmental performance.

The B + S outputs, in which biomimicry and sustainability were requested together, were assessed separately to determine the extent to which the two approaches were integrated within the same design. Unlike the binary presence/absence coding used for the other indicators, this assessment employed an ordinal integration scale ranging from IL0 to IL3 (Table 5). Coding was applied only to the B + S images and distinguished between the mere coexistence of biomimetic and sustainable elements in the same image and their functional integration within the same design system. IL3 was reserved for cases in which the biomimetic principle could be clearly identified in the image as the mechanism through which a sustainability strategy is depicted to operate. Because environmental performance was not measured, all integration levels describe design relationships depicted in the images rather than achieved performance. In borderline cases, the lower integration level was assigned.

Table 5. Combined output integration level (IL0–IL3).

Level	Definition	Criterion
IL0	No integration	Biomimicry and sustainability cannot be identified together, or one of the two dimensions is visually absent
IL1	Additive	Biomimicry and sustainability are present in the same image but are mostly juxtaposed as independent design layers or elements
IL2	Coordinated	The two approaches are coherently interconnected within the same design system, with identifiable functional or spatial coordination
IL3	Synergistic	The biomimetic principle is visually identifiable as the mechanism through which a sustainability strategy is depicted to operate; the two approaches form a single system that is difficult to separate visually (depicted relationship, not measured performance)

After coding was completed, the visual-aesthetic profiles of the B and S conditions were quantitatively compared at the project level using Gower similarity, a general coefficient

that expresses the similarity between two cases as the mean of their similarity scores on the individual variables and that can be applied to binary as well as other types of variables [38]. Only the VA1–VA5 codes were used for this comparison; biomimetic representation (B1–B4), sustainability strategies (S1–S7), and integration level (IL0–IL3) were excluded from the Gower similarity calculation [38]. The resulting Gower value was interpreted as a measure of visual-aesthetic similarity between the two conditions rather than as an indicator of conceptual or performance-based achievement.

Project-level similarity between the B and S outputs was calculated using Gower similarity, with equal weights assigned to the five binary VA variables. For each variable, a score of 1 was assigned when the B and S outputs had the same value and 0 when their values differed. The mean of the five variable-level scores constituted the project-level similarity score:

$$G_i(B, S) = \frac{1}{5}\sum_{k=1}^{5} I\left(\mathrm{VA}_{i,k}^{B} = \mathrm{VA}_{i,k}^{S}\right), \quad (1)$$

where $I(\cdot)$ is an indicator function that takes a value of 1 when the condition in parentheses is satisfied and 0 otherwise. $G_i(B, S)$ ranges from 0 to 1, where 1 indicates a complete match between the B and S visual-aesthetic profiles across all five indicators, whereas 0 indicates that none of the indicators match. Lower values therefore indicate greater visual-aesthetic divergence. All five VA variables were weighted equally in the calculation. Because the score is the proportion of matching decisions across five binary indicators, it can take only six values (0, 0.20, 0.40, 0.60, 0.80, and 1.00); a score of 0.80 therefore means that the B and S profiles matched on four of the five indicators, and a single differing decision changes the score by 0.20.

In Equation (1), the VA variables were treated as symmetric binary variables, that is, as two-state qualitative characters in Gower's formulation [38], so that a joint absence counted as a match; under this treatment, the coefficient is equivalent to the simple matching coefficient. Because a joint absence may say less about shared visual character than a shared presence, the robustness of this measure was examined in three ways. First, the prevalence of each VA feature was determined separately for the B and S outputs, together with the number of projects in which the feature was present in both conditions, absent in both, or present in only one. Second, a sensitivity analysis based on shared presences was performed using the asymmetric treatment that Gower specifies for dichotomous characters, in which joint absences are excluded; this corresponds to the Jaccard coefficient, $J_i(B,S) = a/(a + b + c)$, where a is the number of VA features present in both outputs and b and c are the numbers of features present only in the B or only in the S output. Both measures were also recalculated without the projects with shared or sequential visual inputs (Section 2.2). Third, to examine whether the similarity between paired outputs exceeded the similarity expected from the overall prevalence of these features, each B profile was compared with the S profiles of the other 21 projects, and the observed within-project mean was compared with the means obtained from 20,000 random re-pairings of B and S profiles across projects. The project-level VA profiles are reported in Table S2 of the Supplementary Materials.

For each project, the binary codes were recorded as fixed-order profile strings. For example, the profile V-S = 0101100 corresponds to S1 = 0, S2 = 1, S3 = 0, S4 = 1, S5 = 1, S6 = 0, and S7 = 0, respectively. The same notation was applied to the P-S, P-B, V-B, and VA profiles. This representation enabled direct project-level comparison of the coded prompt and visual-output profiles, including which sustainability strategies were explicitly requested and which could be visually verified in the resulting outputs.

2.5. Prompt–Output Correspondence (P-S/V-S)

For sustainability, the primary aim was to determine the extent to which strategies explicitly requested in the prompts could be visually verified in the resulting outputs. Correspondence was therefore assessed only for strategies coded as $P\text{-}S_k = 1$. The appearance of an additional sustainability element that had not been requested in the prompt was not counted as correspondence with the relevant prompt request.

$$C_i = 100 \times \frac{\sum_{k=1}^{7}(P_{ik} \times V_{ik})}{\sum_{k=1}^{7} P_{ik}} \quad (2)$$

Equation (2) gives the project-level correspondence rate. If the denominator is zero, the project-level percentage is reported as N/A. If the prompt data could not be reliably recovered, NR is used and no rate is calculated; this case did not occur in the present dataset.

$$C_k = 100 \times \frac{\sum_i (P_{ik} \times V_{ik})}{\sum_i P_{ik}} \quad (3)$$

Equation (3) gives the category-level correspondence rate, indicating the percentage of instances in which an explicitly requested sustainability strategy was visually verified in the output.

$$C_{\text{total}} = 100 \times \frac{\sum_i \sum_k (P_{ik} \times V_{ik})}{\sum_i \sum_k P_{ik}} \quad (4)$$

The overall correspondence rate was calculated using an instance-weighted approach in which all explicit prompt-strategy requests were given equal weight, rather than by taking a simple average of the project-level percentages.

To distinguish the limits of visual verification from the limits of image generation, the seven strategy categories were also grouped according to the type of evidence required to verify them, following the visual evidence criteria of the codebook (Supplementary Materials, Section S1). Visually assessable strategies (S2, S4, and S5) are defined by physical components, such as skylights or shading devices, photovoltaic panels or turbines, and cisterns or collection channels, that can be depicted directly in an image. Partially assessable strategies (S3, S6, and S7) can be represented only through visible proxies, such as ventilation openings, reused elements, or acoustic panels, whereas their defining properties, such as airflow, thermal behavior, recycled content, recyclability, air quality, or acoustic and thermal comfort, cannot be observed in an image. Strategies requiring specification or performance evidence (S1) are defined by material properties, such as certification, carbon content, or resource efficiency. Equation (4) was applied separately to each group. This grouping was introduced to interpret the overall rate; it did not change any code.

The visibility of the V-B and V-S codes was calculated from the frequency of each code across the 22 projects.

$$F_k(\%) = 100 \times \frac{\sum_i V_{ik}}{22} \quad (5)$$

These percentages indicate the overall frequency of visual representation across all final images, regardless of whether a strategy was explicitly specified in the prompt.

Three quantities should therefore be distinguished. Equations (2)–(4) measure (a) the visual verifiability of requested attributes, that is, whether a strategy requested in the prompt could be confirmed in the image under the evidence rules of the codebook. Equation (5) measures (b) the depicted presence of a strategy, regardless of whether it was requested. Neither measures (c) actual prompt adherence or the physical sustainability performance of the depicted design: a requested property that cannot be seen, such as a certified or low-carbon material, can be neither confirmed nor excluded from an image. The correspondence

rates are therefore not direct estimates of how often generative AI follows sustainability instructions, and they are interpreted together with the grouping by type of evidence.

All cases were coded by the second author (primary coder) using the same predefined coding guide. The coder was not the instructor of the course from which the dataset was derived, which limited the potential influence of course-assessment expectations or familiarity with individual students on coding decisions. Analytical consistency and reproducibility were supported through explicit operational definitions, inclusion and exclusion rules, binary decision criteria, and an evidence-based documentation system.

For each binary variable, only cases supported by sufficient textual or visual evidence were coded as 1; ambiguous, indirect, or independently unverifiable cases were coded as 0. In particular, a feature specified in a prompt was not assumed to be present in the visual output; prompt (P-) and visual-output (V-) codes were assigned independently. For example, the inclusion of "FSC-certified wood" or "recycled material" in a prompt could provide evidence for P-S1 or P-S6. However, an ordinary wooden surface in an image did not verify these properties and was therefore not automatically coded as 1 for V-S1 or V-S6.

A fixed decision rule was applied in borderline cases. When evidence was insufficient, the more conservative coding decision was adopted. Accordingly, uncertainty in binary variables was coded as 0. For integration levels, IL2 was assigned when the available visual evidence was insufficient to establish the cause-and-effect relationship required for IL3. This applied in particular when the mechanism linking the biological principle to the sustainability strategy was described only in the prompt, as in the case of kinetic or responsive behavior, which cannot be observed in a still image. To ensure the traceability of coding decisions, an audit trail was maintained in which the prompt statement or visual element supporting each code assigned a value of 1 was recorded as an evidence note.

Because a binary code assigns 0 both when a feature is observably absent and when it cannot be determined from an image, the visual codes were interpreted as verified (1) or not visually verified (0) rather than as present or absent, and the meaning of a 0 was specified according to the type of evidence each code requires. For codes defined by visible form or by physical components that can be depicted directly (VA1–VA5, V-B1–V-B3, V-S2, V-S4, and V-S5), a 0 means that the feature was not depicted in the final images (for the VA codes, that it did not characterize their overall appearance); this is an observable absence within the depicted views, although presence outside them cannot be excluded. For codes whose visual evidence consists of proxies (V-S3, V-S6, and V-S7), a 1 means that a qualifying proxy, such as a ventilation opening, a reused element, or an acoustic system, was depicted, not that the defining property itself was verified, and a 0 means that no such proxy was depicted. For V-S1, which depends largely on material specifications, a 0 in most cases indicates indeterminacy, because these properties can rarely be established from an image. For V-B4, a 0 may reflect either situation: no biomimetic element may be depicted in an environmental role, or such an element may be depicted without its mechanism being observable in a still image. Indeterminacy was therefore treated at the level of code categories rather than recorded as a separate value for each case: the grouping of sustainability strategies by evidence type described above reports the strategies whose visual verification is possible in principle separately, and zero frequencies, such as those for V-S1 and V-B4, are reported as the absence of visual verification rather than as evidence that the corresponding features were absent from the students' designs.

The transparency and reproducibility of the coding process were strengthened through a predefined codebook, conservative coding rules, the separation of prompt and visual-output layers, and the systematic recording of evidence notes. The reliability of the coding was further examined through an independent coding check (Section 2.6).

2.6. Independent Coding Check

To examine the reliability of the coding, a second coder, a faculty member in interior architecture who was not involved in the course or in the study, independently coded all 22 projects using the codebook (Supplementary Materials, Section S1), without access to the primary coder's codes. The independent coding covered all binary layers (VA1–VA5 for the B and S outputs, P-B1–P-B4, V-B1–V-B4, P-S1–P-S7, and V-S1–V-S7; 704 decisions) and the integration levels of the B + S outputs (22 decisions). Agreement was assessed using percentage agreement and Cohen's κ, weighted for the ordinal IL scale. Because several codes showed extreme prevalence, under which κ can be low or undefined despite high agreement, Gwet's AC1 was also calculated [39]. All agreement statistics were calculated on the primary coder's original codes, before any revision.

Disagreements were handled as follows. P18, the only project that the primary coder had classified as IL3, had been classified as IL2 by the independent coder; after re-examining the visual evidence, the primary coder concluded that IL3 was not supported, and the code was revised to IL2 (Section 3.5). All other results reported in Section 3 are based on the primary coding. For the classifications most likely to affect the conclusions, namely the boundary between a daylight impression (VA5) and a daylight strategy (V-S2), the visual verification of indoor environmental quality (V-S7), and the integration levels, the results of both coders are reported (Section 3.6). The project-level codes of the independent coder and the agreement statistics are given in Tables S3 and S4 of the Supplementary Materials.

3. Results

This section presents the findings in relation to the four research questions. First, the visual-aesthetic similarity between the outputs generated under the biomimicry (B) and sustainability (S) conditions is examined using Gower similarity calculated from the VA1–VA5 profiles. The frequencies of biomimetic representation and the visual representation of sustainability strategies are then reported. Prompt–output correspondence for sustainability strategies is evaluated at both the project and category levels, followed by the integration levels observed in the B + S outputs, in which biomimicry and sustainability were requested together. Throughout the analysis, the prompt (P-) and visual-output (V-) layers were evaluated independently, and visual-aesthetic similarity was treated as a separate comparative measure rather than as evidence of sustainability performance or the depth of biomimetic representation.

3.1. Visual-Aesthetic Similarity Between B and S Outputs

Gower similarity calculated from the VA1–VA5 profiles showed substantial visual-aesthetic similarity between the images generated separately for biomimicry and sustainability. The mean project-level Gower score was 0.818 (approximately 0.82), and the median was 0.80. Of the 22 projects, 18 (81.8%) had a similarity score of 0.80 or higher (Table 6). Scores ranged from 0.40 to 1.00, with a complete match (1.00) observed in seven projects (31.8%). Among the four projects with similarity scores below 0.80, three scored 0.60 and one scored 0.40.

Table 7 reports the prevalence of each VA feature in the B and S outputs and its distribution across the two conditions. Nature-inspired color palettes (VA3) were present in all 44 outputs, and natural-looking materials (VA2) and daylight/outdoor connection (VA5) were each present in 43. Planting (VA4) was present in 18 B and 20 S outputs. Organic/curved form (VA1) was the only feature that clearly differentiated the two conditions: it was present in all 22 B outputs but in only 8 S outputs. Of the 110 feature-level comparisons, 90 were matches, of which 89 were shared presences and only one was a joint absence (VA4 in P15), whereas 14 of the 20 mismatches involved VA1.

Table 6. Distribution of Gower-based visual similarity scores between biomimicry and sustainability outputs.

Gower Score	*n*	%
1.00	7	31.8
0.80	11	50.0
0.60	3	13.6
0.40	1	4.5
Total	22	100.0

Table 7. Prevalence of the visual-aesthetic features (VA1–VA5) in the B and S outputs and their distribution across the two conditions ($n = 22$ projects).

Code	Feature	B Outputs *n* (%)	S Outputs *n* (%)	Present in Both	Absent in Both	B Only	S Only
VA1	Organic/Curved form	22 (100.0)	8 (36.4)	8	0	14	0
VA2	Natural-looking material	22 (100.0)	21 (95.5)	21	0	1	0
VA3	Nature-inspired color palette	22 (100.0)	22 (100.0)	22	0	0	0
VA4	Planting/Biophilic component	18 (81.8)	20 (90.9)	17	1	1	3
VA5	Daylight/Outdoor connection	21 (95.5)	22 (100.0)	21	0	0	1
Total	Feature-level comparisons ($n = 110$)	-	-	89	1	16	4

Because joint absences were almost entirely absent, the sensitivity analysis based on shared presences yielded nearly identical results. The Jaccard coefficient differed from the Gower score only in P15 (0.50 instead of 0.60); the mean Jaccard similarity was 0.81 (median = 0.80), and 18 projects (81.8%) again scored 0.80 or higher. Excluding the two projects with shared or sequential visual inputs (P05 and P20) left the pattern unchanged (mean = 0.80 under both measures; 16 of 20 projects ≥ 0.80). In the cross-project comparison, B and S profiles from different projects showed a mean similarity of 0.81, and the observed within-project mean did not exceed the means obtained under random re-pairing (permutation $p = 0.34$).

Similarity was also high across the two main platforms: the mean Gower score was 0.83 for the eight projects generated with Gemini 3 Pro and 0.77 for the seven generated with ChatGPT (GPT-5), and 0.92 for the five projects for which the tool was not specified; these subgroups are, however, too small for statistical comparison. These results indicate that the similarity between the two conditions rested on shared presences of natural-looking materials, nature-oriented color palettes, planting, and daylight/outdoor connections rather than on shared absences, whereas organic form remained largely confined to the B outputs. Because these four features were present in nearly all outputs, in both conditions and across projects, the mean score of 0.82 reflects a nature-oriented visual repertoire shared across the dataset rather than a correspondence specific to the paired outputs of individual projects. A high Gower value therefore indicates similarity in visual-aesthetic profiles rather than the achievement of sustainability performance.

3.2. Visual Representation of Biomimetic Features (V-B1–V-B4)

Morphology/pattern transfer was the most prevalent form of biomimetic representation and was observed in all projects (V-B2 = 100%). An identifiable biological reference (V-B1) and spatial/structural transfer (V-B3) were also observed in most projects, each occurring in 19 of the 22 projects (86.4%). In contrast, V-B4, which required a biological function to be visually identifiable as an operational environmental or performance-related mechanism, could not be verified in any project (Table 8). These findings show that

biomimicry was strongly represented at the morphological and spatial/structural levels, whereas representation at the functional-mechanism level was not visually verified.

Table 8. Visual frequency of biomimetic representation codes (V-B1–V-B4).

Code	Indicator	*n*/22	%
V-B1	Identifiable biological reference	19	86.4
V-B2	Morphological/Pattern transfer	22	100.0
V-B3	Spatial/Structural transfer	19	86.4
V-B4	Visually identifiable functional/environmental mechanism	0	0.0

The prompt-level codes show what the B prompts requested. A biological reference (P-B1) was specified in 19 of the 22 prompts, morphological or pattern transfer (P-B2) in 20, spatial or structural transfer (P-B3) in 16, and a biological function (P-B4) in 7 (Supplementary Materials, Table S2). Requests of the first three types were visually verified in almost all cases: all P-B2 and P-B3 requests (20 of 20 and 16 of 16) and 17 of the 19 P-B1 requests. These features also appeared in several projects whose prompts had not requested them (V-B1 in two, V-B2 in two, and V-B3 in three projects). None of the seven functional requests was visually verified. The requested functions concerned the filtering of daylight or solar shading (P01, P04, P09, P14, and P18), water distribution or irrigation (P01 and P04), natural ventilation (P18), acoustic diffusion and thermal comfort (P15), and self-cleaning surfaces (P03). In six of these projects, the B outputs depicted elements associated with the requested function, such as louvers and perforated canopies casting shade, cellular skylight openings, porous ceiling panels, or ribbed walls, but not the mechanism through which the biological principle would produce the effect; responses to the sun, airflow, and water distribution were described only in the prompts. In P03, the requested water-repellent surface appeared only in a material detail collage, which, as a non-perspective output, was not coded (Supplementary Materials, Table S5).

3.3. Visual Representation of Sustainability Strategies (V-S1–V-S7)

The daylight strategy was by far the most frequently represented sustainability dimension, appearing in 18 of the 22 projects (81.8%). Water management was visually represented in 18.2% of projects, renewable energy in 13.6%, passive ventilation/climate control in 9.1%, and circularity in 4.5%. Material/resource efficiency and indoor environmental quality, however, yielded no visually verifiable indicators in any project (Table 9).

Table 9. Visual frequency of sustainability strategy codes (V-S1–V-S7).

Code	Strategy	*n*/22	%
V-S1	Material/Resource efficiency	0	0.0
V-S2	Daylighting strategy	18	81.8
V-S3	Passive ventilation/Climate control	2	9.1
V-S4	Renewable energy	3	13.6
V-S5	Water management	4	18.2
V-S6	Circularity/Reuse/Recycling	1	4.5
V-S7	Indoor environmental quality	0	0.0

Note. Frequencies are calculated across the final images of all 22 projects, regardless of whether the strategy was requested in the prompt (depicted strategy presence; Section 2.5).

3.4. Prompt–Output Correspondence: Comparison of P-S and V-S

Project-level correspondence showed that, although numerous sustainability strategies were explicitly specified in the prompts, only a limited proportion could be visually verified in the resulting outputs. P12 (66.7%) and P13 (60.0%) achieved the highest correspondence rates, whereas none of the sustainability strategies explicitly requested in P05, P07, P15, or P22 could be visually verified (0%). In P06, only one of five explicitly requested strategies was visually verified, resulting in a correspondence rate of 20.0% (Table 10). Overall, four projects (P05, P07, P15, and P22) showed no visual verification of any explicitly requested sustainability strategy. P11 and P20 were not assigned project-level correspondence rates because no S1–S7 strategy was explicitly requested (N/A).

Table 10. Project-level prompt–output correspondence between sustainability prompt and visual profiles (P-S/V-S).

Code	P-S Profile	V-S Profile	Visually Verified/Requested	%
P01	1111011	0100000	1/6	16.7
P02	1110011	0100000	1/5	20.0
P03	1110000	0100000	1/3	33.3
P04	1110111	0100000	1/6	16.7
P05	1110111	0000000	0/6	0.0
P06	1110110	0100000	1/5	20.0
P07	1000010	0100000	0/2	0.0
P08	1110011	0100000	1/5	20.0
P09	0010101	0100100	1/3	33.3
P10	1110011	0100000	1/5	20.0
P11	0000000	0100000	0/0	N/A
P12	1111110	0101110	4/6	66.7
P13	1110110	0110100	3/5	60.0
P14	1110110	0100000	1/5	20.0
P15	1100001	0000000	0/3	0.0
P16	1100110	0100000	1/4	25.0
P17	1101110	0101000	2/5	40.0
P18	1111111	0101100	3/7	42.9
P19	1110010	0100000	1/4	25.0
P20	0000000	0000000	0/0	N/A
P21	1111111	0110000	2/7	28.6
P22	1000011	0000000	0/3	0.0

Note. Project-level rates are calculated only for the 20 projects with at least one explicitly requested S1–S7 strategy; P11 and P20, in which no strategy was requested, are reported as N/A rather than 0%. A strategy coded P-Sk = 0 was not requested, and a strategy coded P-Sk = 1 and V-Sk = 0 was requested but not visually verified. Prompts and final images were available for all projects (Supplementary Materials, Table S6).

Category-Level Correspondence Rates

At the category level, the highest correspondence rate was observed for the daylight strategy (88.2%), followed by renewable energy (60.0%) and water management (36.4%). Passive ventilation/climate control was visually verified in 13.3% of the instances in which it was explicitly requested, while circularity was verified in only 5.9%. Although material/resource efficiency and indoor environmental quality were explicitly requested in 19 and 11 instances, respectively, neither category produced a visually verifiable representation (Table 11).

$$C_{\text{total}} = 100 \times \frac{25}{95} = 26.3\% \quad (6)$$

Table 11. Category-level prompt–output correspondence for sustainability strategies with subtotals by the type of evidence required for visual verification.

Strategy	Description	Request in the Prompt (n)	Visual Representation (n)	%
S1	Material/Resource efficiency	19	0	0.0
S2	Daylighting strategy	17	15	88.2
S3	Passive ventilation/Climate control	15	2	13.3
S4	Renewable energy	5	3	60.0
S5	Water management	11	4	36.4
S6	Circularity/Reuse/Recycling	17	1	5.9
S7	Indoor environmental quality	11	0	0.0
S2, S4, S5	Visually assessable	33	22	66.7
S3, S6, S7	Partially assessable	43	3	7.0
S1	Specification or performance evidence required	19	0	0.0
S1–S7	Total	95	25	26.3

Of the 95 explicit sustainability strategy requests, 25 were visually verified in the resulting outputs, yielding an overall prompt–output correspondence rate of 26.3%. This finding indicates that, although sustainability was specified across multiple dimensions in the prompts, its visually verifiable representation was largely concentrated in a small number of readily representable strategies.

When the requests were grouped by the type of evidence required for their verification (Table 11), correspondence was 66.7% for visually assessable strategies (22 of 33 requests), 7.0% for partially assessable strategies (3 of 43), and 0% for strategies requiring specification or performance evidence (0 of 19). Nearly two-thirds of the requests (62 of 95) thus concerned strategies whose defining properties could be verified only partially or not at all from a rendering, and these accounted for 59 of the 70 requests that were not visually verified. The overall rate of 26.3% therefore depends largely on the composition of the requests: it indicates how much of the requested sustainability content could be visually verified, not how often image generation failed. The independent coding yielded a similar rate for visually assessable strategies (65.2%) and the same rate for material/resource efficiency (0%), whereas the rate for partially assessable strategies was higher (34.1%), mainly because visible acoustic panels were coded as evidence of indoor environmental quality (Section 3.6).

3.5. Integration Levels in B + S Outputs

Most combined B + S outputs exhibited additive integration (IL1) in the primary coding, accounting for 13 of the 22 projects (59.1%). Eight projects (36.4%) exhibited coordinated integration (IL2), one project (4.5%) showed no integration (IL0), and no project reached the synergistic level (IL3) (Table 12). This distribution indicates that, when biomimicry and sustainability were requested together, the two approaches were most often represented as coexisting but largely separate design components. This predominance of IL1 was not replicated in the independent coding, in which 17 of the 22 projects were classified as IL2 (Section 3.6). Coordinated integration was less common, and synergistic integration was not observed. The case closest to this level was P18, whose B + S courtyard image

depicts a canopy of perforated, petal-shaped panels that filters daylight and casts dappled shade (Supplementary Materials, Section B). The primary coder had initially classified this output as IL3, whereas the independent coder, who had no access to this code, classified it as IL2 (Section 2.6). On re-examination, the IL3 classification was not supported: the image shows a biomimetic element coordinated with a daylighting and shading strategy within the same element (IL2), whereas the kinetic, pine-cone-inspired opening and closing that would make the biological principle the mechanism of solar control was specified in the prompt but cannot be observed in a still image.

Table 12. Distribution of biomimicry + sustainability integration levels (IL0–IL3) in the primary and independent codings.

Level of Integration	Type of Integration	Primary Coder *n*/22	%	Independent Coder *n*/22	%
IL0	No integration	1	4.5	3	13.6
IL1	Additive	13	59.1	2	9.1
IL2	Coordinated	8	36.4	17	77.3
IL3	Synergistic	0	0.0	0	0.0

3.6. Results of the Independent Coding Check

Across the 704 binary decisions, the two coders agreed in 90.9% of cases ($\kappa = 0.81$; AC1 = 0.83). Agreement was high for the visual-aesthetic codes (95.0%; AC1 = 0.94), the biomimetic prompt codes (96.6%; $\kappa = 0.92$), and the biomimetic visual codes (94.3%; $\kappa = 0.87$), and both coders coded V-B4 as 0 in all 22 projects. The codes for morphological and spatial/structural transfer were also applied consistently: agreement was 100% for V-B2 and 90.9% for V-B3 ($\kappa = 0.61$; AC1 = 0.88). Agreement was lower for the sustainability prompt codes (86.4%; $\kappa = 0.72$) and visual codes (84.4%; $\kappa = 0.53$). The mean Gower similarity in the independent coding was 0.85, compared with 0.82 in the primary coding (Supplementary Materials, Table S4).

For the classifications identified as consequential, the results were as follows. The daylight impression (VA5) was coded consistently (agreement of 90.9% for the B outputs and 95.5% for the S outputs; AC1 = 0.90 and 0.95), whereas agreement on the daylight strategy (V-S2) was moderate (77.3%; $\kappa = 0.41$; AC1 = 0.64): the independent coder did not assign V-S2 in four projects in which mainly standard windows were visible and assigned it in one project in which the primary coder had not. Indoor environmental quality (V-S7) was coded in no project by the primary coder but in seven projects by the independent coder, who treated visible acoustic panels and baffles as identifiable acoustic systems. In the independent coding, the prompt–output correspondence rate was 65.2% for visually assessable strategies (15 of 23 requests), 34.1% for partially assessable strategies (14 of 41), and 0% for material/resource efficiency (0 of 16), with an overall rate of 36.3% (29 of 80).

Agreement on the integration levels was low (exact agreement in 8 of 22 projects; weighted $\kappa = 0.08$). The independent coder classified 17 projects as IL2, 2 as IL1, and 3 as IL0, whereas the primary coder classified 13 as IL1, 8 as IL2, and 1 as IL0 after the revision of P18 (Table 12). Nine of the 14 disagreements concerned the boundary between IL1 and IL2. The divergence arose mainly from whether a biomimetic element that also admits or filters daylight, such as a branching or cellular skylight, was regarded as two coexisting layers (IL1) or as coordination within the same element (IL2). Both coders agreed that no project reached IL3, and in both codings at least 19 of the 22 projects were classified as IL1 or IL2.

4. Discussion

This study examined whether biomimicry and sustainability, despite their distinct conceptual and performance-related foundations, retain differentiated design intentions when visualized through generative AI. Taken together, the findings reveal a consistent pattern. Even when biomimicry and sustainability were requested separately, their outputs showed substantial similarity in nature-oriented visual-aesthetic features. By contrast, functional and performance-related dimensions were considerably less visible or verifiable in the resulting images. In this dataset, generative AI therefore reproduced visible nature-oriented design cues more consistently than the principles and mechanisms underlying biomimetic and sustainability performance. This pattern bears on the inconsistency in the literature outlined in the Introduction. The images were rich in the nature-oriented cues that biomimicry and sustainability share, which is the level of design potential and creative exploration at which several studies have emphasized the potential of generative AI [35,36]; at the same time, they conveyed little verifiable information about the functional and performance-related intentions that distinguish the two approaches, which is the level at which critical studies have reported limitations [28,29,31]. The findings are therefore consistent with the working interpretation that the two strands of research evaluate different levels of the same kind of output rather than reach contradictory conclusions about it. They are also compatible with alternative explanations, discussed below, that locate part of this pattern in the prompts, in the medium of the still image, and in the evidence rules of the coding. The following subsections interpret each finding in relation to the four research questions and previous research; the corresponding figures are reported in Section 3 and are not repeated here.

4.1. Visual-Aesthetic Convergence Between Biomimicry and Sustainability Outputs

The high Gower similarity calculated from the VA1–VA5 indicators indicates substantial visual-aesthetic similarity between the B and S outputs in this dataset. Although biomimicry and sustainability are based on distinct conceptual and performance-related principles, recurring visible features—natural-looking materials, nature-oriented color palettes, planting, and daylight/outdoor connections—created similar visual profiles across the two conditions, while organic form remained the main feature distinguishing the B outputs. This similarity rested on shared presences rather than shared absences, and the cross-project comparison indicates that it reflects a repertoire common to the dataset rather than a correspondence specific to each project. This finding is consistent with studies suggesting that environmentally responsive design can develop a recognizable aesthetic vocabulary and that text-to-image systems may favor recurring visual patterns, particularly when prompts are not sufficiently differentiated [27,28]. Discussions concerning the possibility that generative AI systems may reproduce similar visual vocabularies and contribute to formal or aesthetic homogenization also provide a relevant context for interpreting this result [29]. The features on which the two conditions converged are, moreover, those that the literature on sustainability aesthetics identifies as the visible vocabulary of environmentally responsive design, such as natural materials, color, planting, and natural light [27], whereas organic form, the only feature that separated the conditions, corresponds to the formal level of biomimetic transfer [13]. In visual terms, sustainability was thus represented through its aesthetic vocabulary and biomimicry largely through its form, so that the two approaches met on the ground they share.

The shared repertoire can be characterized more precisely with reference to biophilic design, a concept that the present study did not use as an analytical category but whose visual-aesthetic indicators closely match. Biophilia denotes an innate human tendency to affiliate with life and lifelike processes [40]. Biophilic design translates this tendency into

the built environment through direct and indirect experiences of nature, such as natural light, plants, views, natural materials and colors, and naturalistic shapes and forms, with the aim of supporting human health and wellbeing [41,42]. Explanations of why people respond to such features draw on environmental psychology and evolved preferences for natural settings [43], and Mehaffy and Salingaros relate biophilia to the structural qualities of environments that people experience as living [44]. All five VA indicators correspond to attributes listed in this literature: natural materials (VA2), natural colors (VA3), plants (VA4), light and views (VA5), and naturalistic shapes and forms (VA1) [42]. The convergence observed here can therefore be described as a convergence of both conditions on a biophilic image vocabulary, from which only organic form, the attribute closest to the formal level of biomimetic transfer, separated the biomimicry outputs.

This distinction matters because biophilia, biomimicry, and sustainability differ in purpose and in their criteria of success. Biophilic design is judged by its effects on human experience and wellbeing, and most of its features are meant to be seen and felt. Biomimicry, as its originators framed it, asks designers to learn from how organisms and ecosystems function and is judged by the performance of the resulting design [11,12], whereas sustainability is judged by the reduction in environmental impacts. The concepts are related: Kellert and Calabrese include biomimicry among the attributes of biophilic design, as one way of evoking nature indirectly, and emphasize that low environmental impact, although essential, is insufficient on its own [42]. A design feature may exhibit recognizable biophilic attributes without demonstrating functional biomimetic transfer or environmental sustainability. A planted, daylit interior with timber surfaces and curving forms meets the visual criteria of biophilic design without showing that any biological process has been transferred or any environmental impact reduced. In this dataset, the biomimicry and sustainability prompts thus appear to have converged on the component of nature-oriented design that a still image conveys most readily, its biophilic appearance, while the functional and performance-related components that distinguish the two approaches remained largely invisible (Sections 4.2 and 4.3).

However, the observed similarity should not be interpreted as evidence that biomimicry and sustainability are conceptually equivalent or that generative AI necessarily produces the same visual language across platforms. Different generative AI systems may vary in prompt adherence and aesthetic interpretation [30], and the present study did not use a controlled single-model design that would allow tool effects to be isolated. The observed similarity may therefore reflect the combined influence of shared nature-oriented prompt content, the shared project concept underlying both conditions, shared or sequential visual inputs in some projects, the students' selection of final images, recurring aesthetic conventions, and tool-specific interpretation (Section 2.2). Because these sources cannot be separated in a course-based dataset, the convergence reported here describes the similarity observed within this image set rather than a general tendency of generative AI tools. The cross-project comparison nevertheless suggests that project-specific factors, such as a shared floor plan or project concept, were not the main source of this similarity, since B and S outputs from different projects were as similar as those from the same project (Section 3.1). Because similarity was high in projects generated with both main platforms, the convergence does not appear to be confined to a single model, although the dataset cannot separate model effects from the other sources listed above. Of these sources, the prompts deserve particular attention: many of the B prompts also referred to daylight, planting, natural materials, or sustainability (Supplementary Materials, Section S2), so the convergence may have been produced jointly by the prompts and the models rather than by the models alone. This would be consistent with the dual use of both terms as a visible design language and as a functional principle described in the Introduction, and the dataset does not allow the two contributions

to be separated. The finding that similarity did not depend on the platform also differs from comparisons that found tool-specific interpretations of the same prompts [30,32]; the two observations are compatible if tools differ in how they render a specific request but draw on a similar repertoire for nature-oriented interiors. Most importantly, a high Gower value indicates similarity in visual-aesthetic profiles rather than the achievement of sustainability performance. A space that appears natural or green should therefore not be assumed to be environmentally sustainable on the basis of visual appearance alone.

4.2. Formal-Functional Asymmetry in Biomimetic Representation

The biomimetic representation findings reveal a marked asymmetry between visible morphological representation and functional mechanisms. Formal, pattern, and spatial transfer were represented in almost all projects, whereas no functional mechanism was visually verified (Section 3.2). This distribution is consistent with the distinction in biomimetic literature between surface-level formal resemblance and deeper forms of biological principle transfer [9,13]. Previous studies have similarly emphasized that biomimicry may range from reductionist to more holistic approaches and that the imitation of biological forms or organism characteristics does not in itself ensure environmental sustainability [14]. The present results also extend the observation that text-to-image models reproduce the geometric analogies of biological forms without conveying the underlying form–function relationship [31]: in this dataset, the same asymmetry appeared in interior images generated with several tools.

Because only seven of the 22 B prompts requested a biological function, a V-B4 value of 0 in the other 15 projects is consistent with what their prompts requested; the representational shortfall concerns the seven functional requests, none of which was visually verified, whereas morphological and spatial requests were almost always represented. The present findings suggest that biological form, pattern, and spatial organization are more readily represented in AI-generated interior images than functional biological mechanisms. This does not necessarily indicate that the underlying design intentions were absent; rather, such mechanisms were not visually verifiable under the criteria applied in this study. Three explanations, which the data cannot fully separate, may account for this result. First, the requested mechanisms, such as airflow, water distribution, and kinetic or seasonal responses to the sun, are processes that unfold over time, whereas a single perspective can depict only a state; part of the asymmetry may therefore lie in the medium rather than in the model. Second, the seven prompts described the intended effect, for example, panels that "naturally circulate air", rather than a configuration that could be drawn, such as inlets, channels, and outlets, which is consistent with evidence that prompts expressing design intent in a structured way reduce the gap between intended and generated images [8]. Third, the V-B4 criterion is demanding by design, because it requires the operation of the mechanism, and not only the presence of the element, to be legible (Section 2.4). This distinction is important because biomimetic design as a problem-solving approach requires biological principles to extend beyond formal inspiration and inform structural, functional, or environmental design decisions [10,12]. From a design-education perspective, the findings suggest a hypothesis, not tested here, that students may need to define explicitly the biological principle, its intended function, and the design mechanism through which that function is expected to operate, rather than relying on visual resemblance alone.

4.3. Limited Visual Verifiability of Sustainability Strategies

Prompt–output correspondence revealed a substantial difference between sustainability strategies that can be represented through visible design elements and those that depend on material, technical, or performance information. Strategies with depictable physical

components, such as skylights, photovoltaic panels, or water-collection elements, were usually verified, whereas strategies defined by material or performance properties were rarely or never verified, and for indoor environmental quality the result depended on whether visible acoustic elements were accepted as evidence (Sections 3.4 and 3.6).

This pattern highlights an important limitation of image-based sustainability assessment. Features such as skylights, photovoltaic panels, shading systems, or visible water-management elements can be represented through recognizable visual components. By contrast, FSC certification, recycled content, low-carbon material properties, indoor air quality, or similar performance-related characteristics cannot generally be verified from appearance alone. The low overall rate should therefore not be read as a failure of image generation or as a direct estimate of prompt adherence (Section 2.5): where a requested strategy had a depictable physical component, it was visually verified in two-thirds of the cases, whereas most of the remaining requests concerned properties that a rendering can verify only partially or not at all. The findings therefore reinforce the distinction between sustainability as an aesthetic expression and sustainability as measurable environmental performance [27]. They also support approaches in which generative AI-assisted design is supplemented with real-world material and carbon data [33]. Where sustainability attributes cannot be visually verified, additional information layers, such as material specifications, environmental data, system diagrams, or performance assessments, may be required. Without such supporting information, visually persuasive images may be interpreted as sustainable despite providing insufficient evidence of environmental performance.

The comparison between requested and depicted strategies also qualifies a homogenization reading of these results [29]. Apart from daylighting, which was also depicted in three of the five projects whose prompts had not requested it, sustainability strategies were visually verified only in projects whose prompts had requested them (Supplementary Materials, Table S2). Depictable components such as photovoltaic panels or water-collection elements thus appeared only when they had been requested, which suggests responsiveness to explicit prompts rather than a default repertoire, whereas daylight appears to belong to the default repertoire of nature-oriented interiors [28]; the high correspondence for daylighting may therefore partly reflect this default rather than the prompts. What the images did not convey was the part of sustainability that cannot be depicted, which is also the part that distinguishes sustainability as environmental performance from sustainability as an aesthetic vocabulary [27].

4.4. Absence of Synergistic Integration in Combined Outputs

The B + S findings further show that requesting biomimicry and sustainability together did not necessarily result in their functional integration. This result complements the findings reported for the separate B and S conditions. Functional biomimetic mechanisms were not visually verified under V-B4, while several sustainability strategies also showed limited visual correspondence. Together, these patterns help explain why synergistic integration was absent in the combined outputs. More specifically, an IL3 classification requires an image to show a biological mechanism producing a sustainability effect and therefore combines the two kinds of information, functional mechanism and performance, that the separate conditions rarely conveyed. The P18 case illustrates this boundary: a biomimetic canopy depicted as a shading device was visible, whereas the biological mechanism intended to drive it was not. The independent coding qualifies the predominance of additive integration: the relative proportion of additive and coordinated integration depended on the coder, whereas the absence of synergistic integration was consistent across both codings. The robust finding is therefore that the combined outputs remained at the level of coexistence or element-level coordination, rather than that they were predominantly additive. The

low agreement at the IL1–IL2 boundary is itself informative: in still images, integration is read mainly through spatial coincidence, so whether a cellular skylight is seen as two coexisting layers or as one coordinated element depends on interpretation rather than on visible operation. Even after the guidance given before the combined stage, which asked students to specify sustainability more explicitly (Section 2.2), no synergistic case emerged, and in P18 a mechanism specified in the prompt could not be shown.

The result is also consistent with the conceptual distinction between using biomimicry as formal inspiration and translating biological principles into processes or systems that contribute to sustainable performance [10,24]. The findings indicate that the simultaneous presence of biomimetic and sustainability-related features does not in itself constitute synergistic integration. Synergy in this sense, in which biological principles inform processes or systems rather than forms, may be difficult to demonstrate with still images alone. From a pedagogical perspective, students may therefore need to establish an explicit relationship between the biological reference, the abstracted principle, the design mechanism, and the intended sustainability outcome. Such a sequence may help move combined designs from additive coexistence toward coordinated or synergistic integration, although this possibility could not be tested with the present data.

4.5. Implications for Design Education

Taken together, the findings have several implications for the use of generative AI in interior architecture education. Two kinds of claims need to be distinguished here: what the dataset shows about the representational properties of the final AI-generated images, and what educators might test in future instructional designs. The present study supports only the first. It did not assess student learning outcomes, cognitive development, changes in prompting competence, or the effects of instructional interventions, and it therefore cannot establish the effectiveness of any pedagogical strategy. The combination of high visual-aesthetic similarity, limited representation of functional biomimetic mechanisms, and the limited visual verifiability of sustainability strategies that depend on specification or performance evidence indicates, within this course-based dataset, the importance of distinguishing visual plausibility from verifiable design performance. Whatever the role of generative AI in ideation and early visual exploration [1,6], the present results indicate that AI-generated images should not be treated as direct evidence of biomimetic functionality or environmental performance. For educational use, this suggests that students may need not only to formulate design intentions through structured prompts but also to critically evaluate whether those intentions are actually represented in the resulting outputs [8]. Whether such critical evaluation can be taught effectively, however, remains to be tested. This interpretation is consistent with broader discussions emphasizing that the creative potential of generative AI in design education should be accompanied by critical and pedagogical evaluation of the resulting work [5,6].

On this basis, we propose, as a hypothesis for future investigation rather than as a demonstrated outcome, that prompt writing could be approached as a form of design literacy rather than merely as a technical instruction for image generation. A possible assessment protocol, which was neither implemented nor evaluated in the present study, would ask students to (1) identify a biological source; (2) specify the transferable principle; (3) explain the intended building-system mechanism; (4) identify which sustainability attributes can be confirmed visually in their images; and (5) provide appropriate technical evidence, such as material specifications, system diagrams, or simulations, for those that cannot. Whether such a protocol improves students' biomimetic reasoning and the critical use of generative AI could be examined in future studies with a baseline assessment, a comparison cohort,

and direct measures of learning, building on studies of generative-AI-assisted biomimetic studios [35,36].

4.6. Image and Process in Sustainability and Biomimicry Education

Sustainability is now widely presented as a core component of architecture and interior design education, and students regard it as central to their training; how they understand its role in design, however, varies with the depth and placement of sustainability content in the curriculum [45]. When sustainability is taught and assessed mainly through presentation drawings and renderings, it can easily slide into what Guy and Farmer called an eco-aesthetic logic, in which environmental values are expressed through form and imagery [46], rather than into an understanding of how a design uses materials, energy, water, and air over its life cycle. The findings illustrate the risk of such a shift when sustainability is represented and assessed primarily through text-to-image outputs. The sustainability outputs converged on the same biophilic vocabulary as the biomimicry outputs (Section 4.1); strategies with depictable components were usually verified, whereas material efficiency, indoor environmental quality, and other strategies defined by processes and performance were rarely or never verifiable (Section 4.3); and students could request a strategy in words without any part of the submission showing how it would work. The course from which the dataset derives is itself an example: its deliverables were prompts and generated images, and the assignment did not require evidence of performance. Teaching sustainability as process would require assessment to ask for such evidence, for example, daylight or energy calculations, material and life-cycle specifications, and system diagrams, and to treat the image as a representation to be checked against this evidence rather than as the outcome of the design task. Whether such a change improves students' understanding of sustainability remains to be examined; the present study shows only that an image-based submission cannot carry this content.

A related problem concerns biomimicry. Benyus framed biomimicry as learning from nature, which serves as model, measure, and mentor [11], and in interior architecture it has likewise been presented as a problem-solving methodology rather than as a source of forms [12]. Pedersen Zari distinguishes biomimicry at the levels of the organism, its behavior, and the ecosystem, and argues that mimicking an organism or its form can leave a building conventional in its environmental impact, whereas emulating behaviors and ecosystem processes offers the greatest potential for sustainable and even regenerative design [47]; biomimetic transfer can similarly be described at formal, structural, mechanical, and functional levels [13]. When biomimicry is introduced mainly through striking biological forms, the present findings indicate what this emphasis produces in combination with text-to-image tools. Formal and spatial transfer was represented in almost every project, whereas none of the seven functional requests was visually verified, and the prompts that requested a function described its intended effect rather than the process through which the effect would be achieved (Sections 3.2 and 4.2). Copied forms then merge into the same biophilic vocabulary as sustainability imagery (Section 4.1): a form copied for its visual impact may have biophilic value for the people who experience it [42], but it does not in itself make a design more sustainable [14]. Teaching biomimicry as a process would mean starting from a design problem, such as ventilation, shading, or water management, identifying biological strategies that address comparable problems, and asking students to show how the abstracted principle operates in the building, with the image serving to communicate this operation rather than to replace it; the assessment protocol proposed in Section 4.5 is one possible way of making this sequence explicit.

The same distinction applies in practice, where AI-generated images increasingly circulate as early design proposals. Such images can convey a biophilic atmosphere persuasively,

but claims that a design is biomimetic or sustainable require evidence of the processes it embodies, such as material, energy, and performance data [33]. Distinguishing between what an image shows and what a design does is therefore not only a methodological requirement of the present study but also a practical one for designers, educators, and clients.

4.7. Limitations and Future Research

The study has a number of limitations. Firstly, the dataset consists of 22 student projects from a single course, and the findings should therefore be interpreted in the context of that specific course rather than generalized to design education. Secondly the visual outputs were generated using different generative AI tools. Given that the various platforms may differ in their degree of prompt adherence and in their aesthetic interpretation [30], the current research design does not allow the visual patterns observed to be ascribed to any one particular model. A controlled study in which the same prompts are used with a single model or a systematic comparison of multiple models would be required to determine the extent to which visual similarity and prompt-to-output correspondence are model-dependent. Thirdly, whilst the independent coding check revealed a high degree of agreement overall, agreement on the boundary between additive and coordinated integration was poor which suggests that the IL1-IL2 distinction requires more clearly defined operational criteria (e.g., for biomimetic elements which also admit or filter daylight); agreement on the daylight strategy (V-S2) and on indoor environmental quality (V-S7) was also low. These criteria could be refined in future work and reliability with multiple coders could be assessed prior to the final coding cycle. Fourthly as described in Section 2.2 the number of generation attempts was left to the students, the final images were selected by the students themselves and a relatively small number of projects involved shared or sequential visual inputs. These are all features of the naturalistic course environment but do mean that the observed similarity cannot be ascribed to a single source; greater procedural in the form of a fixed tool, a fixed number of generations and researcher-controlled image selection would be required in future comparative studies.

A further limitation concerns the scope of visual verification. The study did not measure the actual environmental performance of the proposed interiors; it examined whether sustainability and biomimetic design intentions could be identified in the generated images according to predefined visual criteria. Future studies could therefore combine AI-generated visualizations with material specifications, life-cycle data, daylight or energy simulations, indoor environmental performance indicators, and technical system diagrams. Integrating such evidence would allow researchers to compare visual representation with measurable environmental performance. Future research could also examine whether structured prompts that explicitly connect biological principles, design mechanisms, and sustainability objectives increase functional biomimetic representation and higher levels of B + S integration. Longitudinal and multi-institutional studies would further help determine how these competencies develop across different educational contexts. Finally, the study did not assess student learning outcomes, cognitive development, or changes in prompting competence, and the course did not include a pre-specified intervention, a baseline assessment, or a comparison cohort; the educational implications discussed in Sections 4.5 and 4.6 should therefore be read as hypotheses for future research.

5. Conclusions

This study comparatively examined whether biomimicry and sustainability retain distinct design intentions when visualized through generative AI under separate (B, S) and combined (B + S) conditions in a design-education dataset. Taken together, the findings reveal a consistent pattern within this dataset: nature-oriented visual features were rep-

resented more readily than the functional and performance-related principles associated with biomimicry and sustainability.

The findings across the four research questions illustrate this pattern at different analytical levels. Biomimicry and sustainability outputs showed substantial visual-aesthetic similarity despite their distinct conceptual foundations; this similarity rested on shared natural-looking materials, color palettes, planting, and daylight, whereas organic form remained characteristic of the biomimicry outputs. Biomimetic representation was particularly evident through morphology/pattern transfer and spatial/structural transfer, whereas no project provided visually verifiable evidence of a biological function operating as an environmental or performance-related mechanism. Sustainability was specified across multiple dimensions in the prompts, and 25 of 95 explicit strategy requests were visually verified, resulting in an overall prompt–output correspondence rate of 26.3%. This rate reflects mainly the limits of visual verification rather than unsuccessful image generation: correspondence reached 66.7% for strategies with depictable physical components, such as daylighting, renewable energy, and water management, but was 7.0% for partially assessable strategies and 0% for strategies requiring specification or performance evidence. In the combined B + S condition, the outputs remained at the additive or coordinated level, with the boundary between the two depending on the coder, and no project reached synergistic integration, in which the biomimetic principle would be identifiable as the mechanism through which a sustainability strategy operates. Within the limits of this descriptive, course-based image analysis, these findings indicate that visual plausibility should not be interpreted as evidence of environmental performance and that visual-aesthetic similarity should remain conceptually distinct from functional or performance-based equivalence. For research on generative AI in design, these results imply that evaluations of AI-generated images should state whether they concern visual plausibility or functional and performance-related content, since studies that evaluate different levels may otherwise reach apparently contradictory conclusions (Section 4). For design practice, they imply that sustainability claims based on AI-generated images require supporting material, technical, or performance information.

For sustainable and biomimetic interior design education, the findings identify issues that deserve attention rather than demonstrating the effectiveness of any instructional strategy, since learning outcomes were not assessed. Generative AI can support ideation and visual exploration during early design stages, but AI-generated images should not be treated as direct evidence of sustainability performance or functional biomimetic integration. Treating prompt writing as a structured form of design literacy, in which students distinguish visually represented features from attributes that require material, technical, or performance-based verification, is therefore proposed as a hypothesis for future instructional studies. Beyond the research questions, the findings point to two broader problems discussed in Section 4.6: the risk that sustainability is taught and assessed through its image rather than its processes, and the reduction in biomimicry to the copying of biological forms rather than learning from biological processes. In AI-generated images, both tendencies converge on a biophilic appearance, which has value in its own right but is distinct from the aims of both biomimicry and sustainability. Methodologically, the separation of prompt (P-) and visual-output (V-) layers provides a transparent means of distinguishing between design intentions explicitly specified in prompts and their visually verifiable representation in generated outputs. Future research could extend this approach by integrating material, energy, life-cycle, and other performance data with AI-generated visualizations, systematically comparing different generative AI models, and examining whether more structured prompt strategies improve functional biomimetic representation and higher levels of biomimicry–sustainability integration, for example, by evaluating the assessment protocol proposed in Section 4.5 with baseline and comparison groups.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/su182111364/s1, Section S1: Codebook; Section S2: Project-level prompts and AI-generated visual outputs; Section S3: Project-level generation context and coding profiles; Table S1: Generation context of the projects; Table S2: Project-level coding profiles; Table S3: Project-level coding profiles of the independent coder; Table S4: Agreement between the primary and the independent coder; Table S5: Prompt–output matrix for the projects whose B prompts requested a biological function (P-B4 = 1); Table S6: Project-by-condition inventory of prompts and final images.

Author Contributions: Conceptualization, T.K. and A.K.; methodology, T.K. and A.K.; validation, T.K. and A.K.; formal analysis, A.K.; investigation, T.K.; resources, A.K.; data curation, T.K.; writing—original draft preparation, A.K.; writing—review and editing, T.K.; visualization, T.K. and A.K.; supervision, T.K.; project administration, A.K.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study by the Science and Engineering Sciences Scientific Research and Publication Ethics Board of Kütahya Dumlupınar University (decision date: 28 September 2026; decision no.: 2026/09), because the study involved the retrospective analysis of design outputs produced in a completed course and did not involve surveys, interviews, experiments, or any direct intervention with human subjects. The research is based on the design outputs produced by students as part of a structured elective course, the pedagogical framework of which is described in Section 2.2 of the manuscript.

Informed Consent Statement: Written informed consent was obtained individually from all students whose work is included in the study. Consent was collected by the first author after the final course grades had been announced, and students were informed that participation was voluntary and that declining would have no effect on their grades. The consent covered the publication of the students' prompts and AI-generated images in open access format under a CC BY 4.0 license. The projects are identified only by anonymized codes (P01–P22), and no student names are reported in the article or the Supplementary Material.

Data Availability Statement: The data supporting the findings of this study are provided in the article and the Supplementary Materials. Further inquiries can be directed to the corresponding author.

Acknowledgments: During the preparation of this work, the authors used ChatGPT-5.5 (OpenAI, San Francisco, CA, USA) for language editing and grammar improvement and Claude (Anthropic, San Francisco, CA, USA) for language editing. After using these tools, the authors carefully reviewed and revised the manuscript and take full responsibility for the content of the published article. The AI-generated images analyzed in this study were produced by the students with ChatGPT (GPT-5; OpenAI, San Francisco, CA, USA), Gemini (Gemini 3 Pro; Google, Mountain View, CA, USA), and Nano Banana Pro (Google, Mountain View, CA, USA), as described in Section 2.2. Also, a heartfelt thank you to the students, without whom this research would not have been possible.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ibrahim, I.; Talib, M.A.; Ammar, A.; Aoul, K.A.T.; Abuimara, T. Comparative and experimental analysis of leading text-to-image generative artificial intelligence models for regional residential architectural designs. *Results Eng.* **2025**, *29*, 108835. [CrossRef]
2. Çiftci, S.K.; Erten-Bilgiç, D.; Çetin-Yerlikaya, N. An analysis of seating components created with text-to-image generating ai tools: A case study of MicrosoftDesigner, Clipdrop and Midjourney. *J. ESOGU Eng. Arch. Fac.* **2026**, *34*, 2078–2097. [CrossRef]
3. Güneş, E.; Ünal, B.; Hatunoğlu, D.C. Investigating emotional conveyance in AI-generated interior design: A user perception experiment. *J. Fac. Eng. Archit. Gazi Univ.* **2026**, *41*, 625–638. [CrossRef]
4. Grabska, E.J.; Wen, B. The role of artificial intelligence in architecture and interior design. *Buildings* **2026**, *16*, 2106. [CrossRef]
5. Alamasi, R.; Asfour, O.S. Applications of generative AI in architectural design education: A systematic review and future insights. *Digital* **2026**, *6*, 6. [CrossRef]
6. Huh, M.B.; Miri, M.; Tracy, T. Students' perceptions of generative aı image tools in design education: Insights from architectural education. *Educ. Sci.* **2025**, *15*, 1160. [CrossRef]

7. Kee, T.; Kuys, B.; King, R. Generative artificial intelligence to enhance architecture education to develop digital literacy and holistic competency. *J. Artif. Intell. Archit.* **2024**, *3*, 24–41. [CrossRef]
8. Kwon, J.; Jung, E.C.; Kim, J. Designer-generative AI ideation process: Generating images aligned with designer intent in early-stage concept exploration in product design. *Arch. Des. Res.* **2024**, *37*, 7–23. [CrossRef]
9. Gerola, A.; Robaey, Z.; Blok, V. What does it mean to mimic nature? A typology for biomimetic design. *Philos. Technol.* **2023**, *36*, 65. [CrossRef]
10. Ilieva, L.; Ursano, I.; Traista, L.; Hoffmann, B.; Dahy, H. Biomimicry as a sustainable design methodology—Introducing the 'biomimicry for sustainability' framework. *Biomimetics* **2022**, *7*, 37. [CrossRef]
11. Benyus, J. *Biomimicry—Innovation Inspired by Nature*; Harper Collins Publishers: New York, NY, USA, 1997.
12. El-Zeiny, R.M.A. Biomimicry as a problem solving methodology in interior architecture. *Procedia-Soc. Behav. Sci.* **2012**, *50*, 502–512. [CrossRef]
13. Kandemir, A.; Kalay, T. A Multi-level approach to biomimetic design education: Developing a biomimetic transfer framework and matrix for design analysis. *Biomimetics* **2026**, *11*, 445. [CrossRef]
14. Volstad, N.L.; Boks, C. On the use of biomimicry as a useful tool for the industrial designer. *Sustain. Dev.* **2012**, *20*, 189–199. [CrossRef]
15. Alfuraty, A.B. Sustainable environment in interior design: Design by choosing sustainable materials. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *881*, 012035. [CrossRef]
16. Celadyn, M. Environmental activation of inner space components in sustainable interior design. *Sustainability* **2018**, *10*, 1945. [CrossRef]
17. Celadyn, M. The use of visuals in environmentally responsible interior design. *World Trans. Eng. Technol. Educ.* **2018**, *16*, 212–217.
18. Bulhaz, Ü.Ç. Sürdürülebilirlik kavramının iç mimarlık açısından değerlendirmesi. *Online J. Art Des.* **2020**, *8*, 272–285.
19. Güneş, S.; Demirarslan, D. Sustainability and environmental approaches in furniture design. *Int. J. Humanit. Art Res.* **2020**, *5*, 81–99.
20. Rashdan, W.; Ashour, A.F. Criteria for Sustainable Interior Design Solutions. In Proceedings of the 12th International Conference on Urban Regeneration and Sustainability (Sustainable City), Seville, Spain, 18–20 September 2017; pp. 311–322.
21. Winchip, S.M. *Sustainable Design for Interior Environments*; Fairchild: New York, NY, USA, 2011.
22. Lee, Y.S. Sustainable design re-examined: Integrated approach to knowledge creation for sustainable interior design. *Int. J. Art Des. Educ.* **2014**, *33*, 157–174. [CrossRef]
23. Othmani, N.I.; Mohd Yunos, M.Y.; Ramlee, N.; Abdul Hamid, N.H.; Mohamed, S.A.; Yeo, L.B. Biomimicry levels as design inspiration in design. *Int. J. Acad. Res. Bus. Soc. Sci.* **2022**, *12*, 1094–1107. [CrossRef]
24. Kennedy, E.; Fecheyr-Lippens, D.; Hsiung, B.K.; Niewiarowski, P.H.; Kolodziej, M. Biomimicry: A path to sustainable innovation. *Des. Issues* **2015**, *31*, 66–73. [CrossRef]
25. Raman, R.; Sreenivasan, A.; Suresh, M.; Nedungadi, P. Mapping biomimicry research to sustainable development goals. *Sci. Rep.* **2024**, *14*, 18613. [CrossRef]
26. Ibrahim, I.; Nasreldin, R. Investigating the integration of biomimicry and eco-materials in sustainable interior design education. *Architecture* **2025**, *5*, 39. [CrossRef]
27. Daugelaite, A.; Dogan, H.; Grazuleviciute, I. Characterizing sustainability aesthetics of buildings and environment: Methodological frame and pilot application to the hybrid environments. *Landsc. Archit. Land* **2021**, *19*, 61–72. [CrossRef]
28. Declos, A. AI and aesthetic bias. *J. Aesthet. Art Crit.* **2026**, *84*, 218–233. [CrossRef]
29. Park, S.Y. Architectural design in the age of generative AI: Toward a balance between efficiency and responsibility. *Archit. Res.* **2026**, *28*, 4. [CrossRef]
30. Varghese, J.K.; Rani, P. Creative signatures and systemic bias: A critical analysis of platform affordances in generative diffusion models. *Convergence* **2026**, 1–21. [CrossRef]
31. İnan, N. The potentials of generative artificial intelligence tools in biomimetic architectural design processes: An analysis through the example of DALL-E 3. *Int. J. 3D Print. Technol. Digit. Ind.* **2026**, *10*, 266–283. [CrossRef]
32. Kalay, T.; Söğüt, M.A.; Kandemir, A. The bridge between nature and technology: The impact of artificial intelligence on furniture design. *Res. Mobilis Rev. Int. Investig. Mobil. Objetos Decor.* **2026**, *14*, 110–132. [CrossRef]
33. Gupta, R.; Kyaw, A.H. Insights informed generative AI for design: Incorporating real-world data for text-to-image output. *arXiv* **2025**, arXiv:2506.15008. [CrossRef]
34. AlShkipi, O.A.; Zahran, B. Implementation of artificial intelligence in interior design: Systematic literature review. *Civ. Eng. Archit.* **2024**, *12*, 2889–2906. [CrossRef]
35. Çelik, T. The effect of artificial intelligence on creativity in conceptual design in architectural education: The motion of biomimetics and futurism. *AI Soc.* **2025**, *40*, 2717–2727. [CrossRef]
36. Güler, Ş.; Selçuk, S.A. AI-supported biomorphic design pedagogy. In Proceedings of the 13th International Conference of the Arab Society for Computation in Architecture, Art and Design (ASCAAD), Riyadh, Saudi Arabia, 21–23 October 2025; pp. 551–572.

37. Stefańska, A.; Kurcjusz, M. From nature to neutral networks: AI-driven biomimetic optimization in architectural design and fabrication. *Sustainability* **2025**, *17*, 11333. [CrossRef]
38. Gower, J.C. A general coefficient of similarity and some of its properties. *Biometrics* **1971**, *27*, 857–871. [CrossRef]
39. Gwet, K.L. Computing inter-rater reliability and its variance in the presence of high agreement. *Br. J. Math. Stat. Psychol.* **2008**, *61*, 29–48. [CrossRef]
40. Wilson, E. *Biophilia: The Human Bond with Other Species*; Harvard University Press: Cambridge, MA, USA, 1984.
41. Kellert, S.R.; Heerwagen, J.; Mador, M. *Biophilic Design: The Theory, Science and Practice of Bringing Buildings to Life*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
42. Kellert, S.; Calabrese, E. The Practice of Biophilic Design. 2015. Available online: https://www.biophilic-design.com (accessed on 1 November 2026).
43. Joye, Y. Architectural lessons from environmental psychology: The case of biophilic architecture. *Rev. Gen. Psychol.* **2007**, *11*, 305–328. [CrossRef]
44. Mehaffy, M.W.; Salingaros, N.A. *Design for a Living Planet: Settlement, Science, and the Human Future*; Sustasis Foundation: Portland, OR, USA, 2015.
45. Boarin, P.; Martinez-Molina, A.; Juan-Ferruses, I. Understanding students' perception of sustainability in architecture education: A comparison among universities in three different continents. *J. Clean. Prod.* **2020**, *248*, 119237. [CrossRef]
46. Guy, S.; Farmer, G. Reinterpreting Sustainable Architecture: The Place of Technology. *J. Archit. Educ.* **2001**, *54*, 140–148. [CrossRef]
47. Pedersen Zari, M. Biomimetic approaches to architectural design for increased sustainability. In Proceedings of the Sustainable Building Conference (SB07), Auckland, New Zealand, 14–16 November 2007; Paper 033.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.