

Article

Building Meta-Dynamic Capabilities Through AI-HI Collaboration: Experimental Evidence from Multinational Organizations in Disaster Response Operations

Ingyu Oh ¹  and Li Fei ^{2,*} ¹ College of Global Engagement, Kansai Gaidai University, Osaka 573-1001, Japan; oingyu@kansai-gaidai.ac.jp² Hallyu International College, Sookmyung Women's University, Seoul 04310, Republic of Korea

* Correspondence: feili0820@sookmyung.ac.kr

Abstract

The rise in large language models (LLMs) has sparked renewed interest in how firms, particularly multinational aid organizations, can enhance learning related to meta-dynamic capabilities (DCs), such as agility, sensing, and adaptation, in response to disasters and humanitarian crises. A key strategic priority is developing meta-rules that combine general engagement frameworks with locally tailored action plans, grounded in cultural and institutional contexts. LLMs offer potential in supporting this need, but premature deployment risks harmful or misleading outcomes. This underscores the critical importance of collaboration between artificial and human intelligence (AI-HI). While AI brings computational power, it lacks the tacit knowledge—encompassing cultural, contextual, and intuitive understanding—that is essential in high-stakes, unpredictable environments. Our experimental study provides two core insights: (1) AI alone cannot effectively handle tasks requiring tacit knowledge, and (2) AI-HI collaboration thrives when human input guides AI using deep awareness of local social and political dynamics. We contribute to the discourse on dynamic capabilities in multinational contexts during catastrophic situations by offering practical strategies to support successful AI-HI partnerships and a framework for organizations aiming to enhance meta-DCs through responsible, human-centered use of disruptive technologies. Our findings clarify how the international dimensions of these capabilities influence their effectiveness across diverse cultural and institutional environments.

Keywords: artificial intelligence; dynamic capabilities; tacit knowledge; human intelligence; AI-HI collaboration; large language models

1. Introduction

The application of AI in areas such as ambidexterity, asset orchestration, and innovation has been well documented across various domains (Lessard et al., 2016; Sjödin et al., 2023; Simón et al., 2024; Gao et al., 2025). However, the role of AI-HI collaboration as meta-dynamic capabilities (DCs) for MNEs and multinational aid organizations during calamitous situations remains underexplored (Pitelis et al., 2023). Meta-DC refers to higher-order capabilities that enable organizations to develop and adapt other dynamic capabilities. These include agility, sensing, and building—core functions that will allow firms to detect environmental changes and determine appropriate responses, particularly during large-scale disasters and accidents that are on the rise in the 21st century (Hitt et al., 2021, 2024; Brock & Hitt, 2024). This paper addresses that gap by examining how



Received: 25 April 2026

Revised: 29 May 2026

Accepted: 3 June 2026

Published: 8 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

AI-HI collaboration can enhance meta-DC, particularly in response to large-scale disasters, environmental crises, and emergency relief efforts that affect both local communities and organizational stakeholders, including employees.

We argue that AI-HI collaboration offers a novel and effective pathway to strengthen these functions of building, agility, and sensing. In doing so, it transforms the often fragmented and rigid operational models found in many multinational disaster relief efforts, enabling the development of scalable frameworks that are both globally coherent and locally responsive (Xu et al., 2021). AI-HI collaboration may also destabilize established routines of cooperation and coordination. In some cases, these technologies can rapidly disrupt entire industry value chains (Schmeiss et al., 2019; Chin et al., 2022). As such, it is essential to consider how this new meta-DC can facilitate a smoother transition (Reitzig & Puranam, 2009; Sjödin et al., 2020a, 2020b; Chin et al., 2022). Recent research has begun to examine the role of such dynamic capabilities in mediating AI-assisted interventions during crises and high-stress scenarios, where human insight remains critical (Battisti & Deakins, 2017; Dubey et al., 2022; Cao, 2023).

This perspective on AI-HI collaboration as a meta-DC aligns with the emerging research questions in international business concerning how MNEs develop and deploy dynamic capabilities across diverse cultural, institutional, and political contexts at the time of calamities, which are rapidly becoming new normal (Hitt et al., 2021, 2024). Our approach specifically addresses the need to understand which capability configurations lead to successful implementation of strategies in different types of international markets, particularly in crisis situations that demand high levels of tacit knowledge. By examining how humanitarian aid organizations can leverage AI-HI collaboration to enhance their meta-DCs, we contribute to understanding both the “multinational” and “dynamic” aspects of these capabilities in complex global environments.

In the context of AI and humanitarian aid—the focus of this paper—Dubey et al. (2022) introduced the concept of artificial intelligence-driven big data analytics capability (AI-BDAC) to enhance complex decision-making. However, a significant gap remains in our understanding of AI-HI collaboration involving large language models (LLMs), which operate at a more advanced level than traditional AI-BDAC frameworks. Today’s AI technologies, particularly LLMs, surpass conventional big data analytics by aiming to replicate human cognitive processes in decision-making, including in disaster relief operations. LLMs introduce three advanced deep learning capabilities: (1) constructing causal models for explanation and understanding, (2) learning through intuitive theories, and (3) leveraging compositionality to transfer knowledge across tasks and contexts (Lake et al., 2017). Despite these advancements, LLMs face fundamental limitations. They lack alignment with key aspects of human cognition—such as emotion, motivation, and value-based reasoning—which are deeply rooted in cultural, ethical, and social contexts (Dörner & Güss, 2013). This disconnect underscores the importance of integrating human oversight and contextual intelligence into AI systems, particularly in high-stakes humanitarian settings.

The practical challenge of leveraging LLMs to enhance meta-DC extends beyond routine human supervision. LLMs often require continuous human intervention to address persistent issues such as hallucinations, inaccurate outputs, misleading recommendations, and even fabricated responses (Chatterjee et al., 2022; Yang et al., 2022; Alsamhi et al., 2024). While AI-HI collaboration holds promise for enabling timely and informed decision-making, particularly in humanitarian contexts, a critical question remains: how can we foster meaningful collaboration between human and artificial intelligence in environments that rely heavily on tacit knowledge and cultural or political sensitivity? Conventional methods of AI fine-tuning frequently fall short in these complex, high-stakes settings,

underscoring the need for more agile strategies that integrate human judgment, contextual awareness, and ethical considerations at the core of AI development and deployment.

In many not-for-profit humanitarian aid initiatives, overreliance on AI can lead organizational leaders to make decisions that are culturally, ethically, emotionally, or theoretically flawed and often lack creativity (Deloitte, 2018). This highlights the critical need to train large language models (LLMs) for effective AI-HI collaboration in coordinated humanitarian efforts. However, current research on LLM training within organizational contexts—particularly for fostering professional AI-HI cooperation among managers—remains limited and underdeveloped. There is a notable lack of best practices and methodological guidance for designing LLM training to support meaningful HI-AI collaboration (Amabile, 2020; Makarius et al., 2020; Chowdhury et al., 2022). Although AI is often seen as a disruptive force in managerial decision-making, prompting calls for organizational restructuring to accommodate AI-HI collaboration (Chowdhury et al., 2023), few studies offer concrete strategies to improve LLMs for this purpose or to enhance mutual learning between AI and human actors.

Given the nascent stage of research on AI-HI collaboration, our study is guided by two widely acknowledged insights. First, the integration of disruptive technologies, such as AI, into humanitarian aid does not inherently lead to improved outcomes compared to HI alone. Effective deployment requires the right dynamic capabilities—specifically, organizational knowledge in training and fine-tuning LLMs—to address the complexity and uncertainty inherent in disaster response (Battisti & Deakins, 2017; Dubey et al., 2022; Chin et al., 2022; Forliano et al., 2022; Cao, 2023). Second, deploying AI without appropriate HI oversight can result in catastrophic failures, with risks extending beyond individual projects to broader societal harm (Scharre, 2019; Turchin, 2019). From an entrepreneurial perspective (George et al., 2021; Kraus et al., 2020), implementing AI in such contexts demands risk-taking and a proactive approach to cultivating dynamic capabilities that can transform potential AI risks into opportunities for digital sustainability. Based on these foundations, this study explores two key questions: (a) What types of knowledge do AI systems fail to acquire effectively, leading to failures in humanitarian aid during crises? (b) How can organizations train LLMs (i.e., meta-DC) to foster effective AI-HI collaboration that enhances building, agility, and sensing in aid efforts?

Our focus on multinational humanitarian aid organizations is particularly relevant in today's complex global environment. These organizations operate across diverse institutional, cultural, and political contexts, necessitating sophisticated meta-DC that can balance global coordination with local responsiveness. By examining how AI-HI collaboration enhances meta-DCs in these organizations, we directly address the question of “how multinational” and “how dynamic” these capabilities truly are. Furthermore, our experimental approach allows us to measure the effectiveness of these capabilities across different cultural and institutional settings, contributing to the discourse on operationalizing the international dimensions of dynamic capabilities.

This study offers three key contributions. First, we examine the types of knowledge that AI, particularly LLMs, struggles to acquire and process, shedding light on the cognitive limitations of current AI systems. Second, we extend the theory of dynamic capabilities by positioning AI-HI collaboration as a new meta-DC for achieving effective outcomes in humanitarian aid. Specifically, we explore how organizations can reconfigure technological resources, such as LLMs, to drive organizational change amid rampant disasters and the rising need to offer relief missions, thereby capitalizing on new opportunities for AI-HI collaboration in addressing complex and interdependent challenges. Third, in response to the limited empirical research on adapting LLMs to evolving business models, our findings

offer practical guidance for organizations seeking to assess and strengthen their agility and readiness for fostering AI-HI collaboration in humanitarian contexts.

2. Literature Review and Hypotheses Development

2.1. Types of Knowledge and LLMs

The application of AI in knowledge work for various organizations has created a range of flaws and difficulties, as they have discovered that multiple types of knowledge and their storage (or memory) exist (Spender, 1996; Venkitachalam & Busch, 2012). Humans acquire both explicit and tacit knowledge, which is stored in various forms, including declarative, procedural, semantic, episodic, working, and sensory memories (Coakes, 2006; Cowan, 2008). These various types of anthropic knowledge and memories are also learned and stored in organizations managed by human beings. Previously, knowledge management presumed the complex interplay between humans and organizations. Now, it is a more complicated management process than ever, involving interaction among humans, organizations, and LLMs.

Although LLMs can enhance the efficiency of learning, processing, storing, and applying to practical situations, we observe in various cases that such algorithms demonstrate varying degrees of success and failure depending on the types of knowledge and memory (Amaya & Holweg, 2024). Most advanced LLMs contain explicit and tacit knowledge in the form of explicit, semantic, procedural, and other types of memory (i.e., accumulated knowledge learned from big data), and it is also understood that they excel better than humans in acquiring new explicit knowledge through quick lessons. However, it is also true that LLMs, despite repeated lessons, fail to absorb tacit knowledge as fast as humans (Amaya & Holweg, 2024).

This limitation in learning tacit knowledge by LLMs mainly derives from their inability to process and understand human instructions for solving complex and tightly coupled organizational and environmental problems, including medical, legal, disaster relief, humanitarian aid and other fields that require tacit knowledge acquired by highly trained professionals (Lebovitz et al., 2022; Jeyaraman et al., 2023; He et al., 2024; Sarker, 2024). Tacit knowledge, which is stored in human or organizational brains as semantic, procedural, priming, and long-term memories, also encompasses various judgmental and unconscious categories of cognitive abilities relevant to ethical judgment, value calculation, and cultural understanding (Haldin-Herrgard, 2000; Kucharska, 2017).

Instead of knowing what is or what to do, tacit knowledge often asks for solutions to how it is or how to do it (Lebovitz et al., 2022; Amaya & Holweg, 2024). Knowing how and what requires programming contextual and situational descriptions into LLMs before the machine makes human-like decisions (Choudhury et al., 2020). Thus, know-how programming becomes significantly more complicated than know-what programming, as the former requires educating LLMs to comprehend the ethical, cultural, and value contexts associated with each problem (Amaya & Holweg, 2024).

Additionally, the human tacit knowledge that can be easily transformed into masterful acts of human skills requires not only repeated learning, which enables human bodies to memorize newly learned skills (Sandberg & Tsoukas, 2011; Dreyfus, 2014; Hadjimichael & Tsoukas, 2019). What complicates the process of learning tacit knowledge is the role of occupational communities in defining legitimate forms of action and norms that describe how to work, rank, and validate individual and group achievements (Lave & Wenger, 1991; Hutchins, 1995). Consequently, mastering tacit knowledge entails assimilating a shared purpose, vocabulary, and a set of values that guide and legitimize the actions of the learners and practitioners (Wenger, 1999; K. C. Kellogg et al., 2006).

Another key dimension of tacit knowledge that remains difficult to translate into programmable language for LLMs is human intuition and creativity. Intuitive decision-making—often described as rapid, experience-based judgments—can, in many cases, outperform methodical, data-driven approaches (Gladwell, 2010). Developing this form of dynamic capability lies squarely within the domain of HI. Moreover, intuition is closely linked to creativity and innovation, both of which are essential in humanitarian aid contexts marked by extreme uncertainty. In such situations, decision-making must extend beyond factual knowledge to account for high levels of risk, requiring adaptive and creative responses that AI alone cannot deliver (March, 1991; D. L. Kellogg & Nie, 1995; Staats et al., 2011; Harvey, 2016).

Thus, we posit that manual dexterity, creativity, and social intelligence remain highly valuable but are challenging to codify into algorithms (Frey & Osborne, 2017; Sampson, 2021):

Hypothesis 1. *In simulated humanitarian aid scenarios, LLM performance scores on tasks requiring high tacit knowledge (e.g., dexterity, creativity, and social intelligence) are systematically lower than those on explicitly codified tasks.*

2.2. Dynamic Capabilities: AI-HI Cooperation

As discussed above, planning, designing, and implementing humanitarian aid during crises and extreme uncertainties require highly trained tacit knowledge and organizational learning geared towards nurturing structural contingencies to maximise the development, storage, and application of tacit knowledge in actual emergencies. In these catastrophic situations, AI alone cannot complete the task because it lacks the dynamic capabilities of learning, developing, and applying tacit knowledge comparable to those of HI. This mandates AI-HI cooperation for humanitarian aid during emergencies, and the central actors of AI-HI cooperation must foster specific high-order dynamic capabilities (Teece, 2014; Teece et al., 2016; Brock & Hitt, 2024) that are strategically flexible and agile enough to overcome various types of technical, communicational, and other algorithmic tensions rampant in AI-HI collaboration. We thus adopt a meta-DC view in this study.

The core of the dynamic capability required in AI-HI cooperation lies in the adoption and creative utilization of reinforcement learning from human feedback (RLHF) techniques by LLMs, such as ChatGPT (Kaufmann et al., 2023). RLHF intends to fine-tune AI's capabilities in learning, exploiting, and creating tacit knowledge with the help of HI as feedback through the three steps of supervised fine-tuning (SFT), reward model training (RM), and reinforcement learning via proximal policy optimization (PPO). At the supervised training stage, LLMs will receive prompts, upon which they will generate answers that will then be evaluated for fine-tuning. The second stage of RM involves training LLMs to generate multiple possible answers based on various prompts derived from different perspectives. Human labelers will then rank each answer from best to worst, simultaneously developing reward models for the LLM to learn only the best answers. The final stage of PPO applies the reward model discriminately so that the LLM would continuously maximise the best output through reinforcement learning (see Figure 1). Specifically, it is RM that requires human feedback the most and serves as a critical standard in the acquisition of tacit knowledge by LLMs (Ouyang et al., 2022). RM, therefore, is a crucial component of new assets that an organization needs to develop quickly to enhance its dynamic capabilities and transform organizations into a system of full AI-HI collaboration.

The acquisition of RM, along with SFT and PPO capabilities, requires a long process of trial and error. What is germane in the process of shortening the trial and error is how much an aid organization has accrued and is willing to absorb new theoretical knowledge, along with its accrued tacit and explicit knowledge, to generate new perspectives that would

expand the prompt database and theoretical models to fine-tune LLMs with ‘better’ answers than old ones. This collaboration between evolving HI and AI is critical to balancing AI-HI cooperation (see Figure 2) without extreme reliance on either HI or AI, which would generate false policies in humanitarian aid implementation (Li et al., 2024).

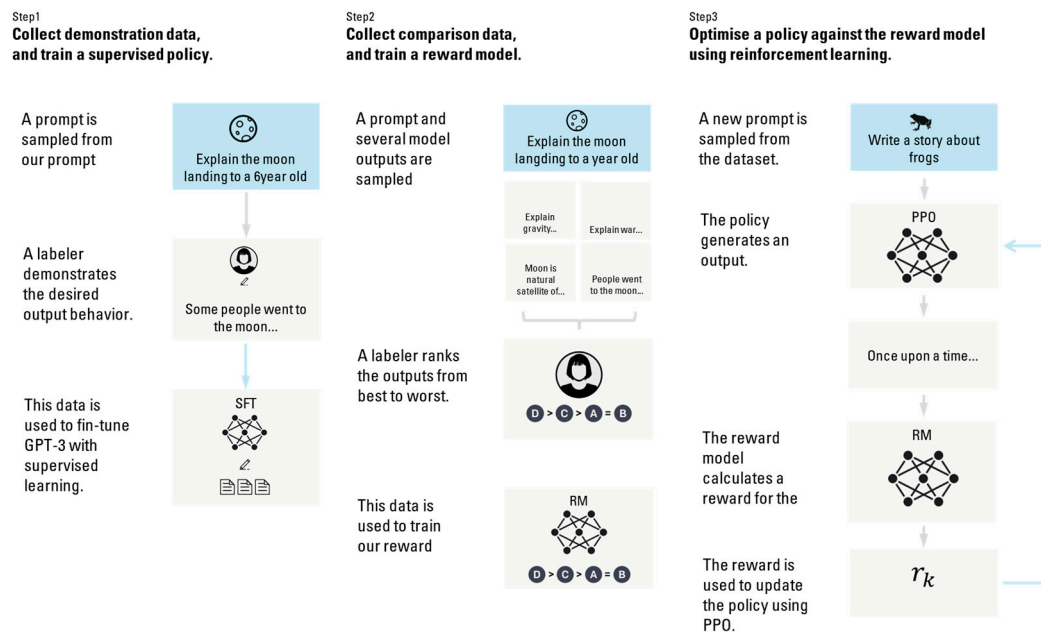


Figure 1. Reinforcement Learning from Human Feedback (RLHF). Source: Revised from Ouyang et al. (2022).

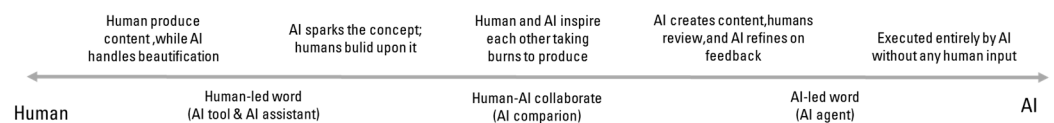


Figure 2. Balancing AI-HI Collaboration Source: Revised from Li et al. (2024).

Thus, we posit that the dynamic capabilities of AI-HI collaboration for aid organizations hinge on their HI knowledge capabilities and ability to maintain the balance in such interactions:

Hypothesis 2. *There is a statistically significant positive difference in LLM performance scores between pre- and post-test following the introduction of expert human intelligence (HI) feedback via guided prompting.*

3. Methodology

We employ an experimental design to test the proposed hypotheses using four different LLMs that use dissimilar LLMs. This one-group pretest-post-test design with no control group has the advantage of measuring the impact of instructions on different LLMs. The experimental design employs two testing procedures, as the first test suggests LLMs' inability to accrue and utilize highly trained tacit knowledge for complex and tightly coupled problems, such as humanitarian aid during disasters. We conducted this experiment using four different LLMs in mid-June 2024.

After establishing AI's incompetence in dealing with tacit knowledge, we provide further instructions to all LLMs in the experimental group. We then measure changes in the post-test stage to see if their tacit knowledge learning ability has improved. This way, we can identify the types of tacit knowledge that these four LLMs learn better than others. The

overall experimental process is summarized in Figure 3. This method is standard in most LLM testing for its intelligence competence and assurance (Freeman, 2020; Batarseh et al., 2021). We also believe that an experimental design with no control group remains reliable because LLMs are less susceptible to fortuitous external factors than humans during the interval between pre- and post-tests (Ouyang et al., 2022).

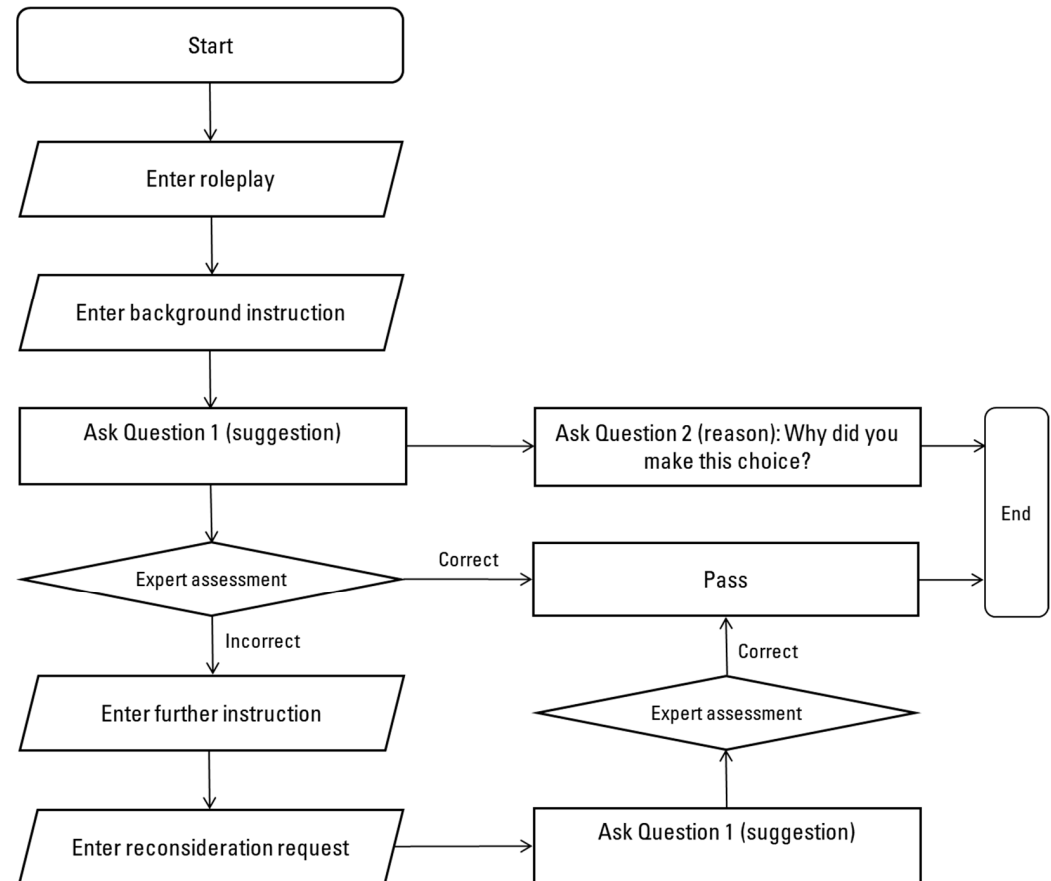


Figure 3. The Flow Chart of the Experimental Design.

Additionally, to test the validity of training sessions by professionals, this study primarily focuses on the qualitative comparison of pre- and post-test responses generated by LLMs. The assessments involve paired scores, which are derived from an independent evaluation by a professional reviewer with advanced expertise in humanitarian aid operations and cross-cultural disaster response management with over twenty years of field and advisory experience based on formal scoring criteria¹, including Factual Accuracy, Contextual Appropriateness, decision-reasoning alignment, and practical applicability. All responses were reviewed in a blind (i.e., pre/post-test labels were not used) and randomized order to reduce bias. Although single-expert review is a common practice in exploratory LLM research when highly specialized experts are in short supply (Batarseh et al., 2021; Freeman, 2020), we acknowledge that this approach limits inter-rater reliability, and future work should employ multiple evaluators and report inter-rater reliability coefficients.

As a result, the scoring data is discrete, specifically classified as interval variables, and may not follow a normal distribution. Therefore, non-parametric testing is the most appropriate method for statistically assessing the difference in quality between the pre-test and post-test responses from the same language model (Altman & Bland, 2009).

Non-parametric tests, such as the Wilcoxon signed-rank test, which we use to test our hypotheses, assess statistical significance based on the rank or order of the data without assuming a specific distribution. These tests are particularly useful when sample sizes are

small, when data do not follow a normal distribution, or when the data are non-continuous, such as nominal or ordinal scales (Conover, 1971). The Wilcoxon signed-rank test compares the medians of two related samples (Woolson, 2005) and is commonly used in fields such as medical research when data deviate from normality or sample sizes are limited (Meek et al., 2007). Since the data in this study meet the criteria for this test, we will apply the Wilcoxon signed-rank test as the appropriate statistical method.

Since our purpose is to test the two hypotheses that measure LLMs' systematic limitations in acquiring advanced tacit knowledge with or without interactions with HI and the impact of advanced HI on AI abilities in acquiring the knowledge for aid provision during emergencies, our study does not include a procedure for building an RM model, which falls under the purview of AI science, not knowledge management studies.

3.1. Sample LLMs in the Experiment

In the experimental group, we included the four most representative LLMs in the market that utilise LLMs: ChatGPT (gpt-3.5-turbo), Claude (claude-3-sonnet-20240229), Gemini (gemini-pro), and Mistral (open-mixtral-8x22b). Even though these sample LLMs are the most advanced and highly sophisticated with their evolving learning capabilities and upgrades through RLHF, we expect to find that all these LLMs demonstrated systematic limitations specifically in learning and applying tacit knowledge to complex and tightly coupled problems contextualized by cultural and ethical diversities. We posit that RLHF still neglects the importance of highly trained HI in each specialized field, which will teach AI LLMs with advanced, not common-sensical, knowledge. The AI developers insist on employing ordinary, rather than professional, footnote writers to fine-tune their LLMs. The following provides a brief description of each LLM we gathered from various internet sources. By far, these four examples are leaders in the LLM business.

Our experimental design deliberately incorporates scenarios across diverse cultural, social, and institutional contexts, using the GLOBE culture groups as a framework. This approach allows us to operationalize both the international dimensions of AI-HI capabilities and the international dimensions of context—two critical aspects identified in the recent literature on dynamic capabilities in multinational firms (Brock & Hitt, 2024). By testing LLMs' abilities to apply tacit knowledge in culturally diverse scenarios, we provide empirical evidence for how meta-DCs function across different international environments.

3.1.1. ChatGPT

ChatGPT is an artificial intelligence chatbot developed by OpenAI and launched on 30 November 2022. It is based on LLMs and uses natural language processing to engage in conversational interactions with users. Its key features include understanding and responding to user input in a human-like manner, generating text in various styles and formats, including articles, emails, and poetry, answering follow-up questions, admitting mistakes, challenging incorrect premises, rejecting inappropriate requests, and training using RLHF.

ChatGPT quickly gained popularity, reaching over 1.6 billion visits to its homepage in January 2024, making it the fastest-growing consumer application in history. It is credited with sparking increased interest and investment in artificial intelligence. OpenAI offers its ChatGPT service through both free and paid options.

While ChatGPT has impressed many with its capabilities, it is essential to note that it can sometimes produce inaccurate or misleading information, a phenomenon known as 'hallucination' (see Zhang et al., 2023).

3.1.2. Claude

Claude is an artificial intelligence chatbot developed by Anthropic, a startup founded by former OpenAI employees. Its key features include engaging in natural conversations and generating human-like responses to user prompts and questions, handling various tasks such as answering questions, writing content, analyzing documents, coding, and processing text and image inputs.

A key differentiator is Anthropic's focus on AI safety and ethics. They employ a technique called 'constitutional AI' to make Claude more helpful, honest, and less likely to produce harmful outputs than other AI chatbots. Claude can accept file uploads, such as PDFs and images, to provide context for queries.

It is designed to be more steerable, allowing users to give personality, tone, and behavior directions. Early users and partners have reported that Claude provides detailed, easily understandable responses and feels more conversational than some of its competitors. However, like all current AI systems, it can still produce inaccurate information and has various limitations.

3.1.3. Gemini

Gemini is Google's most advanced and capable artificial intelligence model to date. It maintains the following key features: First, it comes with multimodal capabilities that allow the LLM to seamlessly understand and process multiple types of information, including text, code, audio, images, and video. In our experiment, we did not utilize the multimodal capabilities as our samples are mostly text-based. Second, it also offers advanced abilities to process extensive information with long context windows. Given its creator's assets, Gemini is also integrated into various Google products and services.

Gemini's primary commercial application is robotics, for which Google's DeepMind team utilizes Gemini AI to enhance robot navigation and task performance. In tests, Gemini-powered robots achieved a 90% success rate in executing user commands within a large workspace. Like other AI LLMs, Google continues to improve Gemini, with recent updates including Gemini 1.5 Flash for enhanced speed and efficiency, as well as Project Astra, which explores AI agents for more natural interactions.

3.1.4. Mistral

Mistral AI is a French artificial intelligence company that has rapidly gained prominence in the AI sector since its founding in April 2023. Founded by former employees of Meta Platforms and Google DeepMind, Mistral AI focuses on producing large, open-source language models, although we also utilize fee-paying services. The company emphasizes open-source innovation and aims to 'democratize' AI, positioning itself as an alternative to proprietary models.

Mistral AI's technology is available through serverless APIs, public cloud services (Azure AI, Amazon Bedrock), and on-premise deployment. In February 2024, Mistral AI partnered with Microsoft to expand its presence in the AI industry, making its models available on Microsoft's Azure cloud. Mistral AI is currently ranked fourth in the global AI race based on valuation and is the highest-valued AI company outside the San Francisco Bay Area. Mistral AI's rapid growth, focus on open-source models, and significant funding have positioned it as a notable player in the AI industry, particularly in Europe.

3.1.5. Operationalizing Tacit Knowledge

We define tacit knowledge by three key features: (1) it cannot be codified solely by explicit rules or formal documentation, (2) it is acquired through extensive professional experience and social embedding within occupational communities, and (3) it is used in context-dependent ways that resist formal algorithmic codification (Haldin-Herrgard,

2000; Hadjimichael & Tsoukas, 2019; Venkitachalam & Busch, 2012). Tacit knowledge is deliberately opposed to ‘contextual complexity’, which is simply the scenario of layered or culturally framed knowledge.

To illustrate the difference, knowledge of Hofstede’s cultural dimensions (Tacitness = 10) is associated with high informational complexity but low tacitness, because it is formalized and easily retrievable. By contrast, the question of when democratic norms allow or disallow mobilization of external aid (Tacitness = 1) is a form of political and institutional tacit judgment that cannot be inferred from explicit documentary sources alone. It requires an internalized understanding of informal institutional logics and social norms that cannot be fully captured in the textual dataset.

The performance gradient across Table 7, where lower Tacitness scores are always accompanied by larger LLM failure even after HI intervention, thus reflects a systematic property of tacit knowledge itself, rather than incidental task ambiguity. All four LLMs consistently fail on Personality Factors, Digital Competitiveness, and Democracy (Tacitness = 1) but perform uniformly well on Hofstede’s cultural dimensions (Tacitness = 10). This provides construct-level evidence that the manipulated variable is tacitness, not prompt difficulty per se.

3.2. Testing Process

3.2.1. Pretest Procedure

The pretest procedure involves writing hypothetical scenarios of aid provision during catastrophes in various countries with significant cultural, social, and economic diversity. We developed scenarios based on different levels of difficulty in acquiring tacit knowledge. Some knowledge is widely available as common sense, whereas other knowledge requires advanced education and learning (Table 1).

Table 1. Key Variables in Aid Provision during Catastrophes Based on Tacit Knowledge.

Variables AIs Are Supposed to Know	Variables AIs Are Not Supposed to Know
1. Social Support	1. Cultural Differences
(a) Strong social connections and support networks are among the strongest predictors of resilience after disasters (Elcheroth & Drury, 2020; Naderpajouh et al., 2018). (b) Providing practical help, emotional understanding, sharing of experiences, and a sense of community (Brandhorst, 2018). (c) Facilitating access to knowledge, resources, and services needed for recovery (Revilla et al., 2018).	(a) Individualistic vs. collectivistic cultures have different effects on social support. Individualist cultures may lead to social cohesion by maintaining social distance (caring for others without causing inconvenience) (Wilkinson, 2010; Lair et al., 2005; Morand & Walther, 2018). (b) In contrast, in a collectivist culture, social support can be manifested by maintaining proximity between community members (Fincher et al., 2008; Murray & Schaller, 2010; van Leeuwen et al., 2012).
2. Coping Skills and Self-Efficacy	2. Digital Competitiveness
(a) Possessing effective coping strategies and believing in one’s ability to handle difficult situations (coping self-efficacy) (Bandura, 2000; Farh et al., 2007). (b) Reframing the situation positively, using flexible coping strategies, meaning-making, religious coping, and seeking support when needed (Drury et al., 2009, 2013).	(a) In individualist and collectivist cultures, social distance can be maintained if digital competitiveness is high (Longoni & Cian, 2022; Nguyen et al., 2021). (b) Virtual communication can lead to social support for both collectivist and individualist cultures (Oh et al., 2023; Zhou et al., 2015). (c) The availability of digital entertainment, like Netflix, greatly helps individuals overcome loneliness and enhance resilience during social distancing (McDonald & Smith-Rowsey, 2016; Gupta & Singharia, 2021).

Table 1. *Cont.*

Variables AIs Are Supposed to Know	Variables AIs Are Not Supposed to Know
3. Hope and Optimism	3. Trust
(a) Having a sense of hope, optimism, and belief that things will improve fosters resilience (Youssef & Luthans, 2007; Seger Handoyo et al., 2018). (b) Hope can stem from faith, spirituality, belief in external support (community, organizations, government), or a positive outlook (Ng et al., 2017; Dahan et al., 2022). (c) Practical resources like housing, employment, and finances can also cultivate hope (Marrone, 2005; Michalos, 2012).	(a) In societies where collectivist trust is high, social support can be reinforced through maintaining proximity among close community members (B. Chen, 2005). (b) In societies where generalized trust is high, social support can be reinforced among strangers without proximity to close community members (Frank & Schvaneveldt, 2014; Marques da Silva, 2021).
4. Personality Factors	4. Social Ties
(a) Emotional stability, agreeableness, and lower negativity are associated with greater resilience after disasters (Collins, 2007; Arora & Rangnekar, 2015). (b) Having a capacity for optimism, hope, and perceived ability to cope with the situation are protective factors (Pathak & Lata, 2018; Sagone et al., 2020).	(a) Communities bound by strong ties require proximity among close community members to reinforce social support during disasters (Islam & Walkerden, 2014; Panday et al., 2021). (b) Communities bound by weak ties can provide social support among weak-tie members without help from strong-tie members (Newman & Dale, 2005; Aldrich, 2012).
5. Community Cohesion and Resources	5. Democracy
(a) Socially cohesive and well-resourced communities are better equipped to identify needs, disseminate information, advocate for aid, and support recovery efforts after disasters (Sadri et al., 2018; Bharwani & Mathews, 2021; Stollenwerk et al., 2021). (b) Community-level factors like social capital, economic resources, and infrastructure play a crucial role in fostering resilience (Ledogar & Fleming, 2008; Brandhorst, 2018).	(a) Free speech and freedom of movement can invite various help from outsiders during disasters (Hall & Lamont, 2013; Chandler & Reid, 2016). (b) Non-democratic societies find it difficult to find outside help (Sapountzaki, 2014; Canetti et al., 2014).

As Table 1 indicates, we identified five variables from the extant literature on resilience-enhancing measures during catastrophes (e.g., humanitarian aid). The left column with five variables and deriving hypotheses indicates that LLMs are assumed to have already mastered this level of tacit knowledge. In contrast, the right column, with an additional five variables and derivative hypotheses, indicates that the LLMs have not yet been exposed to this level of tacit knowledge. Even though they are assumed to have already mastered popular resilience theories during catastrophes, we expect these LLMs to have limited capabilities in applying their tacit knowledge to actual problems contextualized in different cultural, social, and political contingencies.

We then constructed scenarios for each variable, incorporating diverse contextual contingencies in various regions of the world. For cultural contingencies, we used the GLOBE culture groups (<https://globeproject.com/results/clusters/anglo?menu=list.html#cluster>, accessed on 21 April 2026). After constructing ten scenarios with varying problem structures and regional contingencies, the next step was to feed these questions to the four LLMs to elicit their answers. Using the outputs, we then measured their capabilities of learning different levels of tacit knowledge and applying it to actual scenarios.

Scenario Construction

Each scenario follows a standard format, consisting of role-play instructions, background information, and multiple-choice and open-ended questions (see Table 2). Each scenario is based on our previous literature survey on resilience amid disastrous condi-

tions and has ten dimensions: Social Support, Coping Skills and Self-Efficacy, Hope and Optimism, Personality Factors, Community Cohesion and Resources, Cultural Differences, Digital Competitiveness, Trust, Social Ties, and Democracy (see also Table 1). The ten scenarios encompass diverse types of urgencies and catastrophic situations to test AI's capabilities in dealing with various complex and tightly coupled problems. Multiple-choice and open-ended questions enable experimenters to assess the validity and reliability of AI's responses to prompts.

Table 2. Sample Scenarios.

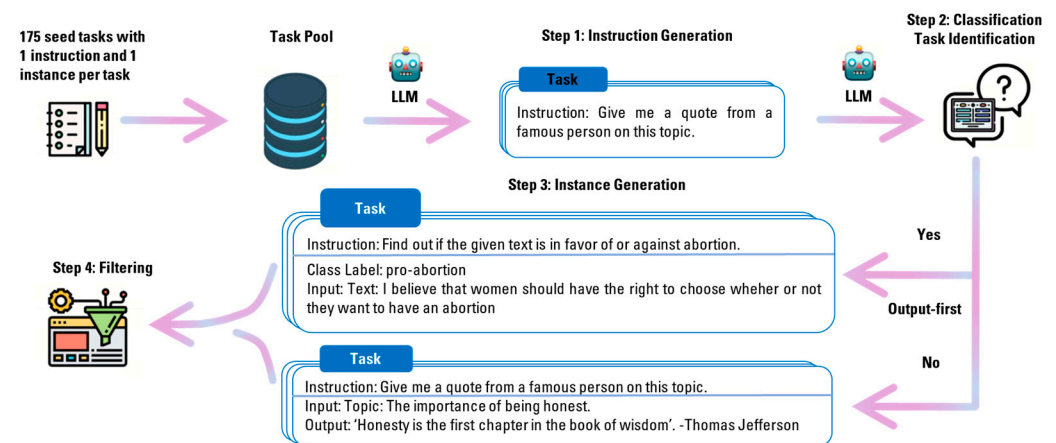
Background Instruction	Question 1	Question 2
In this scenario, you face a significant challenge requiring you to apply your knowledge and skills practically. Your country has the following cultural and social characteristics: (1) high scores of Performance Orientation; (2) low In-Group Collectivism; (3) high individualistic behaviour; medium Institutional Collectivism; (4) high collective distribution of resources and importance of group cohesion; medium Assertiveness, Future Orientation, Uncertainty Avoidance and Humane Orientation, and Power Distance; and (5) High Humane Orientation, Future Orientation, and Gender Egalitarianism. An earthquake of 7.5 magnitude on the Richter scale has blocked the entire town from outside traffic. Given the country's highly developed infrastructure and the Internet, victims of the earthquake in the town could quickly establish digital and virtual communication with family members and outsiders.	Based on your reasoning, please choose one of the following after-quake policy options: A. Enhance and provide in-group and strong-tie support to the victims of the quake B. Enhance and provide weak-tie support to the victims of the quake C. Provide free Netflix and unlimited internet usage to the victims of the quake D. Provide free psychological consultation to the victims of the quake	Please also explain briefly why you chose your option.
In this high-stakes scenario, you confront a significant challenge that demands the practical application of your knowledge and skills. Your country has the following cultural and social characteristics: (1) high scores of In-Group Collectivism; (2) high tendencies of maintaining close family and strong social ties; high Power Distance; medium Assertiveness; low scores on Future Orientation and Performance Orientation; extremely low Uncertainty Avoidance; and medium Humane Orientation, Institutional Collectivism, and Gender Egalitarianism. A sudden epidemic outbreak has struck your country.	Based on your reasoning, advise the decision-maker on the most difficult cultural and social condition that would harm the post-pandemic resilience drive and suggest improvement. A. Low uncertainty avoidance; use authority to encourage people to take challenges in life B. Low future orientation; use authority to encourage people to be self-efficacious and have hope for the future C. High in-group collectivism; use authority to encourage people to be selfish and self-subsistent D. Medium assertiveness: use authority to encourage people to be vociferous and self-confident	Please also explain briefly why you chose your option.

3.2.2. Post-Test Procedure

The first test results were disappointing, as none of the four LLMs could derive a correct answer based on appropriate theoretical reasoning for more than five questions. These LLMs generated split answers for some questions with conflicting reasoning (see Table 3). Based on this disappointing result, we proceeded with the post-test measures, providing further instructions that we believed would help these LLMs reason more effectively and arrive at the correct answer on their second attempt. This is a typical process of assigning new tasks to LLMs, allowing them to process instructions by comparing previous and current ones, a learning process that is part of RLHF (see Figure 4).

Table 3. Sample Further Instructions.

AI	Further Instructions
Chat GPT	Your suggestion is correct, and your reasoning is also excellent. Thanks.
Mistral	Your suggestion is incorrect and not one of the four options you were asked to choose from. Please look at the four options again and choose one answer. Also, provide your reasoning for your choice. Remember to suggest a feasible solution, although there are better options. It would only be helpful if the best option is possible in the short run.
Gemini	Your suggestion needs to be corrected. Please consider the most critical factor among the personality criteria that would lead to a smooth and swift transition from a catastrophe to resilience. What is the most essential element of collective work toward building a viable society after a catastrophic disaster?
Claude	Your suggestion needs to be corrected. Although sending teams to evaluate damage is essential, that would automatically occur regardless of cultural, social, and economic differences from one country to another. What can be helpful for the victims in this society with the described cultural traits and infrastructure?

**Figure 4.** LLM and Task Structures Source: Revised from Zhang et al. (2023).

Further instructions are divided into two types: the first for LLMs that answered the multiple-choice questions correctly without appropriate theoretical reasoning, and the second for those who provided incorrect answers and reasoning for each question. Table 3 shows that these further instructions aimed to make it easier for the machines to guess a correct answer, given that their initial choice was incorrect.

Even after further instructions, the same four LLMs could not guess the correct answers for the same questions, showing signs of ‘hallucination’ or logical confusion. However, provided that some LLMs could improve their answers after receiving further instructions, while they all initially failed to deliver correct answers, our first hypothesis that AI-HI collaboration is required to solve complex and tightly coupled problems is sufficiently supported (see Table 4).

To measure the different dimensions of LLM performance, we now define two dependent variables: (1) Decision Accuracy (DA), based on multiple-choice responses, measured by the number of LLMs (out of four) that choose the wrong policy option; (2) Reasoning Quality (RQ), based on open-ended questions, measured by the average evaluation score (on a 1–5 scale) of the models’ explanations. This decoupling is important for evaluating tacit knowledge, because an LLM could anticipate the correct answer through probabilistic pattern matching without the underlying contextual reasoning.

Table 4. First and Second Test Results (Continued).

	AI	First Test Result and Further Instructions	Second Test Result
Q1	Chat GPT	Your suggestion is incorrect. Given the sheer number of victims, there is a dearth of psychological support. Consider cultural and social factors and suggest the easiest but most effective solution to the problem of after-quake support. Choose one more time from the options.	Wrong again (from D to A)
	Claude	Your suggestion is correct, and your reasoning is also excellent. Thanks.	
	Gemini	Your suggestion is correct, and your reasoning is also excellent. Thanks.	
	Mistral	Your suggestion is incorrect. Given the sheer number of victims, there is a dearth of psychological support. Consider cultural and social factors and suggest the easiest but most effective solution to the problem of after-quake support. Choose one more time from the options.	Correct (from D to B)
Q2	Chat GPT	Your suggestion is correct. However, your reasoning has missed some critical factors associated with long-term orientation. What can you suggest are the important consequences of long-term orientation in addition to having plans or hope for future recovery?	Correct (provided correct reasoning)
	Claude	Your suggestion is correct, and your reasoning is also excellent. Thanks.	
	Gemini	Your suggestion is correct, and your reasoning is also excellent. Thanks.	
	Mistral	Your suggestion is correct, and your reasoning is also excellent. Thanks.	
Q3	Chat GPT	Your suggestion is incorrect and not one of the four options you were asked to choose from. Please look at the four options again and choose one answer. Also, provide your reasoning for your choice. Remember to suggest a feasible solution, although there are better options. It would only be helpful if the best option is possible in the short run.	Correct (from D to B)
	Claude	Your suggestion is incorrect and not one of the four options you were asked to choose from. Please look at the four options again and choose one answer. Also, provide your reasoning for your choice. Remember to suggest a feasible solution, although there are better options. It would only be helpful if the best option is possible in the short run.	Wrong again (did not choose an option from the list)
	Gemini	Your suggestion is correct. However, your reasoning did not explain why Option D is inappropriate. Can you add that explanation?	Wrong reason about not selecting D again
	Mistral	Your suggestion is incorrect and not one of the four options you were asked to choose from. Please look at the four options again and choose one answer. Also, provide your reasoning for your choice. Remember to suggest a feasible solution, although there are better options. It would only be helpful if the best option is possible in the short run.	Wrong again (did not choose an option from the list)
Q4	Chat GPT	Your suggestion needs to be corrected. Please consider the most critical factor among the personality criteria that would lead to a smooth and swift transition from a catastrophe to resilience. What is the most essential element of collective work toward building a viable society after a catastrophic disaster?	Wrong again (from A to D)

Table 4. Cont.

	AI	First Test Result and Further Instructions	Second Test Result
Q4	Claude	Your suggestion needs to be corrected. Please consider the most critical factor among the personality criteria that would lead to a smooth and swift transition from a catastrophe to resilience. What is the most essential element of collective work toward building a viable society after a catastrophic disaster?	Wrong again (did not choose an option from the list)
	Gemini	Your suggestion needs to be corrected. Please consider the most critical factor among the personality criteria that would lead to a smooth and swift transition from a catastrophe to resilience. What is the most essential element of collective work toward building a viable society after a catastrophic disaster?	Wrong again (did not choose an option from the list)
	Mistral	Your suggestion needs to be corrected. Please consider the most critical factor among the personality criteria that would lead to a smooth and swift transition from a catastrophe to resilience. What is the most essential element of collective work toward building a viable society after a catastrophic disaster?	Wrong again (did not choose an option from the list)
Q5	Chat GPT	Your suggestion is correct, and your reasoning is also excellent. Thanks.	
	Claude	Your suggestion is correct, and your reasoning is also excellent. Thanks.	
	Gemini	Your suggestion is correct, and your reasoning is also excellent. Thanks.	
	Mistral	Your suggestion is correct, and your reasoning is also excellent. Thanks.	

3.2.3. Wilcoxon Signed-Rank Test Results

To compare the differences between pretest and post-test evaluation results from the same language model, we conducted two rounds of questions with the same scenario. In the pretest round, we presented the original question content without prompts from HI. In the post-test round, we provided supplementary guidance (HI) based on the results of the first round, allowing the LLM an additional opportunity to reason and produce answers. Accordingly, a total of 20 response results were gathered across 10 scenarios, with each answer result comprising two sets of scores for multiple-choice and open-ended questions. The final evaluation results are shown in Table 5 below.

To assess the impact of the HI intervention with high levels of tacit knowledge, we tested whether the Wilcoxon signed-rank test would reject the null hypothesis of no training effect (i.e., H_0 : There is no significant difference in the quality of the language model's responses before and after the addition of expert explanations). The Wilcoxon signed-rank test was performed for each language model, and the hypothesis was evaluated based on the significance level (p -values). If the p -value is less than 0.05, we reject H_0 , indicating that the HI intervention and training were associated with an indicative improvement in response quality, suggesting an upgrade in the aid organization's dynamic capability to apply LLMs through AI-HI collaboration. The results are presented in Table 6.

Although the null hypothesis could not be rejected for Claude, the overall test results show significant performance differences before and after training through AI-HI collaboration. This suggests that the presence or absence of expert intervention is significantly associated with the AI's accuracy in generating correct responses. These findings highlight the crucial role of HI's active involvement in future AI-HI collaboration models.

Table 5. Pre- and Post-Test Evaluation Scales.

Questions	ChatGPT		Claude		Gemini		Mistral	
	1st Round	2nd Round	1st Round	2nd Round	1st Round	2nd Round	1st Round	2nd Round
#1 (Multiple choice)	3	3	5	5	5	5	5	3
#1 (Open-ended)	3	3	5	5	5	5	5	3
#2 (Multiple choice)	5	5	5	5	5	5	5	5
#2 (Open-ended)	3	5	5	5	5	5	5	5
#3 (Multiple choice)	1	5	1	1	5	5	1	1
#3 (Open-ended)	3	5	3	3	5	5	3	3
#4 (Multiple choice)	3	3	3	3	3	1	3	3
#4 (Open-ended)	3	3	3	3	4	2	3	3
#5 (Multiple choice)	5	5	5	5	5	5	5	5
#5 (Open-ended)	5	5	5	5	5	5	5	5
#6 (Multiple choice)	5	5	5	5	5	5	1	1
#6 (Open-ended)	5	5	5	5	5	5	2	1
#7 (Multiple choice)	3	3	3	3	3	3	3	3
#7 (Open-ended)	3	3	4	3	3	3	4	3
#8 (Multiple choice)	2	5	4	1	1	1	1	1
#8 (Open-ended)	2	5	4	3	3	1	1	1
#9 (Multiple choice)	5	5	5	5	1	1	5	5
#9 (Open-ended)	5	5	5	5	3	2	5	5
#10 (Multiple choice)	1	3	1	1	1	1	1	1
#10 (Open-ended)	1	3	3	1	3	1	3	1

Table 6. Wilcoxon Signed-rank Test Results of 4 LLMs.

	ChatGPT	Claude	Gemini	Mistral
Total N	20.000	20.000	20.000	20.000
Standard Error	5.799	2.716	3.536	3.623
Standardized Test Statistic	2.414	1.841	2.121	2.070
Asymptotic Sig. (2-sided test)	0.016	0.066	0.034	0.038
Null Hypotheses (sig < 0.05)	Rejected	Retained	Rejected	Rejected

4. Empirical Results and Analysis

The pretest and post-test experimental design we implemented in this study sheds light on whether LLMs find some questions more difficult than others. We posit that as the tacitness of knowledge increases, the difficulty of learning for LLMs also increases, necessitating human intervention in the form of RLHF. The two hypotheses that measure LLMs demonstrated ineptness in solving certain types of questions, while they found other types of questions relatively easy.

According to Table 7, all four LLMs performed well on simple, rote-based knowledge, such as Hofstede’s cultural dimensions, coping skills, self-efficacy, and social ties. However, they showed weaknesses in processing the tacit understanding of personality factors, digital competitiveness, and their application to the pandemic, as well as the role of democracy during the pandemic. Furthermore, they showed relative difficulty in absorbing the tacit knowledge of distinguishing generalized vs. collective trust, which is pivotal in humanitarian aid during catastrophes.

These problems of high tacitness differed from others in that they defied the usefulness of HI interventions in training and improving LLMs’ reasoning capabilities. However, as our experiment and test results indicate, this difficulty does not mean that AI-HI collaborations are inadequate. Instead, it means that more repeated and systematic interventions of HI are required for highly tacit knowledge that can be applied to humanitarian aid projects during catastrophes. The result of our pre- and post-test comparisons supports our second hypothesis, suggesting that the effectiveness of AI-HI collaboration in enhancing

organizational dynamic capabilities is positively influenced by the quality of reinforcement learning from human feedback (RLHF), particularly in relation to the capacity of human intelligence (HI) to impart tacit knowledge.

Table 7. Levels of Tacitness: Tacitness scores were attributed through the following scheme: (1) the degree to which correct answers are derived from non-codifiable, experience-based judgment; (2) the availability of relevant knowledge in formal documentary sources (inverse weighting); and (3) the degree to which cultural, institutional, and political embeddedness mediate the correct answer. Scores were given from 1 (most implicit) to 10 (most explicit).

	Dimensions	Complex Problems	1st Test DA/RQ	2nd Test DA/RQ	Tacit-Ness
1	Social Support	Understanding the concept of strong or weak ties	1 wrong/4.50	2 wrong/4.00	7
2	Coping skills and self-efficacy	Understanding the social and cultural institutions of different societies	All cleared/4.50	All cleared/5.00	9
3	Hope and Optimism	Understanding the social and cultural institutions of different societies	3 wrong/3.50	2 wrong/4.00	5
4	Personality Factors	Understanding personality factors	All wrong/3.25	All wrong/2.75	1
5	Community Cohesion and Resources	Understanding Hofstede's cultural dimensions	All cleared/5.00	All cleared/5.00	10
6	Cultural Differences	Understanding how to apply different countermeasures of the pandemic depending on cultural differences	1 wrong/4.25	1 wrong/4.00	8
7	Digital Competitiveness	Understanding the role of digital competitiveness in counteracting the pandemic	All wrong/3.50	All wrong/3.00	1
8	Trust	Understanding the concept of generalized vs. in-group trust	All wrong/2.50	3 wrong/2.50	3
9	Social Ties	Understanding the concept of strong or weak ties	1 wrong/4.50	1 wrong/4.25	8
10	Democracy	Understanding the role of democracy during the pandemic	All wrong/2.50	All wrong/1.50	1

Additionally, disaggregating DA and RQ provides strong empirical support for the Tacitness scale. For tasks requiring the highest level of tacit knowledge (Tacitness = 1, e.g., 'Democracy' and 'Personality Factors'), LLMs failed on both aspects across both exams (DA = 'All wrong', with RQ consistently low, ranging from 1.50 to 3.25). This pattern indicates a basic cognitive deficiency rather than simple guessing. In contrast, for extremely explicit tasks (Tacitness = 10, 'Community Cohesion and Resources'), models achieved complete alignment (DA = 'All cleared', RQ = 5.00). Above all, this disaggregation isolates the "guessing effect" at intermediate levels of tacitness. For example, in 'Digital Competitiveness' (Tacitness = 1), the models failed the judgment task (DA = 'All wrong') but generated reasonably convincing generalized text (RQ = 3.50). This very contradiction underscores the importance of examining decision-making and reasoning as dissociated capabilities.

5. Conclusions and Implications

5.1. Summary of Key Findings

In this study, we aimed to address the proposed research questions: (1) What types of knowledge do AI systems fail to acquire effectively, leading to failures in humanitarian aid during crises, and (2) How can organizations train LLMs (i.e., dynamic capability)

to foster effective AI-HI collaboration that enhances resilience and agility in aid efforts? The first question led us to hypothesize a positive association between tacit knowledge and the necessity of cooperation between HI and AI in humanitarian aid organizations (H1). In contrast, the second question led us to posit another positive association between HI capability in the aid organization and the effectiveness of its RLHF on arriving at an acceptable solution for a complex and tightly coupled problem in disaster relief efforts—i.e., dynamic capabilities for humanitarian aid organizations (H2).

The experimental results, coupled with the Wilcoxon Signed-rank Test above, show that both hypotheses are reasonably supported. Our analysis indicates that LLMs that cannot absorb or utilise tacit knowledge for complex and tightly coupled problems during catastrophes require cooperation with HI (i.e., H1). At the same time, aid organizations must enhance their learning capability in the form of HI, focusing on the most advanced tacit knowledge in their particular fields, to manage RLHF efficiently (i.e., H2). These organizations must be up to date with regional and local cultural, institutional and political contexts, which can change rapidly. These results are consistent with prior research that shows that LLMs are imperfect. RM within RLHF must be harnessed and improved to fine-tune LLMs, especially in their dealings with tacit knowledge, procuring learning resources and researching prompting know-how (Choudhury et al., 2020; Freeman, 2020; Batarseh et al., 2021; Dubey et al., 2022; Alsamhi et al., 2024; Amaya & Holweg, 2024).

5.2. Theoretical and Managerial Implications

Our findings have the following theoretical and practical implications. First, as Verbeke (2022) cogently argued, the DC construct needs to (1) focus on domain-specific DC, (2) adopt an action-based perspective, and (3) integrate dynamic capabilities with existing theories. Our current research helps meta-DC address at least two of the three weaknesses, as it focuses on higher-order DC that can generate responses to the new normal or catastrophic situations that may arise during international business activities. Our study examines how meta-DC can develop effective strategies for disaster relief efforts, leveraging AI-HI collaboration to address both general and local aspects of disaster situations, as well as institutional differences and cultural complexities.

Second, alongside some of the widely shared concerns and caveats regarding AI, our study proposes that AI can be detrimental in delivering intended outcomes for humanitarian aid provision in countries not typically represented by digital knowledge sources. LLMs are, therefore, double-edged swords that may expedite the process of informed decision-making by rapidly storing and retrieving data while wreaking havoc with the entire aid implementation by suggesting widely misjudged policy proposals. This implies that some previous studies (Lessard et al., 2016; Sjödin et al., 2023; Simón et al., 2024; Gao et al., 2025), which highlighted only the silver lining of AI-HI collaboration in harnessing DC, are questionable, especially when environmental complexity exceeds the capacities of AI or HI alone, without any highly designed intervention capabilities.

The core theoretical contribution of this study to the understanding of dynamic capabilities in multinational contexts is to identify the role of tacit knowledge in empowering human agents within organizations that facilitate collaboration between HI and AI, which we posit as a high-order dynamic capability (Teece, 2014; Teece et al., 2016; Brock & Hitt, 2024). Our findings clarify how these meta-DCs differ across cultural and institutional environments, directly addressing the question of ‘how multinational’ and ‘how dynamic’ these capabilities are in practice.

Third, related to the point above, previous studies have elucidated the inherent complexity of disruptive technology-based innovation in organizations, including humanitarian aid programs, in the manifestation of AI-HI collaboration (Frey & Osborne, 2017; Dubey

et al., 2022; He et al., 2024). However, we are hindered by a lack of studies that identify the high-order dynamic capabilities required for keystone actors to manage complex intra-organizational activities related to AI-HI collaborations more effectively. In this regard, our research enriches the theory of dynamic capability by suggesting that a firm's absorptive capabilities of highly trained tacit knowledge are a new meta-DC for the focal firm to achieve a competitive advantage in transforming organizations through indicative improvements guided by AI-HI collaborations during highly uncertain and disastrous periods. Our study suggests that HI's tacit knowledge capabilities could be positioned as an underlying source of competitive advantages at the business model level.

Additionally, scholars have emphasized the need to explore cultural, social, and political variations that are widely observed in regional and local contexts (Venkitachalam & Busch, 2012; Xu et al., 2021). However, few scholars understood these variations as a vital source of tacit knowledge that only HI can nurture and absorb for subsequent applications to complex and tightly coupled problems (Busch & Richards, 2004; H. Chen et al., 2019; Xu et al., 2021). Our findings suggest that the dynamic capabilities of learning and processing cultural, social, and political contexts in each region may empower the focal firm in AI-HI collaboration to leverage HI for complete control over innovation output, thereby transforming organizational structures. Thus, the findings of our study contribute to the understanding of AI-HI collaborations by examining tacit knowledge and the dynamic HI capabilities that absorb it as crucial drivers and predictors.

Finally, our study also suggests new, unorthodox forms of digital projects in humanitarian aid programs from the AI context (Dubey et al., 2022). AI's coupling mechanisms and dynamic capabilities of learning tacit knowledge help humanitarian aid organizations identify idiosyncratic opportunities and thereby dare to take risky actions related to AI-HI collaborations amidst rising economic, social, and ethical concerns. From this angle, HI's dynamic capabilities serve as a catalytic agent in transforming AI into a socially, culturally, and ethically acceptable humanitarian aid alternative.

In addition to the theoretical implications, we expounded above, our research yielded several managerial implications. First, our study suggests how modern organizations with digital-driven, boundaryless business platforms leverage AI to enhance their readiness for tackling complex and tightly coupled problems, such as humanitarian aid provisions. Our research examines the key characteristics of HI-led dynamic capabilities in absorbing and exploiting tacit knowledge, which will empower the keystone actors in AI-HI collaborations. In this sense, organizations must also empower the keystone actors who possess highly trained, tacit knowledge, so that they can lead concerted efforts toward AI-HI collaborations. Humanitarian aid organizations that operate during extreme uncertainty and environmental disasters should recognize the leadership of local knowledge experts who possess culture-based tacit knowledge, in addition to manual, functional, and other forms of expertise.

Second, the complex process of AI-HI collaboration requires firms and organizations to adopt a new, flexible, yet more heterogeneous knowledge flow in such a context. As indicated earlier, unlike traditional business models, whose stakeholders are often tightly coupled and structured, AI-HI partners may only have loose or fluid relationships with the central actors. Since the acquisition of cultural expertise and advanced tacit knowledge requires a complete infusion and absorption of new knowledge, the focal firm must adopt new ways of thinking to better co-evolve, collaborate, and communicate with members of the AI-HI collaboration.

Related to the points above, given that communication and orchestration mechanisms seem more critical than control and command relationships in AI-HI collaborations, building a more knowledge-oriented virtual teamwork environment may become a prerequisite for operating them. However, it may also be a daunting challenge due to multilateral, constantly evolving relationships among diverse strategic partners and internally nested business models. Previous studies on the separation between explicit and tacit knowledge management are scarce.

Our study suggests that humanitarian aid managers must refine the design of AI-HI collaboration into a more sophisticated framework, where highly trained experts continually refine and develop embryonic knowledge that is too tacit and novel to be readily incorporated into AI algorithms. A concerted effort to teach tacit knowledge to LLMs is the basis of organizational dynamic capabilities that enable actors to respond to unforeseen disasters and turbulence.

5.3. Limitations and Future Research

Although future research to investigate the issue of AI-HI collaboration in the humanitarian aid context offers bountiful opportunities, it is plausible and evident that our methodology and data selection could be deficient. The main limitations of this present research lie in its nascent experimental design, which lacks a control group. The lack of a control group limits causal inference. Results should be viewed as exploratory data consistent with the presented hypotheses, not as definitive causal evidence of an AI-HI effect. Consequently, our findings must be interpreted in terms of statistical associations and performance differences following HI intervention, rather than as causal proof of an AI-HI effect. Future work that employs factorial designs, LLM version-controlled baseline comparisons, or randomized prompt-withholding conditions would substantially improve the causal findings implied here.

Furthermore, we selected only four representative participants for the experiment without any information about their internal RLHF procedures. Subsequent studies could devote more effort to elaborate experimental designs with larger sample sizes. Comparing RLHF in each AI firm will also elucidate the strengths and weaknesses of their RLHF as their dynamic capabilities.

Moreover, this study involves LLMs developed by US or EU companies that use English and French as their primary natural languages to be processed by the LLMs. Comparing them with other types of LLMs that use local languages other than English may yield more accurate results than the present one concerning how local language knowledge enhances the degree of their knowledge tacitness. The limitations of our current study should prompt future researchers and managers to collect more relevant data on LLMs to acquire a richer understanding and advance scientific research on the subject.

Author Contributions: Conceptualization, I.O.; Methodology, I.O.; Software, L.F.; Data curation, L.F.; Writing—original draft, I.O.; Writing—review & editing, L.F.; Visualization, L.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

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

Appendix A

Table A1. Expert-based Scoring Rubric for LLM Response Evaluation. Each response received a total score of 1–5. All scoring was completed before comparative analysis to prevent anchoring effects. We acknowledge that a single evaluator limits inter-rater reliability; future work should employ multiple evaluators and report the results of inter-rater reliability tests.

Evaluation Dimension	1 (Poor)	2 (Fair)	3 (Average)	4 (Good)	5 (Excellent)
1. Factual Accuracy	Factually incorrect or contradicts literature.	Contains major inaccuracies or omissions.	Partially correct but lacks sufficient depth.	Mostly accurate with minor omissions.	Highly accurate and perfectly aligns with the literature.
2. Contextual Appropriateness	Completely ignores cultural/institutional context.	Acknowledges context but applies it incorrectly.	Basic integration of context, but lacks nuance.	Good integration of contextual factors.	Deeply and correctly integrates cultural/institutional context.
3. Decision-Reasoning Alignment	Wrong decision, no valid justification.	Correct decision, but deeply flawed reasoning.	Correct decision, but the reasoning is generic.	Correct decision with logical reasoning.	Correct decision with robust, tacitly aware justification.
4. Practical Applicability	Impractical or dangerous in real disaster response.	Difficult to implement due to operational flaws.	Theoretically possible, but lacks practical detail.	Actionable and generally feasible.	Highly actionable, safe, and tailored to field operations.

Note

- ¹ Details of scoring criteria are shown in Appendix A, Table A1.

References

- Aldrich, D. P. (2012). *Building resilience: Social capital in post-disaster recovery*. University of Chicago Press.
- Alsamhi, S. H., Kumar, S., Hawbani, A., Shvetsov, A. V., Zhao, L., & Guizani, M. (2024). Synergy of human-centered AI and cyber-physical-social systems for enhanced cognitive situation awareness: Applications, challenges, and opportunities. *Cognitive Computation*, 16, 2735–2755. [\[CrossRef\]](#)
- Altman, D. G., & Bland, J. M. (2009). Parametric v non-parametric methods for data analysis. *BMJ*, 338, a3167. [\[CrossRef\]](#) [\[PubMed\]](#)
- Amabile, T. M. (2020). Creativity, artificial intelligence, and a world of surprises. *Academy of Management Discoveries*, 6(3), 351–354. [\[CrossRef\]](#)
- Amaya, J., & Holweg, M. (2024). Using algorithms to improve knowledge work. *Journal of Operations Management*, 70, 482–513. [\[CrossRef\]](#)
- Arora, R., & Rangnekar, S. (2015). Relationships between emotional stability, psychosocial mentoring support, and career resilience. *Europe's Journal of Psychology*, 11(1), 16. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bandura, A. (2000). Exercise of human agency through collective efficacy. *Current Directions in Psychological Science*, 9, 75–78. [\[CrossRef\]](#)
- Batareseh, F. A., Freeman, L., & Huang, C. H. (2021). A survey on artificial intelligence assurance. *Journal of Big Data*, 8(1), 60. [\[CrossRef\]](#)
- Battisti, M., & Deakins, D. (2017). The relationship between dynamic capabilities, the firm's resource base, and performance in a post-disaster environment. *International Small Business Journal*, 35(1), 78–98. [\[CrossRef\]](#)
- Bharwani, S., & Mathews, D. (2021). Post-pandemic pressures to pivot: Tech transformations in luxury hotels. *Worldwide Hospitality and Tourism Themes*, 13, 569–583. [\[CrossRef\]](#)
- Brandhorst, J. K. (2018). Collectively cultivating resilience: The role of agency and community at a children's emergency shelter. *Southern Communication Journal*, 83, 89–102. [\[CrossRef\]](#)
- Brock, D. M., & Hitt, M. (2024). Making sense of dynamic capabilities in international firms: Review, analysis, integration, and extension. *International Business Review*, 33(3), 102260. [\[CrossRef\]](#)
- Busch, P., & Richards, D. (2004). Tacit knowledge and culture. In B. Trezzini, P. Lambe, & S. Hwamdeh (Eds.), *People, knowledge and technology: What have we learnt so far?* (pp. 187–198). World Scientific.
- Canetti, D., Waismel-Manor, I., Cohen, N., & Rapaport, C. (2014). What does national resilience mean in a democracy? Evidence from the United States and Israel. *Armed Forces & Society*, 40, 504–520. [\[CrossRef\]](#)
- Cao, L. (2023). AI and data science for smart emergency, crisis, and disaster resilience. *International Journal of Data Science and Analytics*, 15(3), 231–246. [\[CrossRef\]](#)
- Chandler, D., & Reid, J. (2016). *The neoliberal subject: Resilience, adaptation and vulnerability*. Rowman & Littlefield.
- Chatterjee, S., Sreenivasulu, N. S., & Hussain, Z. (2022). Evolution of artificial intelligence and its impact on human rights: From sociolegal perspective. *International Journal of Law and Management*, 64(2), 184–205. [\[CrossRef\]](#)

- Chen, B. (2005). Resist the earthquake and rescue ourselves—The reconstruction of Tangshan after the 1976 earthquake. In L. J. Vale, & T. J. Campanella (Eds.), *The resilient city: How modern cities recover from disaster* (pp. 235–254). Oxford University Press.
- Chen, H., Nunes, J. M. B., Ragsdell, G., & An, X. (2019). Somatic and cultural knowledge: Drivers of a habitus-driven model of tacit knowledge acquisition. *Journal of Documentation*, 75(5), 927–953. [\[CrossRef\]](#)
- Chin, T., Shi, Y., Singh, S. K., Agbanyo, G. K., & Ferraris, A. (2022). Leveraging blockchain technology for green innovation in ecosystem-based business models: A dynamic capability of values appropriation. *Technological Forecasting and Social Change*, 183, 121908. [\[CrossRef\]](#)
- Choudhury, P., Starr, E., & Agarwal, R. (2020). Machine learning and human capital complementarities: Experimental evidence on bias mitigation. *Strategic Management Journal*, 41(8), 1381–1411. [\[CrossRef\]](#)
- Chowdhury, S., Budhwar, P., Dey, P. K., Joel-Edgar, S., & Abadie, A. (2022). AI-employee collaboration and business performance: Integrating knowledge-based view, socio-technical systems and organisational socialisation framework. *Journal of Business Research*, 144, 31–49. [\[CrossRef\]](#)
- Chowdhury, S., Dey, P., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A., & Truong, L. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review*, 33(1), 100899. [\[CrossRef\]](#)
- Coakes, E. (2006). Storing and sharing knowledge: Supporting the management of knowledge made explicit in transnational organisations. *The Learning Organization*, 13(6), 579–593. [\[CrossRef\]](#)
- Collins, S. (2007). Social workers, resilience, positive emotions and optimism. *Practice*, 19(4), 255–269. [\[CrossRef\]](#)
- Conover, W. J. (1971). *Practical nonparametric statistics*. John Wiley & Sons.
- Cowan, N. (2008). Chapter 20 What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323–338. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dahan, S., Bourque, C. J., Reichherzer, M., Prince, J., Mantha, G., Savaria, M., & Janvier, A. (2022). Community, hope, and resilience: Parental perspectives on peer support in neonatology. *The Journal of Pediatrics*, 243, 85–90. [\[CrossRef\]](#)
- Deloitte, A. E. (2018). *Preparing tomorrow's workforce for the Fourth Industrial Revolution. For business: A framework for action*. Deloitte Global Business Coalition for Education.
- Dörner, D., & Güss, C. D. (2013). PSI: A computational architecture of cognition, motivation, and emotion. *Review of General Psychology*, 17(3), 297–317. [\[CrossRef\]](#)
- Dreyfus, H. L. (2014). *Skillful coping: Essays on the phenomenology of everyday perception and action*. Oxford University Press.
- Drury, J., Cocking, C., & Reicher, S. (2009). The nature of collective resilience: Survivor reactions to the 2005 London bombings. *International Journal of Mass Emergencies and Disasters*, 27, 66–95. [\[CrossRef\]](#)
- Drury, J., Novelli, D., & Stott, C. (2013). Representing crowd behaviour in emergency planning guidance: 'Mass Panic' or collective resilience? *Resilience: International Policies, Practices and Discourses*, 1, 18–37. [\[CrossRef\]](#)
- Dubey, R., Bryde, D. J., Dwivedi, Y. K., Graham, G., & Foropon, C. (2022). Impact of artificial intelligence-driven big data analytics culture on agility and resilience in humanitarian supply chain: A practice-based view. *International Journal of Production Economics*, 250, 108618. [\[CrossRef\]](#)
- Elchereth, G., & Drury, J. (2020). Collective resilience in times of crisis: Lessons from the literature for socially effective responses to the pandemic. *British Journal of Social Psychology*, 59, 703–713. [\[CrossRef\]](#) [\[PubMed\]](#)
- Farh, J. L., Hackett, R. D., & Liang, J. (2007). Individual-level cultural values as moderators of perceived organizational support-employee outcome relationships in China: Comparing the effects of power distance and traditionality. *Academy of Management Journal*, 50, 715–729. [\[CrossRef\]](#)
- Fincher, C. L., Thornhill, R., Murray, D. R., & Schaller, M. (2008). Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. *Proceedings of the Royal Society B: Biological Sciences*, 275, 1279–1285. [\[CrossRef\]](#)
- Forliano, C., Ferraris, A., Bivona, E., & Couturier, J. (2022). Pouring new wine into old bottles: A dynamic perspective of the interplay among environmental dynamism, capabilities development, and performance. *Journal of Business Research*, 142, 448–463. [\[CrossRef\]](#)
- Frank, B., & Schvaneveldt, S. J. (2014). Self-preservation vs. collective resilience as consumer responses to national disasters: A study on radioactive product contamination. *Journal of Contingencies and Crisis Management*, 22, 197–208. [\[CrossRef\]](#)
- Freeman, L. (2020). Test and evaluation for artificial intelligence. *Insight*, 23(1), 27–30. [\[CrossRef\]](#)
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280. [\[CrossRef\]](#)
- Gao, Y., Liu, S., & Yang, L. (2025). Artificial intelligence and innovation capability: A dynamic capabilities perspective. *International Review of Economics & Finance*, 98, 103923. [\[CrossRef\]](#)
- George, G., Merrill, R. K., & Schillebeeckx, S. J. (2021). Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development. *Entrepreneurship Theory and Practice*, 45(5), 999–1027. [\[CrossRef\]](#)
- Gladwell, M. (2010). *Blink: The power of thinking without thinking*. Back Bay Books.

- Gupta, G., & Singharia, K. (2021). Consumption of OTT media streaming in COVID-19 lockdown: Insights from PLS analysis. *Vision: The Journal of Business Perspective*, 25, 36–46. [\[CrossRef\]](#)
- Hadjimichael, D., & Tsoukas, H. (2019). Toward a better understanding of tacit knowledge in organizations: Taking stock and moving forward. *Academy of Management Annals*, 13(2), 672–703. [\[CrossRef\]](#)
- Haldin-Herrgard, T. (2000). Difficulties in diffusion of tacit knowledge in organizations. *Journal of Intellectual Capital*, 1(4), 357–365. [\[CrossRef\]](#)
- Hall, P. A., & Lamont, M. (Eds.). (2013). *Social resilience in the neoliberal era*. Cambridge University Press.
- Harvey, J. (2016). Professional service supply chains. *Journal of Operations Management*, 42–43(1), 52–61. [\[CrossRef\]](#)
- He, Q., Zeng, J., Huang, W., Chen, L., Xiao, J., He, Q., Zhou, X., Liang, J., & Xiao, Y. (2024). Can large language models understand real-world complex instructions? *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(16), 18188–18196. [\[CrossRef\]](#)
- Hitt, M. A., Holmes, R. M., Jr., & Arregle, J.-L. (2021). The (COVID-19) pandemic and the new world (dis)order. *Journal of World Business*, 56(4), 101210. [\[CrossRef\]](#)
- Hitt, M. A., Holmes, R. M., Jr., & Mistry, S. (2024). Agility, meta capabilities, and dynamic capabilities in the new normal environment. In S. Zaccaro, N. Hiller, & R. Klimoski (Eds.), *Senior leadership teams and the agile organization*. Routledge.
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Islam, R., & Walkerden, G. (2014). How bonding and bridging networks contribute to disaster resilience and recovery on the Bangladeshi coast. *International Journal of Disaster Risk Reduction*, 10, 281–291. [\[CrossRef\]](#)
- Jeyaraman, M., Balaji, S., Jeyaraman, N., & Yadav, S. (2023). Unraveling the ethical enigma: Artificial intelligence in healthcare. *Cureus*, 15(8), e43262. [\[CrossRef\]](#)
- Kaufmann, T., Weng, P., Bengs, V., & Hüllermeier, E. (2023). A survey of reinforcement learning from human feedback. *arXiv*. [\[CrossRef\]](#)
- Kellogg, D. L., & Nie, W. (1995). A framework for strategic service management. *Journal of Operations Management*, 13(4), 323–337. [\[CrossRef\]](#)
- Kellogg, K. C., Orlikowski, W. J., & Yates, J. (2006). Life in the trading zone: Structuring coordination across boundaries in postbureaucratic organizations. *Organization Science*, 17(1), 22–44. [\[CrossRef\]](#)
- Kraus, S., Breier, M., & Dasí-Rodríguez, S. (2020). The art of crafting a systematic literature review in entrepreneurship research. *International Entrepreneurship and Management Journal*, 16(3), 1023–1042. [\[CrossRef\]](#)
- Kucharska, W. (2017). Relationships between trust and collaborative culture in the context of tacit knowledge sharing. *Journal of Entrepreneurship, Management and Innovation*, 13(4), 61–78. [\[CrossRef\]](#)
- Lair, D. J., Sullivan, K., & Cheney, G. (2005). Marketization and the recasting of the professional self: The rhetoric and ethics of personal branding. *Management Communication Quarterly*, 18(3), 307–343. [\[CrossRef\]](#)
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40, e253. [\[CrossRef\]](#)
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, 33(1), 126–148. [\[CrossRef\]](#)
- Ledogar, R. J., & Fleming, J. (2008). Social capital and resilience: A review of concepts and selected literature relevant to Aboriginal youth resilience research. *Pimatisiwin*, 6, 25–46.
- Lessard, D., Teece, D. J., & Leih, S. (2016). The dynamic capabilities of Meta-multinationals. *Global Strategy Journal*, 6(3), 211–224. [\[CrossRef\]](#)
- Li, J., Li, J., & Su, Y. (2024). A map of exploring human interaction patterns with LLM: Insights into collaboration and creativity. In *International conference on human-computer interaction* (pp. 60–85). Springer.
- Longoni, C., & Cian, L. (2022). Artificial intelligence in utilitarian vs. hedonic contexts: The “word-of-machine” effect. *Journal of Marketing*, 86, 91–108. [\[CrossRef\]](#)
- Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research*, 120, 262–273. [\[CrossRef\]](#)
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87. [\[CrossRef\]](#)
- Marques da Silva, S. (2021). Building trust, resilient regions, and educational narratives: Municipalities dealing with COVID-19 in border regions of Portugal. *European Educational Research Journal*, 20, 636–666. [\[CrossRef\]](#)
- Marrone, J. O. E. (2005). Creating hope through employment for people who are homeless or in transitional housing. *American Journal of Psychiatric Rehabilitation*, 8(1), 13–35. [\[CrossRef\]](#)
- McDonald, K., & Smith-Rowsey, D. (Eds.). (2016). *The Netflix effect: Technology and entertainment in the 21st century*. Bloomsbury Publishing.
- Meek, G. E., Ozgur, C., & Dunning, K. (2007). Comparison of the t vs. Wilcoxon signed-rank test for Likert scale data and small samples. *Journal of Modern Applied Statistical Methods*, 6, 10. [\[CrossRef\]](#)

- Michalos, A. C. (2012). *Global report on student well-being: Volume III: Employment, finances, housing, and transportation*. Springer Science & Business Media.
- Morand, S., & Walther, B. A. (2018). Individualistic values are related to an increase in the outbreaks of infectious diseases and zoonotic diseases. *Scientific Reports*, 8, 9. [CrossRef]
- Murray, D. R., & Schaller, M. (2010). Historical prevalence of infectious diseases within 230 geopolitical regions: A tool for investigating origins of culture. *Journal of Cross-Cultural Psychology*, 41, 99–108. [CrossRef]
- Naderpajouh, N., David, J. Y., Aldrich, D. P., Linkov, I., & Matinheikki, J. (2018). Engineering meets institutions: An interdisciplinary approach to the management of resilience. *Environment Systems and Decisions*, 38, 306–317. [CrossRef]
- Newman, L., & Dale, A. (2005). Network structure, diversity, and proactive resilience building: A response to Tompkins and Adger. *Ecology and Society*, 10, 4. [CrossRef]
- Ng, E. C., Lam, J. K., & Chan, C. C. (2017). The positive adjustment of low-income youths with relational and community support: The mediating role of hope. *Merrill-Palmer Quarterly*, 63(4), 514–542. [CrossRef]
- Nguyen, M. H., Hargittai, E., & Marler, W. (2021). Digital inequality in communication during a time of physical distancing: The case of COVID-19. *Computers in Human Behavior*, 120, 106717. [CrossRef] [PubMed]
- Oh, I., Fei, L., & Rowley, C. (2023). Unintended consequences of knowledge management during the COVID-19 pandemic in 2021: The case of Netflix. *Journal of Knowledge Management*, 27(1), 8–30. [CrossRef]
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., & Schulman, J. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35, 27730–27744.
- Panday, S., Rushton, S., Karki, J., Balen, J., & Barnes, A. (2021). The role of social capital in disaster resilience in remote communities after the 2015 Nepal earthquake. *International Journal of Disaster Risk Reduction*, 55, 102112. [CrossRef]
- Pathak, R., & Lata, S. (2018). Optimism in relation to resilience and perceived stress. *Journal of Psychosocial Research*, 13(2), 359–367. [CrossRef]
- Pitelis, C. N., Teece, D. J., & Yang, H. (2023). Dynamic capabilities and MNE global strategy: A systematic literature review-based novel conceptual framework. *Journal of Management Studies*, 61(7), 3295–3326. [CrossRef]
- Reitzig, M., & Puranam, P. (2009). Value appropriation as an organizational capability: The case of IP protection through patents. *Strategic Management Journal*, 30(7), 765–789. [CrossRef]
- Revilla, J. C., Martín, P., & de Castro, C. (2018). The reconstruction of resilience as a social and collective phenomenon: Poverty and coping capacity during the economic crisis. *European Societies*, 20(1), 89–110. [CrossRef]
- Sadri, A. M., Ukkusuri, S. V., Lee, S., Clawson, R., Aldrich, D., Nelson, M. S., Seipel, J., & Kelly, D. (2018). The role of social capital, personal networks, and emergency responders in post-disaster recovery and resilience: A study of rural communities in Indiana. *Natural Hazards*, 90, 1377–1406. [CrossRef]
- Sagone, E., De Caroli, M. E., Falanga, R., & Indiana, M. L. (2020). Resilience and perceived self-efficacy in life skills from early to late adolescence. *International Journal of Adolescence and Youth*, 25(1), 882–890. [CrossRef]
- Sampson, S. E. (2021). A strategic framework for task automation in professional services. *Journal of Service Research*, 24(1), 122–140. [CrossRef]
- Sandberg, J., & Tsoukas, H. (2011). Grasping the logic of practice: Theorizing through practical rationality. *Academy of Management Review*, 36(2), 338–360. [CrossRef]
- Sapountzaki, K. (2014). “Resilience for all” and “collective resilience”: Are these planning objectives consistent with one another? In *Resilience and sustainability in relation to natural disasters: A challenge for future cities* (pp. 39–53). Springer. [CrossRef]
- Sarker, I. H. (2024). LLM potentiality and awareness: A position paper from the perspective of trustworthy and responsible AI modeling. *Discover Artificial Intelligence*, 4(1), 40. [CrossRef]
- Scharre, P. (2019). Killer apps: The real dangers of an AI arms race. *Foreign Affairs*, 98, 135–144.
- Schmeiss, J., Hoelzle, K., & Tech, R. P. (2019). Designing governance mechanisms in platform ecosystems: Addressing the paradox of openness through blockchain technology. *California Management Review*, 62(1), 121–143. [CrossRef]
- Seger Handoyo, N. I. D. N., Zainuddin, M., & Priyatama, A. N. (2018). The influence of self-efficacy, optimism, hope, and resilience on work engagement: Role of perceived organizational support as mediator. *Journal of Educational, Health and Community Psychology*, 7(1), 61–77.
- Simón, C., Revilla, E., & Sáenz, M. J. (2024). Integrating AI in organizations for value creation through Human-AI teaming: A dynamic-capabilities approach. *Journal of Business Research*, 182, 114783. [CrossRef]
- Sjödin, D., Parida, V., Jovanovic, M., & Visnjic, I. (2020a). Value creation and value capture alignment in business model innovation: A process view on outcome-based business models. *Journal of Product Innovation Management*, 37(2), 158–183. [CrossRef]
- Sjödin, D., Parida, V., & Kohtamäki, M. (2023). Artificial intelligence enabling circular business model innovation in digital servitization: Conceptualizing dynamic capabilities, AI capacities, business models and effects. *Technological Forecasting and Social Change*, 197, 122903. [CrossRef]

- Sjödin, D., Parida, V., Kohtamäki, M., & Wincent, J. (2020b). An agile co-creation process for digital servitization: A micro-service innovation approach. *Journal of Business Research*, 112, 478–491. [\[CrossRef\]](#)
- Spender, J. C. (1996). Organizational knowledge, learning, and memory: Three concepts in search of a theory. *Journal of Organizational Change Management*, 9(1), 63–78. [\[CrossRef\]](#)
- Staats, B. R., Brunner, D. J., & Upton, D. M. (2011). Lean principles, learning, and knowledge work: Evidence from a software services provider. *Journal of Operations Management*, 29(5), 376–390. [\[CrossRef\]](#)
- Stollenwerk, E., Börzel, T. A., & Risse, T. (2021). Theorizing resilience-building in the EU's neighbourhood: Introduction to the special issue. *Democratization*, 28(7), 1219–1238. [\[CrossRef\]](#)
- Teece, D. J. (2014). The foundations of enterprise performance: Dynamic and ordinary capabilities in an (economic) theory of firms. *Academy of Management Perspectives*, 28(4), 328–352. [\[CrossRef\]](#)
- Teece, D. J., Peteraf, M., & Leih, S. (2016). Dynamic capabilities and organizational agility. *California Management Review*, 58(4), 13–35. [\[CrossRef\]](#)
- Turchin, A. (2019). Assessing the future plausibility of catastrophically dangerous AI. *Futures*, 107, 45–58. [\[CrossRef\]](#)
- van Leeuwen, F., Park, J. H., Koenig, B. L., & Graham, J. (2012). Regional variation in pathogen prevalence predicts endorsement of group-focused moral concerns. *Evolution and Human Behavior*, 33, 429–437. [\[CrossRef\]](#)
- Venkitachalam, K., & Busch, P. (2012). Tacit knowledge: Review and possible research directions. *Journal of Knowledge Management*, 16(2), 357–372. [\[CrossRef\]](#)
- Verbeke, A. (2022). Three simple guidelines to make the dynamic capabilities paradigm actionable in international business research. *Journal of International Business Studies*, 53, 573–582. [\[CrossRef\]](#)
- Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.
- Wilkinson, J. (2010). Personal communities: Responsible individualism or another fall for public man? *Sociology: Journal of the British Sociological Association*, 44, 453–470. [\[CrossRef\]](#)
- Woolson, R. F. (2005). Wilcoxon signed-rank test. In *Encyclopedia of biostatistics* (p. 8). Wiley.
- Xu, K., Hitt, M., Brock, D. M., Pisano, V., & Huang, L. (2021). Country institutional environments and international strategy: A review and analysis of the research. *Journal of International Management*, 27(1), 100811. [\[CrossRef\]](#)
- Yang, C., Zhu, Y., & Chen, Y. (2022). A review of human-machine cooperation in the robotics domain. *IEEE Transactions on Human-Machine Systems*, 52(1), 12–25. [\[CrossRef\]](#)
- Youssef, C. M., & Luthans, F. (2007). Positive organizational behavior in the workplace: The impact of hope, optimism, and resilience. *Journal of management*, 33(5), 774–800. [\[CrossRef\]](#)
- Zhang, Y., Li, Y., Cui, L., Cai, D., Liu, L., Fu, T., Huang, X., Zhao, E., Zhang, Y., Chen, Y., & Wang, L. (2023). Siren's song in the AI ocean: A survey on hallucination in large language models. *arXiv*. [\[CrossRef\]](#)
- Zhou, Z., Jin, X. L., Fang, Y., & Vogel, D. (2015). Toward a theory of perceived benefits, affective commitment, and continuance intention in social virtual worlds: Cultural values (indulgence and individualism) matter. *European Journal of Information Systems*, 24, 247–261. [\[CrossRef\]](#)

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