

Beyond Convergent: Reframing Creativity Support Tools through LLM-assisted Decision-Making in Sustainable Contexts.

Ana Rodrigues^{1,2}, David dos Santos^{1,3}, Diogo Cabral^{1,4}, Pedro Campos^{1,2,5},

¹ ITI/LARSyS, Portugal

² University of Madeira, Funchal, Portugal

³ Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal

⁴ ISEL – Polytechnic University of Lisbon, Lisbon, Portugal

⁵ WowSystems, Lda., Funchal, Portugal

2066213@student.uma.pt, david.r.dos.santos@tecnico.ulisboa.pt

diogo.cabral@isel.pt, pedro.campos@iti.larsys.pt

Abstract. Creativity support tools (CSTs) play a fundamental role in supporting creativity in human-computer interaction (HCI). The rapid development of artificial intelligence (AI), particularly Large Language Models (LLMs), has significantly transformed human-computer interactions and, consequently, CSTs. However, little attention has been paid to how LLM-based CSTs reconfigure convergent thinking, decision-making, and agency dynamics throughout the design process, especially in sustainability-oriented contexts. We present a theory-driven conceptual synthesis to analyze how LLM-based CSTs transform AI-assisted convergent thinking. Based on existing research on these dynamics, we propose a conceptual reconfiguration of individual evaluation towards mediated decision-agency. This conceptual framework culminates in five theoretical implications for the design of LLM-based CSTs, proposing principles to support mediated decision-agency, promote situated transparency, and reinforce human autonomy in sustainability-oriented convergence processes.

Keywords: Convergent Thinking, Creativity Support Tools, Sustainability, Human-AI Co-Creativity, AI-assisted Decision-Making, Trust and transparency.

1 Introduction

With the excessive use of natural resources, sustainability has become a central concern for society [1]. It is considered an interdisciplinary domain marked by variability and conflicting values, fitting the definition of "wicked problems" [2] – problems that cannot be solved through traditional analytical methods. This

context requires innovative solutions and careful evaluation of competing criteria, with creativity, as recent research [3] suggests, seen as a problem-solving paradigm for achieving sustainability goals. In this context, creativity is framed as a cognitive process that includes the generation, evaluation, and selection of ideas [4, 5], allowing the reformulation of problems, the exploration of application spaces (i.e., design space) [6], and the progressive narrowing of alternatives through evaluative convergence. Addressing these sustainability challenges requires design processes capable of exploring, evaluating, and refining competing alternatives [7]. However, due to the complexity of the process and the design space as a whole, the need for structured support throughout the design process has arisen. Given this scenario, technology emerges as a central component of the creative process, particularly CSTs [8], which expand human capacity throughout the process. Recent advances in AI are fundamentally reshaping CSTs' capabilities by enabling new forms of Human-AI co-creativity, in which AI actively contributes to both idea generation and the evaluation and selection of alternatives [9-11]. While these capabilities expand creative support, they also raise fundamental questions regarding transparency, trust, human control, and the perception of AI systems as black boxes [12]. These challenges become particularly critical in sustainability-oriented design, where decisions involve competing values, multiple evaluation criteria, and long-term consequences. This paradigm shift requires a deeper understanding of how LLM-based CSTs mediated convergent thinking and decision-making. In these contexts, LLM-based CSTs may influence not only the generation of alternatives but also the structure and prioritization of evaluative criteria during convergence. Despite the growing adoption of LLM-based CSTs, existing research has largely focused on idea generation, leaving the role of LLMs in supporting convergent thinking and decision-making comparatively underexplored. The emergence of LLMs fundamentally reconfigures this process by shifting convergent thinking from an exclusively individual activity toward mediated decision-agency, in which evaluation and selection emerge through the interaction between human judgment and AI support. To address this transformation, this paper offers three main contributions: (1) a structured relational mapping of the conceptual correlations underlying CSTs, examining how control, transparency, and trust develop into dynamic tensions; (2) a reconceptualization of decision-making as a transformative mediating parameter that shapes agency dynamics within the human-AI co-creative process; and (3) five design-oriented implications for LLM-based CSTs to promote human-centered autonomy in sustainable contexts.

2 Background and Related Work

In light of the investigation's goal, we first reviewed the underlying theory and prior work on the impact of creativity on sustainability. We then examined the positioning of creativity as a process within the design process, next explored how computational tools (i.e., CSTs) support the design process, and finally

analyzed the evolution of these tools in light of new technologies (e.g., GenAI/LLMs).

2.1 Sustainability-oriented Creativity

Sustainability is broadly defined as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs,” and is interpreted as a multifaceted challenge involving the interaction of three pillars - environmental impact, economic benefits, and social effects [3, 13]. It is an interdisciplinary domain characterized by variability and conflicting values, which fits the definition of a wicked problem [2].

Within design research, sustainability has established itself as a distinct field through Design for Sustainability (DfS), which extends beyond the development of environmentally enhanced products to encompass product-service systems, organizational innovation, and broader socio-technical transitions [14]. This evolution aligns closely with Transitional Design, which frames sustainability challenges as long-term systemic transformations that require designers to address interconnected social, environmental, and economic concerns at multiple scales [15]. Similarly, perspectives from the Circular Economy reinforce lifecycle thinking and the circularity of resources, demanding that design decisions consider impacts that extend beyond individual artifacts and situate them within broader, interconnected systems [16]. Consequently, design focused on sustainability is characterized not only by the need to generate innovative solutions, but also by the need to evaluate alternatives and balance multiple, often conflicting, criteria across environmental, economic, and social dimensions.

Given this scenario, sustainability challenges cannot be addressed through traditional analytical methods alone. In this context, creativity (i.e., defined as the manifestation of new and valuable ideas—valuable in their context of application [4, 5]) plays a crucial role in addressing sustainability challenges, involving two distinct and interrelated cognitive processes [4]: divergent thinking that supports the generation of ideas and convergent thinking that supports the evaluation and selection of these ideas. However, research has focused on the divergent phase rather than the convergent phase [5, 17]. In sustainability-oriented creative processes, the convergent phase is particularly critical because designers must compare alternatives while balancing multiple, often competing, environmental, economic, and social criteria [18, 19]. Instead of seeking a single optimal solution, sustainability-oriented decision-making requires negotiating trade-offs among conflicting values throughout the creative workflow and relies on structured evaluation mechanisms that support the systematic comparison and justification of design alternatives. Recent research has explored different forms of computational support for this purpose. For example, Rodrigues et al. [20] proposed integrating sustainability-oriented evaluation metrics directly into the idea selection process, while Corujeira et al. [21] developed the interactive tool FreeCAD-LCA, which incorporates Life Cycle Assessment (LCA) methods into a CAD environment to support environmental evaluation during

the early stages of product development. More recently, Preuss et al. [22] explored the potential of LLMs to support data structuring and decision-making tasks within LCA workflows, while highlighting important challenges related to data quality, transparency, and potential bias. Despite growing availability of computational support for sustainability assessment, decision-making remains characterized by conflicting values, uncertainty, and a lack of transparency, underscoring the need for mechanisms that support human judgment alongside computational capabilities—a perspective this article seeks to develop.

2.2 Creativity as a Process in Design

In this investigation, design is understood as a cognitive and iterative process in which problems and solutions co-evolve through cycles of construction, reformulation, and evaluation under conditions of uncertainty [23]. Design is therefore not limited to the application of search or optimization techniques to defined problems. Instead, it involves continuous reformulation of both design goals and the means available to achieve them, making creativity a fundamental dimension of the design process [6, 23]. Creativity extends beyond idea generation, encompassing the continuous negotiation between novelty, appropriateness, and the evolving understanding of the design problem [24]. Accordingly, contemporary perspectives describe creativity as a co-evolution between the problem and solution spaces, unfolding through iterative cycles that progressively shape both the problem framing and the direction of emerging solutions [13,23,25]. This process-oriented perspective suggests that creativity should be understood and supported as a continuous process throughout the design lifecycle. Building on this perspective, creativity support should extend across both the problem and solution spaces, encompassing not only ideation but also the evaluative activities that characterize convergent thinking [10, 26]. Consequently, the need for support mechanisms that can accompany the continuous evolution of both problems and solutions throughout the design process becomes evident.

2.3 Creativity Support Tools (CSTs) in the Design Process

To address the need for continuous support throughout the design process, CSTs have been developed to assist creative work while preserving user control, transparency, and iterative exploration across the design space. Originally introduced by Shneiderman [8], CSTs are currently defined as “tools run on one or more digital systems, fulfill one or more creativity-focused capabilities, and are employed to positively influence users of varying experience in one or more distinct phases of the creative process” [27]. The CSTs ecosystem has been systematically mapped in recent years, synthesizing research published between 1999 and 2018 [27]. The studies highlight the diversity of existing approaches and the different forms of support offered throughout the creative process, particularly for ideation, brainstorming, and the early exploration of alternatives.

More recent research [17] has sought to address this gap by proposing design recommendations that incorporate contextual information, decision criteria, and process evolution to support creativity. Collectively, these studies indicate that CSTs can progressively evolve a more holistic approach to supporting creativity in the design process.

2.4 LLM-based CSTs and Human-AI Co-Creativity

Traditional CSTs focus on divergent thinking, offering limited support for the evaluation and selection of ideas [10, 17]. In this context, AI has become a major tool promising deeper support [28]. Unlike traditional CST approaches [8], LLM-based CSTs can engage in evaluative, interpretive, and reflective activities, acting as co-creative partners in the design process [11]. Recent work has explored this paradigm [28-30], investigated how an LLM-based CST supports the idea evaluation and selection process in different scenarios and levels, and the system's reasoning, perceptions of opacity, and tensions related to users' trust in the ideas generated by the system. However, they reveal that the tool does not always provide results that are explicitly legible or controllable for the user. Glassberg et al. [31] also demonstrated that perceived transparency depends not only on access to information but also on how the system's reasoning is shared. In this sense, LLM-based CSTs, while expanding support for the creative process, have raised fundamental challenges related to process transparency, human control over the system, and human trust in decision-making. These challenges reinforce the perception of co-creative systems as a "black box" in the design process, highlighting the need to address them through models that are more explainable and auditable for users. However, these issues go beyond the interface or technical explainability. They reflect a deeper transformation in how users interact with LLM-based CSTs, making particularly evident the role of technology in mediating evaluation, justification, and decision-making processes during the convergent phase of the creative process. This interpretation finds support in various theoretical traditions that conceive of human activity as a relational process among individuals, artifacts, and contexts. Among them are Distributed Cognition [32], which presents cognition as distributed between humans and artifacts; Mediated Action [33], which emphasizes the mediating role of tools in human action; and Sociotechnical Systems Theory [34], which analyzes the interdependence between social and technological systems. Together, these perspectives offer a framework for understanding that technology does not merely play an instrumental role but actively shapes human activity. In our research, we adopt these perspectives as a framework for understanding how LLM-based CSTs mediate the evaluation processes and decision justification during the convergent phase of creativity. It is in this context that we propose the concept of mediated decision-agency, which we develop in the following sections.

3 Relational Mapping of LLM-based CSTs

In this section, we provide a conceptual and relational analysis that structures the creative process driven by LLMs, in which tensions over control, transparency, and trust give rise to the black box perception observed in LLM-based CSTs.

3.1 Analysis Approach

The objective of this article was to develop a relational framework for how LLM-based CSTs transform convergence, the agency dynamics of the creative process, and the reconfiguration of user control, transparency, and trust during AI-assisted design. To this end, we adopted a theory-driven conceptual synthesis approach, informed by a review of the relevant literature rather than an exhaustive systematic review [35]. Consistent with this approach, the purpose of the review was not to identify all available studies but to construct a purposively selected conceptual corpus to support the development of the proposed framework.

The literature review followed an iterative process throughout the manuscript's development. Relevant publications were iteratively identified through searches in widely recognized databases, including ACM Digital Library, IEEE Xplore, Scopus, ScienceDirect, and Google Scholar, complemented by backward citation tracking of influential publications. Searches combined terms such as co-creative human-AI, convergent thinking, creativity, CSTs, LLM-based decision-making, sustainability-oriented creativity, and black box perception.

The literature was selected according to its conceptual relevance to four interrelated themes: (1) creativity and the creative process; (2) LLM-based Creativity Support Tools; (3) AI-assisted idea evaluation and selection; and (4) human-AI interaction mechanisms related to control, transparency, and trust. The final conceptual corpus comprised 41 publications, which were comparatively analyzed to identify recurring concepts, relationships, and tensions underpinning the proposed framework. Additional references were incorporated to contextualize the research, support the methodological approach, and further discuss the implications of the proposed conceptual framework. Rather than producing a quantitative synthesis, the selected literature was examined through comparative and interpretative analyses to identify recurring concepts, relationships, and tensions across the reviewed studies. Recurring concepts identified in the literature were then iteratively grouped by conceptual affinity into broader analytical dimensions. The analytical relationships identified through the conceptual synthesis (Table 1) are interpreted through four analytical layers. Their conceptual correlations are subsequently synthesized in Table 2 and integrated into the conceptual framework presented in Figure 1.

Table 1. Development of the analytical labels through the conceptual synthesis.

Analytical Relationship	Recurring concepts identified in the literature	Representative references	Analytical Label
Sustainability - Creativity	Wicked problems; conflicting values; long-term impact; multi-criteria evaluation; sustainability trade-offs	[2], [3], [22], [36]	Contextual dependency
Creativity - Creative Process	Problem framing; idea generation; iterative development; reformulation; co-evaluation;	[6], [23], [24], [37]	Processual grounding
Creative Process - Decision-making	Evaluation; idea selection; convergence; trade-offs; judgment; progressive refinement	[17], [20], [26], [29], [38]	Progressive convergence
Decision-Making - Control	Human oversight; shared agency; autonomy; intervention; human-centred AI	[12], [39], [40], [41]	Agency tension
Decision-Making - Transparency	Explainability; reasoning visibility; justification; interpretability; explainable AI	[12], [31], [41], [42], [43]	Interpretability requirement
Decision-Making - Trust	Trust calibration; reliance; confidence; responsibility; acceptance of AI; recommendations	[39], [40], [41], [44]	Reliance formation
Control - Black Box Perception	Reduce oversight; automation; loss of recommendations	[45], [46], [47]	Inverse relationship
Transparency - Black Box Perception	Explainability; legibility; reasoning disclosure; user understanding	[12], [31], [45], [47]	Mediating relationship
Trust - Black Box Perception	Over-reliance; uncertainty; confidence despite opacity; trust calibration	[39], [45], [46], [47]	Ambivalent relationship
Sustainability - Decision-Making	Long-term consequences; competing values; ethical responsibility; environmental impact; multi-criteria judgment; trade-offs	[2], [3], [20], [22], [29], [36]	Normative amplifications

The resulting relational mapping synthesizes the conceptual architecture of mediated decision-agency in LLM-assisted convergent thinking, forming the basis of the discussion presented in Section 4. The identified relationships should not be read as independent analytical categories. Instead, they function as interdependent explanatory mechanisms that show how LLM-based CSTs progressively reshape convergent decision-making. Taken together, these

mechanisms explain how control, transparency, trust, and various forms of opacity interact to mediate decision-agency in sustainability-oriented design.

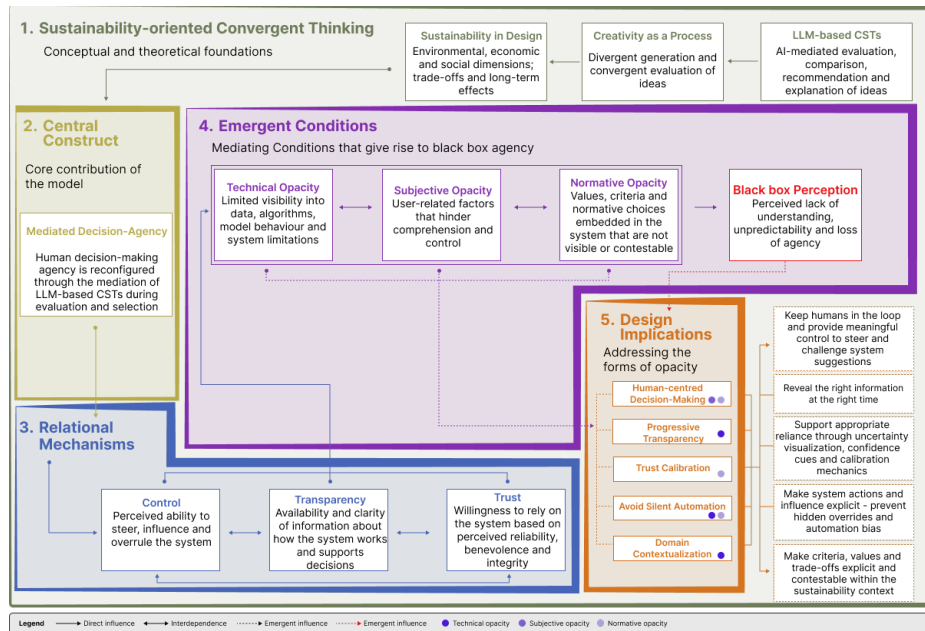


Fig. 1. Conceptual model, organized into five conceptual levels: (1) the context of convergent thinking oriented towards sustainability, (2) the central construct of mediated decision-making agency, (3) the relational mechanisms (control, transparency and trust), (4) the forms of opacity that emerge during LLM-assisted decision-making (technical, subjective and normative opacity), and (5) the design implications that respond to each of these forms of opacity.

3.2 Results and Synthesis

We report findings from a four-layer analysis examining transparency, control, trust, and perceptions of black boxes in relation to CST-based LLMs. Each layer addresses the emergence of one dimension, and their interactions are visually represented in Table 2. We discuss how these dimensions conceptually correlate to characterize the mediated agency in LLM-assisted decision-making.

Table 2. Conceptual correlations underlying LLM-based CSTs.

Conceptual Origin	Relationship with	Conceptual Correlation	Analytical Interpretation
Sustainability	Creativity	Contextual	The complexity, ambiguity, and

		dependency	conflict of values surrounding sustainability shape the type of creativity required [3, 36].
Creativity	Creative Process	Processual grounding	Creativity manifests itself through an iterative process, not in a single moment [5, 20].
Creative Process	Decision-Making	Progressive convergence	The creative process inevitably leads to moments of decision and compromise [20, 38].
Decision-Making	Control	Agency tension	The system's involvement in evaluation and selection affects the perception of human control [29].
Decision-Making	Transparency	Interpretability requirement	Decisions require an understanding of the underlying criteria and reasoning [17, 20].
Decision-Making	Trust	Reliance formation	Trust emerges when the user accepts or delegates judgment to the system [39].
Control	Perceived black box	Inverse relationship	Perceived lack of control intensifies the perception of opacity [45].
Transparency	Perceived black box	Mediating relationship	Transparency acts as a mediator of the perception of a black box [31, 46].
Trust	Perceived black box	Ambivalent relationship	High trust can mask opacity, reducing questioning [31, 46].
Sustainability	Decision-Making	Normative amplifications	Sustainability decisions are normatively amplified due to their ethical implications and long-term temporal impact [36].

Framing Correlations. Framing what counts as a problem, what counts as a solution, and how sustainability is assessed [6]. This relationship operates at the contextual and normative level, considering sustainability as complex, with conflicting values and a systemic nature [3]. In this sense, creativity assumes a central role in improving correlation, reframing sustainability problems through informed decision-making throughout the process [20]. According to Bollen et al. [36], sustainable decisions are influenced by factors (e.g., individual attitudes and attention mechanisms), which highlights the importance of information. In this context, a layered creative framework serves as the conceptual basis for analyzing AI-based creativity-support tools and the tensions surrounding trust, transparency, and black box perception.

Internal Dynamics Correlations. Internal dynamic correlations describe the procedural grounding throughout the design process, working towards progressive convergence through the structure of the creative process. In light of design, creativity is framed as a situated practice, composed of iterative phases

[37]. It is described not only as moments of idea generation, but also as moments of evaluation and respective selection of these ideas [26]. Widely adopted models in these practices, such as the Double Diamond [48, 49], illustrate the divergence-convergence dynamic throughout the design process and highlight the gradual progression. In this sense, creativity is structured by processes, decisions, and constraints, without being reduced to a merely internal cognitive phenomenon [38]. Recent research on creativity support tools [26] has demonstrated how these tools support both the generation and the evaluation and selection of ideas. Demonstrating that as the creative process progresses, the iterative cycles progressively narrow the space of possibilities, leading to moments of choice – reinforcing the understanding of the cognitive process as a case of procedural grounding.

AI-assisted Decisions Correlations. These correlations associated with AI-assisted decision-making describe the transition of the creative process as a socio-technical space, where decisions remain human but are mediated by structured information from the computational system. This mediation not only informs human judgment but also reorganizes the conditions under which judgment is exercised. By structuring evaluative criteria, LLM-based systems help shape decisional agency. Convergent thinking, traditionally understood as an internal cognitive narrowing of alternatives, is now partially structured through computational mediation. Normally, the decision is interpreted as the final phase of the creative process, resting solely on personal opinion [50]. However, with the introduction of Generative Artificial Intelligence (GenAI), this phase has become more autonomous and critical, with visible tensions emerging related to control, transparency, and trust [12, 25, 46]. Given the research context – sustainability – it is necessary to develop structured support for decision-making, as both the evaluative and selective processes involve multiple criteria. In this sense, the introduction of contextualized metrics transforms creative decision-making into a process mediated by structured information [20]. In recent years, AI-assisted decision-making has become a driving force, where decisions have real impacts on people's lives [40]. It is precisely because AI enters the decision-making phase that the tension of agency emerges. This situation became evident in AI-based creativity support tools [29, 51, 52, 53], where the system intervenes in the convergent phase, traditionally associated with human judgment. Transparency emerges as a fundamental requirement. Rodrigues et al. [29] emphasize that acceptance of the system's recommendations depends on users' ability to understand the criteria underlying them. However, according to [41], transparency is not limited to the technical exposition of AI's internal workings, but rather is achieved through a design element in the interface, situated at the moment of decision [42]. It is from this causal relationship between perceived control and situated transparency that human trust in the system's recommendations arises. In AI-assisted decision-making scenarios, trust results from the perceived consistency and alignment with the initially defined objectives [47]. In addition to concerns regarding transparency and trust, this analysis indicates that decision-making

serves as a locus for the articulation of authority. Although formal decisional power may reside with the human user, the organization of evaluative criteria, comparisons, and rationales establishes a mediated layer that partially configures epistemic and justificatory authority. This insight provides a foundation for re-examining the transformation of convergence in LLM-based computer-supported technologies (CSTs).

Emergent and Non-linear Correlations. These correlations result from the dynamic interplay among control over outcomes, process transparency, and trust in AI-assisted decision-making, shaping users' perceptions of LLM-based CSTs and contributing to black box perceptions [46]. Although the black box is often treated as a single challenge, the reviewed literature suggests that different forms of opacity underlie it, requiring distinct design responses. We distinguish three complementary forms: technical opacity, relating to the limited visibility of the system's internal reasoning processes, which reinforces the need for explainability and transparency [12, 41]; subjective opacity, relating to users' ability to interpret and trust AI recommendations, influencing trust calibration and perceived understanding [31, 39, 40]; and normative opacity, referring to the values, assumptions, and commitments implicit in AI-assisted decision-making, especially in sustainability contexts where environmental, social, and economic criteria coexist [2, 3, 22, 36]. This distinction underlies the interpretation of the following relationships. The relationship between control and black box perception is inverse: the less human control, the greater the perception of opacity. Systems that generate autonomous recommendations, without allowing intervention or critical evaluation, limit users' ability to understand and justify the decision-making process, reinforcing this perception [45, 46]. Transparency mediates perceptions of black boxes, making AI reasoning more accessible and interpretable. However, it should not be understood merely as the exposure of technical information, but as the ability to allow users to understand the criteria underlying the recommendations in the context of the design task [12, 47]. Thus, explainability reduces technical opacity and favors more informed interaction. The relationship between trust and black box perception is more complex: trust does not increase simply because opacity decreases, nor does it necessarily eliminate it. Trust emerges before repeated interaction, users' understanding of the system, and the perceived legitimacy of recommendations [44, 54]. Therefore, subjective opacity depends not only on access to explanations but also on users' ability to properly interpret and calibrate this trust. These relationships indicate that black box perception is an emergent phenomenon, resulting from the interaction among the technical, subjective, and normative dimensions of opacity—a multidimensional perspective that underlies the design implications presented in Section 4.

4 Discussion

In the context of this investigation, sustainability-oriented creativity, the mechanisms underlying the perception of the black box – control, transparency, and trust – are reconfigured in an interdependent, situational manner, taking into account AI's role in the convergent phase of the creative process. Specifically, we seek to frame the impacts and effects of these relationships as design requirements for developing new LLM-based CSTs. Much of the current discussion on AI-assisted creativity focuses on idea generation and expansion. However, our analysis suggests that the potentially most significant shift may occur during convergence. It is at the moment when alternatives are evaluated, justified, and selected that responsibility becomes concrete. In sustainability-oriented contexts, this phase carries particular weight, as decisions often involve long-term environmental, economic, social, and ethical consequences. When LLM-based CSTs operate at this stage, they do not simply add information; they also help shape how commitments are formed and defended.

4.1 Reframing Control, Transparency, and Trust

The analysis suggests the need to reformulate the classical notions of mechanisms in the context of LLM-based CSTs. These mechanisms are understood as relational rather than static, emerging as dynamic relationships situated particularly in the convergent phase of the creative process [12]. In this sense, it is fundamental that LLMs and the applications that use them (e.g., CSTs) are developed and implemented responsibly. The proposed reframing suggests that offering control is different from fully controlling the process, process transparency is different from opening the black box, and trust is different from understanding everything. These mechanisms are then understood as attributes of Human-AI collaboration, depending on the context of use and the moment at which the decision is made [43]. In this framework between control, transparency, and trust, opacity emerges as a design requirement. The design of the new CSTs must enable them to manage their levels to respond to deliberate decisions about what is made visible, when this information is presented, and how it supports decision-making.

4.2 Decision-Making as a Mediating Core

Based on the analysis presented, decision-making can be understood as the phase in which value conflicts become operational, evaluative criteria are articulated, and user responsibility emerges. When LLM-based CSTs are incorporated at this stage of convergent thinking, human-AI co-creativity, turning decision-making into a mediating core that organizes user-system interaction, aligning with Human-Centered AI approaches [55]. This mediating

core positions decision-making as a regulator of tensions, balancing the co-agency that LLM-based CSTs should support.

This reconfiguration signals a transition from delegated agency, in which the tool merely executes commands, to mediated decision-agency, where both the user and the LLM participate in shaping how decisions are evaluated and justified. This aligns with recent frameworks that position prompting not merely as a technical command, but as a form of cognitive mediation in human-AI interaction [56]. However, mediated decision-agency is not inherently benign. It carries the risk of ethical erosion resulting from automation bias, in which the system's structured rationales may lead users to passively accept algorithmic trade-offs. To address this concern, mediated decision-agency should be designed to promote productive friction, encouraging deliberate interactions that prompt users to question the AI's framing. In this way, the critical re-evaluation of ethical boundaries remains an intentional human responsibility rather than an unnoticed consequence of system mediation.

Within this context, decision-making serves as a mediating core, since the system not only assists users but also actively structures the evaluative scenario. By emphasizing specific criteria or offering structured justifications, LLM-based CSTs become co-shapers of the ethical and technical boundaries of design decisions, requiring a new form of critical oversight by the human agent (e.g., if the system assumes excessive responsibility for evaluation, trust can become passive). If, on the contrary, humans hold all the power, the potential of LLMs remains underutilized [25].

This productive friction, therefore essential. As LLM-based CSTs increasingly structure the evaluative landscape, preserving opportunities for critical oversight helps prevent trust from becoming passive while maintaining meaningful human responsibility. In a sustainability-oriented context, where decisions involve competing environmental, economic, and social factors [17], AI-assisted decision-making should be understood as a transformative socio-technical balancing mechanism that actively supports the negotiation of these competing values while preserving human responsibility. These findings suggest that LLM-based CSTs should not be understood merely as recommendation systems, but as normative mediation systems that actively shape how design decisions are interpreted and justified. From this perspective, explainability goes beyond simply presenting model results and becomes a fundamental design requirement to support informed, responsible, and value-conscious decision-making in a sustainability-oriented convergent thinking.

4.3 Implications for the Design of LLM-based CSTs

Based on the conceptual analysis, five design implications are proposed for the development of LLM-based CSTs. These principles derive directly from the tensions identified among perceptions of control, transparency, trust, and black box perception in AI-assisted decision-making. As presented in Section 3.2.4, black box perception emerges from complementary forms of technical,

subjective, and normative opacity. Consequently, the following design implications address different sources of opacity, offering distinct responses to the challenges introduced by LLM-assisted convergent thinking in sustainability-oriented design.

Illustrative Design Scenario. To illustrate how the proposed conceptual framework can guide the design of LLM-based CSTs, we developed an illustrative scenario situated in a sustainability-oriented design context. The scenario demonstrates how mediated decision agency and the proposed design implications can support AI-assisted convergent thinking during idea evaluation and selection. Imagine a product designer responsible for developing a new sustainable packaging solution for a food product. After an initial ideation session, four design alternatives remain under consideration. Instead of generating additional ideas, the LLM-based CST supports the convergent phase of the process, assisting the designer in evaluating alternatives against the pillars of sustainability (environmental impact, economic benefit, and social effects) [22]. Instead of recommending a single “optimal” solution, the system presents the underlying reasoning for each evaluation, highlights the uncertainties associated with life cycle predictions, and exposes the trade-offs between competing sustainability criteria. Throughout the process, the designer remains responsible for interpreting these recommendations, questioning the proposed evaluations, and selecting the final concept.

Design Implications. Designing for human-centered decision support - The central objective of this principle is to mitigate normative opacity, ensuring that evaluative control remains the designer's responsibility. As LLMs integrate into the convergent phase, agency shifts from individual evaluation to mediated decision-agency, where the system supports rather than replaces human judgment. Consequently, LLM-based CSTs should be designed to structure evaluative reasoning without determining the final decision. Based on the previously presented scenario, instead of recommending a single “best” solution, the system could present multiple design alternatives along with evaluation criteria, underlying justification, and associated trade-offs, allowing designers to compare, question, and justify their final selection. In this way, the authority to evaluate, justify, and decide remains with the human, avoiding the silent transfer of decision-making agency to the system. Designing for progressive and situated transparency - This principle addresses technical opacity by making the underlying logic of AI-assisted evaluations understandable within the context of the design task. Transparency should allow designers to understand why alternatives are evaluated differently and what criteria influence those evaluations. In the illustrative scenario presented above, transparency can be achieved through interface mechanisms, such as contrasting panels that compare the performance of alternatives against the pillars of sustainability and uncertainty visualizations that communicate the confidence associated with life-cycle predictions or environmental assessments. These mechanisms transform transparency into a contextual resource that supports informed evaluation

during the convergence phase rather than treating it as a purely technical requirement. Designing for trust as an emergent phenomenon - By fostering a relationship of trust built through progressive understanding of the system's behavior, rather than the automatic acceptance of its recommendations, this principle addresses subjective opacity. In the illustrative scenario, trust results from the explicit communication of the criteria used to evaluate alternatives, the system's limitations, and the uncertainties associated with the presented predictions. In this sense, LLM-based CSTs should communicate the degree of confidence in their evaluations, explicitly state the limitations of the available data, and encourage the user to critically question the recommendations before making a final decision. Thus, trust ceases to depend on the authority attributed to AI and begins to emerge from the interaction between the designer and the system throughout the convergent process. Designing to avoid silent automation - This principle simultaneously addresses subjective and normative opacity, thereby preventing LLM-based CSTs from implicitly assuming authority in the evaluation and selection of ideas. Based on the illustrative scenario, we can see this implication: the system neither automatically recommends a single solution nor hides the criteria that support its evaluations. By requiring the designer to critically review recommendations before final selection, the system can introduce productive friction that promotes reflection rather than silent automation during the decision process. In this way, automation stops replacing human judgment and starts supporting the evaluation process, preserving creative autonomy and responsibility for the final decision. Designing for domain contextualization - By making explicit the values, assumptions, and trade-offs that influence decision-making in sustainability-oriented contexts, this principle partially addresses normative opacity. According to the illustrative scenario presented, this implication manifests when the LLM-based CST evaluates each design alternative against multiple sustainability criteria rather than offering generic or decontextualized recommendations. Thus, the system should highlight the conflicts inherent in wicked problems, such as the balance between environmental impact, economic benefits, and social effects, allowing the designer to understand how different priorities lead to different decisions. In this way, the LLM ceases to obscure the problem's normative complexity and begins to support value-conscious decision-making that guides the convergence process.

5 Conclusions

The integration of LLM-based CSTs is fundamentally reshaping the creative process, particularly within complex domains such as sustainability. Rather than simply assisting with the convergent phase of creativity, these systems reconfigure decision-making through the mediated decision-agency concept, in which human judgment and AI jointly shape the evaluation and selection of alternatives. By articulating the conceptual relationships among sustainability, creativity, and the design process, this work provides a structured analytical

framework for understanding how control, transparency, trust, and different forms of opacity interact in AI-assisted decision-making. Building on this framework, five design implications were proposed to support the development of LLM-based CSTs that foster informed, transparent, and value-conscious convergent thinking while preserving human responsibility throughout the decision-making process. Beyond sustainability-oriented design, the proposed conceptual framework provides design guidance for AI-assisted CSTs in the creative industries, where transparent, accountable, and human-centred evaluation processes are increasingly required to support innovation. Future work should empirically validate the proposed conceptual relationships, the illustrative design scenario, and the practical applicability of the design implications across different sustainability-oriented domains.

CRedit author statement. **Ana Rodrigues:** Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **David dos Santos:** Investigation, Writing – review, and editing. **Diogo Cabral:** Supervision, Writing – review and editing. **Pedro Campos:** Supervision, Writing – review and editing.

Funding. This research was funded by FCT – Fundação para a Ciência e Tecnologia, by project reference 2021.06734.BD and DOI identifier 10.54499/2021.06734.BD, by Portuguese Recovery and Resilience Program (PRR).

Disclaimer/Publisher's Note: The statements, opinions and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. ASLERD and/or the member of the editorial board of IxD&A disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

References

1. Ashraf M.Z., Wei W., Usman M., Mushtaq S.: How can natural resources, dependence, Environmental-related technologies, and digital trade protect the environment: Redesigning SDGs policies for a sustainable environment, *Journal of Resources Policy*, 88 (2024) <https://doi.org/10.1016/j.resourpol.2023.104456>
2. Hanson L.L.: Wicked Problems and Sustainable Development, in *Encyclopedia of Sustainability in Higher Education*, pp. 2090–2098, Springer International Publishing (2019) https://doi.org/10.1007/978-3-030-11352-0_522
3. Saleh R., Brem A.: Creativity for Sustainability: An integrative literature review, *Journal of Cleaner Production*, 388 (2023) <https://doi.org/10.1016/j.jclepro.2023.135848>
4. Guilford J.P.: Creativity, *American Psychologist*, 5, pp. 444–454 (1950) <https://doi.org/10.1037/h0063487>

5. Runco M.A., Jaeger G.J.: The Standard Definition of Creativity, *Creativity Research Journal*, 24, pp. 92–96 (2012) <https://doi.org/10.1080/10400419.2012.650092>
6. Dorst K.: The core of ‘design thinking’ and its application, *Design Studies*, 32, pp. 521–532 (2011) <https://doi.org/10.1016/j.destud.2011.07.006>
7. de Vries H.B., Lubart T.I.: Scientific Creativity: Divergent and Convergent Thinking and the Impact of Culture, *Journal of Creative Behavior*, 53, pp. 145–155 (2017) <https://doi.org/10.1002/jocb.184>
8. Shneiderman B.: Creating creativity: user interfaces for supporting innovation, *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7, pp. 114–138 (2000) <https://doi.org/10.1145/344949.345077>
9. Hwang A.H.C.: Too Late to be Creative? AI-Empowered Tools in Creative Processes, in *CHI EA 2022*, pp. 1–9 (2022) <https://doi.org/10.1145/3491101.3503549>
10. Singh S., Hindriks K., Heylen D., Baraka K.: A Systematic Review of Human-AI Co-Creativity, *ArXiv.org* (2025)
11. Rezwana J., Maher M.L.: Designing Creative AI Partners with COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems, *ACM Transactions on Computer-Human Interaction*, 30, pp. 1–28 (2023) <https://doi.org/10.1145/3519026>
12. Liao V., Voughan J.W.: AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap, *Harvard Data Science Review*, Special Issue 5 (2024) <https://doi.org/10.1162/99608f92.8036d03b>
13. DiSalvo C., Sengers P., Brynjarsdóttir H.: Mapping the landscape of sustainable HCI, in *CHI 2010*, pp. 1975–1984 (2010) <https://doi.org/10.1145/1753326.1753625>
14. Ceschin F., Gaziulusoy I.: *Design for Sustainability: A Multi-level Framework from Products to Socio-technical Systems*, 1st ed., Routledge (2019) <https://doi.org/10.4324/9780429456510-1>
15. Irwin T.: Transition Design: A Proposal for a New Area of Design Practice, Study, and Research, *The Journal of the Design Studies Forum*, 7, pp. 229–246 (2015) <https://doi.org/10.1080/17547075.2015.1051829>
16. Geissdoerfer M., Savaget P., Bocken N.M.P., Hultink E.J.: The Circular Economy - A new sustainability paradigm?, *Journal of Cleaner Production*, 143, pp. 757–768 (2017) <https://doi.org/10.1016/j.jclepro.2016.12.048>
17. Rodrigues A., Cabral D., Campos P.F.: Creativity Support Tools and Convergent Thinking: A Preliminary Review on Idea Evaluation and Selection, in *C&C 2023*, pp. 305–311 (2023) <https://doi.org/10.1145/3591196.3596821>
18. Xu T., Liu W., Zhu Z.: Multi-Criteria Decision-Making Methods Applied To Achieve Sustainable Design: A Systematic Review, *Proceedings of the Design Society*, pp. 2725–2734 (2023) <https://doi.org/10.1017/pds.2023.273>
19. Parolin G., McAloone T.C., Pigosso D.C.A.: What’s the catch? Trade-off challenges in early design for sustainability, *Proceedings of the Design Society*, 4, pp. 1399–1408 (2024) <https://doi.org/10.1017/pds.2024.142>
20. Rodrigues A., Cabral D., Campos P.: Fostering Informed Decision-Making: A Visual and Textual Approach to Sustainability Metrics for Idea Selection, in *ICGI 2024*, pp. 1–8 (2024) <https://doi.org/10.1109/ICGI64003.2024.10923803>
21. Corujeira J., Cabral D.: FreeCAD-LCA Interactive Tool: integrating and visualizing a Life-Cycle Assessment analysis in a CAD system, in *UIST Adjunct 2025*, pp. 1–3 (2025) <https://doi.org/10.1145/3746058.3758356>
22. Preuss N., Alshehri A.S., You F.: Large language models for life cycle assessments: Opportunities, challenges, and risks, *Journal of Cleaner Production*, 466 (2024) <https://doi.org/10.1016/j.jclepro.2024.142824>
23. Dorst K., Cross N.: Creativity in the design process: co-evolution of problem-solution, *Design Studies*, 22, pp. 425–437 (2001)

- [https://doi.org/10.1016/S0142-694X\(01\)00009-6](https://doi.org/10.1016/S0142-694X(01)00009-6)
24. Sternberg R.J.: The Nature of Creativity, *Creativity Research Journal*, 18, pp. 87–98 (2006) https://doi.org/10.1207/s15326934crj1801_10
 25. Salma Z., Hijón-Neira R., Pizarro C.: Designing Co-Creative Systems: Five Paradoxes in Human-AI Collaboration, *Information - Emerging Research in Computational Creativity and Creative Robotics*, 16 (2025) <https://doi.org/10.20944/preprints202509.1541.v1>
 26. Rodrigues A., Cabral D., Campos P.: Inside the Solution Diamond: The Role of CSTs Supporting Divergent and Convergent Thinking, in *ECCE 2025*, pp. 1–11 (2025) <https://doi.org/10.1145/3746175.3746194>
 27. Frich J., Vermeulen L.M., Remy C., Biskjaer M.M., Dalsgaard P.: Mapping the Landscape of Creativity Support Tools in HCI, in *CHI 2019*, pp. 1–18 (2019) <https://doi.org/10.1145/3290605.3300619>
 28. Ivcevic Z., Grandinetti M.: Artificial intelligence as a tool for creativity, *Journal of Creativity*, 34 (2024) <https://doi.org/10.1016/j.yjoc.2024.100079>
 29. Rodrigues A., Cabral D., Campos P.: LLMs-Assisted Decision-making in Sustainable Contexts, in *ECCE 2025*, pp. 1–12 (2025) <https://doi.org/10.1145/3746175.3746188>
 30. Chakrabarty T., Padmakumar V., Brahman F., Muresan S.: Creativity Support in the Age of Large Language Models: An Empirical Study Involving Professional Writers, in *C&C 2024*, pp. 132–155 (2024) <https://doi.org/10.1145/3635636.3656201>
 31. Glassberg I., Brender-Ilan Y., Zwilling M.: The key role of design and transparency in enhancing trust in AI-powered digital agents, *Journal of Innovation & Knowledge*, 10 (2025) <https://doi.org/10.1016/j.jik.2025.100770>
 32. Hollan J., Hutchins E., Kirsh D.: Distributed cognition: toward a new foundation for human-computer interaction research, *ACM Transactions on Computer-Human Interaction*, 7, pp. 174–196 (2000) <https://doi.org/10.1145/353485.353487>
 33. Wertsch J.V.: *Mind as Action*, Oxford University Press (1998) <https://doi.org/10.1093/acprof:oso/9780195117530.001.0001>
 34. Baxter G., Sommerville I.: Socio-technical systems: From design methods to systems engineering, *Interacting with Computers*, 23, pp. 4–17 (2011) <https://doi.org/10.1016/j.intcom.2010.07.003>
 35. Schreiber F., Cramer C.: Towards a conceptual systematic review: proposing a methodological framework, *Educational Review*, 76, pp. 1458–1479 (2024)
 36. Bollen Z., Wyss A.M., Rosales E.G., Corbi Z.L.H.: The role of attitudes and attention in pro-environmental decision-making: An eye-tracking study, *Journal of Environmental Psychology*, 107 (2025) <https://doi.org/10.1016/j.jenvp.2025.102776>
 37. Howard T.J., Culley S.J., Dekoninck E.: Describing the creative design process by the integration of engineering design and cognitive psychology literature, *Design Studies*, 29, pp. 160–180 (2008) <https://doi.org/10.1016/j.destud.2008.01.001>
 38. Lubart T.: Creativity and Convergent Thinking: Reflections, Connections and Practical Considerations, *RUDN Journal of Psychology and Pedagogics* (2016)
 39. Humr S., Canan M.: Intermediate Judgments and Trust in Artificial Intelligence-Supported Decision-Making, *Entropy*, 26 (2024) <https://doi.org/10.20944/preprints202404.1107.v1>
 40. Vereschak O., Alizadeh F., Bailly G., Caramiaux B.: Trust in AI-assisted Decision Making: Perspectives from Those Behind the System and Those for Whom the Decision is Made, in *CHI 2024*, pp. 1–14 (2024) <https://doi.org/10.1145/3613904.3642018>
 41. Bansal G., Wu T., Zhou J., Fok R., Nushi B., Kamar E., Ribeiro M.T., Weld D.: Does the Whole Exceed its Parts? The Effect of AI Explanations on Complementary Team Performance, in *CHI 2021*, pp. 1–16 (2021) <https://doi.org/10.1145/3411764.3445717>

42. Zhang Y., Liao Q.V., Bellamy R.K.E.: Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making, in FAT* 2020, pp. 295–305 (2020) <https://doi.org/10.1145/3351095.3372852>
43. Ehsan U., Watkins E.A., Wintersberger P., Manger C., Kim S.S.Y., Van Berkel N., Riener A., Riedl M.O.: Human-Centered Explainable AI (HCXAI): Reloading Explainability in the Era of Large Language Models (LLMs), in CHI EA 2024, pp. 1–6 (2024) <https://doi.org/10.1145/3613905.3636311>
44. Palazzolo P., Stahl B., Webb H.: Measurable Trust: The Key to Unlocking User Confidence in Black-Box AI, in TAS 2024, pp. 1–7 (2024) <https://doi.org/10.1145/3686038.3686058>
45. Herrera F., Calderón R.: Opacity as a feature, not a flaw: Role-sensitive explainability, institutional trust, and the LoBOX ethics governance framework for AI, *Technology in Society*, 86 (2026) <https://doi.org/10.1016/j.techsoc.2026.103302>
46. Mehrotra S., Degachi C., Vereschak O., Jonker C.M., Tielman M.L.: A Systematic Review on Fostering Appropriate Trust in Human-AI Interaction: Trends, Opportunities and Challenges, *ACM Journal on Responsible Computing*, 1, Article 26, 45 pages (2024) <https://doi.org/10.1145/3696449>
47. Li M., Zhou M., Li Y., Wei W., Zhu T., Xu X., Ren L., Zhang N., Xu R., Li J.: From Black Box to Transparency: The Impact of Multi-Level Visualization on User Trust in Autonomous Driving, *Sensors*, 25 (2025) <https://doi.org/10.3390/s25216725>
48. Banathy B.H.: *Designing Social Systems in a Changing World*, 1st ed., Springer US (1996) https://doi.org/10.1007/978-1-4757-9981-1_1
49. Jun G.T., Hignett S., Clarkson P.J.: *Design Creativity*, Cambridge University Press (2024) <https://doi.org/10.1017/9781009325318>
50. Csikszentmihalyi M.: *Creativity: Flow and the Psychology of Discovery and Invention*, Harper Collins Publishers (2013)
51. Kim J., Suh S., Chilton L.B., Xia H.: Metaphorian: Leveraging Large Language Models to Support Extended Metaphor Creation for Science Writing, in DIS 2023, pp. 115–135 (2023) <https://doi.org/10.1145/3563657.3595996>
52. Suh S., Chen M., Min B., Li T.J.-J., Xia H.: Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation, in CHI 2024, pp. 1–26 (2024) <https://doi.org/10.1145/3613904.3642400>
53. Wang W.-F., Lu C.-T., Campanyà N.P., Chen B.-Y., Chen M.Y.: Aldeation: Designing a Human-AI Collaborative Ideation System for Concept Designers, in CHI 2025, pp. 1–28 (2025) <https://doi.org/10.1145/3706598.3714148>
54. Scharowski N., Perrig S.A.C., von Felten N., Aeschbach L.F., Opwis K., Wintersberger P., Brühlmann F.: To Trust or Distrust AI: A Questionnaire Validation Study, in FAccT 2025, pp. 361–374 (2025) <https://doi.org/10.1145/3715275.3732025>
55. Shneiderman B.: Human-centered AI: ensuring human control while increasing automation, in HUMAN 2022, pp. 1–2 (2022) <https://doi.org/10.1145/3538882.3542790>
56. Dalsgaard P.: Thinking through Prompting: Cognitive Mediation in Human-AI Interaction, in ECCE 2025, pp. 1–6 (2025) <https://doi.org/10.1145/3746175.3747192>