Meet, Greet or Tweet: Empowering Employees with Role-Optimized AI Assistance

Claus Hegmann-Napp
University of Hamburg
Tijmen Jansen
University of Hamburg
Mark Heitmann
University of Hamburg

Cite as:
Hegmann-Napp Claus, Jansen Tijmen, Heitmann Mark (2025), Meet, Greet or Tweet: Empowering Employees with Role-Optimized AI Assistance. *Proceedings of the European Marketing Academy*, 54th, (126091)

Paper from the 54th Annual EMAC Conference, Madrid, Spain, May 25-30, 2025

Meet, Greet or Tweet: Empowering Employees with Role-Optimized AI Assistance

This paper investigates the transformative potential of fine-tuned large language models (LLMs) in product design and innovation, focusing on their role in analyzing customer reviews and fostering human-AI collaboration. By leveraging a dataset of Amazon reviews, the research fine-tuned an LLM to synthesize customer insights and generate innovative product ideas. The study includes three empirical investigations: the comparative quality of ideas from fine-tuned LLMs, base LLMs, and human participants, the effectiveness of presenting AI as expert collaborators in enhancing human creativity, and the impact of different communication modalities (text, audio, and video) on the phases of idea creation. Results demonstrate the superior ability of fine-tuned LLMs in addressing consumer issues, enhancing purchase intent, and supporting human ideation processes. Participants expressed positive perceptions of AI collaboration, particularly in structured and innovative workflows. This research advances the integration of AI in design and innovation, emphasizing the synergy of human expertise and AI tools to redefine creativity and productivity.

Keywords: *Generative AI, AI-Human Collaboration, Product Development*

Proposed Track: *Innovation Management & New Product Development*

1. Introduction

LLMs are revolutionizing innovation and collaboration by offering powerful tools for generating insights and enhancing creativity. Fine-tuning these models for specific tasks has proven highly effective, particularly in analyzing unstructured data to uncover actionable opportunities (Flory et al. 2017; Zhou et al. 2018; Samha et al. 2014). While LLMs have been extensively applied in technical domains, their potential to enhance collaboration and drive innovation in non-technical contexts remains underexplored.

This study evaluates the dual role of fine-tuned LLMs: their ability to generate innovative ideas and their impact on human-AI collaboration. Beyond their analytical capabilities, we examine how different communication modalities (text, audio, and video) affect collaboration. Drawing on media richness theory (Kahai and Randolph 2003), we demonstrate that video fosters trust and alignment in rich discussions, audio supports brainstorming, and text excels in documentation and asynchronous refinement. These modality-specific strengths align with different phases of the innovation process, showcasing the potential of AI-human synergy to improve productivity and creativity.

By integrating LLMs into collaborative workflows, this research highlights how businesses can leverage AI to transform innovation processes. The findings demonstrate that tailoring AI tools to specific tasks and communication needs, enhances both the quality of ideas and the effectiveness of human-AI collaboration.

2. Related Literature

2.1 Applications of AI and AI-Human collaboration in Marketing Literature

The integration of artificial intelligence (AI) into professional domains is transforming workflows by enhancing efficiency, enabling deeper analysis, and unlocking creative potential. In marketing, AI excels in processing large datasets and automating repetitive tasks, offering capabilities unmatched by traditional methods (Arora et al. 2024). Generative AI models and LLMs enable users to rapidly generate and analyze content, making complex tasks faster and simpler. However, this efficiency has increased competition in repetitive tasks, particularly affecting freelancers (Demirci et al. 2024). Professionals must embrace AI as a collaborative tool to maximize their value and output.

AI-human collaboration is particularly impactful in customer engagement, where human intelligence (HI) adds empathy and contextual understanding to AI's data-driven automation. While AI offers efficiency, human oversight ensures meaningful interactions, avoiding customer disconnect (Huang and Rust 2021). Emerging technologies like "Feeling AI," which recognizes emotions, enhance customer care but require human involvement for ethical and context-sensitive applications (Huang and Rust 2024). Hybrid approaches that combine AI's analytical strengths with human contextual understanding consistently outperform standalone methods (Arora et al. 2024).

This synergy between AI and HI also drives innovation in domains like marketing research. AI accelerates problem-solving by processing vast amounts of data and identifying patterns, while humans contextualize these insights within broader strategies (Verganti et al. 2020). As a decision-support tool, AI addresses complex challenges, but its success depends on complementing, rather than replacing, human decision-makers to ensure trust, ethics, and contextual relevance (Mariani et al. 2023; Longoni et al. 2019). By integrating AI's analytical power with human expertise, organizations can enhance operational efficiency, foster strategic innovation, and remain competitive in an evolving global landscape.

2.2 Virtual Communication Modalities in Idea Creation

In today's remote workforce, employees rely on text, audio, and video communication tools to foster collaboration and innovation. Each modality offers varying levels of richness, defined by its ability to deliver cues and provide immediate feedback, making it suited to different phases of idea creation: ideation, development, and finalization (Kahai and Randolph 2003; Agulnik et al. 2024).

The ideation phase, focused on brainstorming and exploration, benefits most from audio communication. Audio's immediacy and ability to enable real-time interactions foster creativity and the rapid exchange of ideas, making it ideal for divergent thinking (Sundaravej et al. 2015; Skootsky et al. 2023). Text complements audio by capturing and organizing ideas for further refinement (Morrison-Smith and Ruiz 2020).

In the development phase, where ideas are refined and aligned with goals, video conferencing is essential. Its ability to deliver visual, auditory, and non-verbal cues fosters trust, resolves ambiguities, and enhances collaboration, making it ideal for consensus-building and transforming ideas into actionable plans (Purvanova 2013). Text remains vital for documenting decisions and maintaining alignment, particularly in asynchronous contexts (Hosseinkashi et al. 2024).

The finalization phase, focused on preparing ideas for implementation, is dominated by text-based communication. Text provides clarity, permanence, and supports asynchronous collaboration, enabling teams to formalize concepts and create comprehensive records for use (Morrison-Smith and Ruiz 2020; Hosseinkashi et al. 2024). While video or audio may occasionally present finalized ideas, text ensures outcomes are precise and accessible.

Aligning communication tools with each phase enhances creativity, collaboration, and decision-making: audio supports ideation, video facilitates development, and text ensures clarity in finalization.

3. Empirical Analysis

We conduct a series of studies to explore the role of our fine-tuned LLM in workplace idea generation and collaboration. First, we examine its ability to address challenges compared to off-the-shelf LLMs and humans. Next, we investigate how presenting the LLM as an AI expert influences human performance and perceptions. Finally, we explore which communication modalities optimize AI-human collaboration across idea creation phases.

3.1 Study 1: Idea Quality of Fine-tuned LLMs

This first study demonstrates the fine-tuning process of the LLM and evaluates the performance of this fine-tuned LLM in generating innovative ideas compared to off the shelf LLMs and human participants. The focus is on the LLM's ability to address consumer issues effectively, leveraging customer reviews. The study consists of two stages: collecting and grouping ideas from humans and base LLMs and then evaluating these ideas to determine their quality, commercial potential, and problem-solving ability.

Method

We developed a structured pipeline to analyze customer reviews and generate innovative ideas, using a publicly available dataset of 230 million Amazon reviews. After narrowing the focus to non-technical products in the "Office Products" category, we selected a portable lap desk with 2,427 reviews as the target. Relevant reviews were filtered based on criteria such as recentness, readability, and sentiment extremes, using SetFit, a fine-tuned Transformer-based classifier with 90.47% accuracy, yielding 1,186 useful reviews.

These relevant reviews were used to fine-tune OpenAI's GPT-4-Turbo, enabling the model to synthesize insights and generate ideas tailored to recurring themes and user problems. By incorporating product descriptions and GPT-4 Vision-analyzed images, the model generated 15 innovative ideas based on the extracted insights. This pipeline formed the basis for evaluating the LLM's performance in idea generation compared to humans and base LLMs.

In the first stage, ideas were collected from 50 non-expert participants recruited via Prolific (25 female, 25 male, $\text{age}_{\text{range}} = 19\text{–}61$). Participants were divided into two groups: 26 participants received review summaries generated by the LLM, while 24 participants worked without any review information. Both groups were tasked with generating as many innovative ideas as possible, producing 93 and 74 ideas, respectively. Additionally, a off the shelf LLM generated a set of ideas. From these, the four most commonly mentioned ideas were manually grouped into four sets, each containing one idea from the fine-tuned LLM, the off the shelf LLM, and the two human groups.

Consequently, in the second stage, 65 participants (36 female, 29 male, ages 20–70), recruited via Prolific, evaluated these ideas. Participants were randomly assigned one set of four ideas, with the origins of the ideas concealed to prevent bias. After reviewing the product description, participants' initial purchase intent is asked on a 1-7 Likert scale. Subsequently, they are provided with the key issues derived from reviews. Participants ranked the ideas from 1 (best) to 4 (worst) and rated each idea on four criteria (liking, innovativeness, purchase intent and ability to address issues) using a 7-point Likert scale. This evaluation resulted in 1,040 ratings across 16 ideas.

Results & Discussion

We find that the ideas generated by our fine-tuned LLM significantly outperformed all other groups in idea quality. Mann-Whitney U tests revealed that our LLM's ideas (M = 2.18) were ranked better than those of the base LLM (M = 2.52, $p < .05$), humans with review summaries (M = 2.71, $p < .01$), and humans without summaries (M = 2.59, $p < .05$).

Participants rated the ideas of our fine-tuned LLM's higher on liking (M = 5.50) compared to the base LLM (M = 4.80, $p < .01$), humans with summaries (M = 4.80, $p < .01$), and humans without summaries (M = 4.98, $p < .05$). Additionally, To inspect the performance on purchase intention, we calculate the difference in purchase intention, measures before and after the ideas are presented. Our LLM's ideas showed a greater increase in purchase intent compared to other groups ($p < .05$ for all groups). Importantly, our LLM performed best in addressing consumer issues (M = 5.30) compared to the base LLM (M = 3.45, $p < .01$), humans with summaries (M = 4.73, $p < .01$), and humans without summaries (M = 4.42, $p < .01$). No significant differences were found between groups regarding innovativeness, aligning with prior findings that LLMs do not outperform humans in creating novel ideas (Boussioux et al. 2024). These results demonstrate our LLM's ability to generate commercially viable and consumer-focused solutions, validating its potential as a helpful tool for innovation.

3.2 Study 2: Performance and Perceptions of Working with AI Experts

Building on the findings of Study 1, which demonstrated the effectiveness of our fine-tuned LLM in generating innovative ideas, this study investigates whether humans benefit from collaborating with such AI experts. The study focuses on whether AI experts improve human idea creation and explores participants' perceptions of working with AI, including their comfort, engagement, and belief in AI's impact on workplace productivity.

Method

To simulate a realistic workplace task, 59 participants (25 females, 29 males, $\text{age}_{\text{range}} = 25\text{–}58$) were recruited via Prolific. Participants, all with interests in marketing or entrepreneurship and holding at least a bachelor's degree, were tasked with improving a product based on a description and 1,186 product reviews. During a 10-minute case study, participants generated as many improvement ideas as possible and evaluated these ideas on a 7-point Likert scale for innovativeness (Boussioux et al. 2024), including the ability to address issues, the likelihood of implementation, financial value, and overall quality.

After this initial case study and survey, participants were introduced to an AI expert, fine-tuned on the content of consumer reviews, which provided a summary of the most mentioned product issues and areas for improvement. Participants were additionally asked to generate new ideas and evaluated these in the same manner. Following this, participants rated the AI expert on engagement, clarity, and usefulness, and indicated their preference for future collaboration with such tools. Finally, participants shared their overall opinions on AI, including interest, comfort, usefulness, and potential impact on workplace productivity.

Results & Discussion

Participants expressed a strong willingness to work with AI experts, with ratings for the AI expert significantly exceeding the scale midpoint (M = 5.29, p ¡ .01). The reliability of idea evaluations was confirmed with a Cronbach's α of .75 before and .83 after the AI expert intervention, validating the consistency of participant responses across the four evaluation criteria. Using these reliable measures, we found that the evaluation of participants' ideas improved after seeing the AI expert ($M_{before} = 6.17$ vs. $M_{after} = 6.36$), though this difference was not statistically significant ($p = .26$), likely due to the small sample size or self-assessment bias. To address these limitations, increasing the sample size and involving independent experts for idea evaluation will provide a more accurate and objective representation of idea quality. Nevertheless, participants strongly agreed that their initial ideas would have been of higher quality if they had access to the AI expert beforehand ($M = 5.20$, $p < .01$).

Participants evaluated the AI-expert positively on engagement, clarity, and future preference, with a Cronbach's α of .82 and an average score significantly above the scale midpoint ($M = 6.06$, $p < .01$). Additionally, participants expressed a favorable overall opinion of AI-experts, reflected in a high reliability score (Cronbach's $\alpha = .93$) and an average rating of $M = 5.98$ ($p < .01$), indicating interest, comfort, and recognition of their usefulness and potential to enhance workplace productivity. These findings suggest that AI-human collaboration improves performance and is well-received in the context of idea creation.

3.3 Study 3: Matching Interactive AI with Correct Tasks and Formats

Building on the findings of previous studies, this study will examine how different modalities (text, audio, and video) of an interactive AI-expert impact performance during various phases of the idea creation process. Study 1 demonstrated the superiority of fine-tuned LLMs in idea generation, while Study 2 highlighted the benefits of AI-human collaboration. Study 3 will focus on identifying which communication format best supports specific tasks such as brainstorming, problem-solving, and documentation.

Method

An interactive AI-expert system will be created using our fine-tuned LLM, offering three modalities: text chat, audio call, and video call with an AI-avatar. The text format will integrate the fine-tuned LLM, while the audio format will use real-time Text-to-Speech (TTS) to convert outputs into spoken responses. For video, TTS will be paired with an AI-avatar using lip synchronization. A third-party platform will enable real-time interactivity for audio and video formats. Participants will be assigned to case studies designed to represent the three phases of idea creation—ideation, development, and finalization. Each group will initially collaborate without AI assistance before repeating the case study with one randomly assigned AI-expert format (text, audio, or video). Performance will be evaluated across all phases on metrics such as idea quality, time efficiency, and discussion content. This approach will provide insights into which modality is best suited to each phase of idea creation, offering practical guidance on integrating AI-experts effectively into collaborative workflows.

4. Conclusion and Outlook

This research highlights how fine-tuned LLMs leverage customer reviews to generate innovative, consumer-focused ideas. Fine-tuned LLMs consistently outperformed untrained models and human participants, demonstrating their ability to address key consumer concerns with greater precision and quality. Integrating AI into collaborative workflows also underscored the importance of presentation formats in enhancing innovation cycles.

In our future study, the evaluation of text, audio, and video modalities can show that each format has unique strengths depending on the phase of idea creation. Video is expected to have most effective for discussions requiring non-verbal cues, audio can excel in brainstorming and dynamic exchanges, and text proved optimal for documentation and refinement. In general, we find participants to express positive perceptions of working with interactive AI systems, emphasizing their potential to improve idea quality, engagement, and productivity.

These outcomes pave the way for a new era of workforce innovation, where AI-driven insights and human collaboration revolutionize productivity, creativity, and engagement. The framework proposed in this paper serves as a cornerstone for leveraging AI in innovation workflows, transforming how ideas are generated and refined. By seamlessly integrating AI capabilities into collaborative processes, this approach can redefine efficiency and user satisfaction, setting a new standard for product design and development in the age of AI.

References

Agulnik, A., Boykin, D., O'Mally, D. M., Price, J., Yang, M., McKone, M., Curran, G., and Ritchie, M. J. (2024). Virtual facilitation best practices and research priorities: a scoping review. *Implement Sci Commun*, 5(16).

Arora, N., Chakraborty, I., and Nishimura, Y. (2024). EXPRESS: AI-human hybrids for marketing research: Leveraging LLMs as collaborators. *Journal of Marketing*.

Boussioux, L., Lane, J. N., Zhang, M., Jacimovic, V., and Lakhani, K. R. (2024). The crowdless future? generative ai and creative problem-solving. *Organization Science*, 35(5).

Demirci, O., Hannane, J., and Zhu, X. (2024). Who is AI replacing? the impact of generative AI on online freelancing platforms. *SSRN Electronic Journal*.

Flory, L., Osei-Bryson, K.-M., and Thomas, M. (2017). A new web personalization decision-support artifact for utility-sensitive customer review analysis. *Decision Support Systems*, 94.

Hosseinkashi, Y., Tankelevitch, L., Pool, J., Cutler, R., and Madan, C. (2024). Meeting effectiveness and inclusiveness: Large-scale measurement, identification of key features, and prediction in real-world remote meetings. *Proceedings of the ACM on Human-Computer Interaction*, 8(93).

Huang, M.-H. and Rust, R. T. (2021). Engaged to a robot? the role of AI in service. *Journal of Service Research*, 24(1).

Huang, M.-H. and Rust, R. T. (2024). The caring machine: Feeling AI for customer care. *Journal of Marketing*, 88(5).

Kahai, S. S. and Randolph, B. C. (2003). Exploring the core concepts of media richness theory: The impact of cue multiplicity and feedback immediacy on decision quality. *Journal of Management Information Systems*, 20(1).

Longoni, C., Bonezzi, A., and Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4).

Mariani, M. M., Machado, I., and Nambisan, S. (2023). Types of innovation and artificial intelligence: A systematic quantitative literature review and research agenda. *Journal of Business Research*, 155.

Morrison-Smith, S. and Ruiz, J. (2020). Challenges and barriers in virtual teams: a literature review. *SN Appl. Sci.*, 2(1096).

Purvanova, R. K. (2013). Virtual versus face-to-face teams: What have we really learned? *The Psychologist-Manager Journal*, 17(1).

Samha, A. K., Li, Y., and Zhang, J. (2014). Aspect-based opinion extraction from customer reviews. In *Computer Science Information Technology (CS IT)*.

Skootsky, T. J., Sanchez, D. R., and Kawasaki, K. (2023). Feeling connected: Technology-mediated communication and the relationship between modality and affective outcomes. *Multimodal Technologies and Interaction*, 7(11).

Sundaravej, T., Mirchandani, D., and Lederer, A. (2015). Synchronous collaboration technology use in teamwork. *Proceedings of the 48th Hawaii International Conference on System Sciences*.

Verganti, R., Vendraminelli, L., and Iansiti, M. (2020). Innovation and design in the age of artificial intelligence. *Journal of Product Innovation Management*, 37(3).

Zhou, S., Qiao, Z., Du, Q., Wang, G. A., Fan, W., and Yan, X. (2018). Measuring customer agility from online reviews using big data text analytics. *Journal of Management Information Systems*, 35(2).