Augmented Learning for Joint Creativity in Human-GenAI Co-Creation

Abstract

Our research introduces a novel theory of augmented learning by reconceptualizing the concept and providing both theoretical insights and empirical evidence to explain its role in human-GenAI co-creation for enhancing joint creativity over time. We propose shifting the focus of augmented learning from traditional human cognitive learning to a collective learning process, where humans and GenAI collaboratively rearrange their levels of involvement in co-creation activities to continuously improve joint creativity over time. To examine this reconceptualized augmented learning, we adopted a mixed-methods approach across three studies. Studies 1 and 2 inductively revealed that human-GenAI co-creation does not automatically achieve augmented learning or enhance their joint creativity even after multiple rounds of co-creation experiences. A deeper qualitative analysis of human-GenAI dialogues identified a decline in *Idea Co-development*—a co-creation activity characterized by feedback exchanges and iterative idea refinement—as the primary reason for this failure. Based on this finding, Study 3 deductively demonstrated that providing instructions and guidance on *Idea Co-development* improved joint creativity over time. Building on these findings, we further refine our theory of augmented learning, offering actionable insights for improving human-GenAI co-creation to achieve greater joint creativity.

Keywords: theory of augmented learning, human-GenAI co-creation, joint creativity, *Idea Co-development*

1. Introduction

The introduction of Generative Artificial Intelligence (GenAI) in recent years has sparked substantial research interest due to its potential for creativity, setting it apart from traditional technologies (Epstein and Hertzmann 2023, Noy and Zhang 2023, Gordetzki et al. 2024, Wu et al. 2024, Bauer et al. 2024, Vardi and Choudhary 2024, Doshi and Hauser 2024, Hou et al. 2024). Creativity, defined as the generation of new and useful ideas regarding products, services, processes, or work methods (Anderson et al. 2014), has long been seen as a uniquely human capability (Noy and Zhang 2023). Traditional technologies, ranging from simple machines to advanced information systems, have only reinforced this view. They excel in performing routine and well-defined tasks but are limited in their ability to engage in creative processes, such as the flexible exploration of new ideas (Schwab 2017). Thus, they have been viewed as tools for humans rather than independent contributors to the creative processes (Anthony et al. 2023). In contrast, GenAI holds considerable promise to collaborate more independently and closely with humans in creative work (Guzik et al., 2023). This is reflected in its definition: a class of artificial intelligence systems capable of generating new content (e.g., text, images, audio, or video) in response to human prompts, based on patterns learned from existing data (Chen and Chan 2024). Unlike traditional technologies, GenAI's adaptability and interactive responsiveness also enable it to collaborate dynamically and iteratively with humans.

This advancement opens up new opportunities for human-GenAI co-creation, where humans and GenAI collaborate to create creative outcomes (van den Broek et al. 2021). In recent years, several scholars have begun investigating whether such collaboration affects joint creativity, i.e., the level of creativity in outcomes (e.g., ideas) co-created by humans and GenAI (e.g., Chen and Chan 2024, Jia et al. 2024). Operating under the implicit assumption that humans and GenAI can achieve *instantaneous learning* of effective strategies for co-creation, these studies have primarily focused on single-round creative tasks, where humans and GenAI collaborate only once. The findings, however, have been largely inconclusive, reporting positive, negative, and mixed effects of human-GenAI co-creation on joint creativity (Gao and Jiang 2021, Dell'Acqua et al. 2023, Hitsuwari et al. 2023, Noy and Zhang 2023, Messer 2024, Jia et al. 2024, Doshi and Hauser 2024, McGuire et al. 2024, Boussioux et al. 2024, Chen and Chan 2024, Sun et al. 2025). These inconsistencies show that the existing approach to understanding the relationship between human-GenAI co-creation and joint creativity is limited, pointing to the need for a new theoretical lens.

Our research develops a novel theory on human-GenAI co-creation, anchoring it around the concept of augmented learning. We reconceptualize it as *an evolutionary process in which humans and GenAI continuously adjust their levels of involvement in co-creation activities within a task over time to achieve ongoing joint creativity enhancement*. This reconceptualization reflects our new theoretical insights that the primary goal of human-GenAI co-creation is not limited to producing a single creative outcome but rather to achieving continuous improvement in their joint creativity via augmented learning. To capture this ongoing improvement, human-GenAI co-creation must be studied across broader spatial and temporal scales (Raisch and Krakowski 2021). Expanding the spatial scale entails examining a variety of interactive activities required to complete a task (i.e., co-creation activities) rather than focusing solely on the task as a whole at a high level. Expanding the temporal scale involves studying human-GenAI co-creation over multiple rounds rather than limiting investigations to a single round. By expanding these two scales, the strengths and limitations of both humans and GenAI become more apparent, allowing the human-GenAI dyads to progressively adjust their roles in different co-creation activities to determine effective adjustments for fostering joint creativity.

Building on our theory of augmented learning, we have a twofold objective. First, we aim to investigate whether human-GenAI co-creation can spontaneously achieve augmented learning, resulting in the continuous improvement of joint creativity over multiple rounds. Previous studies have overlooked the evolving dynamics between humans and GenAI over time, offering only a partial view of the impact of human-GenAI co-creation on joint creativity. Second, we aim to identify which co-creation activities humans and GenAI should prioritize to enhance joint creativity over time. A few studies have investigated some modes of human-GenAI co-creation in creative tasks—e.g., GenAI taking the lead in writing tasks versus merely offering feedback (Chen and Chan 2024). However, none have investigated how diverse co-creation activities emerge and evolve across multiple rounds, nor how humans and GenAI adjust their levels of involvement in these activities over time.

To achieve our objectives, we conducted three laboratory and online experiments, adopting an abductive research approach that combines inductive and deductive methods (Timmermans and Tavory 2012). This mixed-methods approach enables us to both discover new phenomena inductively and validate our findings deductively. Study 1 inductively observed that human-GenAI co-creation did not automatically lead to improved joint creativity over time, even after 10 rounds of creative tasks, suggesting the difficulty of achieving augmented learning. Building on these insights, the goal of Study 2 was to explore why humans and GenAI struggle to improve joint creativity, even after multiple rounds of collaboration, and to identify specific co-creation activities that could facilitate improvements in joint creativity during human-GenAI co-creation. Through qualitative and quantitative analyses, we identified three distinct co-creation activities employed by human-GenAI dyads during the co-creation process: *Idea Generation-Response, Idea Request-Idea Generation,* and *Idea Co-development*. Among these, a gradual increase in collaborative engagement in *Idea Co-development* emerged as the most effective approach for fostering continuous improvement in joint creativity over time. However, we observed that human-GenAI dyads failed to recognize the importance of their engagement in *Idea Co-development*, even after multiple rounds of co-creation. Such failure was the primary reason why human-GenAI co-creation struggled to achieve consistent improvements in joint creativity. Finally, in Study 3, we conducted an empirical test to examine whether the provision of instructions encouraging participants to engage more in *Idea Co-development* could effectively enhance joint creativity. The findings demonstrated that emphasizing this activity led to significant improvements in joint creativity across rounds, confirming its critical role in fostering creative co-creation between humans and GenAI.

Our research significantly contributes to the literature on human-GenAI co-creation for joint creativity. First, we introduce a novel theory of augmented learning in human-GenAI co-creation, providing new insights into human-GenAI co-creation processes related to augmented learning. Unlike the traditional definition of augmented learning that emphasizes technology's role in enhancing human cognition as a supportive tool, our reconceptualization recognizes GenAI as a potential co-creator, enhancing joint creativity synergistically with humans. We then demonstrate that augmented learning does not occur naturally; rather, it requires deliberate interventions, such as structured templates, to guide human-GenAI co-creation and foster continuous improvements in joint creativity over time. Second, our research sheds light on the processes underlying human-GenAI co-creation by identifying three emerging co-creation activities: *Idea Generation-Response, Idea Request-Idea Generation,* and *Idea Co-development.*

Among these, *Idea Co-development* proves critical for enhancing joint creativity, emphasizing the importance of continuous, iterative feedback exchanges and refinements of generated ideas by both humans and GenAI. Finally, our theory of augmented learning suggests that previously inconsistent findings in the literature on human-GenAI co-creation and joint creativity may be attributed to insufficient consideration of the extended temporal scale. Overall, our research provides a robust framework for advancing research on optimizing human-GenAI synergy in creative contexts. It also offers valuable practical insights for organizations and GenAI developers seeking to improve the effectiveness of human-GenAI co-creation.

2. The Theory of Augmented Learning in Human-GenAI Co-creation

2.1. Reconceptualization of Augmented Learning

The concept of augmented learning was first introduced by Engelbart (1962) to describe the phenomenon in which technology enhances humans' ability to acquire knowledge more effectively and efficiently. Initially, this idea treated technology as a supportive tool, primarily designed to provide existing information in a more accessible manner. At its most advanced, it incorporated limited adaptiveness, where technology adjusts to human needs based on preset rules (Elam and Mead 1990, Althuizen and Reichel 2016, Brynjolfsson and Mitchell 2017). The advent of GenAI, however, has dramatically reshaped how humans interact with technology. Unlike earlier tools, GenAI is capable not only of retrieving knowledge but also of generating new content based on human prompts and context (Epstein and Hertzmann, 2023). As a result, people often form the impression that GenAI possesses human-like agency, referred to as perceived agency, or the perception that GenAI has the capacity to think, reason, plan, and act on intentions (Bigman and Gray 2018, Vanneste and Puranam 2024). This elevated status positions GenAI not just as a passive tool, but as a legitimate cognitive partner, capable of participating meaningfully in complex, joint tasks (Anthony et al. 2023). This evolution calls for a redefinition of augmented learning in the context of human-GenAI interaction. Rather than focusing solely on human cognition being extended by passive tools, augmented learning must now account for collaborative division of labor between two agents—humans and GenAI—who collectively shape the learning process.

Our research thus suggests that the concept of augmented learning should shift its focus from the cognitive learning of an individual human to the *collective learning* of humans and GenAI. In the team literature, particularly research on transactive memory systems, collective learning no longer accounts for individuals' cognitive learning. Instead, it relates to the adjustment of team members' collective activities (Argote and Guo 2016). Transactive memory systems are collective systems of "who knows what," allowing team members to access more information collectively than they possess individually when completing their team tasks (Wegner 1987, Lewis and Herndon 2011). At the heart of transactive memory systems is the concept of collective learning, which involves two core processes: identifying each team member's strengths and weaknesses, and continuously reorganizing the team's task assignments to maximize those strengths while minimizing weaknesses (Ren and Argote 2011, Argote and Guo 2016). Given the elevated role of GenAI as a potential collaborator, we argue that augmented learning must now account for the collective learning of humans and GenAI. This involves understanding the capabilities of both parties and restructuring their respective roles and involvement in collective activities to complete tasks successfully.

However, augmented learning does not occur instantaneously; certain key conditions must be met for humans and GenAI to achieve it. We argue that expanding spatial and temporal scales should be an integral part of augmented learning. The importance of such expansion in the context of human-GenAI co-creation is also highlighted in Raisch and Krakowski (2021)'s paradoxical theory of automation-augmentation. This theory begins with the question of whether organizations should make a trade-off decision about AI use, choosing either automation, where AI independently takes over organizational tasks, or augmentation, where humans and AI collaborate to jointly carry out these tasks. Organizations have traditionally approached human-GenAI co-creation at the level of a single task at a specific point in time, neglecting the diverse activities within the task (narrow spatial scale) and their evolution over extended time periods (narrow temporal scale). Consequently, this trade-off decision became necessary, as automation and augmentation cannot be implemented simultaneously for the same task at a given moment. However, "one-sided orientations toward either automation or augmentation cause vicious cycles, because they neglect the dynamic interdependencies between AI's dual applications in management…While managers initially perceive automation and augmentation as a trade-off, they may eventually recognize that they cannot simply choose between these dual AI applications, because either choice intensifies the need for its opposite" (Raisch and Krakowski, 2021; p. 200). That is, in organizations, automation and augmentation should not be viewed as an either-or choice. Instead, both approaches can coexist if organizations consider the diverse activities within each task (an expanded spatial scale) over time (an expanded temporal scale).

The expansion of the spatial scale involves examining a broader set of activities within a task. Organizations are complex information processors, and completing a task within them typically involves multiple interdependent activities. An illustrative example can be found in Netflix's creative production workflows, where humans and AI collaborate in script and content generation. Instead of treating script development as a single task, Netflix breaks it down into sub-activities such as idea generation, evaluation, and selection, each involving different degrees of human and AI participation. For example, human writers remain central to generating storylines and early drafts, while AI plays a larger role in analyzing narrative pacing, character development, and audience demand patterns, providing insights that guide refinement and positioning decisions about how the content should be targeted and marketed to audiences (DigitalDefynd 2025). This example demonstrates that expanding the spatial scale involves breaking down a broader task into smaller, discrete activities. When researchers analyze a task at a high level, they may perceive only a dilemmatic tension between automation and augmentation. However, by subdividing the task into multiple activities, they can resolve the tension and uncover strategies that optimize human and GenAI contributions across activities.

Extending the temporal scale involves allowing for multiple rounds of human-GenAI co-creation. Through these repeated interactions, both humans and GenAI have ample opportunities to understand each other's strengths and limitations (Brynjolfsson and McAfee 2016, Calabretta et al. 2017). For example, over multiple co-creation experiences, they may develop an understanding of humans' intuition and insights (a human strength) alongside their limited cognitive capacity (a human limitation). Similarly, they may recognize GenAI's vast information processing and generative power (a GenAI strength) as well as its inability to interpret contextual nuances (a GenAI limitation). By identifying each party's strengths and weaknesses, they can gradually determine which party should be more involved in specific activities within the task. Consequently, they may uncover more effective levels of human and GenAI involvement across different activities over time (van den Broek et al. 2021).

Building on theoretical insights from the concept of collective learning in the literature of transactive memory systems and the paradox theory of automation-augmentation, we define augmented learning *as an evolutionary process in which humans and GenAI continuously adjust* *their levels of involvement in multiple co-creation activities within a task over time to achieve ongoing joint creativity enhancement.* This definition shifts attention away from the focus of the original concept of augmented learning—i.e., the augmentation of human cognitive capabilities. Instead, our reconceptualization emphasizes the augmentation of joint outcomes (in our case, joint creativity) through synergetic co-creation processes that emerge from collective learning between humans and GenAI and unfold over time across multiple activities within a task.

2.2. The Necessity of Human in the Loop for Augmented Learning

Importantly, we suggest that humans still need to take on an active and agentic role in augmented learning (Zanzotto 2019, Gnewuch et al. 2023). Even activities designated for automation should be regularly revisited by humans to reassess and update the respective roles of humans and GenAI (Raisch and Krakowski 2021). This means that the rearrangements of humans' and GenAI's involvement in different co-creation activities over time (i.e., augmented learning) should be predominantly driven by humans. Humans are responsible for monitoring the effectiveness of human-GenAI co-creation and actively refining each party's contributions to ensure optimal levels of human and GenAI involvement across various co-creation activities. Although GenAI appears highly autonomous compared to traditional technology, its function is designed to be initiated by a human—by definition, it generates outputs in response to human prompts (Grimes et al. 2023, Chen and Chan 2024). It differs from agentic AI, which independently initiates actions and adapts to goals (Baird and Maruping 2021). Although GenAI displays increasing agency (i.e., the capacity to think, reason, plan, and act on intentions), it is not fully agentic in the sense of autonomously directing or managing collaborative processes. Therefore, augmented learning must remain centered on human judgment and oversight.

Looking ahead, it is conceivable that future AI systems could be fully agentic, learning autonomously from repeated human-AI interactions and agentically supervising role adjustments for both themselves and human collaborators. However, such capabilities remain limited in practice for several reasons (Davenport and Kirby 2016). First, human involvement is essential to ensure that the human-AI co-creation process aligns with broader organizational or societal objectives (Bankins et al. 2024). AI systems are not inherently designed to account for higher-level organizational and societal priorities, which are often critical to the success of human-AI collaboration. Second, delegating such responsibilities to AI systems poses ethical risks that might go unnoticed without human oversight, potentially compromising the co-creation process (Köbis et al. 2021). While AI decisions are inherently data-driven and potentially more objective, they can inadvertently overlook the ethical implications of role adjustments. Lastly, AI systems often lack the contextual and socio-emotional understanding necessary to produce outcomes that are fully attuned to human needs (Tschang and Almirall 2021). Taken together, even if we assume the emergence of fully agentic AI in the future, these limitations show the indispensable role of human judgment in human-AI co-creation. For these reasons, augmented learning is, and should remain, primarily human-directed.

3. Extant Literature on Human-GenAI Co-creation and Joint Creativity

A few researchers have investigated the effects of human-GenAI co-creation on joint creativity, but their findings have been inconclusive (see Table 1 for an overview of empirical research on how human-GenAI co-creation affects joint creativity). For example, Noy and Zhang (2023) found that participants using ChatGPT in a writing task produced more creative and original writings compared to those who did not use ChatGPT. In contrast, Chen and Chan (2024) reported the opposite: in an advertisement ideation task, human-GenAI co-generated ideas were less creative than those produced solely by humans. A recent meta-analysis of past studies conducted between January 2020 and June 2023 further reflects this inconsistency. Although this study examined human-AI synergy more broadly (across a range of AI systems), it found that on performance in creative tasks, the pooled effect of human-AI co-creation was not statistically significant (Vaccaro et al. 2024), reflecting the inconsistency in the literature. Importantly, these studies focused on single-round creative tasks and paid limited attention to how the dynamics of human-GenAI co-creation evolve over time.

Our research suggests that the predominant use of single-round creative tasks in research designs, which limit exploration of augmented learning, may be a key reason for the inconsistent findings. Examining a single round may only capture the chance-based or serendipitous effects of human-GenAI co-creation on joint creativity. For instance, during a single interaction with GenAI, humans might randomly discover—or fail to discover—an effective arrangement of humans' and GenAI's involvement in co-creation activities that can enhance their joint creativity. However, this scenario does not reflect the full potential of human-GenAI co-creation, whose true value lies in continuous co-creation and the resulting augmented learning. We speculate that such serendipity might contribute to the inconsistencies observed in previous findings. Moreover, most past studies have predominantly focused on understanding whether human-GenAI co-creation increases or decreases joint creativity, rather than investigating the underlying mechanisms or exploring ways to enhance joint creativity. This limited approach hinders a deeper understanding of how humans and GenAI can co-create to achieve continuous improvements in joint creativity over time. Furthermore, without exploring the underlying mechanisms, it remains challenging to identify actionable strategies for optimizing human-GenAI co-creation in creative tasks.

4. Applying the Theory of Augmented Learning to Human-GenAI Co-Creation for Joint Creativity

Our research aims to achieve two research objectives by introducing a fresh theoretical perspective on augmented learning to the literature on human-GenAI co-creation and joint creativity. The first objective is to investigate whether humans and GenAI can naturally achieve augmented learning in the co-creation process. Unlike past studies, our research expands both spatial and temporal scales by examining multiple co-creation activities (expanded spatial scale) across multiple rounds (expanded temporal scale). This approach enables us to observe whether and how humans and GenAI adjust their levels of involvement in different co-creation activities over rounds to achieve greater joint creativity.

Figure 1 illustrates an exemplary process of augmented learning in human-GenAI co-creation for joint creativity. In task round 1, the human-GenAI dyad may begin with a balanced approach, engaging evenly across all co-creation activities. As they move into task round 2, the dyad may adjust their levels of involvement in these activities based on their experiences from the first round. As a strategy to enhance joint creativity, they might choose to increase their focus on Co-creation Activity 1 while reducing involvement in the other two activities. This process of reflection, adjustment, and iteration continues over time, with each task round bringing new insights. By task round 10, the dyad may develop an effective co-creation framework that leverages their combined strengths—e.g., the dyad engages more in Co-creation Activities 2 and 3 while minimizing their efforts in Co-creation Activity 1. Through this active adjustment of their participation in co-creation activities, humans and GenAI could potentially achieve continued improvements in their joint creativity.

Although this example demonstrates a successful instance of augmented learning for joint creativity, such an ideal outcome may not be guaranteed. The human-GenAI dyad might not make such effective adjustments to their involvement in co-creation activities, thereby failing to achieve continuous improvement in joint creativity. In fact, past studies have shown that people often over-rely on technologies, which leads to reduced efforts to understand these technologies in depth and explore alternative uses—a phenomenon known as *automation bias* (Parasuraman and Manzey 2010, Goddard et al. 2012, Endsley 2017, Bogert et al. 2021, Jussupow et al. 2021). If such over-reliance is observed in humans' interactions with GenAI, it is possible that humans and GenAI may not achieve augmented learning. Therefore, our first research objective is to examine whether human-GenAI co-creation can naturally achieve augmented learning, leading to the continuous improvement of joint creativity over time.

Our second objective is to identify one or a few effective co-creation activities that can potentially enhance joint creativity without requiring extensive lead time. Augmented learning typically takes time, as both humans and GenAI need to engage in multiple rounds of co-creation to effectively refine their respective roles in co-creation activities. However, if they can initiate their collaboration by adopting an effective set of co-creation activities, they may achieve performance benefits right from the start. Furthermore, as discussed above, in cases where the spontaneous achievement of augmented learning is challenging, the provision of exemplary co-creation activities can support improvement in joint creativity and/or stimulate further insights for subsequent augmented learning. Here, we do not claim that there is a single best way for humans and GenAI to co-create for joint creativity, as multiple effective approaches for achieving joint creativity likely exist. Instead, our goal is to identify one or a few such effective activities that facilitate improvement in joint creativity. To achieve this objective, we first identify the emerging co-creation activities that are beneficial for the continuous improvement of joint creativity and then assess their validity and generalizability through multiple studies. Therefore, we set two research questions:

Research Question 1: Can humans and GenAI naturally achieve augmented learning over multiple rounds of co-creation, leading to continuous improvement in joint creativity? If not, why?

Research Question 2: What specific co-creation activities can help human-GenAI dyads improve their joint creativity over time?

To answer these research questions, we adopted an abductive approach without having specific hypotheses for several reasons. First, the application of GenAI to creativity is a relatively new area of study, with limited research and a lack of established theories (Amabile 2020). To make theorization more difficult, the existing findings are inconclusive, as we reviewed earlier (see Table 1). The scarcity of prior knowledge, coupled with empirical inconsistencies, makes it challenging to formulate hypotheses based solely on existing literature (Aguinis and Vandenberg 2014, Behfar and Okhuysen 2018). Second, no prior research has investigated the effects of human-GenAI co-creation on joint creativity over time. The absence of multi-round studies means there is little understanding of how joint creativity evolves through multiple rounds of interaction between humans and GenAI. Lastly, we currently have no theoretical framework that identifies which specific co-creation activities between humans and GenAI are most effective. Without such insights, it is challenging to predict phenomena using a purely deductive approach.

Given these limitations, our mixed-methods approach, which allows for dynamic interplay between inductive and deductive reasoning, is particularly suitable for our research

(Timmermans and Tavory 2012, Golden-Biddle 2020). This methodology is well-suited to exploratory studies where existing theories are insufficient, enabling us to generate new insights through a flexible and iterative process. Specifically, our research comprises three studies that strategically integrate both inductive and deductive approaches. Studies 1 and 2 adopt inductive methods to explore and understand the phenomenon of human-GenAI co-creation, while Study 3 employs a deductive method to test a hypothesis developed based on the findings of the earlier studies.[1]

In Study 1, we conduct a quantitative exploration to observe whether humans and GenAI naturally achieve augmented learning and continuous improvement in joint creativity over multiple rounds of co-creation. This study aims to observe patterns and trends without imposing preconceived theories, allowing the data to inform our understanding of the co-creation process (Aguinis and Vandenberg 2014). Study 2 consists of three parts. Part 1 begins with a quantitative replication of the findings from Study 1. Afterward, using the human-GenAI interaction data obtained in Part 1, we conduct a qualitative investigation in Part 2 to identify emerging human-GenAI co-creation activities (Locke 2007). In Part 3, we return to quantitative analyses, evaluating which of these activities are most effective for enhancing joint creativity and investigating why and how human-GenAI co-creation does (or does not) consistently lead to continuous improvement in joint creativity. This mixed-methods approach enables us to delve deeper into the nuances of the co-creation process, identifying key co-creation activities that contribute to or hinder the improvement of joint creativity. Based on the insights gained from Studies 1 and 2, Study 3 adopts a deductive approach (Behfar and Okhuysen 2018). We

[1] All research designs and procedures were approved by the Ethics Review Group of the first author's institution (Reference Number: 23-30). All studies were preregistered prior to data collection—Study 1: https://aspredicted.org/6QJ_V16; Study 2: https://osf.io/akhc8/; Study 3: https://aspredicted.org/69bq-v4d2.pdf.

formulate a hypothesis that the provision of instructions on effective co-creation activities identified in Study 2 could improve joint creativity over time. This hypothesis is then rigorously tested in an experiment.

By structuring our research in this way, we leverage the strengths of both inductive and deductive reasoning within an abductive framework, allowing us to generate new theories from empirical observations and subsequently test these theories to establish their validity (Timmermans and Tavory 2012).

5. Study 1

5.1. Sample and Procedure

A total of 162 participants (99 females; M_{age} = 27.02, SD_{age} = 11.09) were recruited for a laboratory experiment conducted at a UK university's Behavioral Laboratory. Participants received £10 as compensation for their participation.[2] Upon arrival at the laboratory, each participant was led to a private booth equipped with a computer. After providing informed consent, participants were randomly assigned to one of two experimental conditions: human-GenAI co-creation (N = 81) or human-only ideation (N = 81). Participants in both conditions were informed that they would complete a total of 10 rounds of ideation tasks. The tasks were about contemporary societal and environmental issues, such as climate change. The full list of our ideation tasks can be found in Supplementary Text C. In the human-GenAI co-creation condition, participants co-generated creative solutions for the ideation tasks with GenAI. In the human-only condition, participants independently generated creative solutions for the tasks.

[2] The initial sample consisted of 172 participants. We excluded 10 participants who withdrew from the study midway, participants who failed to adhere to study instructions, participants who did not complete all three tasks, and participants who were in the human-GenAI condition but did not collaborate with GenAI for task completion.

Since our focus is on examining augmented learning, which is inherently a within-individual phenomenon, our research design does not necessarily require a comparison group, i.e., the human-only condition. However, we included the human-only condition as a comparison group only in Study 1 for two specific informational reasons. First, past findings have shown that humans can self-learn through repeated engagement in similar creative tasks (Parnes 1961, Chua and Iyengar 2008, Lucas and Nordgren 2015). By comparing these two conditions, we can examine whether the progression of joint creativity in human-GenAI co-creation is comparable to, better than, or worse than the natural progression of human creativity over time. Second, as shown in Table 1, previous studies on human-GenAI co-creation often use a human-only condition as a comparison group (e.g., Hitsuwari et al. 2023, Jia et al. 2024, Doshi and Hauser 2024). Study 1 adopted the same approach.

The GenAI was a chatbot developed using the GPT-3.5 API (temperature = 1), prompted with "You are going to work with a participant to generate some creative solutions," and hosted on Hugging Face (https://huggingface.co/). This custom-built chatbot, which mirrors ChatGPT (OpenAI 2024) in functionality, was selected for its ability to simulate human-like dialogue. We chose GPT as the GenAI for this study because of its broader conversational applications and advanced functionality (OpenAI 2024). The study concluded with the collection of demographic data and a thorough debriefing. In total, 1,620 data points (162 participants × 10 rounds) were collected. An illustration of the Study 1 design can be found in Supplementary Figure S1.

5.2. Measures

The creativity of the solutions for the 10 tasks (hereafter referred to as 'output creativity' within Study 1)[3] was assessed using the consensual assessment technique (Amabile 1983, Kaufman et

[3] In our paper, we refer to the creativity of an outcome produced jointly by humans and GenAI as "joint creativity." However, in Study 1, we label the dependent variable as "output creativity" because Study 1 includes both human-

al. 2009). We recruited three evaluators who independently rated the novelty and usefulness of each solution on a 7-point Likert scale ('1' = Completely NOT novel/useful, '7' = Completely novel/useful; see Supplementary Text A for evaluation instructions). Inter-rater agreement tests (ICC1, ICC2, and Rwg) were conducted to assess inter-rater reliability and agreement across evaluators, and the results (see Supplementary Table S3) were above the standard recommended by LeBreton and Senter (2008). In cases of low agreement, the three evaluators revisited the corresponding solutions to resolve discrepancies in their evaluations. To calculate the final scores of output creativity, we multiplied the novelty and usefulness ratings, as is standard practice in creativity research (Zhou et al. 2019, Harvey and Berry 2022). In the main analysis, we used this score as our dependent variable.

Lastly, we measured participants' age, gender, their own creativity (i.e., human creativity), daily use of AI (Pok et al. 2022), and trust of AI (Liu 2021), as well as controlling for these variables in our main analyses. Sample items include "I spend most of the time working with AI" for daily use of AI and "If a decision is made by AI, it must be precise" for trust of AI. Human creativity was assessed using the alternative uses test (AUT; Guilford, 1967a). In the AUT, participants generated as many novel and useful uses as possible for a coat hanger within two minutes. Responses were evaluated for creativity using the consensual assessment technique (Amabile 1983, Kaufman et al. 2009) by six evaluators and demonstrated good reliability (ICC1 = 0.39, ICC2 = 0.80), according to the standard proposed by LeBreton and Senter (2008). The inclusion of these controls helps account for individual differences that may impact the output creativity of human-GenAI co-creation. For instance, more creative participants, younger participants, or those more familiar with GenAI may be better equipped to engage creatively

GenAI and human-only conditions. In contrast, in Studies 2 and 3, we use the term "joint creativity" because these studies focus exclusively on human-GenAI co-creation.

with it. Prior research has also shown that women are less likely to use GenAI than men (Otis et al. 2024), so we controlled for gender in our study. Additionally, participants with higher levels of trust of AI may be more likely to accept GenAI-generated ideas without engaging in the co-creation process. By controlling for these variables, we can more accurately assess changes in output creativity—particularly in the human-GenAI condition—without potential confounding influences. Notably, our interpretation of the results did not change when these control variables were removed from our analyses.

5.3. Analytical Strategy

The data had a multilevel structure, comprising both between-individual (L2) and within-individual (L1) levels. The between-individual level (L2) included experimental condition and control variables, while the output creativity of each task was measured at the within-individual level (L1). To validate this structure, we assessed the degree of non-independence (intraclass correlation) and found that the two-level model explained a significant portion of the variance in output creativity (ICC1 = 0.18). This indicates that 18% of the variability in output creativity could be attributed to between-individual differences, while the remaining 82% was explained by within-individual differences. All analyses were conducted in R using the *lmerTest* and *lme4* packages (Kuznetsova et al. 2020, Bates et al. 2024).

In Study 1, we investigate whether human-GenAI co-creation, compared to humans working alone, can lead to continuous improvement in the creativity of their joint outcomes. Specifically, we focus not on the simple effect of experimental conditions on output creativity but rather on their impact on *changes* in output creativity over 10 rounds. To assess these changes across the experimental conditions, we tested the interaction between experimental condition and task round number in predicting output creativity. Supplementary Figure S2 illustrates our analytical model.

5.4. Results

The means, standard deviations, and correlations for all variables in Study 1 are presented in Table 2. Our results showed that the interaction between the experimental condition and task round number in predicting output creativity was significant (b = -0.34, SE = 0.16, p < 0.05; Table 3). This means that the change in output creativity in the human-GenAI condition over 10 rounds was significantly different from the change in output creativity in the human-only condition. A follow-up simple slopes analysis revealed that participants in the human-only condition showed a significant improvement in output creativity over 10 rounds (b = 0.38, SE = 0.11, p < 0.001; Figure 2). However, no such improvement was observed in the human-GenAI condition (b = 0.04, SE = 0.11, p = 0.70; Figure 2).

5.5. Additional Analysis

Furthermore, we conducted additional analyses comparing the two conditions in each round to further understand the lack of improvement in the human-GenAI condition. For the first six rounds, participants in the human-GenAI condition ($M_{R1\text{-}R6}$ = 22.22, $SD_{R1\text{-}R6}$ = 6.26) generated significantly more creative output than those in the human-only condition ($M_{R1\text{-}R6}$ = 19.94, $SD_{R1\text{-}R6}$ = 5.17; $F(1, 160)$ = 6.37, p < 0.05). However, this difference became non-significant starting in the 7th round ($M_{\text{human-GenAI in R7-R10}}$ = 22.55, $SD_{\text{human-GenAI in R7-R10}}$ = 6.67; $M_{\text{human-only in R7-R10}}$ = 22.33, $SD_{\text{human-only in R7-R10}}$ = 5.23; $F(1, 160)$ = 0.06, p = 0.81). These results indicate that while participants in the human-GenAI condition initially demonstrated higher output creativity compared to those in the human-only condition, this advantage did not persist. Due to the lack of creativity improvement in the human-GenAI condition and the creativity gains achieved by participants in the human-only condition, the latter group was able to catch up with the former from the 7th round onward.

5.6. Study 1 Discussion

Results and additional analysis from Study 1 indicate that human-GenAI co-creation has both advantages and disadvantages compared to the human-only condition. In the initial rounds of the ideation tasks, participants co-creating with GenAI generated more creative output than those in the human-only condition. However, this advantage disappeared in later rounds, as participants in the human-only condition demonstrated continuous improvement in output creativity over time, while those in the human-GenAI condition did not show such improvement. This indicates that human-GenAI co-creation has a unique limitation: the lack of improvement in creativity over time, implying that participants in this condition did not spontaneously achieve augmented learning.

The lack of creativity improvement in the human-GenAI condition raises important questions about the underlying dynamics of human-GenAI co-creation, specifically around whether they may inadvertently engage in ineffective co-creation activities that hinder augmented learning. Therefore, the main objective of Study 2 is twofold: first, we examine why humans and GenAI struggle to improve joint creativity, even after multiple rounds of co-creation. Second, we aim to uncover a specific set of co-creation activities that could lead to improvement in joint creativity of human-GenAI co-creation. In Study 2, we focus exclusively on the human-GenAI co-creation condition to enable a deeper exploration of the collaborative dynamics between humans and GenAI. In Study 1, although a human-only control group was not essential for addressing our core research questions, we included it to compare human-only learning with augmented learning in the context of human-GenAI co-creation. With this foundational comparison established, Study 2 concentrated solely on the human-GenAI co-creation condition. Furthermore, we reduced the number of rounds from 10 to 3 in Study 2 to allow participants and GenAI more time per task to engage in richer, more detailed dialogues. Rather than prioritizing the quantity of rounds, our goal was to capture more elaborate interaction patterns within each round, which were critical for identifying nuanced co-creation activities.[4]

6. Study 2

Study 2 consists of three parts, adopting both qualitative and quantitative methods. In Part 1, similar to Study 1, we asked all participants to engage in human-GenAI co-creation over multiple rounds of ideation tasks. The main purposes of Part 1 are to (1) replicate the findings of Study 1, demonstrating no improvement in joint creativity in human-GenAI co-creation over time, and (2) acquire elaborate dialogues between participants and GenAI. In Part 2, we conducted a qualitative analysis of all human-GenAI dialogues to identify emerging co-creation activities adopted by the human-GenAI dyads. In Part 3, building on the activities inductively identified in Part 2, we coded the dialogues accordingly and performed quantitative analyses to investigate: (1) why humans and GenAI fail to enhance joint creativity over time, and (2) what adjustments to the set of co-creation activities might effectively improve joint creativity over time. These insights will directly inform the design of Study 3, where the most effective set of co-creation activities identified across the three parts of Study 2 will be tested experimentally to evaluate its potential to foster augmented learning and continuous improvement in joint creativity within human-GenAI co-creation.

[4] We also conducted a sensitivity analysis using the Study 1 dataset to test the robustness of joint creativity trajectories across fewer rounds in the human-GenAI condition. Even with fewer rounds (e.g., two rounds and three rounds), we still found no evidence that output creativity improves over time. Therefore, the results of our sensitivity analysis support the use of fewer rounds in Study 2. We report the sensitivity analysis in Supplementary Text B.

6.1. Part 1: Quantitative Replication

6.1.1. Sample and Procedure. One hundred sixty-six participants (93 females; M_{age} = 29.53, SD_{age} = 11.94) took part in a one-hour laboratory study, receiving £10 in compensation.[5] Participants were recruited from the Behavioral Laboratory participant pool at a UK university. Upon arrival, each participant was led to an individual booth equipped with a computer. All participants began by reading a tutorial on how to collaborate with GenAI. After completing the tutorial, they were instructed to interact with ChatGPT (OpenAI 2024) to co-generate one creative idea for each of three business-related issues (in-basket tasks, see Supplementary Text D for details; Shalley, 1991). Upon completion of the tasks, participants reflected on their attitudes toward GenAI and provided demographic information. The illustration of our Study 2 design is shown in Supplementary Figure S3. In total, 498 data points (166 participants × three rounds) were collected.

6.1.2. Measure. To evaluate the creativity of solutions generated through human-GenAI co-creation (hereafter referred to as 'joint creativity'), we adopted the consensual assessment technique (Amabile 1983, Kaufman et al. 2009). We recruited three independent evaluators to rate the novelty and usefulness of each solution on a 7-point Likert scale ('1' = Completely NOT novel/useful, '7' = Completely novel/useful) and multiplied them to form the final creativity score (see Supplementary Table S5 for inter-rater reliability scores). This variable was used as the dependent variable in our main analysis. Lastly, we measured participants' age, gender, their own creativity (i.e., human creativity), daily use of AI (Pok et al. 2022), and trust of AI (Liu 2021) to control for the influence of demographics, prior experience, and attitudes toward GenAI

[5] The initial sample consisted of 186 participants. We excluded 20 participants who withdrew from the study midway, those who failed to adhere to study instructions, those who did not complete all three tasks, and those who did not collaborate with GenAI for task completion.

on joint creativity. Human creativity was measured again using AUT (Guilford 1967) and evaluated using the consensual assessment technique (Amabile 1983, Kaufman et al. 2009) by three evaluators and ensured good reliability (ICC1 = 0.59, ICC2 = 0.81). The results remained consistent when these control variables were removed from our analyses.

6.1.3. Analytical Strategy. We conducted a multilevel regression analysis to examine changes in joint creativity of solutions generated by human-GenAI collaboration over time. This approach accounted for both within-individual and between-individual variability. At the between-individual level (L2), we included control variables, while the joint creativity of each task was measured at the within-individual level (L1). To assess the degree of non-independence in joint creativity, we calculated the intraclass correlation coefficient (ICC1) using a one-way random effects model. The ICC1 was 0.54, indicating that 54% of the total variability in joint creativity levels could be attributed to between-individual differences, while the remaining 46% was due to within-individual variability.

6.1.4. Results and Discussion. The results of the multilevel analyses supported our previous findings from Study 1, showing that human-GenAI co-creation failed to produce continuous improvements in joint creativity over time (b = -0.20, SE = 0.27, p = 0.45; Supplementary Table S4; Supplementary Figure S5). These findings reinforce the observation that human-GenAI co-creation does not naturally evolve into higher levels of joint creativity and imply the challenges inherent in achieving augmented learning in human-GenAI co-creation.

6.2. Part 2: Qualitative Analyses and Discovery of Emerging Co-creation Activities

6.2.1. Data Analysis. This lack of improvement in joint creativity over time implies that humans and GenAI failed to identify effective rearrangements of their involvement in co-creation activities across rounds. To further explore this issue, it is necessary to first identify the types of co-creation activities in which humans and GenAI generally engage (Part 2). This will then enable us to investigate what adjustments to the set of co-creation activities are (dys)functional for enhancing joint creativity (Part 3). Therefore, the main goal of Part 2 is to identify emerging types of co-creation activities.

To achieve this goal, we conducted a qualitative analysis of 498 human-GenAI co-creation dialogues collected from three rounds of human-GenAI co-creation by 166 participants in Part 1. On average, participants and GenAI engaged in 6.13 interactions (SD = 3.35) per task, which yielded sufficient data for our qualitative analysis. Given the absence of studies on the evolution of human-GenAI co-creation over time, we aimed to allow the set of co-creation activities to emerge organically from the data, without the constraints of predefined theoretical frameworks. Following the inductive qualitative approach outlined by Corbin and Strauss (2008), we began our analysis with a comprehensive review of all collected dialogues. Three researchers—two authors and one research assistant (RA)—independently reviewed the dialogues. The research assistant was blind to our research questions and study design, which helped mitigate bias and promote objective observation. Each researcher employed open coding to categorize emerging themes without prior assumptions. Through this process, patterns reflecting distinct co-creation activities emerged, offering initial insights into how humans and GenAI co-created solutions. After this independent coding, the researchers convened to compare findings and discuss the co-creation activities identified. These discussions, informed by the literature on the creative process (Amabile 1983, Perry-Smith and Mannucci 2017), allowed us to extrapolate overarching co-creation activities.

6.2.2. Initial Findings of Co-creation Activities. As a result of our initial coding to discover emerging co-creation activities, we found 11 co-creation activities, such as “Proposing new ideas", "Requesting GenAI to generate new ideas", and "Expanding on a generated idea". In the column of "Initial Codes of Co-creation Activities" in Table 5, we provide the full list of these initially discovered co-creation activities with examples.

An interesting finding was that both participants and GenAI were actively adapting to and influencing each other. Human participants were proactive in seeking out and receptive to GenAI's ideas and opinions. While some research suggests that people often disregard or downplay information provided by GenAI (Zhang and Gosline 2023, Turel and Kalhan 2023), our study found no evidence of such discrimination against GenAI-generated information. Instead, participants actively sought GenAI's ideas, evaluations, and feedback, making deliberate efforts to incorporate GenAI-provided insights into their final solutions. For instance, in one exchange, a participant asked, "*What do you think is the best option?*"—demonstrating an openness to GenAI's evaluative role. In another case, a participant revised their proposed solution, incorporating GenAI's feedback, stating, "*Here is an amended solution: I will strengthen security with more cameras and security staff, who are rewarded for identifying thieves among employees.*"

In addition, GenAI demonstrated careful consideration of human requests. Based on its understanding, it made significant creative contributions throughout the process of human-GenAI co-creation. As noted in our theory of augmented learning, and consistent with the definition of GenAI (i.e., a type of AI that generates content in response to human prompts, based on patterns learned from data), the decision about which types of co-creation activities to pursue is primarily made by humans (Raisch and Krakowski 2021). While human initiation is required to begin these activities, once initiated, GenAI contributes significantly by providing useful and often critical information, ultimately leading to the generation of better solutions in human-GenAI co-creation. In fact, we observed several instances where GenAI provides critical feedback to human participants' ideas. For example, in one interaction focused on addressing employee theft, a participant suggested holding employees financially liable for monetary repercussions and asked GenAI for feedback. The GenAI disagreed, noting that this idea was not particularly creative, and instead proposed an alternative: a positive incentive program to foster a culture of awareness. The participant built upon this concept, further developing ideas centered around cultivating such a culture. In other cases, participants introduced an idea, which GenAI then evaluated and further developed, ultimately shaping the final solution. For example, in one interaction, a participant suggested enforcing a code of conduct as a strategy for addressing employee theft. They then asked GenAI, *"Could you please provide me with a detailed solution about the code of conduct?"* The GenAI's response helped refine and elaborate on the proposed approach. These examples demonstrate that while the interactions are initiated by humans, GenAI's responses are neither submissive nor passive. Instead, GenAI actively contributes to the co-creation process.

In summary, we found that, given the definition and current design of GenAI, human participants primarily initiate the specific types of co-creation activities. However, once a co-creation activity is initiated, both human participants and GenAI become actively interdependent, working together to generate final solutions for the ideation tasks.

6.2.3. Final Set of Co-creation Activities. After the initial findings of 11 emerging co-creation activities, we drew on the creativity literature to determine whether these co-creation activities can be meaningfully consolidated into broader categories. As a result, we found four broader co-creation activities (see the column of "Final Set of Co-creation Activities" in Table 5). These activities were broadly aligned with existing literature on creativity but were also adapted to the unique context of human-GenAI co-creation.

The first co-creation activity, *Idea Generation-Response*, refers to a co-creation activity in which humans initially generate new ideas and propose them to GenAI, which then offers responses to these ideas (Guilford 1967). For instance, a participant said, *"Here is an idea I have..."* to which GenAI replied, *"Your idea is a good starting point…"*[6] Moreover, GenAI memorized the human-generated ideas during subsequent interactions within each round. Upon the participant's request, GenAI effectively incorporated these ideas into later interactions.

A second co-creation activity identified was *Idea Request-Idea Generation*, where humans request GenAI to generate new ideas, and GenAI generates new ideas. For example, participants asked GenAI, *"Can you suggest 10 ideas for this problem?"* to which GenAI responded, *"Here are 10 ideas…"*

Both *Idea Generation-Response* and *Idea Request-Idea Generation* involve idea generation, but they differ regarding the source of the idea. In *Idea Generation-Response*, the human provides the idea, whereas in *Idea Request-Idea Generation*, the GenAI does. We separated these two co-creation activities in light of previous studies showing heterogeneous effects of human-generated versus GenAI-generated ideas on creativity (e.g., Koivisto and Grassini 2023, Magni et al. 2023, Guzik et al. 2023, Haase and Hanel 2023, Hubert et al. 2024). These studies show that the two sources—human versus GenAI—produce qualitatively distinct ideas, which may ultimately lead to differences in creative outcomes. They have demonstrated that human-generated ideas are often more creative than GenAI-generated ideas (e.g., Magni et

[6] Note that when humans simply generated their own ideas and presented them to GenAI, GenAI's responses were generally positive. However, when humans not only generated their own ideas but also exchanged feedback with GenAI, the GenAI became highly evaluative. In our coding, the former scenario was categorized as *Idea Generation-Response*, while the latter was coded as both *Idea Generation-Response* and *Idea Co-development*.

al. 2023), although GenAI could sometimes generate more creative ideas under certain conditions (e.g., Guzik et al. 2023, Hubert et al. 2024). Given the distinct effects associated with the idea sources, we treated these two forms of co-creation activities separately.[7]

The third co-creation activity, *Idea Co-development*, refers to a co-creation activity in which humans and AI engage in critical feedback exchanges and the joint refinement of generated ideas (Lubart 2005, Cropley 2006, Grønsund and Aanestad 2020). We identify two tightly interwoven sub-activities of *Idea Co-development*: feedback and idea refinement. While distinct in function, these sub-activities are inherently interdependent and operate in tandem during *Idea Co-development*—e.g., feedback naturally gives rise to idea refinement, and idea refinement, in turn, prompts further feedback.

Feedback, in this context, refers to co-creation activities involving the exchange of critical information that assesses how well an idea aligns with standards of creativity (Kim and Kim 2020). It includes the evaluation of generated ideas and the sharing of insights about their strengths and limitations. For example, a participant stated, "*Idea 1 seems more innovative, but Idea 2 is more practical*," and GenAI responded, "*That's true—maybe we build on Idea 2 first.*" Feedback can also be explicitly solicited, as in the case of a participant asking, "*What do you think of this revised idea?*" and GenAI replying, "*This idea is better, but...*"

Idea refinement refers to co-creation activities aimed at strengthening or further developing a generated idea through elaboration, modification, or integration with other ideas (McMahon et al. 2016, Perry-Smith and Mannucci 2017). In our analysis, we identified five initial codes of co-creation activities that fall under idea refinement. First, expanding on a

[7] In our "6.3. Part 3: Quantitative Analyses," we also conducted a multilevel mediation analysis by combining *Idea Generation-Response* and *Idea Request-Idea Generation* into a single "idea generation" variable. However, this analysis did not change the interpretation of our findings. Specifically, neither of the two co-creation activities individually showed a significant relationship with joint creativity, nor did the combined idea generation variable.

generated idea refers to co-creation activities that further develop or elaborate on an idea by adding specific features, descriptions, or clarifications to enhance it (e.g., *"Please expand on this idea"* with a response, *"Here's a more detailed version of the idea"*). Second, integrating different ideas refers to co-creation activities that combine multiple distinct ideas into a cohesive idea, retaining key elements from each original idea (e.g., "*How about we combine ideas 2 and 3?*" with a response, "*Here is a combined idea*"). Third, reframing a generated idea refers to co-creation activities that change how an idea is presented or interpreted, such as shifting its perspective, tone, or intended audience, without modifying its core content (e.g., "*What if we frame this as employee empowerment instead of surveillance?*" with a response, "*That would make it more appealing*"). Fourth, streamlining a generated idea refers to co-creation activities simplifying an idea by removing unnecessary components or reducing complexity, making it clearer, more efficient, or easier to understand (e.g., "*We should cut down a few parts of the idea to make it clearer*" followed by "*Okay, here's a more concise version*"). Finally, setting constraints on a generated idea refers to co-creation activities that revise an idea to ensure it aligns with specific constraints or guidelines ("*Let's make sure this idea stays cost-effective*" with a response of *"We'll need to use cheaper machines*").

Feedback and idea refinement are fundamentally interdependent activities that work in tandem to enable effective creative collaboration between humans and GenAI. They rarely function as independent activities in human-GenAI co-creation. Feedback offers evaluative direction, indicating whether an idea meets standards of creativity, while idea refinement leverages this guidance to improve a generated idea through elaboration, revision, or integration. When separated, each activity risks becoming ineffective. Feedback without idea refinement results in unproductive critique: evaluations may identify shortcomings or offer suggestions, but without follow-through, the creative process stalls. On the other hand, idea refinement without feedback risks becoming misguided or directionless. Without evaluative input to clarify which ideas are worth developing or how they might fall short of creative standards, idea refinement efforts may lead to trivial modifications, overelaboration, or a drift away from the creative process. This interdependence is well supported in the creativity literature, which consistently frames creative work as an iterative process involving cycles of idea evaluation, feedback, and refinement (Csikszentmihalyi 1997, Amabile and Pratt 2016). In this view, feedback functions as a critical input that triggers the evolution of ideas via idea refinement (Ashford et al. 2003, Hargadon and Bechky 2006). Idea refinement involves creative problem-solving to act on that feedback, transforming initially abstract ideas into more complete and viable ones (Paulus and Nijstad 2003, Perry-Smith and Mannucci 2017). Rather than operating in a linear sequence, feedback and refinement typically unfold in recursive loops, where each feeds and reshapes the other over time (Amabile and Pratt 2016). This interdependence underscores why we conceptualize feedback and idea refinement not as separate co-creation activities, but as inherently interdependent sub-activities that together constitute the *Idea Co-development* activity.

These dynamics were evident in our empirical data. In one instance, a participant provided feedback on a proposed whistleblower program, stating, "*I'm afraid the whistleblower program, while helpful, might erode the confidence the employees have on the company and on themselves. It still feels like a very useful option though.*" The participant then invited a reframing of the idea to address this concern: "*Can you provide ideas on how to implement it such that employees see it as a benefit and not as a threat?*" GenAI responded by transforming the original idea into a more positively framed alternative: "*Here are some ideas on how to* *implement a whistleblower program in a way that employees see it as a benefit rather than a threat.*" This interaction exemplifies how evaluative input triggered the reworking of the idea into a more acceptable and potentially creative direction. In another case, a participant expressed a preference for specific GenAI suggestions—"*I'm interested a bit more in options number 4 and 5*"—and invited elaboration: "*Do you think you could expand on them a little bit?*" GenAI proceeded to develop these options further: "*Sure, here are some ideas to expand on options 4 and 5.*" Here, the participant's feedback served as a selective filter, guiding GenAI to refine the most promising ideas. In both cases, feedback directly shaped refinement, and refinement in turn opened space for further evaluation, illustrating the bidirectional, integrated nature of the co-development activity.

The final co-creation activity, *Miscellaneous Activity*, involved interactions that helped maintain engagement in the dialogue but were not directly related to the creative output. These included simple exchanges such as greetings (e.g., saying "Hi" to each other) and expressions of gratitude.

In summary, our qualitative analysis of 498 human-GenAI co-creation dialogues identified three core co-creation activities: (1) *Idea Generation-Response*, where humans propose new ideas and GenAI responds; (2) *Idea Request-Idea Generation*, where humans request GenAI to generate new ideas; and (3) *Idea Co-development*, where humans and GenAI collaboratively exchange feedback and refine generated ideas. In addition, we identified a category of *Miscellaneous Activity*, which maintains the flow of dialogue without directly contributing to creative outcomes.

6.2.4. Discussion. Through our inductive analysis, we identified three key co-creation activities that contribute to the joint creative process: *Idea Generation-Response, Idea Request-Idea*

Generation, and *Idea Co-development*. Participants and GenAI actively engaged in these activities and contributed to a dynamic co-creation process. Unlike previous studies that showed people sometimes downplay information provided by GenAI (Zhang and Gosline 2023, Turel and Kalhan 2023), our participants proactively interacted with GenAI and were receptive to GenAI's inputs. GenAI also displayed active and critical engagement throughout the co-creation process. In the next part, we quantitatively analyze whether the qualitatively identified co-creation activities explain the lack of improvement in joint creativity over time during human-GenAI collaboration. Furthermore, we assess whether any specific co-creation activities can help humans and GenAI improve joint creativity over time.

6.3. Part 3: Quantitative Analyses to Discover Limitations of Human-GenAI Co-creation in Achieving Augmented Learning and Identify Effective Co-creation Activities

6.3.1. Analytical Strategy. In Part 3, our objectives are twofold: to understand why human-GenAI co-creation generally does not lead to improvements in joint creativity over time and to discover effective co-creation activities that can help humans and GenAI improve their joint creativity. To attain these goals, we employed a two-step approach. In the first step, we implemented a structured coding process to operationalize four variables, each representing one of the identified co-creation activities. Three researchers independently coded the dialogues using the four identified co-creation activities as a framework. Each dialogue was systematically reviewed to classify each human-GenAI interaction into one(s) of the identified four co-creation activities, as multiple activities could occur within a single interaction. Regular meetings were held to calibrate the coding scheme, ensuring consistency and reliability in the coding process. Based on the coding results, we then created variables representing the relative percentage of each co-creation activity. This was calculated as the number of instances of each co-creation activity divided by the total number of interactions.

In the second step, we conducted statistical analyses. Using the four variables created, we conducted a multilevel mediation analysis with 10,000 bias-corrected bootstrap resampling (Hayes 2013) to explore why human-GenAI co-creation did not improve joint creativity over three rounds. For this, we used task round number as the independent variable, the three key co-creation activities as mediators, and joint creativity as the dependent variable. This approach enabled us to assess how the set of four co-creation activities evolved over time and how these changes impacted joint creativity across rounds. In this analysis, we controlled for *Miscellaneous Activity*, as it does not directly contribute to the creativity process. Supplementary Figure S4 illustrates our analytical model.

6.3.2. Result. The means, standard deviations, and correlations for all variables in Study 2 are presented in Table 4. Before conducting our multilevel mediation analysis, we first examined each link individually. Specifically, we tested (1) the effects of ideation task rounds on the three key co-creation activities and (2) the effects of the three key co-creation activities on joint creativity. In the first set of analyses, we found that as the ideation task rounds progressed, *Idea Co-development* significantly decreased (b = -0.05, SE = 0.01, p < 0.001; Table 6), while *Idea Generation-Response* (b = 0.01, SE = 0.01, p < 0.05) and *Idea Request-Idea Generation* (b = 0.03, SE = 0.01, p < 0.01) significantly increased over time. This indicates that participants and GenAI reduced their engagement in *Idea Co-development* across rounds but increased levels of other co-creation activities across rounds. In the second set of analyses, results showed that *Idea Co-development* significantly increased joint creativity (b = 2.77, SE = 1.28, p < 0.05; Table 6), while *Idea Generation-Response* (b = -0.46, SE = 2.09, p = 0.83) and *Idea Request-Idea*

Generation (b = -0.77, SE = 1.73, p = 0.66) did not have significant impacts on joint creativity. Overall, these results indicate that participants and GenAI progressively reduced their use of *Idea Co-development*, which was the only co-creation activity with a positive influence on joint creativity over time.

Our multilevel mediation analysis confirmed the findings from the discrete analyses, showing that the primary reason for the absence of continuous improvement in joint creativity over rounds was the decline in *Idea Co-development* over time. The results indicated that over time, participants and GenAI reduced their engagement in *Idea Co-development*, which in turn reduced joint creativity over time (b = -0.14, 95% CI [-0.31, -0.01], $p < 0.05$). Neither *Idea Generation-Response* (b = -0.01, 95% CI [-0.08, 0.06], p = 0.83) nor *Idea Request-Idea Generation* (b = -0.02, 95% CI [-0.12, 0.07], p = 0.65) significantly mediated the relationship between task round number and joint creativity. These results suggest that collaborative idea generation alone (i.e., *Idea Generation-Response* or *Idea Request-Idea Generation*), without subsequent refinement of the generated ideas (i.e., *Idea Co-development*), does not significantly contribute to improvements in joint creativity over time.[8]

6.4. Study 2 Discussion

The findings from the three parts of Study 2 provide insights into the key issue of the lack of augmented learning in human-GenAI co-creation. Our analysis revealed that co-creation activities focused solely on generating new ideas, whether initiated by humans (*Idea Generation-Response*) or by GenAI upon request (*Idea Request-Idea Generation*), do not significantly

[8] After completing Study 2, we applied the same coding and statistical analyses to the Study 1 dataset and observed consistent findings: (1) *Idea Co-development* decreased over time in the human-GenAI condition, (2) *Idea Co-development* positively impacted joint creativity, and (3) *Idea Co-development* mediated the relationship between time and joint creativity. These results again demonstrated that the lack of improvement in joint creativity over time was due to the decreased level of *Idea Co-development*. The complete mediation analysis is provided in Supplementary Tables S1 and S2.

improve joint creativity over time. Instead, it is *Idea Co-development*, a collaborative activity involving further development of already generated ideas via both feedback and iterative idea refinement, that significantly contributes to continuous improvement in human-GenAI creative performance. Compared to existing creativity literature (Amabile 1983, Perry-Smith and Mannucci 2017), our findings show both convergence and divergence. Consistent with prior work, we found that idea generation and development are both critical components of creative collaboration.

The main difference between our findings and the creativity literature is the disproportionate importance of *Idea Co-development*. While the literature does not explicitly differentiate the importance of idea generation and idea development, it often implicitly assumes that idea generation holds greater importance. This assumption likely results from the recognition that generating diverse and numerous ideas is one of the most challenging tasks for humans during the ideation process (Anderson et al. 2014). Consequently, idea generation has received much more attention in the literature than idea development (Zhou et al. 2019). In the context of human-GenAI co-creation, however, our findings suggest that humans and GenAI should engage in *Idea Co-development* more than in the other two co-creation activities related to idea generation.

Importantly, we do not argue that idea generation becomes unimportant in human-GenAI co-creation. It remains a crucial activity, as the ideation process relies on an initial pool of generated ideas (Berg 2019). What remains more critical is how humans and GenAI collaboratively develop the generated ideas. By investing effort into exchanging feedback and refining ideas—activities constituting *Idea Co-development*—human-GenAI dyads can progressively improve their creative outcomes. This insight signals the necessity of a shift in the roles of both human and GenAI actors: rather than being most valuable as generators of ideas, they contribute most meaningfully by actively refining and steering the creative trajectory through engaged, iterative co-development.

The insights gained from Study 2 provide a foundation for Study 3, where we will experimentally test the effectiveness of *Idea Co-development* as a targeted co-creation activity to improve joint creativity in human-GenAI co-creation. By instructing participants on how to employ *Idea Co-development* in their co-creation process, Study 3 aims to validate whether this structured, strategic instruction for human-GenAI co-creation can overcome the limitations observed in Studies 1 and 2 and ultimately enhance joint creativity.

> **Hypothesis**: The provision of instructions on *Idea Co-development* will enhance joint creativity in human-GenAI co-creation over time. Specifically, human-GenAI dyads that receive the instruction will exhibit significantly greater improvement in joint creativity across rounds compared to dyads that do not receive the instruction.

7. Study 3

7.1. Sample and Procedure

A total of 166 participants (81 females; M_{age} = 40.16, SD_{age} = 13.34) were recruited through Prolific (www.prolific.com).[9] The study was advertised as focusing on GenAI, and participants received £2 for their participation. After providing informed consent, participants were asked to generate a creative solution for a first ideation task with GenAI. Following this first task, participants were randomly assigned to one of two experimental conditions: a treatment condition (N = 81) or a control condition (N = 85). In the treatment condition, participants

[9] The initial sample consisted of 200 participants. We excluded participants who failed to adhere to study instructions, participants who did not complete both tasks, and participants who did not collaborate with GenAI for task completion.

received instructions on the advantages of *Idea Co-development* and how to apply it to enhance joint creativity with GenAI. The instructions presented to the treatment group can be found in Supplementary Text E. The instructions emphasized two core components of *Idea Co-development*: feedback and idea refinement. These components were presented as key activities for enhancing joint creativity in human-GenAI co-creation. After reading the instructions, participants were asked to complete a second ideation task with GenAI. In the control condition, participants received no additional instructions and proceeded directly to the second ideation task. The study concluded with the collection of demographic information and a thorough debriefing. An illustration of the Study 3 design can be found in Supplementary Figure S6.

7.2. Experimental Task and Measures

For the ideation tasks, we selected two topics from the list of societal and environmental issues used in Study 1: chemical waste management and clothing recycling policy. The task order was randomized to mitigate any potential influence of the order on joint creativity. Our analysis showed that the task order did not influence our results (b = -1.52, SE = 1.11, p = 0.17). The GenAI chatbot used in the experiment was developed with the GPT-4o-mini API (temperature = 1), prompted with "You are going to work with a participant to generate some creative solutions," and hosted on Hugging Face (https://huggingface.co/). Participants in the treatment group ($M_{\text{interactions per task}}$ = 3.48, $SD_{\text{interactions per task}}$ = 2.02) did not differ in the number of interactions they had with GenAI compared to the control group ($M_{\text{interactions per task}}$ = 3.15, $SD_{\text{interactions per task}}$ = 1.95; $F(1, 161)$ = 1.50, p = 0.22).

Joint creativity was assessed using the consensual assessment technique (Amabile 1983, Kaufman et al. 2009), following the steps outlined in Studies 1 and 2. We recruited three independent evaluators, and their evaluation demonstrated good reliability (for task 1, ICC1 =

0.90, ICC2 = 0.96; for task 2, ICC1 = 0.82, ICC2 = 0.93) according to the standard proposed by LeBreton and Senter (2008). Since participants were randomly assigned to either the treatment or control condition, we did not control for individual differences or demographic variables, as these factors were accounted for by the randomization design. However, controlling for these variables—i.e., participants' age, gender, daily use of AI (Pok et al. 2022), and trust of AI (Liu 2021)—did not alter the interpretation of our results. Additionally, we controlled for participants' boredom (van Tilburg and Igou 2012) and intrinsic motivation (Grant 2008) to address the possibility that boredom and a lack of intrinsic motivation could serve as alternative explanations for the lack of improvement in joint creativity.[10] Sample items included "I felt restless and unchallenged" for boredom and "I was motivated to work with GenAI because I find the work with GenAI engaging" for intrinsic motivation. Our results remain consistent after controlling for these variables.

7.3. Analytical Strategy

The data were structured at two levels: between-individual (L2) and within-individual (L1). At the between-individual level (L2), we included experimental condition ('0' = control group; '1' = treatment group) and control variables. Joint creativity was measured at the within-individual level (L1). To validate the multilevel structure, we calculated the intraclass correlation coefficient (ICC1 = 0.09), indicating that 9% of the variability in joint creativity was due to between-individual differences and 91% to within-individual differences.

7.4. Results

[10] We thank the reviewers and the associate editor for highlighting this issue in our previous manuscript. In our Study 3 analysis, neither boredom ($b = 0.19$, $SE = 0.42$, $p = 0.65$) nor intrinsic motivation ($b = 0.60$, $SE = 0.54$, $p = 0.28$) significantly influenced joint creativity.

The means, standard deviations, and correlations for all variables in Study 3 are presented in Table 7. To ensure our manipulation was effective, participants were asked to indicate to what extent they recalled receiving instructions on how to engage in *Idea Co-development* for improving creativity. Participants in the treatment condition reported a higher level of recall than participants in the control condition ($t(112.5) = -6.82$, $p < 0.001$). In addition, one of the authors and one RA coded the occurrences of *Idea Co-development* activities, analyzing entire dialogues produced by all human-GenAI dyads. Compared to dyads in the control condition ($M = 1.20$, $SD = 1.52$), human-GenAI dyads in the treatment condition engaged in more *Idea Co-development* activities in Task 2 after receiving instructions ($M = 2.47$, $SD = 1.95$, $t(147.21) = -4.60$, $p < 0.001$).

To examine changes in joint creativity across task rounds in both experimental conditions, we tested the interaction between condition and task round number in predicting joint creativity. The interaction was significant ($b = 3.97$, $SE = 1.92$, $p < 0.05$; Table 8), indicating that the relationship between task round number and joint creativity differed by condition. A simple slopes analysis revealed that participants in the treatment condition showed a significant improvement in joint creativity from Task 1 to Task 2 ($b = 2.74$, $SE = 1.37$, $p < 0.05$; Figure 3). However, no significant improvement was observed in the control condition ($b = -1.24$, $SE = 1.34$, $p = 0.36$; Figure 3). Furthermore, at Task 1, joint creativity in both conditions was not different ($b = -0.71$, $SE = 1.43$, $p = 0.62$; Figure 3). Yet at Task 2, joint creativity in the treatment condition was significantly higher than in the control condition ($b = 3.26$, $SE = 1.43$, $p < 0.05$; Figure 3), indicating that the instruction on *Idea Co-development* effectively enhanced joint creativity.

Although not hypothesized, we tested whether *Idea Co-development* mediated the effect of the instructions on joint creativity. A mediation analysis suggested that the effect of the instruction on joint creativity was partially mediated by increased *Idea Co-development*. The indirect effect was marginally significant (indirect effect = 1.49, p = .06), accounting for approximately 35% of the total effect, while the direct effect was not statistically significant after accounting for the indirect effect (direct effect = 2.80, p = .18). These results suggest that increased *Idea Co-development* is a key mechanism underlying the observed improvement in joint creativity.

8. General Discussion and Directions of Future Research

Our research provides critical insights into the creative process of human-GenAI co-creation, highlighting both its potential and its limitations. Across Studies 1 and 2, we observed that while human-GenAI co-creation initially generated more creative outputs than human-only ideation, it failed to sustain the improvement in creativity over time. This stagnation resulted primarily from a lack of augmented learning. Specifically, we found that *Idea Co-development*—a co-creation activity in which humans and GenAI engage in feedback exchanges and iterative refinement of generated ideas—was particularly beneficial. However, humans and GenAI failed to increase their engagement in this activity over successive rounds, leading to the observed stagnation. In Study 3, an intervention targeted at *Idea Co-development* demonstrated that instructing participants on the importance of *Idea Co-development* significantly improved their joint creativity over rounds. These findings show the need to rethink how human-GenAI co-creation should be structured, emphasizing the importance of iterative and reflective *Idea Co-development* activity in augmented learning.

8.1. New Theorization of Augmented Learning in Human-GenAI Co-creation

Based on our findings, we propose a novel theory of augmented learning, reconceptualizing it to emphasize the need to move beyond its traditional definition. Our theory suggests that achieving augmented learning is not a spontaneous process but instead requires deliberate interventions—such as introducing an exemplary template—to stimulate and guide the collective learning process between humans and GenAI. In the following sections, we elaborate on this new theory, outline its core tenets, and propose directions for future research.

8.1.1. Reconceptualization of Augmented Learning. Our research reconceptualizes augmented learning *as an evolutionary process in which humans and GenAI continuously adjust their levels of involvement in multiple co-creation activities within a task over time to achieve ongoing joint creativity enhancement*. The original concept of augmented learning, introduced by Engelbart (1962), focused on human cognitive learning aided by technology, treating technology as a supportive tool that enhanced human knowledge acquisition by making existing information more accessible. His definition described augmented learning as a cognitive process within an individual human and did not fundamentally differ from general human learning, apart from the assistance provided by technology (Thees et al. 2020).

The advent of GenAI, however, has transformed technology from a passive assistant to a potential co-creator (Anthony et al. 2023). Although GenAI, by definition, must be initiated by humans, once collaboration begins, it can actively participate in tasks and decision-making. In this way, GenAI moves beyond the traditional role of technology as a mere tool for supporting human cognitive learning. Therefore, we propose that the focus of augmented learning between humans and GenAI should shift from an intrapersonal cognitive phenomenon to a collective, intersubjective process. In this reconceptualization, augmented learning can be understood through *collective learning* between humans and GenAI. According to the team literature, specifically the research on transactive memory systems (Wegner 1987, Lewis and Herndon 2011), collective learning is no longer understood solely as a cognitive process occurring within an individual. Instead, it involves the dynamic rearrangement of team members' roles and involvement in team activities based on a shared understanding of each member's strengths and weaknesses. Our theory suggests augmented learning is a specific form of collective learning unique to the collaborative interactions between humans and GenAI.

Our research further suggests that achieving augmented learning requires broadening both the spatial and temporal scales of human-GenAI interactions (Raisch and Krakowski, 2021). Instead of engaging in the traditional dilemma between automation (delegating the entire task to AI) and augmentation (collaborating on the task as a whole), we argue that the whole task should be broken down into distinct, manageable co-creation activities, thereby broadening the spatial scale. This decomposition allows for varying levels of involvement from humans and GenAI across different activities, enabling more flexible and effective human-GenAI co-creation. In parallel, our research emphasizes the importance of multiple rounds of engagement over time (expanding the temporal scale), as understanding each other's strengths and limitations cannot be achieved in a single interaction. Through repeated rounds of collaboration, humans and GenAI can iteratively adjust their respective roles, gradually discovering more effective arrangements of co-creation activities. Over time, these adaptive adjustments can enhance joint task performance. In essence, the expansion of these spatial and temporal scales forms a foundation for enabling augmented learning.

Therefore, our definition of augmented learning incorporates two essential components: collective learning between humans and GenAI, and the expansion of spatial and temporal scales. We believe this definition can more effectively guide future research efforts toward optimizing the complementary strengths of humans and GenAI while addressing their respective weaknesses, ultimately improving the outcomes of human-GenAI co-creation over time.

8.1.2. Unveiling Co-creation Activities of Augmented Learning. Building on our new conceptualization of augmented learning, we aimed to investigate whether human-GenAI co-creation can naturally lead to augmented learning and, in turn, support the continuous improvement of joint creativity over time. Because augmented learning fundamentally involves the dynamic rearrangement of human and GenAI involvement across different co-creation activities over time, our first step was to identify the range of co-creation activities that occur between humans and GenAI during their creative collaboration process. Given the lack of prior theoretical and empirical evidence on these co-creation activities, we adopted an inductive approach to identify various emerging co-creation activities during the creative process. As a result, we identified three final categories of co-creation activities closely related to the joint creative process (excluding *Miscellaneous Activity*): *Idea Generation-Response*, *Idea Request-Idea Generation*, and *Idea Co-development*. Building on our discovery of these co-creation activities, we examined the mechanisms underlying augmented learning in human-GenAI co-creation, which will be discussed in the next section.

Although these co-creation activities were discovered to understand augmented learning, future research needs not limit their application to this context alone. One promising application of our findings may pertain to research on GenAI discrimination, where humans systematically devalue ideas and artworks generated by GenAI (Dietvorst et al. 2015, Zhang and Gosline 2023, Horton Jr et al. 2023, Turel and Kalhan 2023). While most studies on GenAI discrimination focus solely on human reactions to a final GenAI-generated idea, the reactions elicited during the ongoing human-GenAI co-creation activities could differ substantially. For example, how humans respond to GenAI-generated ideas presented during one co-creation activity (e.g., *Idea Co-development*) may not mirror their reactions to ideas introduced during another co-creation activity (e.g., *Idea Generation-Response*). In other words, examining human responses across the entire co-creation process—and at various co-creation activities—may provide valuable new insights into how and when GenAI discrimination emerges. We encourage future research to explore and apply the discovered co-creation activities to these and other domains.

8.1.3. Difficulty in Achieving Augmented Learning and the Importance of Idea Co-development. One of our most intriguing findings is that augmented learning does not occur automatically. Merely introducing GenAI into the creative process is insufficient for the continuous improvement of joint creativity over time. This indicates a lack of spontaneous augmented learning. The failure of augmented learning may be partly due to the difficulty humans have in recognizing GenAI's strengths and limitations. Research in social psychology has shown that people possess an innate desire to understand the causes and processes underlying others' behaviors (Ybarra 2002, Epley et al. 2007). In the absence of such understanding, people experience heightened uncertainty about others' future actions, hindering effective interactions (FeldmanHall and Shenhav 2019). Given that GenAI's information-processing mechanisms often operate as a "black box," humans frequently struggle to comprehend GenAI's thinking processes or predict its future behavior (Cadario et al. 2021). This challenge likely explains why humans find it difficult to recognize GenAI's strengths and limitations during interactions. To address this challenge, providing humans with a clear, effective template for starting their co-creation process with GenAI can be highly beneficial. Hence, our Study 3 introduced such a template, designed to reduce the perceived ambiguity of interacting with GenAI. Our results revealed that the provision of this initial template significantly enhanced joint creativity of human-GenAI co-creation compared to a control group where participants did not receive the template.

More specifically, our research identifies *Idea Co-development* as a critical co-creation activity that has significant potential for improving joint creativity in human-GenAI collaboration. *Idea Co-development* involves feedback exchanges and the iterative refinement of generated ideas through dynamic interactions between humans and GenAI. Findings from Study 2 showed that *Idea Co-development* significantly contributes to continuous improvement in human-GenAI creative performance compared to *Idea Generation-Response* and *Idea Request-Idea Generation*, and Study 3 revealed that providing instructions on *Idea Co-development* to human-GenAI dyads led to improvement in joint creativity over time. Supplementary Tables S7 and S8 present exemplary dialogues between humans and GenAI across rounds, illustrating cases where joint creativity improved and declined over time. Specifically, Supplementary Table S7 provides full dialogues from a human-GenAI dyad in Study 2 that experienced a decline in joint creativity over time. At Round 1, the human and GenAI first engaged in *Idea Request-Idea Generation*, and quickly moved to co-developing a generated idea. By Round 3, however, the collaboration shifted predominantly toward *Idea Request-Idea Generation*, leading to a wider but less in-depth idea pool. This imbalance prevented deeper idea refinement and demonstrated the human-GenAI dyad's struggle to adjust their levels of involvement effectively in co-creation activities. Supplementary Table S8 provides full dialogues from a human-GenAI dyad in Study 3 that experienced an increase in joint creativity over time. In Round 1, the human and GenAI only engaged in the co-creation activity of *Idea Request-Idea Generation* multiple times. However, in Round 2, the human-GenAI dyad significantly reduced idea generation activities and focused more on *Idea Co-development*. The human participant requested an idea and then refined the idea with GenAI. The participant further refined the idea by providing critical feedback, fostering deeper collaboration, and producing a more developed outcome. Overall, these results show that engaging in *Idea Co-development* is one of the effective strategies for augmented learning between humans and GenAI.

This finding also offers new insights into existing paradigms in creativity research, which often emphasize the importance of divergent thinking—the generation of a wide array of ideas—based on the premise that humans struggle to generate a high volume of diverse ideas (Runco 2004). However, in the context of human-GenAI co-creation, the generation of numerous ideas can now be achieved rapidly through both human and GenAI contributions, owing in part to GenAI's powerful generative capabilities. Moreover, a recent meta-analysis has shown that the quality of GenAI-generated ideas is comparable to that of human-generated ideas (de Rooij and Biskjaer 2025). As a result, the creative bottleneck shifts from idea generation to idea development in human-GenAI co-creation. A more effective co-creation strategy may therefore involve strengthening collaborative efforts in *Idea Co-development*, where humans and GenAI work together to evaluate, refine, and improve generated ideas. Emphasizing *Idea Co-development* allows humans to leverage GenAI's generative capabilities while applying their unique insights, contextual understanding, and critical thinking skills to enhance their joint creativity. The importance of *Idea Co-development* thus highlights the necessity of keeping humans in the loop, as humans' unique strengths still play a critical role in human-GenAI co-creation (Grønsund and Aanestad 2020). We believe that, even in the future, our findings about the crucial role of humans will remain relevant, because humans possess superior insights and understanding of social values, ethics, human needs and desires, and other contextual information critical to the success of new ideas.

8.2. Theoretical Implications to the Current Literature on Human-GenAI Co-creation and Joint Creativity

We believe that our theory of augmented learning offers a theoretical explanation for the inconsistencies in the current literature on human-GenAI co-creation for joint creativity. Our literature review highlights that prior research has reported mixed findings regarding the effectiveness of human-GenAI co-creation on joint creativity (see Table 1). Notably, these studies have predominantly focused on single-round creative tasks, where humans and GenAI had only one opportunity to interact. However, single-round tasks provide only momentary snapshots of human-GenAI interactions and may conflate serendipitous advantages or disadvantages in finding effective arrangements of human and GenAI involvement in co-creation activities. Since humans and GenAI are likely to start with random strategies in their initial interactions, single-round studies risk producing random results, which could explain the inconsistencies observed in prior research. Our theory of augmented learning suggests that a single-round creative task may be insufficient to capture the iterative nature of augmented learning, which requires multiple rounds of human-GenAI interactions. We therefore encourage future researchers to expand the spatial and temporal scales of their studies on human-GenAI co-creation. By doing so, they can better capture the iterative and dynamic nature of augmented learning and produce more reliable and insightful findings.

Furthermore, the co-creation activities identified in our study both encompass and extend the findings of prior research on human-GenAI co-creation. Recent studies have proposed valuable models for understanding how humans and GenAI collaborate in creative or problem-solving contexts (Dell'Acqua et al. 2023, Jia et al. 2024, Chen and Chan 2024). However, our findings differ in meaningful ways from this prior work, offering a more comprehensive and organic view of the collaborative activities that unfold during human-GenAI co-creation. First, we conceptualize human-GenAI co-creation as a bi-directional and dynamic process, rather than a fixed or one-sided interaction. GenAI delegation models, for example, typically emphasize assigning tasks to one party, either human or GenAI, that is expected to perform better, and often treat these delegation decisions as static or one-off choices (e.g., Jussupow et al. 2021, Fügener et al. 2022). For another example, Chen and Chan (2024) emphasized specific collaboration modalities of "LLM as Sounding Board" (i.e., feedback provision by LLM) and "LLM as Ghostwriter" (i.e., full delegation of ideation to LLM), which imply a clear division of labor between human and GenAI and one-directional flow of input. While valuable, these studies do not fully capture the continuous and iterative nature of human-GenAI co-creation as it unfolds over time. In their models, tasks are either performed entirely by the human (no delegation) or by GenAI (full delegation), leaving little room for observing dynamic collaboration over time. Furthermore, the flow of feedback in their research is unidirectional, always from LLM to participants, while idea generation and refinement are conducted exclusively by the participants. Similarly, Jia et al. (2024) adopt a sequential collaboration design in which GenAI initiates conversations through scripted protocols and then hands the interaction over to human agents, with little to no further dynamic interplay. This approach reflects fixed, asymmetrical patterns of collaboration. In contrast, our conceptualization of augmented learning moves beyond static delegation and one-directional interaction. It reveals that co-creation between humans and GenAI often unfolded as an iterative process, with both parties generating ideas, providing feedback, and refining ideas. This bi-directionality supports a more flexible and interactive form of ideation than what is captured by one-way or handoff-based models.

Second, our findings offer a comprehensive taxonomy of co-creation activities, whereas prior studies often adopt simpler approaches, focusing on full AI delegation or high-level human-GenAI co-creation, without examining the specific collaborative activities that occur during co-creation. For example, Dell'Acqua et al. (2023) introduce two archetypes of human-GenAI co-creation: Centaurs, representing a clear division of labor between human and AI tasks (i.e., no true co-creation), and Cyborgs, reflecting high-level human-GenAI co-creation. While the Cyborgs archetype offers useful conceptual distinctions, it represents simplistic human-GenAI co-creation and does not specify whether and how they collaborate on specific co-creation activities such as *Idea Generation-Response, Idea Request-Idea Generation*, and *Idea Co-development*. Similarly, Raisch and Fomina (2024) offer a conceptual typology of hybrid problem-solving, identifying autonomous search*,* sequential search*,* and interactive search as core modes of human-AI collaboration. These modes capture important variations in how responsibilities may be distributed but remain abstract and do not specify the detailed activities that constitute the co-creation process. In contrast, our research offers a detailed taxonomy of co-creation activities grounded in observed human-GenAI co-creations. This taxonomy identifies specific co-creation activities, such as proposing, critiquing, reframing, integrating, and expanding on ideas, that offer a more granular and activity-oriented understanding of how human-GenAI co-creation unfolds over time.

8.3. Practical Implications

Our research offers practical implications for enhancing the effectiveness of human-GenAI co-creation in organizations. Since the release of ChatGPT in 2022, many companies have rushed to adopt GenAI systems, driven by the belief that GenAI will empower employees to generate more creative ideas and thereby enhance overall firm performance (Kemp et al. 2024). However, our research shows that the mere implementation of GenAI does not automatically offer these benefits. Particularly with GenAI, the effectiveness of its integration depends significantly on how well human users understand and interact effectively with it. We propose a two-step approach. As a first step, companies should conduct a critical assessment of GenAI's strengths and weaknesses to develop a deeper understanding of its capabilities. GenAI excels at processing large volumes of data and identifying patterns, which form the foundation for its ability to generate diverse ideas. However, it lacks the ability to comprehend nuanced human contexts, needs, and desires. Thus, GenAI should be seen as a complementary agent, rather than a standalone one, that facilitates more effective and creative co-creation processes.

As the second step, based on their understanding of GenAI's strengths and weaknesses, companies should identify effective human-GenAI co-creation strategies to enhance joint creativity and invest in comprehensive training and development programs to educate employees on these strategies. Our findings emphasize the importance of *Idea Co-development*, a co-creation activity where humans and GenAI collaboratively engage in feedback exchanges and iterative refinements of generated ideas. To support this, training programs should include methodologies and practical exercises designed to enhance skills required for *Idea Co-development*. In our Study 3, we demonstrated that even the simple provision of explanations about the significance of *Idea Co-development*, along with examples, effectively improved joint creativity. We believe that formal training programs featuring more contextualized and detailed examples of *Idea Co-development* would be even more effective in enhancing joint creativity of humans and GenAI within organizations.

From the perspective of GenAI system design, our findings highlight the need for advanced GenAI systems that actively support and promote *Idea Co-development*. Developers should aim to create GenAI systems that go beyond simply generating ideas and instead engage users in deeper, more interactive forms of collaboration. These systems should incorporate features that facilitate iterative co-development of ideas. Examples include interactive feedback loops, suggestions for improvement, and mechanisms that promote both users and GenAI to critically assess their collaborative outputs. For instance, GenAI systems could provide rationales for their suggestions, pose questions to encourage further elaboration, or highlight specific aspects of an idea that could benefit from additional refinement. Such features would foster a more dynamic and productive environment for *Idea Co-development*, enabling humans and GenAI to collaborate more effectively and achieve continuous improvement in their joint creativity over time.

8.4. Limitations and Future Research Agendas

While our study provides significant insights into the dynamics of human-GenAI co-creation for augmented learning aimed at improving joint creativity over time, it is important to acknowledge some limitations that may influence the interpretation and generalizability of our findings.

One limitation is the exclusive focus on ideation tasks, which may not encompass the full range of creative tasks involved in human-GenAI co-creation. We selected ideation tasks because they offer a beneficial context for identifying co-creation activities. The interactive and conversational nature of ideation makes human-GenAI co-creation more natural, enabling participants to engage in dynamic exchanges with GenAI. This setting also allows for the systematic analysis of all dialogues between humans and GenAI in written form, providing a rich and analyzable dataset to examine co-creation activities. However, other creative domains, such as artistic creation, technical problem-solving, or complex decision-making, might involve different types of creative tasks with different dynamics and mechanisms of human-GenAI co-creation. These domains may present unique challenges and opportunities that were not captured in our study. Therefore, future research should examine a broader array of creative tasks to better understand how the principles of augmented learning and co-creation activities, identified in our research, can be applied across diverse creativity contexts.

Second, our Studies 1 and 2 were conducted via a UK university behavioral lab. Although participation was not limited to students, and local community members also actively participated in our studies, it is likely that our sample skewed toward younger individuals with relatively high technological similarity. However, we believe the key mechanisms uncovered—particularly the dynamics of augmented learning and the role of *Idea Co-development*—are not limited to younger generations. Both Study 1 (ages 18–63) and Study 2 (ages 18–77) featured broad age distributions, and the average participant ages (Study 1: M = 27.02; Study 2: M = 29.53) are consistent with prior research on AI (e.g., Dietvorst et al. 2015, Turel and Kalhan 2023, Messer 2024, Jia et al. 2024, McGuire et al. 2024, Sun et al. 2025). We also controlled for participants' age and daily use of AI as a proxy for AI familiarity in all analyses. These controls did not alter the interpretation of our results and were not influential factors in shaping joint creativity outcomes. To further enhance generalizability, Study 3 incorporated a more diverse sample, spanning a wide range of age groups (18-69, M = 40.16, SD = 13.34), gender groups (49% female), and work experiences (88% have work experiences). Across all three studies, the consistency of results provides confidence in the robustness of our findings. We also acknowledge that our sample sizes, while consistent with prior experimental research on creativity and human-GenAI collaboration (e.g., Doshi and Hauser 2024, Chen and Chan 2024, Sun et al. 2025), may limit the statistical power to detect smaller effects. Larger-scale replications would offer opportunities to validate and extend our findings, particularly in diverse organizational contexts.

Third, our findings indicate that increased human-GenAI co-creation efforts in *Idea Co-development*, compared to the other two co-creation activities (*Idea Generation-Response* and *Idea Request-Idea Generation*), are particularly beneficial for enhancing joint creativity over time. However, this does not imply that the latter two co-creation activities are not important. Our results suggest that in human-GenAI co-creation, excessive collective efforts in idea generation do not necessarily lead to improved joint creativity over time. For example, we show no significant difference in joint creativity between two scenarios: one where humans and GenAI repeatedly interacted to generate diverse and new ideas (i.e., multiple engagements in idea generation activities), and another where humans engage only in a single interaction with GenAI, asking it to generate numerous ideas on the ideation issue (i.e., a single engagement in idea generation activity). For this reason, in the context of human-GenAI co-creation, a single interaction with GenAI requesting idea variations can be as productive as multiple engagements in idea generation activities involving both humans and GenAI. Moreover, we observe that final ideas often could not be clearly attributed to either the human or GenAI. As both engaged in back-and-forth *Idea Co-development*, initial contributions of idea generation were transformed through mutual development, resulting in final ideas that were truly collaborative in nature. This attributional ambiguity highlights the integrative dynamics of *Idea Co-development* and further underscores its importance in effective human-GenAI co-creation. Together, our research shows the importance of shifting the focus of co-creation activities from idea generation to *Idea Co-development* to achieve greater joint creativity over time. Future research could also build on this insight and examine whether instructing GenAI systems to facilitate *Idea Co-development* before interaction with humans enhances joint creativity in human-GenAI co-creation.

Although our findings show that in human-GenAI co-creation, emphasizing *Idea Co-development* is generally more effective in improving joint creativity, it is important to acknowledge that this approach may not be equally effective for individuals with different preferences and creativity skills. For example, individuals who are proficient at further developing generated ideas but require initial inspiration may benefit more from engaging in idea generation activities with GenAI, such as *Idea Generation-Response* and *Idea Request-Idea Generation*. Writers often exemplify this approach; they seek a spark of inspiration and, once they have an initial idea, can compose the writing quickly (Thrash et al. 2010). For such individuals, increased engagement in idea generation activities with GenAI might be more beneficial than focusing primarily on *Idea Co-development*. This suggests that the effectiveness of different human-GenAI co-creation activities on joint creativity may depend on individual differences of humans in creative processes and preferences. Indeed, our findings from Study 2 showed that the decline in *Idea Co-development* was the negative mechanism hindering joint creativity over time. However, the other two co-creation activities (*Idea Generation-Response* and *Idea Request-Idea Generation*) increased over time but did not function as positive mechanisms that could offset the negative impact of reduced *Idea Co-development*. One plausible explanation is that *Idea Co-development* functions as a consistent and salient mechanism, whereas the effectiveness of the other two co-creation activities may vary across participants. Some individuals may benefit from increased idea generation, while others may not, suggesting the presence of meaningful heterogeneity in how people co-create with GenAI. Therefore, future research should investigate how individual differences and preferences impact the efficacy of various human-GenAI co-creation activities for joint creativity over time. Exploring these individual differences could lead to more tailored strategies for human-GenAI co-creation, optimizing creative outcomes for a range of users.

9. Conclusion

In conclusion, we introduce a novel theory of augmented learning by reconceptualizing it and providing both theoretical insights and empirical evidence to understand its role in human-GenAI co-creation for joint creativity over time. Through two inductive studies (Studies 1 and 2), we discovered that human-GenAI co-creation does not spontaneously result in augmented learning. While increasing collective efforts in the *Idea Co-development* activity proved effective for enhancing joint creativity over time, human-GenAI dyads could not readily discover this strategy. Building upon these insights, Study 3 deductively confirmed that explicit instructions on *Idea Co-development* were effective for improving joint creativity over time, suggesting that augmented learning can be facilitated through deliberate strategies and educational interventions. Our theory provides a foundational framework for future exploration into the dynamics of human-GenAI co-creation, with significant theoretical and practical implications for improving joint creativity between humans and GenAI.

References

Aguinis H, Vandenberg RJ (2014) An ounce of prevention is worth a pound of cure: Improving research quality before data collection. *Annual Review of Organizational Psychology and Organizational Behavior* 1:569–595.

Althuizen N, Reichel A (2016) The Effects of IT-Enabled Cognitive Stimulation Tools on Creative Problem Solving: A Dual Pathway to Creativity. *Journal of Management Information Systems* 33(1):11–44.

Amabile T, Pratt MG (2016) The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior* 36:157–183.

Amabile TM (1983) The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology* 45:357–376.

Amabile TM (2020) Creativity, Artificial Intelligence, and a World of Surprises. *AMD* 6(3):351–354.

Anderson N, Potočnik K, Zhou J (2014) Innovation and Creativity in Organizations: A State-of-the-Science Review, Prospective Commentary, and Guiding Framework. *Journal of Management* 40(5):1297–1333.

Anthony C, Bechky BA, Fayard AL (2023) "Collaborating" with AI: Taking a System View to Explore the Future of Work. *Organization Science*.

Argote L, Guo JM (2016) Routines and transactive memory systems: Creating, coordinating, retaining, and transferring knowledge in organizations. *Research in Organizational Behavior* 36:65–84.

Ashford SJ, Blatt R, VandeWalle D (2003) Reflections on the Looking Glass: A Review of Research on Feedback-Seeking Behavior in Organizations. *Journal of Management* 29(6):773–799.

Baird A, Maruping LM (2021) The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MISQ* 45(1):315–341.

Bankins S, Ocampo AC, Marrone M, Restubog SLD, Woo SE (2024) A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior* 45(2):159–182.

Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, Dai B, et al. (2024) lme4: Linear Mixed-Effects Models using "Eigen" and S4.

Bauer K, Jussupow E, Heigl R, Vogt B, Hinz O (2024) All Just in Your Head? Unraveling the Side Effects of Generative AI Disclosure in Creative Tasks. (February 29) https://papers.ssrn.com/abstract=4782554.

Behfar K, Okhuysen GA (2018) Perspective—Discovery Within Validation Logic: Deliberately Surfacing, Complementing, and Substituting Abductive Reasoning in Hypothetico-Deductive Inquiry. *Organization Science* 29(2):323–340.

Berg JM (2019) When silver is gold: Forecasting the potential creativity of initial ideas. *Organizational Behavior and Human Decision Processes* 154:96–117.

Bigman YE, Gray K (2018) People are averse to machines making moral decisions. *Cognition* 181:21–34.

Bogert E, Schecter A, Watson RT (2021) Humans rely more on algorithms than social influence as a task becomes more difficult. *Sci Rep* 11(1):8028.

Boussioux L, Lane JN, Zhang M, Jacimovic V, Lakhani KR (2024) The Crowdless Future? Generative AI and Creative Problem-Solving. *Organization Science*.

van den Broek E, Sergeeva A, Huysman M (2021) When the Machine Meets the Expert: An Ethnography of Developing Ai for Hiring. *MIS Quarterly* 45(3):1557–1580.

Brynjolfsson E, McAfee A (2016) *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* Reprint edition. (W. W. Norton & Company, New York London).

Brynjolfsson E, Mitchell T (2017) What can machine learning do? Workforce implications. *Science* 358(6370):1530–1534.

Cadario R, Longoni C, Morewedge CK (2021) Understanding, explaining, and utilizing medical artificial intelligence. *Nat Hum Behav* 5(12):1636–1642.

Calabretta G, Gemser G, Wijnberg NM (2017) The Interplay between Intuition and Rationality in Strategic Decision Making: A Paradox Perspective. *Organization Studies* 38(3–4):365–401.

Chen Z, Chan J (2024) Large Language Model in Creative Work: The Role of Collaboration Modality and User Expertise. *Management Science*:mnsc.2023.03014.

Chua RYJ, Iyengar SS (2008) Creativity as a Matter of Choice: Prior Experience and Task Instruction as Boundary Conditions for the Positive Effect of Choice on Creativity. *The Journal of Creative Behavior* 42(3):164–180.

Corbin J, Strauss A (2008) *Basics of qualitative research: Techniques and procedures for developing grounded theory, 3rd ed* (Sage Publications, Inc, Thousand Oaks, CA, US).

Cropley A (2006) In Praise of Convergent Thinking. *Creativity Research Journal* 18(3):391–404.

Csikszentmihalyi M (1997) *Creativity: Flow and the psychology of discovery and invention* (HarperCollins Publishers, New York, NY, US).

Davenport TH, Kirby J (2016) *Only Humans Need Apply: Winners and Losers in the Age of Smart Machines* First Edition. (Harper Business, Erscheinungsort nicht ermittelbar).

Dell'Acqua F, McFowland E, Mollick ER, Lifshitz-Assaf H, Kellogg K, Rajendran S, Krayer L, Candelon F, Lakhani KR (2023) Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *SSRN Journal*.

Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General* 144(1):114–126.

DigitalDefynd T (2025) 20 Ways Netflix Is Using Artificial Intelligence [In Depth Analysis][2025]. *DigitalDefynd*.

Doshi AR, Hauser OP (2024) Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances* 10(28):eadn5290.

Elam JJ, Mead M (1990) Can Software Influence Creativity? *Information Systems Research* 1(1):1–22.

Endsley MR (2017) From here to autonomy: Lessons learned from human–automation research. *Human Factors* 59(1):5–27.

Engelbart DC (1962) *Augmenting Human Intellect: A Conceptual Framework*

Epley N, Waytz A, Cacioppo JT (2007) On seeing human: a three-factor theory of anthropomorphism. *Psychol Rev* 114(4):864–886.

Epstein Z, Hertzmann A (2023) Art and the science of generative AI. *Science* 380(6650):1110–1111.

FeldmanHall O, Shenhav A (2019) Resolving uncertainty in a social world. *Nat Hum Behav* 3(5):426–435.
Fügener A, Grahl J, Gupta A, Ketter W (2022) Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research* 33(2):678–696.
Gao Z, Jiang J (2021) Evaluating Human-AI Hybrid Conversational Systems with Chatbot Message Suggestions. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (ACM, Virtual Event Queensland Australia), 534–544.
Gnewuch U, Morana S, Hinz O, Kellner R, Maedche A (2023) More Than a Bot? The Impact of Disclosing Human Involvement on Customer Interactions with Hybrid Service Agents. *Information Systems Research*.
Goddard K, Roudsari A, Wyatt JC (2012) Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc* 19(1):121–127.
Golden-Biddle K (2020) Discovery as an Abductive Mechanism for Reorienting Habits within Organizational Change. *AMJ* 63(6):1951–1975.
Gordetzki P, Blohm I, Clegg M, Schakols F, Hofstetter R (2024) Refining Ideas with Generative AI: How Text- and Image-Based Scenarios Influence Idea Refinement. *Information Systems Research* (Special Issue on Analytical Creativity).
Grant AM (2008) Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity. *Journal of Applied Psychology* 93(1):48–58.
Grimes M, von Krogh G, Feuerriegel S, Rink F, Gruber M (2023) From Scarcity to Abundance: Scholars and Scholarship in an Age of Generative Artificial Intelligence. *AMJ* 66(6):1617–1624.
Grønsund T, Aanestad M (2020) Augmenting the algorithm: Emerging human-in-the-loop work configurations. *The Journal of Strategic Information Systems* 29(2):101614.
Guilford JP (1967) Creativity: Yesterday, Today and Tomorrow. *The Journal of Creative Behavior* 1(1):3–14.
Guzik EE, Byrge C, Gilde C (2023) The originality of machines: AI takes the Torrance Test. *Journal of Creativity* 33(3):100065.
Haase J, Hanel PHP (2023) Artificial muses: Generative artificial intelligence chatbots have risen to human-level creativity. *Journal of Creativity* 33(3):100066.
Hargadon AB, Bechky BA (2006) When collections of creatives become a creative collectives: A field study of problem solving at work. *Organization Science* 17(4):484–500.
Harvey S, Berry J (2022) Toward a meta-theory of creativity forms: How novelty and usefulness shape creativity. *AMR*.
Hayes AF (2013) *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach* (The Guilford Press, New York).
Hitsuwari J, Ueda Y, Yun W, Nomura M (2023) Does human–AI collaboration lead to more creative art? Aesthetic evaluation of human-made and AI-generated haiku poetry. *Computers in Human Behavior* 139:107502.
Horton Jr CB, White MW, Iyengar SS (2023) Bias against AI art can enhance perceptions of human creativity. *Sci Rep* 13(1):19001.

Hou J (Jove), Wang L, Wang G, Wang H, Yang S (2024) The Double-Edged Roles of Generative AI in the Creative Process: Experiments on Design Work. (November 30) https://papers.ssrn.com/abstract=4739471.
Hubert KF, Awa KN, Zabelina DL (2024) The current state of artificial intelligence generative language models is more creative than humans on divergent thinking tasks. *Sci Rep* 14(1):3440.
Jia N, Luo X, Fang Z, Liao C (2024) When and How Artificial Intelligence Augments Employee Creativity. *AMJ* 67(1):5–32.
Jussupow E, Spohrer K, Heinzl A, Gawlitza J (2021) Augmenting Medical Diagnosis Decisions? An Investigation into Physicians' Decision-Making Process with Artificial Intelligence. *Information Systems Research* 32(3):713–735.
Kaufman JC, Baer J, Cole JC (2009) Expertise, Domains, and the Consensual Assessment Technique. *The Journal of Creative Behavior* 43(4):223–233.
Kemp A, Bramwell M, Evans R, Rizk N, Rugeroni L, Sabherwal N, Townley M (2024) *AI-Powered Employee Experience: How Organisations Can Unlock Higher Engagement and Productivity* (Deloitte).
Kim YJ, Kim J (2020) Does Negative Feedback Benefit (or Harm) Recipient Creativity? The Role of the Direction of Feedback Flow. *Academy of Management Journal* 63(2):584–612.
Köbis N, Bonnefon JF, Rahwan I (2021) Bad machines corrupt good morals. *Nat Hum Behav* 5(6):679–685.
Koivisto M, Grassini S (2023) Best humans still outperform artificial intelligence in a creative divergent thinking task. *Sci Rep* 13(1):13601.
Kuznetsova A, Brockhoff PB, Christensen RHB, Jensen SP (2020) lmerTest: Tests in Linear Mixed Effects Models.
LeBreton JM, Senter JL (2008) Answers to 20 Questions About Interrater Reliability and Interrater Agreement. *Organizational Research Methods* 11(4):815–852.
Lewis K, Herndon B (2011) Transactive Memory Systems: Current Issues and Future Research Directions. *Organization Science* 22(5):1254–1265.
Liu B (2021) In AI We Trust? Effects of Agency Locus and Transparency on Uncertainty Reduction in Human–AI Interaction. *Journal of Computer-Mediated Communication* 26(6):384–402.
Locke EA (2007) The Case for Inductive Theory Building†. *Journal of Management*.
Lubart T (2005) How can computers be partners in the creative process: Classification and commentary on the Special Issue. *International Journal of Human-Computer Studies* 63(4):365–369.
Lucas BJ, Nordgren LF (2015) People underestimate the value of persistence for creative performance. *J Pers Soc Psychol* 109(2):232–243.
Magni F, Park J, Chao MM (2023) Humans as Creativity Gatekeepers: Are We Biased Against AI Creativity? *J Bus Psychol*.
McGuire J, De Cremer D, Van de Cruys T (2024) Establishing the importance of co-creation and self-efficacy in creative collaboration with artificial intelligence. *Sci Rep* 14(1):18525.
McMahon K, Ruggeri A, Kämmer JE, Katsikopoulos KV (2016) Beyond Idea Generation: The Power of Groups in Developing Ideas. *Creativity Research Journal*.
Messer U (2024) Co-creating art with generative artificial intelligence: Implications for artworks and artists. *Computers in Human Behavior: Artificial Humans* 2(1):100056.

Noy S, Zhang W (2023) Experimental evidence on the productivity effects of generative artificial intelligence. *Science* 381(6654):187–192.

OpenAI (2024) ChatGPT [Large language model]. Retrieved https://chatgpt.com.

Otis NG, Haas B, Cranney K, Koning R (2024) Global Evidence on Gender Gaps and Generative AI.

Parasuraman R, Manzey DH (2010) Complacency and bias in human use of automation: an attentional integration. *Hum Factors* 52(3):381–410.

Parnes SJ (1961) Effects of extended effort in creative problem solving. *Journal of Educational Psychology* 52(3):117–122.

Paulus PB, Nijstad BA eds. (2003) *Group creativity: Innovation through collaboration* (Oxford University Press, New York, NY, US).

Perry-Smith JE, Mannucci PV (2017) From Creativity to Innovation: The Social Network Drivers of the Four Phases of the Idea Journey. *The Academy of Management Review* 42(1):53–79.

Pok MT, Koopman J, Mcclean ST, Zhang JH, Chi Hon Li, De Cremer D, Yizhen Lu, Chin Tung Stewart Ng (2022) When Conscientious Employees Meet Intelligent Machines: An Integrative Approach Inspired by Complementarity Theory and Role Theory. *Academy of Management Journal* 65(3):1019–1054.

Raisch S, Krakowski S (2021) Artificial Intelligence and Management: The Automation–Augmentation Paradox. *Academy of Management Review* 46(1):192–210.

Ren Y, Argote L (2011) Transactive Memory Systems 1985–2010: An Integrative Framework of Key Dimensions, Antecedents, and Consequences. *ANNALS* 5(1):189–229.

de Rooij A, Biskjaer MM (2025) Has AI Surpassed Humans in Creative Idea Generation? A Meta-Analysis: The European Conference on Cognitive Ergonomics.

Runco MA (2004) Creativity. *Annu. Rev. Psychol.* 55(1):657–687.

Schwab K (2017) *The Fourth Industrial Revolution* (Crown).

Shalley CE (1991) Effects of productivity goals, creativity goals, and personal discretion on individual creativity. *Journal of Applied Psychology* 76(2):179–185.

Sun S, Li A, Foo M, Zhou J, Lu J (2025) How and For Whom Using Generative AI Affects Creativity: A Field Experiment. *Journal of Applied Psychology*.

Thees M, Kapp S, Strzys MP, Beil F, Lukowicz P, Kuhn J (2020) Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior* 108:106316.

Thrash TM, Maruskin LA, Cassidy SE, Fryer JW, Ryan RM (2010) Mediating between the muse and the masses: inspiration and the actualization of creative ideas. *J Pers Soc Psychol* 98(3):469–487.

van Tilburg WAP, Igou ER (2012) On boredom: Lack of challenge and meaning as distinct boredom experiences. *Motivation and Emotion* 36(2):181–194.

Timmermans S, Tavory I (2012) Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. *Sociological Theory* 30(3):167–186.

Tschang FT, Almirall E (2021) Artificial Intelligence as Augmenting Automation: Implications for Employment. *Academy of Management Perspectives* 35(4):642–659.

Turel O, Kalhan S (2023) Prejudiced Against the Machine? Implicit Associations and the Transience of Algorithm Aversion. *MIS Quarterly* 47(4):1369–1394.

Vaccaro M, Almaatouq A, Malone T (2024) When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nat Hum Behav* 8(12):2293–2303.

Vanneste BS, Puranam P (2024) Artificial Intelligence, Trust, and Perceptions of Agency. *AMR*:amr.2022.0041.
Vardi S, Choudhary V (2024) A Random Walk Modeling Framework for Boosting the Creativity of Humans and AI. (April 8) https://papers.ssrn.com/abstract=4787223.
Wegner DM (1987) Transactive Memory: A Contemporary Analysis of the Group Mind. Mullen B, Goethals GR, eds. *Theories of Group Behavior* (Springer, New York, NY), 185–208.
Wu L, Lou B, Ma X (2024) Artificial Intelligence, CEO Turnover, and Exploration Orientation in Firm Innovation. *Information Systems Research* (Special Issue on Analytical Creativity).
Ybarra O (2002) Naive causal understanding of valenced behaviors and its implications for social information processing. *Psychological Bulletin* 128(3):421–441.
Zanzotto FM (2019) Viewpoint: Human-in-the-loop Artificial Intelligence. *Journal of Artificial Intelligence Research* 64:243–252.
Zhang Y, Gosline R (2023) Human favoritism, not AI aversion: People's perceptions (and bias) toward generative AI, human experts, and human–GAI collaboration in persuasive content generation. *Judgment and Decision Making* 18:e41.
Zhou J, Wang XM, Bavato D, Tasselli S, Wu J (2019) Understanding the Receiving Side of Creativity: A Multidisciplinary Review and Implications for Management Research. *Journal of Management* 45(6):2570–2595.

Table 1. Related Empirical Studies on the Joint Creativity of Human-GenAI Co-creation

Study	# of Rounds	Creative Task Type	Comparison Group	Main Finding
Boussioux et al. (2024)	1	Ideation	Human-only	[Mixed] Human-GenAI solutions demonstrated higher financial value but lower novelty compared to human-generated solutions.
Chen and Chan (2024)	1	Ideation	Human-only	[Negative] Human-GenAI produced marginally lower creativity than human-only.
Dell'Acqua et al. (2023)	1	Ideation	Human-only	[Positive] Human-GenAI collaboration had higher solution quality for problem-solving tasks, compared to the human-only condition.
Doshi and Hauser (2024)	1	Story writing	Human-only	[Mixed] Human-GenAI co-created stories were more creative, better written, and more enjoyable. However, these stories were more similar to each other and less diverse compared to those created by humans alone.
Gao and Jiang (2021)	1	Ideation	Human-only and GenAI-only	[Negative] Human-GenAI hybrid systems produced lower-quality suggestions compared with the human-only baseline.
Hitsuwari et al. (2023)	1	Poem writing	Human-only and GenAI-only	[Positive] Human-GenAI co-creation achieved higher creativity than human-only and GenAI-only creations.
Jia et al. (2024)	1	Ideation	Human-only	[Positive] The human-GenAI condition had higher employee' creativity in answering customers' questions than the human-only condition.
McGuire et al. (2024)		Poem writing	Human-only and GenAI-only	[Negative] Poems in the human-only condition were regarded as more creative than poems in the human-GenAI condition.
Messer (2024)	1	Art creation	Human-only	[Mixed] Human-GenAI co-created art was more novel but less liked compared to the arts generated in the human-only condition.
Noy and Zhang (2023)	1	Ideation	Human-only	[Positive] The human-GenAI condition, compared to the human-only condition, received higher scores for creativity, measured by writing quality, content quality, and originality.
Sun et al. (2025)	1	Ideation	Human-only	[Positive] Human-GenAI produced better employee creativity than the human-only condition.
Our Study 1	10	Ideation	Human-only	There was no improvement in joint creativity in human-GenAI co-creation over time, but there was improvement in creativity in the human-only condition.
Our Study 2	3	Ideation	-	The lack of improvement in joint creativity in human-GenAI co-creation was explained by the failure to focus on *Idea Co-development* activity between humans and GenAI.
Our Study 3	2	Ideation	-	Human-GenAI groups receiving instructions on *Idea Co-development* enhanced joint creativity compared to human-GenAI groups receiving no instruction.

Note. In the three studies with GenAI-only comparison groups, GenAI-only outputs underperformed compared to the human-GenAI condition and either underperformed or matched human-only performance.

Table 2. Means, Standard Deviations, and Correlations in Study 1

Variable	*M*	*SD*	1	2	3	4
L2						
1. Age	27.02	11.09				
2. Trust of AI (α = .83)	3.11	1.76	0.23**			
3. Daily Use of AI (α = .87)	2.28	1.39	-0.03	0.06		
4. Human Creativity	3.50	0.42	-0.01	-0.10**	-0.11**	
L1						
1. Output Creativity	22.31	10.16				
2. Miscellaneous Activity	0.01	0.09	0.02			
3. Idea Co-development	0.33	0.4	0.21**	-0.01		
4. Idea Request-Idea Generation	0.64	0.35	-0.06	0.04	-0.45**	
5. Idea Generation-Response	0.19	0.34	-0.03	-0.05	-0.17**	-0.61**

Note. $N_{between\text{-}individual\ (L2)}$ = 162. $N_{within\text{-}individual\ (L1)}$ = 1620. *M* and *SD* are used to represent mean and standard deviation, respectively. * indicates $p < .05$. ** indicates $p < .01$.

Table 3. Multilevel Analysis of Output Creativity across Task Rounds in Human-GenAI vs. Human-Only Conditions in Study 1

	Output Creativity			**Output Creativity**		
Predictors	*Estimates*	*SE*	*p*	*Estimates*	*SE*	*p*
Intercept	18.82	0.83	**<0.001**	19.14	3.39	**<0.001**
Task Round	0.38	0.11	**0.001**	0.38	0.11	**0.001**
Condition: Human-AI	3.27	1.17	**0.005**	3.27	1.15	**0.004**
Task Round × Condition	-0.34	0.16	**0.030**	-0.34	0.16	**0.030**
Age				-0.11	0.04	**0.003**
Gender: Female				-1.40	0.84	0.097
Gender: Non-binary				3.81	3.44	0.268
Gender: Non-disclosure				-5.57	2.23	**0.013**
Daily Use of AI				-0.88	0.29	**0.002**
Trust of AI				-0.02	0.22	0.941
Human Creativity				1.63	0.84	0.054
Random Effects						
σ^2	79.33			79.33		
τ_{00}	17.61 Participants			14.43 Participants		
ICC	0.18			0.15		
N	162 Participants			162 Participants		
Observations	1607			1607		
Marginal R^2 / Conditional R^2	0.011 / 0.191			0.053 / 0.198		

Note. $N_{between\text{-}individual\ (L2)} = 162$. $N_{within\text{-}individual\ (L1)} = 1607$. Removing all control variables did not alter our interpretation of the results.

Table 4. Means, Standard Deviations, and Correlations in Study 2

Variable	*M*	*SD*	1	2	3	4
L2						
1. Age	29.53	11.94				
2. Daily Use of AI (α = .90)	2.12	1.18	-0.17**			
3. Trust of AI (α = .94)	3.80	1.72	0.01	0.18**		
4. Human Creativity	14.37	2.32	0.05	-0.13	-0.16*	
L1						
1. Output Creativity	17.86	6.87				
2. Miscellaneous Activity	0.09	0.15	-0.00			
3. Idea Co-development	0.63	0.30	0.19**	-0.16**		
4. Idea Request-Idea Generation	0.28	0.23	-0.12*	-0.08	-0.54**	
5. Idea Generation-Response	0.10	0.18	-0.03	0.04	-0.22**	-0.25**

Note. $N_{between\text{-}individual\ (L2)}$ = 166. $N_{within\text{-}individual\ (L1)}$ = 498. The mean percentages of the four co-creation activities (i.e., *Idea Generation-Response, Idea Request-Idea Generation, Idea Co-development,* and *Miscellaneous Activity)* do not add up to one because a single interaction can be categorized into multiple activities—e.g., *Idea Co-development* and *Idea Request-Idea Generation*. *M* and *SD* are used to represent mean and standard deviation, respectively. * indicates p < .05. ** indicates p < .01.

Table 5. Co-creation Activities from Part 2 of Study 2

Final Set of Co-creation Activities	Definition	Initial Codes of Co-creation Activities	
		Codes	**Example Interaction**
Idea Generation-Response	Humans initially generate new ideas and propose them to GenAI, which then offers responses to these ideas (Guilford 1967)	Proposing new ideas	"Here is an idea I have..." - "Your idea is a good starting point…"
Idea Request-Idea Generation	Humans request GenAI to generate new ideas, and GenAI generates new ideas	Requesting GenAI to generate new ideas	"Can you suggest 10 ideas for this problem?" - "Here are 10 ideas…"
Idea Co-development	A collaborative activity in which humans and GenAI engage in critical feedback exchanges and the joint refinement of generated ideas (Lubart 2005, Cropley 2006, Grønsund and Aanestad 2020)	**Feedback**	
		Evaluating and exchanging feedback on generated ideas	"Idea 1 seems more innovative, but Idea 2 is more practical." - "That's true—maybe we build on idea 2 first."
		Idea Refinement	
		Expanding on a generated idea	"Please expand on this idea." - "Here's a more detailed version of the idea…"
		Integrating different ideas	"How about we combine ideas 2 and 3?" - "Here is a combined idea…"
		Reframing a generated idea	"What if we frame this as employee empowerment instead of surveillance?" - "That would make it more appealing."
		Streamlining a generated idea	"We should cut down a few parts of the idea to make it clearer." – "Okay, here's a more concise version."
		Setting constraints on a generated idea	"Let's make sure this idea stays cost-effective." - "We'll need to use cheaper machines."
Miscellaneous Activity	Maintaining engagement	Greeting	"Hi!" - "Hello!"
		Expressing gratitude	"Thank you!" - "You're welcome!"
		Formatting the conversation	"Let's use a numbered list for clarity." - "Understood! Here's a numbered list of ideas…"

Table 6. Mediation Analysis in Study 2

	Joint Creativity			**Idea Request-Idea Generation**			**Idea Generation-Response**			**Idea Co-development**			**Joint Creativity**		
Predictors	*Estimates*	*SE*	*p*	*Estimates*	*SE*	*p*	*Estimates*	*SE*	*p*	*Estimates*	*SE*	*p*	*Estimates*	*SE*	*p*
Intercept	21.10	3.86	**<0.001**	0.27	0.13	**0.038**	0.10	0.10	0.288	0.71	0.15	**<0.001**	19.40	4.01	**<0.001**
Task Round	-0.20	0.27	0.466	0.03	0.01	**0.004**	0.01	0.01	**0.043**	-0.05	0.01	**<0.001**	-0.03	0.28	0.916
Miscellaneous Activity	-0.74	1.78	0.679	-0.03	0.06	0.576	0.02	0.05	0.666	-0.33	0.08	**<0.001**	0.17	1.84	0.925
Age	-0.05	0.04	0.225	0.00	0.00	0.763	0.00	0.00	0.560	-0.00	0.00	0.260	-0.04	0.04	0.272
Gender: Female	-0.58	0.96	0.548	-0.01	0.03	0.697	-0.03	0.02	0.273	0.02	0.04	0.558	-0.66	0.94	0.484
Daily Use of AI	-0.18	0.44	0.685	0.01	0.02	0.321	-0.02	0.01	**0.032**	-0.00	0.02	0.913	-0.17	0.44	0.690
Trust of AI	0.33	0.28	0.247	0.00	0.01	0.957	-0.00	0.01	0.619	-0.01	0.01	0.415	0.35	0.28	0.206
Human Creativity	-0.13	0.22	0.552	-0.01	0.01	0.464	0.00	0.01	0.712	0.01	0.01	0.291	-0.16	0.21	0.460
Idea Generation-Response													-0.46	2.09	0.826
Idea Request-Idea Generation													-0.77	1.73	0.657
Idea Co-development													2.77	1.28	**0.031**
Random Effects															
σ^2	21.67			0.02			0.01			0.05			21.70		
τ_{00}	26.08 Participants			0.03 Participants			0.02 Participants			0.03 Participants			24.69 Participants		
ICC	0.55			0.58			0.53			0.41			0.53		
N	148 Participants			148 Participants			148 Participants			148 Participants			148 Participants		
Observations	444			444			444			444			444		

Marginal R^2 / Conditional R^2	0.017 / 0.554	0.018 / 0.584	0.040 / 0.547	0.061 / 0.450	0.035 / 0.549

Note. $N_{between\text{-}individual\ (L2)} = 148$. $N_{within\text{-}individual\ (L1)} = 444$. Removing all control variables did not alter our interpretation of the results (see Supplementary Table S6).

Table 7. Means, Standard Deviations, and Correlations in Study 3

Variable	*M*	*SD*	1	2	3	4
L2						
1. Age	40.16	13.34				
2. Boredom (α = .92)	3.11	1.41	-0.24**			
3. Intrinsic Motivation (α = .97)	5.66	1.33	0.05	-0.15**		
4. Daily Use of AI (α = .98)	4.21	1.93	-0.23**	0.21**	0.52**	
5. Trust of AI (α = .97)	5.02	1.56	0.11	0.05	0.47**	0.31**
L1						
1. Joint Creativity	19.71	9.25				

Note. $N_{between\text{-}individual\ (L2)} = 166$. $N_{within\text{-}individual\ (L1)} = 332$. *M* and *SD* are used to represent mean and standard deviation, respectively. * indicates $p < .05$. ** indicates $p < .01$.

Table 8. The Changes in Joint Creativity in Treatment and Control Conditions in Study 3

	Joint Creativity			Joint Creativity		
Predictors	*Estimates*	*SE*	*p*	*Estimates*	*SE*	*p*
Intercept	19.71	1.00	**<0.001**	22.42	3.72	**<0.001**
Second Task	-1.24	1.34	0.357	-1.24	1.34	0.357
Treatment Condition	-0.71	1.43	0.620	-0.68	1.48	0.645
Second Task × Condition	3.97	1.92	**0.039**	3.97	1.92	**0.039**
Age				-0.02	0.04	0.623
Gender: Female				-1.26	1.12	0.259
Gender: Non-binary				-0.41	5.20	0.937
Task Order				-1.52	1.11	0.172
Boredom				0.19	0.42	0.654
Intrinsic Motivation				0.60	0.55	0.277
Daily Use of AI				-0.49	0.36	0.177
Trust of AI				-0.48	0.41	0.241
Random Effects						
σ^2	76.22			76.22		
τ_{00}	8.52 $_{\text{Participants}}$			9.07 $_{\text{Participants}}$		
ICC	0.10			0.11		
N	166 $_{\text{Participants}}$			166 $_{\text{Participants}}$		
Observations	332			332		
Marginal R^2 / Conditional R^2	0.018 / 0.116			0.037 / 0.139		

Note. $N_{between\text{-}individual\ (L2)} = 166$. $N_{within\text{-}individual\ (L1)} = 332$. Including all control variables did not alter our interpretation of the results.

Figure 1. Theoretical Model of Augmented Learning

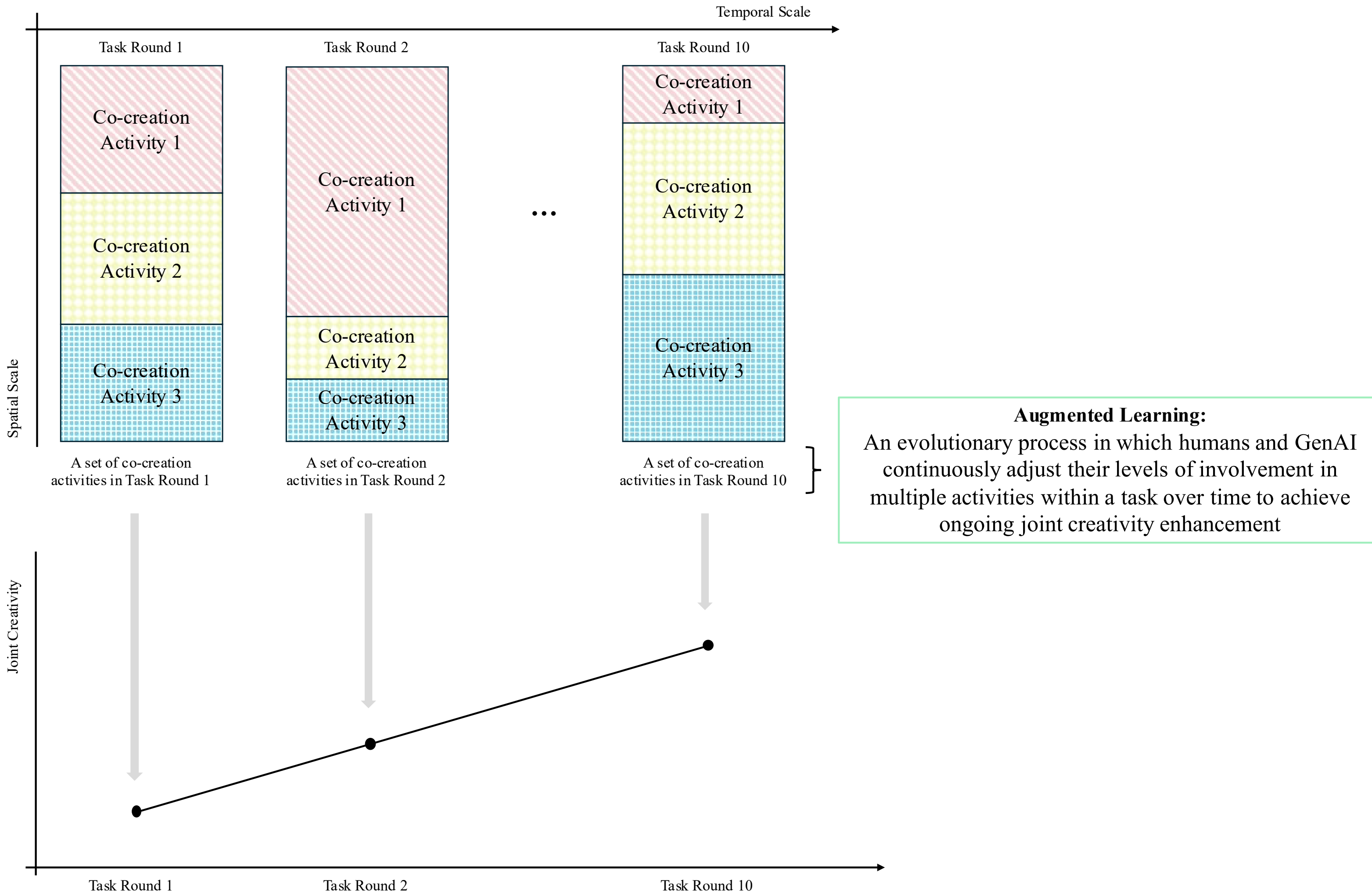

Figure 2. The Changes in Output Creativity across Task Rounds in Human-GenAI vs. Human-Only Conditions in Study 1

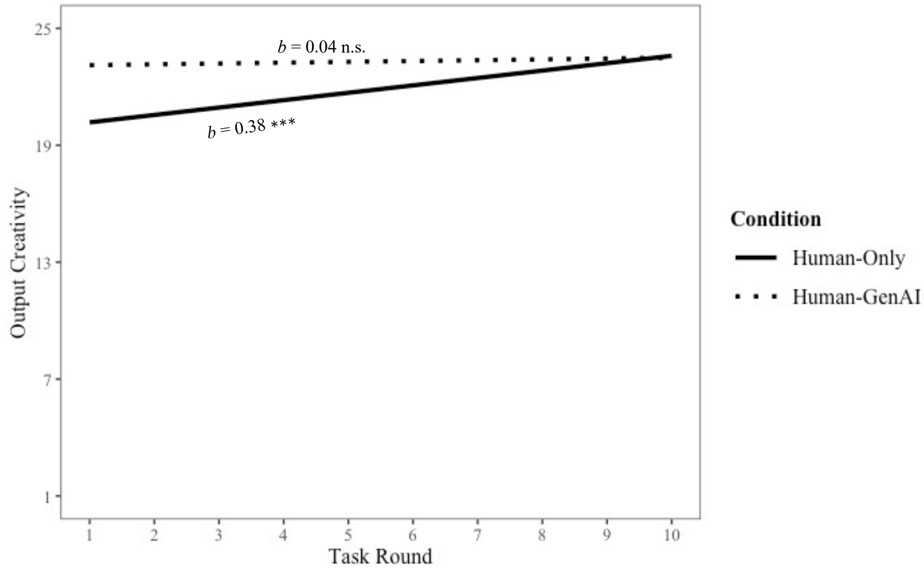

Figure 3. The Changes in Joint Creativity in Treatment vs. Control in Study 3