

Maximizing Generative AI Benefits with Task Creativity and Human Validation

Charu Sinha

Chapman University

Veselina Vacheva

Chapman University

Cristina Nistor

Chapman University

Much of the existing literature on generative AI applications is conflicting, with findings suggesting that investing in AI will lead to better organizational outcomes but also pointing out that incorporating AI may be a wasteful even counterproductive initiative. We develop a conceptual framework to characterize generative AI benefits based on the types of tasks that generative AI may be used for in management. Our work suggests that task creativity plays a key role in successful generative AI outcomes, but human validation - the extent to which a human engages in a supervisory role - is required to reap the benefits. Our conceptual framework is focused on white collar jobs and suggests that the management of generative AI is a strategic choice with important managerial implications.

Generative artificial intelligence (GAI) is a new form of artificial intelligence (AI) that applies machine learning algorithms to existing data in a training dataset to generate new data, and has the potential to disrupt current business models and human labor (McKinsey, 2024a). The increase in computing power and development of new predictive models have led to an impressive increase in the number of GAI companies like ChatGPT (Dilmegani, 2024; Zia, 2023). Popular press articles suggest either wild optimism for the potential of AI in management (Zia, 2023) or unqualified optimism that humans cannot be replaced by AI bots (Barnett, 2024). As such, the role of GAI in management and the business world is still uncertain: will the new technology lead to more productivity, higher overall returns, and change the business world or will it be another automation tool that can be used to marginally improve current business outcomes?

Recent research has focused on AI use as a tool to assist humans in their tasks. For example, Brynjolfsson et al. (2023) find that an AI bot which helps call center assistants identify answers for customer conversations can foster higher productivity, increased customer satisfaction for the call, and faster on-the-job learning for human workers. However, experienced workers benefit the least from AI tool implementation, which suggests that the impact of AI may be overall limited. Acemoglu (2024) estimates that AI will contribute to the US economy in a “nontrivial but modest—no more than a 0.66% increase in total factor productivity (TFP) over 10 years” (p. 1), while listing several ways that this small estimate is still an optimistic measure of the impact of AI on the business world. Korinek and Suh (2024) arrive at more optimistic estimates but rely on the assumption that business processes

include “human work [that] can be decomposed into atomistic tasks that differ in their complexity” (p. 2). Thus, there is still considerable uncertainty over the effect of AI in the business world and how managers should allocate tasks between AI tools and human workers.

We propose a conceptual framework that examines two attributes, task creativity and human validation, to anticipate the level of benefits achieved from incorporating GAI in management tasks. Our conceptual framework focuses on GAI benefits such as lowering labor costs, faster time to completion for tasks, or improved quality in the execution of the task. Generative AI is considered a potential disruptor for white collar jobs (Eisfeldt et al., 2024; Noy & Zhang, 2023). Thus, we focus on this part of the economy in our theorizing. Moreover, we characterize situations where tasks performed by GAI can substitute for human labor rather than instances where GAI may be used to supplement it.

Generative AI in Business Applications

AI has been making its impact felt in various functional areas in business over the last decade, yet the excitement around the recent introduction of GAI parallels few other moments in the history of business. NVIDIA’s CEO Jensen Huang has described GAI as the “iPhone moment” for AI (Caufield, 2023). JP Morgan Chase CEO Jamie Dimon has predicted that “AI might require future workers to work only 3.5 days”, while McKinsey Global Institute has predicted that GAI and other tools will automate 30% of work by 2030 (Ellingrud et al., 2023) and constitute the new productivity frontier for businesses (Chui et al., 2023). Current estimates put AI adoption by companies much lower: McElharan et al. (2024) use data from the

2018 Annual Business Survey of firms across the United States to estimate that fewer than 6% of companies use AI-related technologies in any capacity.

As companies have adopted GAI tools in various facets of business, the results have been mixed even for tasks that are similar such as customer-facing bots implemented in different industries and companies. Klarna, a financial technology company, is reaping significant benefits by adopting an AI assistant capable of doing work equivalent to 700 full-time customer service agents (Klarna, 2024). The AI assistant resolves customer queries in about one-fifth of the time and does so more accurately than the humans. As a result, customer satisfaction with the AI assistant has been similar to that with human agents and the overall productivity gains thus far are projected to result in a \$40 million profit improvement in 2024. However, not all chatbots have been equally well-received. Koko is an online mental health support company whose recent experiment with using GPT-3 to respond to customers seeking mental health support turned controversial (Ingram, 2023). Strong criticism followed when news about the experimental use of a generative AI bot was made public by Koko's co-founder. This was despite the fact that the AI-generated responses were deemed to be at least as good as responses that would have been written entirely by humans.

The implementation of generative AI has been uneven, with some seemingly straightforward tasks leading to public and costly failures while other difficult tasks enjoyed resounding successes. For example, Air Canada's use of a chatbot that provided incorrect advice to a customer seeking information on the airline's bereavement policy eventually landed the airlines in court (Proctor, 2024). The airlines chose to not stand by the pricing information offered by the chatbot and ultimately lost the lawsuit brought about by the customer who was able to provide the details of the interaction with the chatbot as evidence of the inaccurate information. However, GAI has been used to create new drugs with promising initial results. Starting with patient-focused problems and using GAI to find a molecule that targets the problem without harming the patient can lead to breakthroughs. In particular cases of proteins with a specific function, the need for AI generated solutions has been validated at the proof-of-concept stage by tech companies like Generate Biomedicines (Heaven, 2023). The regulatory process is currently incorporating changes that would allow AI discovered drugs to enter testing, a complex system of innovation with a focus on patient safety (FDA, 2024).

Adoption of new technologies is a costly and difficult decision for most businesses and organizational leaders (Ataman et al., 2023). If the technology is capable of replacing human labor partially or completely, the stakes become even higher not only for business leaders who decide whether to adopt, but also for the workers who may have to learn the new technology or even risk being replaced. GAI has the potential to disrupt the labor market (Frank et al., 2019), which leads to considerable concerns for workers (Yehiav, 2023).

There is thus a need to characterize tasks where generative AI may yield the highest benefits in terms of cost reduction and/or increased revenue, thereby improving the bottom line. As managers struggle to decide to what degree to adopt and deploy generative AI in their organizations, our conceptual framework is a call for more empirical research to measure the boundaries of the proposed theory with a focus on task creativity and the degree of human validation. We focus on jobs where there is a need for human validation after GAI is implemented.

AI, Management, and Organizational Outcomes

Management involves coordinating tasks and activities to achieve specific objectives and characterizes creatively solving problems that arise during the planning, organizing, leading, and controlling of resources. Simply put, it is the "judicious use of means to accomplish an end" (Merriam-Webster, n.d.). Accordingly, the application of generative artificial intelligence to real-world management problems has profound implications for the business sector.

It should come as no surprise that the use of AI is expanding as companies increasingly leverage AI to enhance the organization's profitability and efficiency. McKinsey's 2022 State of AI survey finds that AI adoption has more than doubled since 2017 (McKinsey, 2022). The survey also estimates that, though originally used to manage manufacturing (e.g., simulations like 3-D modeling) and risk (e.g., fraud and debt analytics), AI now generates the most significant impact on revenues in the management of marketing and sales (e.g., customer service analytics and segmentation), product and service development (e.g., creation of new AI products), and strategy and corporate finance (e.g., capital allocation and M&A support) and on costs in supply chain management (e.g., sales and demand forecasting and logistic networks optimization) (McKinsey, 2022). Thus, the promise of GAI benefits has shifted from automation of simple discrete tasks to creative and complex tasks that can increase productivity and supplement human labor. The benefits of GAI can come in the form of reduced costs, time savings, higher labor productivity, or higher customer satisfaction.

The 2024 State of AI survey (McKinsey, 2024b) suggests a shift in this trend, with the most significant cost decreases coming from generative AI use in the human resource management function and revenue increases in supply chain management. The proportion of companies that spend most aggressively (more than 20% of their digital budget according to the survey) on generative AI is the greatest in energy and materials (17%), technology (11%), and financial services (11%) (McKinsey, 2024b). Healthcare, manufacturing, and retail operations, however, have traditionally seen the strongest impact of AI innovation (Dogru & Keskin, 2020). The capacity of AI to enhance business operations is based largely on enhanced productivity by reducing search time and processing faster data used in decision making, and freeing employees to engage in creative and adaptive work (Tarafdar, 2019).

A vast body of academic research corroborates the growing significance of AI in the management of organizational processes to achieve organizational goals. Kobbacy et al. (2007) review 1200 papers published in the period 1995-2004, and Kobbacy & Vadera (2011) review 1400 papers published between 2005-2009 to classify AI applications in 4 areas: (1) design, (2) scheduling, (3) process planning and control, and (4) quality, maintenance and fault diagnosis. More recently, Jarrahi et al. (2023) assess the potential of AI in supporting aspects of knowledge management, including creation, storage and retrieval, sharing, and application of knowledge in organizations. AI is also applied in the analysis of customer data to personalize the company's interaction with customers, which has been linked to increased satisfaction, brand loyalty, and business growth (Odejide & Edunjobi, 2024). AI is increasingly used to manage credit, market and operational risk, as well as compliance (Aziz & Dowling, 2019) and is even transforming educational management in streamlining administrative tasks, improving student outcomes and enhancing the learning process (Igbokwe, 2023).

Several studies have investigated the use of AI and AI-driven tools on the financial performance of the firm. Mullangi (2017) argues for a positive long-term correlation between AI application and the firm's economic outcomes when AI is used for predicting market trends and more efficient resource allocation to guide decision-making aimed at optimizing relationships across units and between the organization and its environment. AI can improve the firm's economic outcomes in several ways, including enhanced efficiency, automated and fast adaptation to a rapidly changing external environment, detection of fraud, optimization of sales processes, and automating quality management (Wamba-Taguimdje, 2020). GAI is also shown to affect firm value and future profitability vis-à-vis employees' exposure to GAI - the degree to which the workforce can be replaced by GAI - where an increase in firm value is notable in companies when the workforce is more exposed to GAI (Eisfeldt et al., 2024).

Different occupations entail varying levels of nonroutine cognitive (analytical and interpersonal), routine cognitive and manual, and nonroutine manual (physical and interpersonal) skills and research shows that technology substitutes for workers in routine cognitive and manual tasks and complements workers in non-routine problem solving and complex communication tasks (e.g., Acemoglu & Autor, 2011; Autor et al., 2003). Furthermore, 40% of white-collar occupations characterize high GAI exposure in comparison to 9% for blue-collar and service occupations (Eisfeldt et al., 2024). Also, there are distinctive differences in how managers approach AI-based innovation management falling in one of four distinctive clusters according to Füller et al. (2022) - AI-Frontrunners, AI-Practitioners, AI-Occasional innovators and Non-AI innovators. Such variation of occupational and task exposure to GAI, degree of managerial openness to AI-based innovation management, and employee skill requirements suggests that the association between the

application of GAI in the management of organizational problems and the realized organizational benefits may be non-linear.

Task Creativity and AI

Individual differences in creativity are conventionally measured using divergent thinking tests (Silvia et al., 2008). Creativity has traditionally been viewed as an ability mainly attributable to human beings and the common consensus is that the notion of independently creative AI is still many years away (Hwang, 2022). In that sense, AI can be used as a substitute for routine tasks which can unlock employee creative potential. In support of this, Koivisto and Grassini (2023) argue that the "best humans still outperform artificial intelligence in a creative divergent thinking task" (p. 13601). Specifically, they conduct a popular divergent thinking test to compare the creativity of 256 humans to that of three current AI chatbots. Findings suggest that on average, humans often include poor-quality ideas relative to the chatbots, but the top human ideas are comparable to, or better than those, produced by AI. More recently, Castelo et al. (2024) find that new product ideas generated by GPT4 are rated higher than those produced by laypeople and professionals. Similarly, Girotra et al. (2023) find that ChatGPT outperforms MBA students in coming up with new product ideas, a surprising result for Prof. Terwiesch, one of the researchers, who said "I had naively believed that creative work would be the last area in which we humans would be superior at solving problems [...]" (Basiouny, 2003).

Such findings have caused some to question the usefulness of a comparison between humans and AI and to call for a shift in focus from competing to collaborating with AI for enhanced creativity (e.g., Elfa & Dawood, 2023). Accordingly, Wu et al. (2021) propose the term "AI creativity" - "the ability for human and AI to co-live and co-create by playing to each other's strengths to achieve more" (p. 171) and Elfa and Dawood (2023) suggest a choice board of options to use AI in helping people overcome the limitations of the human brain and enhance creativity by providing feedback on unfinished work and producing supplementary artwork. Empirical research supports this trend. Using a dataset of more than 4 million artworks from over 50,000 unique users Zhou et al. (2024) show that text-to-image GAI increases human creative productivity by 25%.

Relying on previous research on creativity in general, we envision the concept of task creativity as the degree to which creativity should be applied for successful task completion. Thus, tasks may be on a continuum from low creativity tasks, such as summarizing legal documents, summarizing customer reviews, aggregating search results, to high creativity tasks such as creating new protein molecules for new drugs or writing a movie script. Table 1 provides some examples of white-collar occupations classified by task creativity taking into consideration the dominant set of tasks in each occupation. As can be observed, certain occupations are inherently more creative than others.

Table 1

A Potential Classification of White-Collar Jobs based on Task Creativity

White Collar Occupations	Task Creativity
Accountant	Low
Interpreter	Low
Customer service representative	Low
Data engineer	Low
Drug innovation professional	High
Writer	High
Coder	High
Data Scientist	High
Music creation	High
Marketing content creation	High
Employee training videos creation	High

Given the present work investigates the role of creativity at the level of the task, Table 2 includes examples across several industries of tasks that vary in their degree of creativity, the benefit to the company from using GAI and whether the task was successfully completed at the time GAI was implemented.

Human Validation and AI

As has been noted earlier, use of GAI has led to mixed results, with some use cases yielding higher productivity gains through faster task completion and reduced labor needs, and others resulting in an array of outcomes including no efficiencies, lawsuits and negative publicity. Unlike human intelligence, GAI lacks contextual understanding of the application. This can result in an inauthentic or untrustworthy GAI-created output (Robinson, 2024), which could fail to satisfy the task requirements or customer specification. To derive usefulness from the automation of a GAI-created output, there is a need for human supervision that can validate the task output created by GAI. Human validation can ensure that GAI output is relevant, as well as useful and safe for the specific task performed and in the context of the larger managerial application.

The degree of human validation needed for the GAI output may vary by task. It could range from a cursory scan to validate a routine task in a customer service application to a longer and more thoughtful validation process in a critical marketing campaign. Deploying human vali-

ation of the GAI task output will be a necessary overlay to maximize the benefits of GAI through various actions. These actions could include creating accurate training prompts (Metcalf, 2024), mitigating fakes or lies (Stimson, 2023) which are euphemistically called hallucinations but should be called “bullshit” as suggested by Hicks et al. (2024), eliminating out-of-context results (Hoffman, 2023), incorporating cultural perspective for the application (Tsanni, 2023) and ensuring bias free results (Trifilo & Blau, 2024; Unesco, 2024a). AI bias has been a growing concern associated with risks in deploying AI systems. In general research, bias is defined as the generation, intentionally or unintentionally, of systematic errors by selecting one outcome or answer over another (Pannucci & Wilkins, 2010). In AI, this can take the form of, for example, a face-recognition algorithm being trained with more photos of light-skinned faces than dark-skinned faces, leading to poor performance in recognizing darker-skinned faces (Srinivasan & Chander, 2021). Biases can occur at every stage of AI development, including data collection, preparation and annotation as well as model development, deployment, and evaluation (Gichoya et al., 2023). Hence, the human validation aspect can be integral in the successful adoption of GAI systems.

A Conceptual Framework of GAI Benefits

Creativity, long believed to be a feature mainly attributable to humans, is now beginning to be simulated by generative AI models. However, applying GAI to the business world is still in a nascent phase. We suggest that a potential ramification of the development of GAI is to predict how it would perform tasks usually undertaken in white collar jobs by humans. These tasks may vary in their creativity level, potentially leading to different levels of GAI substitution, as seen in the examples in Table 2. Thus, we propose that:

Proposition 1: Generative AI can be used to substitute human labor for low and highly creative tasks in white collar jobs to derive GAI benefits.

To build on current research, future work might focus on a systematic analysis of jobs or occupations that would be amenable to GAI substitution for human labor. The current research landscape suggests examples of white collar jobs without fully defining the boundaries of what might be possible for GAI in the future.

Moreover, despite the promise that GAI will be fully automated to take over complex business processes, there is considerable uncertainty that organizational processes are amenable to that level of GAI integration. Instead, we expect that GAI will substitute for some tasks currently performed by human white-collar workers and that humans will need to operate in a supervisory role. Thus, human workers may be involved with the GAI output in different degrees: some tasks and roles may be a simple check for context while others may require detailed validation. Therefore,

Table 2*Examples of Task Creativity in Real World Business Applications*

Company	GAI Task	Creativity	Benefit to Company	Outcome
Qualcomm	make TikTok videos	High	reduce hiring of people with video editing skills	G
Ecolab	analyze rival earnings reports	Low	help Ecolab prepare their own earnings call	60% G, 40% B
Cisco	employee conflict management	High	successful conflict diagnosis	G
Klarna	customer service for a Fin Tech company	Low	handle workload of 700 customer service agents	G
Koko	send mental health support messages to users	Low	Controversial, caused bad press for firm	B
Mind Meld	send public relation pitches to journalists	Low	no efficiencies since humans had to clean up the gen AI output	B
Air Canada	give bereavement policy info	Low	company had to refund the customer due to incorrect info being provided	B
Trivago	create new ad campaign for global markets	High	reduced from 35 different productions to one by using AI to translate to 12 other languages.	G
Coca Cola	Create “Masterpiece” commercial	High	created a unique commercial drawing a lot of attention	G
Generate Biomedicines	software twists strands of amino acids to form new proteins	High	creating specific drugs for a particular condition	G

Note. Outcomes can be Good (G) or Bad (B).

Proposition 2: Generative AI can be used with low and high human validation to derive GAI benefits.

As the GAI technology improves, we see a need for more research to determine the uses for human validation for GAI. So far, managers and researchers are deploying GAI with ad-hoc levels of human validation and checking. Future field studies could more precisely determine the human validation needs for particular tasks.

Organizations that aim to deploy GAI must be prepared to determine the extent to which the technology should be used. Thus, human validation of GAI becomes a strategic choice for investment. We propose that to maximize GAI benefits managers should match the level of human validation to the task creativity when GAI is used to substitute for white collar labor tasks.

Proposition 3: The degree of human validation is a strategic management choice that organizational leaders must align with the task creativity needs to maximize GAI benefits.

Management scholars might investigate the organizational capabilities that allow organizational leaders to deploy GAI as a strategic tool rather than a timely new technology or fad. The degree of GAI adoption along with human validation are two important variables that organizations will have to consider in addition to all the other strategic decisions, thus opening a new research perspective in the future.

Several business occupations such as accountants, paralegals, and customer service agents perform largely routine cognitive tasks that may also be repetitive and rule-based. These tasks include preparing financial statements from input data, summarizing large amounts of text or providing customer service in the form of answering customers' questions or providing informational assistance to customers. The degree of creativity required in their daily tasks is relatively low as their tasks are governed by discipline-specific rules and does not offer opportunity for much flexibility or imagination to the task. Utilizing GAI in these tasks presents an opportunity to reduce the time

to complete the task while using fewer employees, thereby reducing costs, increasing productivity, and ultimately improving the overall profitability of the company. This would lead to high GAI benefits. The output of the tasks needs to be aligned with well-defined rules and parameters within the task problem. Thus, the degree of human validation needed for any error correction in the output is expected to be minimal. For example, a task like text summarization/annotation involves generating concise summaries of lengthy documents or articles. GAI algorithms can rapidly extract key points and summarize content automatically, with humans reviewing summaries for accuracy. Thus, we posit that:

Proposition 4: GAI used to perform low creativity tasks can yield high GAI benefits when the GAI-created output requires low human validation.

In this proposition we theorize that for tasks with low creativity and low human validation needs, GAI adoption is likely to yield high benefits, which implies that businesses should adopt GAI for these specific conditions first. Thus, future research might investigate empirically if businesses are finding these types of tasks as their first for GAI adoption. Moreover, as GAI technology improves, we expect that such studies would capture the changing business decision-making in further GAI adoption across organizations.

GAI may be less suited for certain low creativity tasks. Some low creativity tasks such as compliance audits, which ensure adherence to financial regulations and standards, require significant human oversight as a crucial validation to interpret nuanced regulations and ensure ethical standards. Though the tasks may be repetitive or routine, the ability to interpret complicated yet flexible accounting regulations requires a high level of human validation. Thus, any efficiencies gained from deploying GAI in these tasks could be largely negated by the costs of an increased level of human validation, which may lead to low overall benefits of GAI adoption. Therefore,

Proposition 5: GAI used to perform low creativity tasks can yield low GAI benefits when the GAI-created output requires high human validation.

One of the most promising areas of GAI applications, tasks with high creativity that require very little human validation, have been studied in academic research experimentally (Peng et al., 2023) and already applied in business settings (Soper, 2023). For example, code generation and automated programming involving writing code and developing software applications are particularly well suited for GAI applications. GAI tools can generate code snippets, automate repetitive programming tasks, and even assist in algorithm development with minimal human validation, as the resulting algorithms can then pass automated tests to ensure they are successful in achieving the desired results. We thus expect that:

Proposition 6: GAI used to perform high creativity tasks can yield high GAI benefits when the GAI-created output requires low human validation.

High creativity tasks in business may encompass a range of activities including writing (scripts, novels, song lyrics), producing music, producing imagery (photos, videos, art), creation of marketing content, and creation of employee training videos. Each of these tasks requires a specific artistic skill set, usually acquired and honed through many years of learning and practice. Most of these high creative tasks also consume a significant amount of time to produce the desired output. Substituting the human effort on these tasks by applying GAI promises to bring to the endeavor a level of imagination at least equivalent to humans but with the ability to produce the creative output in a much shorter time and with many output options to choose from (Swant, 2023). This would greatly reduce the time expended on, for example, producing a creative advertisement, while also reducing the need to hire several employees with different skill sets to produce the advertisement (Swant, 2023). The drawback with the GAI application in this case is the possibility of too fanciful or unrealistic GAI-produced output (Weekman, 2024). The creative endeavor may also require a deep understanding of human emotions, cultural nuances or contextual factors that are challenging for GAI to replicate. This necessitates extensive human validation to lend human judgment on the appropriateness and effectiveness of the GAI output. We expect that despite this human overlay to the machine-produced creative output, the overall savings in labor and time could be substantial. Therefore,

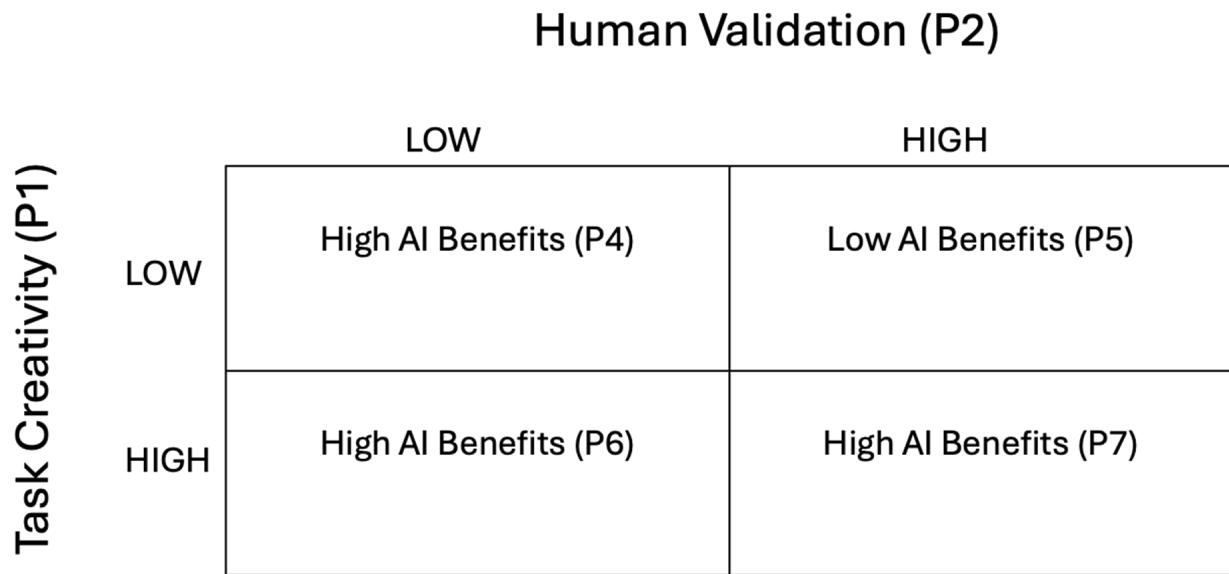
Proposition 7: GAI used to perform high creativity tasks can yield high GAI benefits when the GAI-created output requires high human validation.

Future research might investigate the degree to which humans need to validate the GAI output for each industry and domain. Since different areas of business have different requirements for creativity, such empirical research would characterize the nuanced levels of high human validation needed in each industry or area.

Figure 1 summarizes the conceptual framework. Our framework focuses on white collar jobs that may benefit from substituting GAI for some tasks. We suggest that the level of benefits organizations may derive from GAI implementation will depend on the level of task creativity and human validation needed. In most combinations of task creativity and human validation, GAI is expected to yield a high benefit if deployed correctly. However, Proposition 5 describes that for low creativity tasks that require high validation the expected benefit from GAI is low. This level is relative to other returns and includes benefits that are greater than zero. Companies might still choose to use GAI in this scenario, because a low or moderate benefit to GAI might still be preferable from a cost perspective.

Practical Implications for Managers

Our framework has immediate practical and managerial implications. While it is estimated that a quarter of Fortune 2000 companies have a VP or higher position to

Figure 1*Conceptual Framework for GAI benefits*

Note. P3 is the overarching proposition, which captures the entire conceptual framework and is implied in the above figure.

manage AI in the organization (and the number is expected to grow to 80% in 2024), the role of this leadership position is not yet well defined (Molla, 2023). One of the core tasks of an AI manager should include how to best make use of GAI to ensure maximum benefits from the technology. We suggest that one of the ways a manager could determine how to deploy GAI is to match the creativity level of the task to the level of human validation that a successful implementation of the task would require. A thoughtful planning of how to use GAI tools is a core function of an AI manager who can then deploy human and GAI assets accordingly.

Moreover, managers should consider legal and regulatory aspects of using GAI. Human validation is crucial for businesses: if the level of human validation does not match the needs of the task, it can lead to costly mistakes for the brand, antagonize loyal and prospective customers, and expose the company to legal challenges. Thus, managers should carefully decide if the benefit of using GAI is worth the risk, and whether the company has the capacity to correctly deploy GAI.

Limitations and Future Research

The present work has several limitations. It is conceptual, which limits its ability to reflect real-world scenarios and to provide concrete evidence or validate the propositions. There is no standardized measure for creativity across different tasks and contexts and creativity is inherently subjective and difficult to quantify uniformly across disciplines and industries. Also, focusing specifically on task creativity and human validation might overlook other critical factors influencing AI economic outcomes, such

as the firm's technological infrastructure, strategic alignment, scalability, and AI integration with existing systems. Taken together, these factors limit the generalizability of the proposed framework.

Nevertheless, future research can explore several avenues based on the present conceptual framework. As GAI is increasingly integrated into various management functions, the potential for biases in decision-making when creative and complex tasks are concerned, including for which AI is used to augment or automate decisions, becomes more prominent. Ohlheiser (2024) suggests that "Those dominant patterns might show up in the training data an AI system learns from, in the tasks it is asked to complete, and in the algorithms that power its learning process." Bias in GAI introduced through training datasets that might contain skewed, incorrect or biased information is difficult to identify accurately and timely, and to correct (Robertson, 2024). Biased AI systems can lead to adverse economic outcomes for the firm vis-à-vis decisions regarding unfair and discriminatory practices, strategic resource allocation, the development of competitive advantages, talent search, employee promotion, and customer satisfaction among others. Unesco (2024b) identified gender biases that might be particularly damaging for hiring decisions for example "Open-source LLMs in particular tended to assign more diverse, high-status jobs to men, such as engineer, teacher and doctor, while frequently relegating women to roles that are traditionally undervalued or socially-stigmatized, such as "domestic servant", "cook" and "prostitute"." Decisions influenced by biased AI can have profound consequences for individuals and society. Understanding and mitigating such bias-

es will affect the quality of managerial decision-making. Thus, conducting empirical studies on the impact of biased AI decisions in real-world scenarios, and evaluating the effectiveness of mitigation practices is an important area for future research on the economic outcomes of AI. Anchoring bias in decision-making – a biased perception caused by an anchoring effect from AI decisions, which restricts the consideration of alternative scenarios – is particularly problematic (Rastogi et al., 2022) and could be prioritized in future investigations on how to improve the benefits of AI systems that managers consider for integration into business operations.

The expectation is that organizations will continue to replace human employees with intelligent machines, especially in the performance of creative tasks, potentially making the workforce totally unrecognizable by 2040 (De Cremer & Kasparov, 2021). Thus, a key managerial consideration will be the degree to which and ways in which the integration of AI alters the roles and responsibilities of human managers. Questions like how AI helps and hinders managerial work and the issue of trust between managers and employees and its subsequent impact on AI adoption rates can be examined. Concomitantly, if human work is increasingly replaced by AI, how will AI influence organizational culture – one of the most powerful levers for successful implementation of a firm's strategic plan – is a viable research question. Even more importantly, culture itself may change dramatically. The promotion of important behaviors like innovation, risk-taking, and collaboration all rooted in the company's organizational culture has managerial and organizational implications ripe for investigation.

Future research might focus on the ethical considerations of GAI adoption that substitutes completely for human labor. As nascent technologies rely on human labor to develop (Grynbaum & Mac, 2024) and then have the potential to completely replace the humans that created the inputs, it is important to consider the boundaries of the ethics of such substitution processes and the potential worker protections or compensations that should be put in place.

Workforce skill transition and mobility within the organization are major AI-related concerns. To that end, understanding how AI impacts career progression and job roles within firms and identifying key skills that employees need to effectively collaborate with AI will be a significant challenge for managers. AI-based machines are fast, versatile, and can be trained to perform a variety of human tasks. As of now, humans excel in intuition, emotion, moral reasoning, sensory experiences, and cultural sensitivity. One area for research is identifying the kind of opportunities for collaborations between organic and inorganic intelligence associated with optimal firm outcomes.

Conclusion

To quote Dr. Harrick Vin, CTO of Tata Consultancy Services on AI innovation, “This is not a one-time transformation” (Jacobs, 2024). As generative AI improves

and produces more authentic output, the benefits of implementing GAI in management tasks will need to be reevaluated. Our current framework suggests that decision makers will continuously update their strategic decisions regarding the degree of human validation deployed relative to the task creativity requirements. The degree of human validation will also depend on algorithm quality and availability of data for training, and will in turn affect the GAI benefit derived by the organization. Potential empirical validations of the theory will have to consider contemporaneous technological advances.

References

- Acemoglu, D. (2024). *The simple macroeconomics of AI*. NBER, Working Paper 32487. 10.3386/w32487. <http://www.nber.org/papers/w32487>
- Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043-1171). Elsevier.
- Ataman, K., Myhr, N., & Nistor, C. (2023). Disruptive innovation as a network dilemma: A conceptual model. *Journal of Behavioral and Applied Management*, 23(2), 104-113. <https://doi.org/10.21818/001c.84555>
- Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4), 1279-1333. <https://doi.org/10.1162/003355303322552801>
- Aziz, S., & Dowling, M. (2019). *Machine learning and AI for risk management* (pp. 33-50). Springer International Publishing.
- Barnett, T. (2024). *Here's why AI probably isn't coming for your job anytime soon*. Fast Company. https://www.fastcompany.com/91130124/heres-why-ai-probably-isnt-coming-for-your-job-anytime-soon?utm_source=pocket-newtab-en-us
- Basiouny, A. (2023). *Is ChatGPT a better entrepreneur than most?* Knowledge at Wharton. <https://knowledge.wharton.upenn.edu/article/is-chatgpt-a-better-entrepreneur-than-most>
- Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work*. NBER, Working Paper No. 3116. <http://dx.doi.org/10.3386/w31161>
- Castelo, N., Katona, Z., Li, P., & Sarvary, M. (2024). How AI outperforms humans at creative idea generation. *SSRN*, <http://dx.doi.org/10.2139/ssrn.4751779>
- Caufield, B. (2023). *SIGGRAPH Special Address: NVIDIA CEO brings generative AI to LA show*. Nvidia. <https://blogs.nvidia.com/blog/siggraph-2023-special-address/>
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smajc, K., Sukharevsky, A., Yee, L., & Zemmle, R. (2023). *The economic potential of generative AI: The next productivity frontier*. McKinsey & Co. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-AI-the-next-productivity-frontier#business-value>
- De Cremer, D., & Kasparov, G. (2021). AI should augment human intelligence, not replace it. *Harvard Business Review*, 18(1). <https://www.daviddecremer.com/wp-content/uploads/>

- HBR2021_AI-Should-Augment-Human-Intelligence-Not-Replace-It.pdf
- Dilmegani, C. (2024). *Top 35+ generative AI tools by category* (Text, Image...), AIMultipleResearch. <https://research.aimultiple.com/generative-ai-tools/>
- Dogru, A. K., & Keskin, B. B. (2020). AI in operations management: Applications, challenges and opportunities. *Journal of Data, Information and Management*, 2(2), 67-74. <https://doi.org/10.1007/s42488-020-00023-1>
- Eisfeldt, A. L., Schubert, G., Zhang, M. B., & Taska, B. (2024). *The labor impact of generative AI on firm values*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4436627
- Elfa, M. A., & Dawood, M. E. T. (2023). Using artificial intelligence for enhancing human creativity. *Journal of Art, Design and Music*, 2(2), 106-120. <https://doi.org/10.55554/2785-9649.1017>
- Ellingrud, K., Sanghvi, S., Gurneet Singh Dandona, G.H., Madgavkar, A., Chui, M., White, O., & Paige Hasebe, P. (2023). *Generative AI and the future of work in America*. McKinsey Global Institute. <https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america>
- FDA (2024). Artificial intelligence and machine learning (AI/ML) for drug development. FDA. <https://www.fda.gov/science-research/science-and-research-special-topics/artificial-intelligence-and-machine-learning-aiml-drug-development>
- Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., & Rahwan, I. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences of the United States of America*, 116(14), 6531-6539. <https://doi.org/10.1073/pnas.1900949116>
- Füller, J., Hutter, K., Wahl, J., Bilgram, V., & Tekic, Z. (2022). How AI revolutionizes innovation management—Perceptions and implementation preferences of AI-based innovators. *Technological Forecasting and Social Change*, 178, 121598. <https://doi.org/10.1016/j.techfore.2022.121598>
- Gichoya, J. W., Thomas, K., Celi, L. A., Safdar, N., Banerjee, I., Banja, J. D., Seyyed-Kalantari, L., Trivedi, H., & Purkayastha, S. (2023). AI pitfalls and what not to do: mitigating bias in AI. *The British Journal of Radiology*, 96(1150), 20230023. <https://doi.org/10.1259/bjr.20230023>
- Girotra, K., Meincke, L., Terwiesch, C., & Ulrich, K. T. (2023). Ideas are dime a dozen: Large language models for idea generation in innovation. The Wharton School Research Paper Forthcoming. <https://ssrn.com/abstract=4526071> or <http://dx.doi.org/10.2139/ssrn.4526071>
- Grynbaum, M. M. & Mac, R. (2023, December 27). The Times sues OpenAI and Microsoft over A.I. use of copyrighted work. *New York Times*. <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>
- Heaven, W. D. (2023). AI is dreaming up drugs that no one has ever seen. Now we've got to see if they work. *MIT Technology Review*. <https://www.technologyreview.com/2023/02/15/1067904/ai-automation-drug-development/>
- Hicks, M.T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technol.*, 26(2). <https://doi.org/10.1007/s10676-024-09775-5>
- Hoffman, B. (2023, March 31). Leaders looking to leverage AI need to think about context. *Forbes*. <https://www.forbes.com/sites/brycehoffman/2023/03/31/leaders-looking-to-leverage-ai-need-to-think-about-context/>
- Hwang, A. H. C. (2022). Too late to be creative? AI-empowered tools in creative processes. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (pp. 1-9).
- Igbokwe, I. C. (2023). Application of artificial intelligence (AI) in educational management. *International Journal of Scientific and Research Publications*, 13(3), 300-307.
- Ingram, D. (2023). *A mental health tech company ran an AI experiment on real users. Nothing's stopping apps from conducting more*. NBC News. <https://www.nbcnews.com/tech/internet/chatgpt-ai-experiment-mental-health-tech-app-koko-rcna65110>
- Jacobs, E. (2024, April 13). The rise of the chief AI officer. *Financial Times*. <https://www.ft.com/content/4c5d823e-004b-4f00-9536-9fea93330453>
- Jarrah, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*, 66(1), 87-99.
- Klarna (2024). Klarna AI assistant handles two-thirds of customer service chats in its first month. <https://www.klarna.com/international/press/klarna-ai-assistant-handles-two-thirds-of-customer-service-chats-in-its-first-month/>
- Kobbacy, K.A.H. & Vadera, S. (2011). A survey of AI in operations management from 2005 to 2009. *Journal of Manufacturing Technology Management*, 22(6), 706-733. <https://doi.org/10.1108/17410381111149602>
- Kobbacy, K.A., Vadera, S., & Rasmy, M. H. (2007). AI and OR in management of operations: History and trends. *Journal of the Operational Research Society*, 58, 10-28.
- Koivisto, M., & Grassini, S. (2023). Best humans still outperform artificial intelligence in a creative divergent thinking task. *Scientific Reports*, 13(1), 13601.
- Korinek, A., & Suh, D. (2024). Scenarios for the transition to AGI. National Bureau of Economic Research. *Working Paper* 32255. https://www.nber.org/system/files/working_papers/w32255/w32255.pdf
- McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. (2024). AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy*, 33, 375-415. <https://doi.org/10.1111/jems.12576>
- McKinsey (2022). *The state of AI in 2022 – and a half decade in review*. QuantumBlack. <https://www.mckinsey.com/~media/mckinsey/business%20functions/quantumblack/our%20insights/the%20state%20of%20ai%20in%202022%20and%20a%20half%20decade%20in%20review/the-state-of-ai-in-2022-and-a-half-decade-in-review.pdf>
- McKinsey (2024a). What is generative AI? <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>

- McKinsey (2024b). *The state of AI in 2022 – and a half decade in review*. QuantumBlack. <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai#/>
- Merriam-Webster (n.d.) Dictionary. <https://www.merriam-webster.com/dictionary/management>
- Metcalfe, C. (2024, April 24). Job titles of the future: AI prompt engineer. *MIT Technology Review*. <https://www.technologyreview.com/2024/04/24/1091125/ai-prompt-engineer-generative-ai-job-titles/>
- Molla, R. (2023). The hottest new job is “head of AI” and nobody knows what they do. *Vox*. <https://www.vox.com/technology/2023/7/19/23799255/head-of-ai-leadership-jobs>
- Mullangi, K. (2017). Enhancing financial performance through AI-driven predictive analytics and reciprocal symmetry. *Asian Accounting and Auditing Advancement*, 8(1), 57-66.
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381, 187-192. 10.1126/science.adh2586
- Odejide, O. A., & Edunjobi, T. E. (2024). AI in project management: Exploring theoretical models for decision-making and risk management. *Engineering Science & Technology Journal*, 5(3), 1072-1085.
- Ohlheiser, A.W. (2023, June 14). AI automated discrimination. Here’s how to spot it. *Vox*. <https://www.vox.com/technology/23738987/racism-ai-automated-bias-discrimination-algorithm>
- Pannucci, C. J., & Wilkins, E. G. (2010). Identifying and avoiding bias in research. *Plastic and reconstructive surgery*, 126(2), 619-625.
- Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of AI on developer productivity: Evidence from GitHub Copilot. <https://doi.org/10.48550/arXiv.2302.06590>
- Proctor, J. (2024, February 15). Air Canada found liable for chatbot’s bad advice on plane tickets. *CBC News*. <https://www.cbc.ca/news/canada/british-columbia/air-canada-chatbot-lawsuit-1.7116416>
- Rastogi, C., Zhang, Y., Wei, D., Varshney, K. R., Dhurandhar, A., & Tomsett, R. (2022). Finding fast and slow: The role of cognitive biases in AI-assisted decision-making. *Proceedings of the ACM on Human-computer Interaction*, 6(CSCW1), 1-22.
- Robertson, A. (2024). Google apologizes for ‘missing the mark’ after Gemini generated racially diverse Nazis. *The Verge*. <https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical>
- Robinson, K. (2024, May 30). Google defends AI search results after they told us to put glue on pizza. *The Verge*. <https://www.theverge.com/2024/5/30/24168344/google-defends-ai-overviews-search-results>
- Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, 2(2), 68–85. <https://doi.org/10.1037/1931-3896.2.2.68>
- Soper, T. (2023, June 13). AI and coding: How these tech companies are using generative AI for programming. *GeekWire*. <https://www.geekwire.com/2023/ai-and-coding-how-seattle-tech-companies-are-using-generative-ai-for-programming/>
- Srinivasan, R., & Chander, A. (2021). Biases in AI systems. *Communications of the ACM*, 64(8), 44-49.
- Stimson, B. (2023, December 29). Michael Cohen admits to inadvertently citing fake cases generated by AI in legal motion. *Fox News*. <https://www.foxnews.com/politics/michael-cohen-admits-inadvertently-citing-fake-cases-generated-ai-legal-motion>
- Swant, M. (2023, December 18). AI Briefing: Creative ways companies used AI in advertising in 2023. *DigiDay*. <https://digiday.com/media-buying/ai-briefing-creative-ways-companies-used-ai-in-advertising-in-2023/>
- Tarafdar, M., Beath, C. M., & Ross, J. W. (2019). Using AI to enhance business operations. *MIT Sloan Management Review*, 60(4), 37-44.
- Trifilo, A., & Blau, G. (2024). Developing racially ambiguous job candidate avatars with strong versus ambiguous job candidate backgrounds in preparation for a field experiment for enhancing inclusive hiring and reducing implicit bias. *Journal of Behavioral and Applied Management*, 24(1), 14–25. <https://doi.org/10.21818/001c.115890>
- Tsanni, A. (2023, November 17). This company is building AI for African languages. *MIT Technology Review*. <https://www.technologyreview.com/2023/11/17/1083637/lalapa-ai-african-languages-vulavula/>
- Unesco (2024a). Challenging systematic prejudices: An investigation into bias against women and girls in large language models (Catalogue Number 0000388971). <https://unesdoc.unesco.org/ark:/48223/pf0000388971.locale=en>
- Unesco (2024b). Generative AI: UNESCO study reveals alarming evidence of regressive gender stereotypes. <https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes>
- Wamba-Taguimdje, S. L., Wamba, S. F., Kamdjoug, J. R. K., & Wanko, C. E. T. (2020). Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects. *Business Process Management Journal*, 26(7), 1893-1924.
- Weekman, K. (2024, March 1). 'Terrifying' Willy Wonka event used AI images to mislead guests. Here's how to avoid getting duped by fake marketing photos. *Yahoo News*. <https://www.yahoo.com/news/terrifying-willy-wonka-event-used-ai-images-to-mislead-guests-heres-how-to-avoid-getting-duped-by-fake-marketing-photos-191205992.html>
- Wu, Z., Ji, D., Yu, K., Zeng, X., Wu, D., & Shidujaman, M. (2021). AI creativity and the human-AI co-creation model. In *Human-Computer Interaction. Theory, Methods and Tools: Thematic Area, HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I 23* (pp. 171-190). Springer International Publishing.
- Yehiav, G. (2023, August 8). Will AI Augment Or Replace Workers?. *Forbes*. <https://www.forbes.com/sites/forbestechcouncil/2023/08/08/will-ai-augment-or-replace-workers/>

Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS nexus*, 3(3), 1-8.

Zia, A. (2023, September 3). 18 biggest generative AI companies in the world. *Yahoo Finance*. <https://finance.yahoo.com/news/18-biggest-generative-ai-companies-140447246.html>

Charu Sinha (csinha@chapman.edu)

Veselina Vacheva (vracheva@chapman.edu)

Cristina Nistor (nistor@chapman.edu)
