The Leadership Quarterly xxx (xxxx) xxx

Contents lists available at ScienceDirect

The Leadership Quarterly

journal homepage: www.elsevier.com/locate/leaqua

Beyond efficiency: How artificial intelligence (AI) will reshape scientific inquiry and the publication process

Niels Van Quaquebeke [a,b,*], Scott Tonidandel [c], George C. Banks [c]

[a] Kühne Logistics University, Germany
[b] University of Exeter Business School, UK
[c] Belk College of Business, UNC Charlotte, USA

ARTICLE INFO

Keywords:
Artificial intelligence
AI
Authorship
Philosophy of science
Scholarship
Publishing

ABSTRACT

Artificial intelligence is rapidly transforming how research is conceived, executed, published, and shared. This editorial examines the "elephant in the room": the integration of AI across every stage of the research life cycle and publication pipeline. We trace AI's expanding footprint on the author side, from sparking novel research ideas, mapping literature, and designing studies, to simulating data, analyzing results, and drafting manuscripts. We also consider AI's growing role on the journal side, including automated manuscript triage, AI-assisted peer review, decision synthesis, and revision checks. And we discuss AI's impact on research dissemination. Throughout, we highlight not only the unprecedented opportunities for creativity, efficiency, and accessibility, but also the ethical risks, such as epistemic homogenization, challenges to accountability, and the loss of scholarly craft. We urge researchers, editors, and institutions not to fall into the false binary of blind optimism or blanket skepticism. Instead, we call for deliberate engagement: a principled, transparent, and reflexive partnership between human scholars and machine collaborators. The question is no longer whether we want AI to shape the future of scholarship—it already is. The challenge now is to ensure that what it amplifies is not only our productivity, but our judgment, imagination, and collective responsibility in knowledge creation and dissemination.

Let us address the "elephant in the room:" the research process in organizational behavior—and across the social sciences or any science for that matter—is being quietly transformed by artificial intelligence (AI). While publishers try to stay on top with ever new guidelines and regulations regarding the use of AI in the research process (e.g., Elsevier, 2025), scholars are already brainstorming with large language models, reviewing literatures through AI-curated maps, and refining their prose with the help of generative writing tools. And once authors receive their work back with an editorial decision and attached reviews, they may wonder how much of that was (co)written by an AI (Singh Chawla, 2024). What was once a curiosity or gimmick is rapidly becoming explicitly or implicitly part of the scholarly workflow as also noted in various other editorials (e.g., Binz et al., 2025; Gatrell et al., 2024; Grimes et al., 2023; Kulkarni et al., 2024; Stollberger et al., 2025).

But with this transformation comes a deeper, more urgent question: what kind of scholarship are we producing—and what kind of scholars are we becoming? And just as importantly, what kinds of guardrails must we put in place to ensure that speed and scale do not come at the expense of reflection, originality, and accountability?

We approach this conversation not as skeptics or evangelists, but as researchers with a deep commitment to the values of rigorous and relevant scholarship—underpinned by our own curiosity and reflexivity. This is therefore not a manifesto. Nor is it a warning. It is an attempt to map a rapidly shifting terrain—and an invitation to navigate it together. As such, we hope this editorial will serve both as a catalyst for new use cases and a guide to the pitfalls that may accompany them. At the same time, we imagine it sparking robust debate in graduate seminars and research teams alike—prompting critical reflection on how best to engage with AI in the evolving landscape of social science.

In the following, we focus on the role of AI in the research and publication process itself—not as a subject of study (such as in, e.g., Bankins et al., 2024; Van Quaquebeke & Gerpott, 2023) but as a collaborator or catalyst in how we design, execute, evaluate, and communicate our research. On the author side, we follow the full arc of inquiry from idea generation, literature review, study design, data generation, analysis, to writing the actual manuscript. On the publishing

* Corresponding author.
E-mail address: niels.quaquebeke@klu.org (N. Van Quaquebeke).

https://doi.org/10.1016/j.leaqua.2025.101895

1048-9843/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

side, we examine emerging uses of AI in pre-submission evaluation, desk decisions, reviewing, and revisions. We also explore new frontiers in dissemination, where AI enables post-publication transformation, tracking, and interpretation of research in real time. The promise is immense: AI can amplify creativity, scale our understanding, and democratize analysis and writing quality. However, the risks are just as prominent: overreliance, ethical drift, loss of craft, and the creeping optimization and homogenization of ideas. At each stage, we therefore ask not just what AI can do, but what it should do. We weigh its benefits and risks. Against this analysis, we propose a set of principles for a research culture that is enhanced—but not eclipsed—by artificial intelligence (see Table 1 and 2).

Using AI in the research process

Idea generation/inception

Benefits of working with AI

The use of generative AI in sparking research ideas is no longer novel (Amabile, 2019; Von Krogh et al., 2023). Scholars across disciplines now routinely "spar" with large language models (LLMs) to generate research questions, explore underexamined themes, and find novel intersections between literatures. It has become a new method of invention (Bianchini et al., 2022). And not a bad one. Case in point, a study pitting the ideas of 100 expert natural language processing (NLP) researchers against those of an LLM ideation agent, finds that the LLM-generated ideas are judged as more novel than the human expert ideas (Si et al., 2024). Ask an AI for input or to challenge your ideas and it will offer its perspective. Ask it to present "three overlooked angles on abusive supervision," and within seconds, it delivers responses that would have taken hours and often long discussion amongst authors to assemble[1](Cornelissen et al., 2024). And afterwards one may double check with an AI to see whether the suggested project has not been done before (Skarlinski et al., 2025).

Interestingly, research suggests that these benefits are not limited to newcomers. While some evidence shows that AI boosts novelty most for novices or those who typically struggle with creative output (Doshi et al., 2025; Doshi & Hauser, 2024), other studies indicate that skilled experts may benefit even more from using AI as a strategic sparring partner (Jia et al., 2024). Either way, AI's ability to push thinking beyond conventional boundaries appears to hold promise across levels of expertise. Though a recent *meta*-analysis also argues that effects may depend on the used LLM, i.e. older models surprisingly seem to help human creativity more but newer models are potentially more creative themselves (Holzner et al., 2025).

Some scholars perhaps are already expanding the use of AI further, taking note of the influential "That's interesting" article (Davis, 1971) by tasking the AI to propose interesting shifts to an established field. As a result, it may then, for instance, delineate from extant research on transformational leadership a first inkling of a theory on "transformational companies". Or it may try to cross reference literatures in innovative ways, for example, when prompting it with "what are some overlooked aspects in the literature on abusive supervision when simultaneously considering the clinical psychological literature on domestic abuse".

At this point, we may also speculate that researchers are already beginning to leverage AI to scan and summarize entire archives such as recent *Leadership Quarterly* issues. To this end, as a starting point for their research journey, they may—illegally—upload all papers and then ask the AI to identify from the discussion sections where methodological rigor is lacking or which theoretical debates remain unresolved.

Now, imagine a future in which journals publicly share reviewer comments—not just for accepted papers, but for rejected ones as well (i.e., a complete open peer review system). A scenario not unlikely given that on June 16, 2025, the journal Nature has introduced a policy that all published papers will include a peer-review file that documents the reviewer comments and as well as author responses over the various revision rounds. When AI systems access this corpus, they could begin to simulate what counts as a "hot topic" in real time—analyzing which ideas consistently excite reviewers, where enthusiasm clusters, and how editorial preferences evolve. Such systems could offer dynamically updated guidance to authors about what the field is currently primed to value—and perhaps even make this intelligence public, serving as a living orientation map for those navigating the research landscape.

Additionally, we can envision AI-driven citation simulators, forecasting which themes are likely to be popular in the community. Simply take the list of articles accepted over the last five years and gather not only their citation trends but also corresponding altmetrics. Think of it as trading on Wall Street trying to identify big trends and weak signals. Relatedly, scholars may also try to use AI to scrape real-time signals from Reddit, YouTube, or industry reports to spot practice-relevant problems awaiting scientific scrutiny. But it does not always have to be about trying to predict future trends. As translation models improve or rather as LLMs are also fed with more obscure, non-English, or historical texts, researchers can now ask AI to mine these for interesting ideas that would resonate with today's themes, thereby reviving dormant concepts for contemporary research. How good any of these topic suggestions currently are is another question. Though, given the trajectory we have witnessed in the AI space, we should not take comfort in the current state of things but rather assume that answers will rapidly—as the OpenAI developer roadmap details—reach "PhD level reasoning skills" and beyond.

Risks of working with AI

Yet, as generative AI democratizes idea generation, new risks are emerging—and some are already manifest. Most notably, as recent studies show (Bail, 2024; Doshi & Hauser, 2024), AI can amplify the average but mute the unconventional. When many scholars use similar tools trained on the same corpora, the field risks "algorithmic monoculture" in its ideas (Holzner et al., 2025). Collective diversity—so crucial for innovation—declines as more minds are shaped by the same invisible hand.

Additionally, over-reliance on predictive "hot topics" or citation forecasts can also shift incentives away from curiosity-driven inquiry toward strategic trend-chasing. The risk is that scholars choose questions not for their potential to disrupt or enlighten, but for their likelihood of becoming popular and thereby fall into the trap of Goodhart's Law, according to which, "when a measure becomes a target, it ceases to be a good measure."

Perhaps most quietly, AI's efficiency can undermine serendipity. Wandering through tangential literatures, encountering oddball references, or having unplanned conversations—these moments have sparked many foundational ideas in our field. If the process becomes too streamlined, we may lose the unexpected intersections that only undirected exploration can provide (Bechky & Davis, 2025).

Synthesis/recommendation

Our recommendation is therefore threefold. First, a recent field experiment (Doshi & Hauser, 2024) found that the most innovative proposals came when researchers combined AI suggestions with their own curiosity—deliberately counter-prompting the AI for outlier or dissenting themes. As such, scholars should use AI as a creative amplifier, but never as the sole engine of novelty. Feed AI with diverse, even dissenting, sources because otherwise it will gravitate towards

[1] Here are the abbreviated answers generated for the first author when he queried ChatGPT-4o with the cited prompt: "(1) *How do teams and organizations co-construct the definition of "abusive" behavior, and how do these social processes shape outcomes? (2) Abusive supervision as a product of collective blind spots or silent complicity within teams and organizations, not just a "bad apple" leader. (3) Narratives of resilience, learning, and identity reconstruction after surviving abusive supervision—often absent from quantitative studies."* .

Table 1
Author side – AI across the research pipeline.

Stage	AI Enables	Risks / What to Watch For	Guiding Principle
Idea Generation	Literature-crossing prompts, propose new theoretical pathways, topic novelty checks	Trend amplification, algorithmic monoculture, loss of serendipity	Prompt for outliers and dissent, plan human-only creativity sessions
Literature Review	Summaries, mind maps, detect overlapping constructs and theories, cross-disciplinary connections	Shallow synthesis, prioritization bias, illusion of comprehensiveness, hallucinations	Use AI for breadth, critcially double check results; human insight for depth
Study Design	Protocol writing, item generation, better translation, virtual pilots	Overconfident designs, ethical blind spots, disconnected methods	Ground in context, challenge assumptions, get your hands dirty to understand the reality where it happens
Data Generation	Synthetic populations, AI SMEs; passive data collection, multimodal inputs	Epistemic contamination, unclear consent, deceptive AI–AI data loops	View simulated data as a supplement, not a replacement; ensure participant clarity regarding AI usage
Data Analysis	Code-free analysis, pattern discovery, multiverse modeling, automatic visualizations	Overconfidence, lack of understanding, reduced interpretive nuance	Cross-check with logic, different methods, and theory; know your methods
Writing	Draft generation, tone/grammar refinement, style alignment, implication scaffolding	Loss of author voice, hallucinated citations, gaming for reviewers or metrics	Retain narrative ownership, validate references, master copy-editing

Table 2
Publisher Side – AI in Peer Review and Dissemination.

Stage	AI Enables	Risks / What to Watch For	Guiding Principle
Peer Review & Editorial Decisions	Manuscript triage, reviewer matching, improved tone and structure in reviews and decision letters, revision comparison	Desk-bias entrenchment, flattened review tone, accountability opacity	Keep editorial teams in the loop; maintain critical dialogue
Revisions	Structured response letters, review interpretation, rebuttal drafting	Loss of creative friction, AI-shaped arguments rather than unique author response	Use AI for revision structure, not strategy
Dissemination	Auto-summaries, AI podcasts/videos, cross-platform content, *meta*-science tracking agents	Over-customization, performance over substance, filter bubbles	Prioritize clarity, fidelity, and public trust

incrementally instead of radically new ideas (Lee & Chung, 2024). To that end, deliberately prompt for the unconventional: "What research questions would surprise the field?", "What would be the contrarian viewpoint.", "How could the idea be critiqued?", or "Which perspectives have been systematically overlooked?" or "what are some unstated assumptions in this area?" Further, such questions can be answered with competing AI models.

Second, cultivate space for human intuition and serendipity—-through unscripted reading, interdisciplinary conversations, and time spent outside algorithmic suggestion loops. Then, spend time to train your LLMs on your mind wanderings. Treat it as your personal journal that you feed with your thoughts and with links you see between subjects as unconventional and tenuous as they might be. Because the future of research idea generation will belong to those who balance the machine's reach with the human capacity for surprise. It is then that we might be in for Cambrian explosion of ideas that truly break new ground.

Last, while the aforementioned field experiment (Doshi & Hauser, 2024) found that AI assistance made it easier for individuals to propose creative research ideas, group-level outputs became more homogenous and less surprising. Our suggestion would thus be to take a page out of our long experience with offline brainstorming and translate it to this current age: Individuals should brainstorm creatively in isolation, then with an AI, and only then meet in groups to collectively discuss the ideas before continuing with individual brainstorming again (Paulus & Kenworthy, 2019).

Understanding the state of and reviewing the literature

Of course, evaluating the quality of new ideas requires a solid grasp of the field itself. This is another area where AI shows promise—helping authors reach new levels of breadth and depth in mapping existing knowledge. Yet, much like with idea generation, relying on AI alone to understand the state of a field carries its own risks.

Benefits of working with AI

The task of mapping a field, once the domain of painstaking manual reading and years of scholarly accumulation, is now increasingly shaped—and sometimes upended—by artificial intelligence. See, for instance, a study that with the help of AI-enhanced workflows reproduced and updated an entire issue of Cochrane reviews ($n = 12$) in two days, representing—in their estimate—approximately 12 work-years of traditional systematic review work (Cao et al., 2025). Where once researchers switched for convenience or access reasons from Web of Science to Google Scholar, emerging scholars are developing an understanding of the literature via AI.

Presently, AI allows for more than simply the instantaneous synthesis of an article; authors can now listen to AIs have a conversation about the article in a podcast style, or they can have a conversation with the AI about the article. With tools like notebookLM or AlphaXiv, you can query a paper's logic, ask for clarifications, or request a mind map that lays out its arguments, methods, and findings—sometimes revealing hidden assumptions or subtle contradictions that might escape a quick read. At the very least, you can simply ask AI to explain the article to you like you are a 10-year-old, and maybe for the first time ever you will understand some of those denser articles that are otherwise hard to unlock.

But AI does not stop at the boundaries of a single article. If you provide AI with a handful of seminal papers, you can prompt it to trace outward, mapping networks of influence and surfacing connections that once took months of work. As another example, for those of us working in leadership studies, this also means we might finally get a handle on the jingle-jangle problem. That is, with its capacity to for pattern recognition, AI can sort through overlapping concepts and rival definitions, helping to clarify where constructs align—and where we are just renaming existing constructs in the literature (Banks, Gooty, et al., 2018; Banks, Woznyj, & Mansfield, 2021; Cronin et al., 2021). Paired with statistics from the respective papers, such an undertaking would be more than just a conceptual sorting, but it might actually help our field to become more diligent in how we handle constructs and their relations (Cronin et al., 2022; Podsakoff et al., 2016; Shaffer et al., 2016). Indeed, the European Commission also finds that research fields that are "rougher"—defined by the combinatorial complexity of a field's knowledge space, as one might suppose is true for the domain of leadership research—stand to gain more from AI use (Bianchini et al., 2025). While experienced researchers may already have such an understanding of their field, having such a service available will prove especially

valuable to newcomers to a topic who might be overwhelmed by the number of constructs.

Another promise emerges when AI breaks down disciplinary walls. Want to understand how other disciplines talk about dominance and status or want to get inspired about how a mechanism of fear may impact followers? Prompt the system to look in fields such as biology, political science, or economics, and it will surface related debates, theories, and empirical findings that may cross-pollinate our understanding. As we upload more articles into these systems, whether as authors or via a publisher agreement, an AI could also review the entire set with a guiding question in mind: Where have we been? Where are we headed? What are we missing? The resulting syntheses—maybe even accompanied by frameworks or visual maps—can reveal patterns and directions that would be invisible to the siloed reader or otherwise would take years to properly digest.

Once the AI is trained accordingly, it also becomes easier to look at the data and interpretations from different angles. Ever had that nagging feeling about the field thinking that all leadership studies can just be boiled down to relationship quality (Eva et al., 2024) or that transformational leadership theorizing is in need of a redo (van Knippenberg & Sitkin, 2013)? It has just become a whole lot easier to get at least a first draft of such a "review with an attitude".

Meanwhile, our old standby—the reference manager—is positioned to receive major upgrades. For instance, mark a set of papers and ask the integrated AI to summarize or synthesize the findings. Furthermore, what once used to be a static collection may soon serve as a foundation for auto-suggested new articles that support, refute or extend the theories authors engage with. Some reference managers such as Zotero can already flag retracted work; soon it may proactively identify studies with a high probability of *p*-hacking or other questionable research practices employed, making it easier to separate the signal from the noise (e.g., see the SCORE project: https://www.cos.io/score).

Risks of working with AI

Yet even as we marvel at these advances, we should be clear-eyed about the risks. First and foremost, the saying of "garbage-in-garbage-out" is particularly true in this domain. As of today, many LLMs were not (legally) trained on the scientific literature behind paywalls. As a result, today, many crawls will be faced with the paywalls of publishers. That said, future agreements between AI companies and publishers with wider access to high-quality published articles are on the horizon.[2] More generally, however, it is an issue of opacity: AI models may draw on sources without clear inclusion criteria, making it difficult to know whose voices are being amplified or excluded. As such, as an academic community, we may trust that these models are up to date with state of scientific research, when in fact they are not.

Second, even the best AI can sometimes offer a shallow synthesis—producing neat summaries that gloss over the deep philosophical divides or unresolved tensions that define the most vibrant debates. For example, AI might collapse interpretivistic and positivistic approaches to leadership into a single narrative, missing crucial nuance. As such, even if AI had access to all kinds of paywalled articles, depending on the prompt, there's a real risk that all literature reviews start to sound alike (i.e. homogenization)—converging on mainstream safe consensus instead of engaging with the nuances or challenging the field.

Finally, we must also acknowledge that AI systems are not immune to bias in how they evaluate and prioritize literature. A recent large-scale analysis of 274,951 references generated by GPT-4o across 10,000 papers found that the model consistently favored certain types of citations—preferring shorter titles, more recent publications, fewer authors, and papers with higher visibility (Algaba et al., 2025). While this potentially amplifies the well-known Matthew effect in scholarly recognition, the deeper concern lies elsewhere. When researchers conduct a literature review using traditional tools like Web of Science or Google Scholar, they are at least somewhat aware of the limitations of their search. By contrast, relying on AI may create a false sense of confidence that the field was comprehensively reviewed—obscuring the fact that the model's suggestions reflect hidden biases in what it has seen and how it has learned to prioritize.

Synthesis/recommendation

The opportunity before us is profound: AI can make literature reviews broader, deeper, and more versatile than ever before. But only if we treat these tools as starting points, not endpoints—partners in sensemaking rather than judges of consensus (Malik & Terzidis, 2025). AI should be regarded as an incredibly powerful research assistant, but just like any research assistant, the AI might not be as diligent when sourcing or when summarizing. It is the task of its supervisor (i.e. the author or user) to firstly guide its search and viewpoints and secondly critically check the results before engaging with them. The future of reviewing the literature belongs to those who use AI to cultivate new lines of inquiry, not eradicate them.

Planning study design

Benefits of working with AI

AIs can not only be trained with an "understanding" of the field, but also in methods. In fact, many LLMs have already been trained on various methods textbooks as well as tutorials and methodological discussions available online. As such, the planning of research studies that once demanded hours of author team brainstorming, deep dives into methods papers, or trial-and-error drafting, can now happen in tandem with a machine.

On the most practical level, generative AI is already proving invaluable for protocol and speech writing. Whether you need a subtle workplace task or a description of leadership behavior, you can prompt an AI to generate variations, holding everything constant but the variable you seek to manipulate—then refining and contextualizing them. In the future, through tools such as VEO3 and Flow or Runway, such vignettes can also easily be produced in the form of videos and thus provide a level of engagement that regular text-based scenarios often lack. By a similar token, AI can also be used to assist in scale development (Fyffe et al., 2024; Götz et al., 2024; S. Hoffmann et al., 2024). This is not just a shortcut; it means richer, more diverse stimulus materials, potentially better reflecting cross-cultural or other context specific elements.

Truly globalizing research, AI also promises to help researchers move beyond the outdated Brislin method for translating items (Brislin, 1970)—a method that is either championed by those who are not fluent in two languages and who therefore do not see the extreme shortcoming of its current practice or those who have massive resources at their hand to follow the full protocol. Using AI-powered translation engines (like DeepL) with subsequent human oversight can ensure that items retain conceptual nuance across languages and contexts, while ensuring a high level of translation fidelity.

Yet, the AI's role does not end with wording. Feed it the latest methodological articles or best-practice guidelines, and you can consult it for planning your research design—from operationalizing constructs to selecting controls and anticipating sources of bias. As its training on actual scholarly literature expands, this capacity is only growing. So far so, that some researchers are already running "virtual pilot studies," having AI simulate responses from synthetic populations to estimate likely outcomes, effect sizes, or even unanticipated confounds before a single study is launched (Argyle et al., 2023; Ashokkumar et al., 2024; Lippert et al., 2024; Meinhart, 1966).

Looking ahead, we are approaching a point where researchers might submit everything *but* their methods section to an AI and have the AI

[2] At the time of writing this editorial, already Wiley and Taylor & Francis have publicly announced partnerships with AI companies, notably without opt-out options for authors.

draft the entire design and analysis plan, guided by the latest standards and tailored to the specifics of the research question. This process would essentially equate to the procedure of a registered report, but instead of proposing and subsequently asking reviewers, authors would ask the AI for sensible tests of their theorizing, including the proper statistical power or context for examination and testing. So, there is an opportunity for peers and maybe even journals themselves to share prompts that would lead an AI to generate rigorous and comprehensive test and analysis protocols of presented theory.

Risks of working with AI

Before we get swept away by these possibilities, it's worth pausing to consider the risks. Chief among them is the illusion of optimization: AI-designed studies may seem ideal in theory or in simulation, but might fail spectacularly in the messy, contextual realities of real-world organizations. In the lab, for instance, an AI may act more cooperative and altruistic than humans would (Mei et al., 2024). When too much weight is placed on virtual pilots or machine recommendations, we risk missing human or contextual nuances.

Additionally, much of organizational behavior research is already conducted online. While there are many good reasons to do so (Van Quaquebeke et al., 2022), such an approach also runs the risk of not properly understanding the populations one is studying because one has, in fact, never met them. This is even more so the case when AI is in the mix. Where researchers once had to "get their hands dirty" to gather data where it is happening, in the future we run the risk of becoming complete desk researchers, confidently mapping a research plan for a world that neither we nor our AI partner have ever visited.

There is also the ethical blind spot. Even well-trained AI may not recognize whose voices are excluded or when a design quietly reinforces bias or structural inequality and neglects to consider protections for human subjects. If left unchecked, the risk is that our studies become technically advanced but socially shallow. Therefore, as we automate more of the design process, there is the real danger of losing researcher agency. If AI drafts the methods and analysis plan, what becomes of the critical judgment, ethical reasoning, and creativity that define methodological craft?

Synthesis/recommendation

The lesson here is not to reject AI as a study design partner, but to use it intentionally—with eyes wide open. AI can spark new ideas, catch overlooked details, and stress-test designs before we commit resources. But its recommendations should be a starting point for collective deliberation, not the final word. The best research will come from partnerships where human judgment, ethical reflection, and contextual sensitivity are sharpened—not dulled—by our use of AI (Resnik & Hosseini, 2025).

Study execution/data generation

Benefits of working with AI

In the future we are detailing, AI is not simply a design tool but may become an active participant in how we execute studies and generate data (Musslick et al., 2025). As noted in the previous section, researchers are already using AI to simulate human responses—both quantitative and qualitative (Cui et al., 2024). This can dramatically accelerate pre-testing and help refine instruments before involving real participants. Moreover, the boundaries keep expanding. Some researchers are inserting AI agents into entirely fictitious scenarios—such as time travel (Gans, 2025) or alternate economies—to test "what if" questions beyond the limits of actual experimentation.

AI can do more than mimic survey-takers. Increasingly, experiments are using AI as standardized role-players—virtual employees or stakeholders who interact with human subjects in controlled, repeatable ways (Krafft et al., 2017). These AI "confederates" can be programmed with consistent personalities, making it possible to manipulate leader–follower dynamics with a precision but also naturalism that is difficult to achieve with pre-programmed scripts. Consider, for instance, an AI that is trained to behave like a bad actor in team settings but does so not along a static script but in a more contextually aware manner. That is, logically responding to what other experiment participants say. Not only does it provide a more naturalistic counterpart compared to what scripted confederates would be able to do, but prompts could easily be shared amongst researchers trying to build on the experiment. And any interaction transcripts are subsequently available for further richer conversational analysis.

AI may also begin to challenge the traditional role of subject-matter experts in coding participant responses. Recent findings suggest, for instance, that, when trained on task-specific criteria, AI can outperform individual human raters in evaluating job candidate responses (Thompson et al., 2023). Even more striking, recent work shows that zero-shot large language models—without any fine-tuning—can surpass independent human evaluators across multiple datasets (Speer et al., 2024).

Meanwhile, automated digital ethnography may move from science fiction to practice: bots can (with consent) enter digital workspaces—Slack channels, Zoom meetings, LinkedIn groups—passively collecting nuanced interactional data that would be nearly impossible to capture at scale with traditional methods. That, however, does not need to be the end. Bots can, for instance, also be more actively inserted in such channels to see whether and how they can infuence people, such as was the case in a recent experiment where AI bots were submitted into a subreddit on "Change my view" (notably, unbeknownst to and therefore without consent of the audience).[3] Here researchers had the bots post over 1500 statements with made-up bios (e.g., a sexual assault survivor, a trauma counselor, and a Black man opposed to Black Lives Matter) to see what type of responses they were met with by the community and whether these statements were, in fact, persuasive.

On top of all this, AI's ability to capture and synchronize multimodal data—voice, facial expression, biometric signals, chat transcripts, including inferences about personality (Fan et al., 2023)—promises unprecedented richness in understanding leadership interactions in real time, in particular as more virtual working contexts will allow collecting all these data at the same time (Wenzel & Van Quaquebeke, 2018). With the right tools, researchers can now contextualize decisions, emotions, and group dynamics with a granularity previously unimaginable. Think of interaction coding (Güntner et al., 2023), but on the next level.

On a broader scale, AI could enable a shift from static *meta*-analyses to continuous, real-time synthesis, much like a stock ticker. As new studies appear, an AI could update cumulative findings on a rolling basis, keeping knowledge current without the delays and bottlenecks of traditional *meta*-analytic processes—not just aggregating the statistics but also interpreting and writing about them with the aim of guiding practice and identifying new theoretical pathways (e.g., unexplored moderators).

Risks of working with AI

Yet, these innovations come with deep and unresolved risks. Chief among them is epistemic contamination: if part of your data is generated by AI—whether synthetic employees or simulated teams—where do we draw the line between simulation and empirical reality? There is a real risk of circular reasoning: AI learns from past data and generates "new" responses based on those same patterns, potentially reinforcing existing biases (Bisbee et al., 2023).

Moreover, participants may also be affected by the use of AI in

[3] The title of the paper was "Can AI change your view? Evidence from a Large-Scale Online field experiment" and was led by a group of researchers at the University of Zurich. Following extreme backlash from the Reddit community and beyond, individual author names are not identifiable neither does it seem to be listed as part of the ACM proceedings anymore.

studies as confederates. It is policy at *The Leadership Quarterly* that deception can only be used as a last resort (and may be a cause for a desk rejection). One of the reasons is that deception opens up threats to internal validity and negative externalities may occur for other researchers (i.e., pollution of subject pool). Thus, as our community continues to rely heavily on online panels, it is reasonable to expect that study participants—already wary of whether they are interacting with real people—will become even more skeptical when responses by a supposed other human participant appear unusually quick or polished. At some point, we may need to ask: Can we still assume these are human-to-human interactions, or must we begin treating such data as reflective of human–AI exchanges as respondents assume there is an AI on the other side?

Relatedly, who is to say that respondents may not turn to AI themselves. As such, researchers might not even be aware of the contamination of their data when their panel respondents use AI operator models that take control of their desktops, letting the AI fill in online surveys and experiments in the thousands.

As AI is increasingly used to engage with human participants, the ethical landscape will become ever more complex. Synthetic participants and AI agents interacting with human subjects will blur lines of informed consent, deception, emotional impact, and autonomy. If participants do not know they are engaging with AI, or if AI subtly nudges behavior, we wade into murky ethical waters.

Synthesis/recommendation

The challenge ahead is clear: Do not let efficiency or technological novelty distract from the core mission of understanding real, complex, context-dependent human behavior in organizations. In this new era, the best research will be the work that treats AI not just as a source of speed or novelty, but as a force for deeper, more critical reflection on what counts as (human) data, reality, and evidence (Larson et al., 2024).

We should also all note that AI-generated, i.e. synthetic data are a tool for exploration and refinement, not as a replacement for human experience. It is for this reason that data simulated by AI should only be treated as such, i.e. simulated data. It can serve to finetune your methods or reasoning, but it should not be mistaken as empirical data. It is theoretical data (Von Krogh et al., 2023).

Finally, we will only receive data we can trust, when the people we query trust us that what we say is true. Therefore, it is advisable to be transparent about when and how AI agents are used within our designs. That is, we should follow the no-deception rule championed by experimental economists and also this journal.

Data analysis and interpretation

Benefits of working with AI

The work of analysis is no longer limited to those with advanced coding skills. Today, AI is lowering the technical barriers, facilitating analysis and interpretation (for instance, when using more sophisticated NLP to analyse text data, Feuerriegel et al., 2025), and opening new possibilities for integrating and visualizing results. One of the most tangible benefits is in sensemaking and error-checking (Gibney, 2025). You can now upload tables, statistical outputs, or even raw datasets to an AI-powered assistant and receive instant feedback on correctness, patterns, or anomalies that warrant a closer look (cf. yesnoerror.com). Numbers that were previously abstract can now be explained to you. This not only accelerates the vetting process but can help catch mistakes or outliers that might otherwise slip through the cracks. In the end, it holds the promise that people actually understand the statistics better than they do today because there is someone, i.e. the AI, that can explain it to them.

AI is also democratizing more sophisticated analysis. With "Python at your fingertips," researchers who once relied on a friendly statistician or weeks of tutorials can now generate advanced models, create custom or even interactive visualizations, or attempt more obscure analytical techniques by describing their analytical goals in natural language. And because the threshold for conducting advanced analyses is now so low, researchers are at liberty to explore their data from multiple angles—without feeling locked into a single analytical path simply because it was costly or time-consuming to begin with (cf. escalation of commitment).

Risks of working with AI

Alongside these advances, we must acknowledge the shadows they cast. There is the risk of reductionism, particularly in qualitative analysis. While AI can summarize, categorize, or code massive volumes of text, it may flatten the layered meanings that characterize human discourse at work. Subtle signals—of power, identity, or emotion—can be lost when human interpretation gives way to statistical efficiency.

On the flipside, there is the danger that AI confidently produces results that do not hold up to scrutiny. As such, it may detect patterns that are statistically robust but contextually misleading—overfitting to idiosyncratic data or generating plausible but irrelevant correlations (Wenzel & Van Quaquebeke, 2018). Essentially this is not that different to the point-and-click way of how many operate their statistical software today, but because AI so confidently also interprets or even defends the results, a statistics novice might question the output even less.

Synthesis/recommendation

Researchers must remain vigilant—treating AI as powerful analysis tools, not substitutes for theoretical grounding or methodological competence. Transparent reporting, mixed-method validation, and humility about what automated analysis can—and cannot—capture will be key. In its essence, AI is a tool much like previous software tools we have come to embrace over the last decades. As such, the quality of the output will correspond to the competence of the user. A solid foundation in methods and analytical approaches therefore continues to be a must for all researchers. That said, the structure of the required foundation might change. Raher then knowing methods in detail, we need to trust (after some vetting) that the software will develop the results we asked for—in the same way we trust SPSS or STATA to reliably execute behind the scenes what we would obtain when we calculate a certain test by hand. As with prior software, however, results must be reproducable. Here, the application of multiple AI tools may serve as a form of robustness check.

In the future, researchers will be able to make the most out of AI in their data analysis when they are aware of the breadth of options available and can critically cross-validate and check whether core principles of the respective statistics were adhered to.

Reporting/write-up

Benefits of working with AI

A blank page no longer creates the sense of writer's block it once did. With generative AI, scholars are rarely at a loss for words—even if those words do not always sound like their own. Whether you are voice-recording argument fragments, freewriting in a late-night burst, or uploading notes in random order, today's tools can stitch these ideas together, structure them into a narrative, suggest references, check for logical gaps, and even draft counter-arguments to keep your reasoning sharp. Consider, for instance, an experiment amongst college-educated professionals showing that for those who worked on a writing task with ChatGPT the average time taken decreased by 40 % and output quality rose by 18 % (Noy & Zhang, 2023) with similar results for consultants from the Boston Consulting Group (BCG) on a variety of tasks (Dell'Acqua et al., 2023). This collaborative writing model is more than a convenience. For many, it makes the hardest part of writing—getting started and letting go—significantly easier. When AI assembles the first draft, it is easier to "kill your darlings" and rewrite without attachment, adding precision and clarity where needed. The process shifts from solitary struggle to iterative refinement.

Writing with AI doesn't just reduce effort—it can also help focus.

Quasi-experimental evidence from the coding platform GitHub (Hoffmann et al., 2024) shows that generative AI significantly improves both the speed and quality of knowledge work, particularly for those with lower baseline performance. AI assistance allows people to devote themselves more to the core of their (own) authorship process rather than having to coordinate and manage the process with others—as the authors describe it in their abstract: it leads to "an increase in independent rather than collaborative work, and an increase in exploration activities rather than exploitation".

The impact of writing with AI is perhaps most profound for those writing in a second or third language. AI does not just correct vocabulary and grammar—it adapts tone, polishes style, and can even mimic the conventions of academic English, making the field more accessible to non-native speakers and helping to level the playing field in global scholarship. AI can also bridge the gap between theory and practice. When prompted, it can propose plausible practical implications, frame research findings for specific audiences, and translate complex arguments into actionable recommendations—a perennial struggle for many academics.

Formatting headaches are receding as well. AI can instantly adapt manuscripts to meet the shifting requirements of different journals, freeing scholars from the grind of adjusting to specific styles. Maybe one day, we will see "Journal Optimizer" tools, where you can upload a draft and receive tailored suggestions for which journals might be most receptive, complemented with some advice on how to tweak the writing once more to fully fit the usual journal style.

Risks of working with AI

Yet, as we lean into these new possibilities, the risks come sharply into focus. Chief among them is the loss of scholarly voice. Automated drafting can lead to prose that is polished but generic—flattening the idiosyncrasies, passions, and provocations that make for memorable academic writing, especially in interpretive or critical work. There is also the matter of references. As already outlined above in the section on reviewing and understanding the literature, AI-generated citations are not infallible: hallucinations, misattributions, and entirely fake papers can creep in if the system is trusted too much. Vigilance in checking every citation and its content becomes more—not less—important as automation grows.

Equally on the horizon: When authors with AI assistance become more capable of producing solid scientific texts and analyses, we may see a decline in collaborative writing (Hoffmann et al., 2024). This raises a concern: the co-authorship process—so often a space for intellectual exchange, critique, mutual learning, and initiation for the next generation of researchers—could quietly erode, with implications for the social fabric of scholarship itself.

Perhaps most subtly, there is the risk of "reviewer gaming." If AI can predict not only which journals but what a reviewer might favor in a manuscript, scholars may be tempted to optimize for acceptance, tailoring arguments to expected preferences rather than to the demands of rigor, originality, or intellectual honesty. The temptation to (have AI) write for the algorithm, rather than for the advancement of knowledge, is a risk the field must face head-on.

Synthesis/recommendation

Having outlined the risks, let us also be candid. A lot of this is already happening. We train PhDs to adopt a certain scholarly style. In fact, the authors of this paper know supervisors who encourage their PhDs to adopt the style of specific successful academics. Part of the reason is that every author hopes for acceptance of their submitted work, thus also at times adhering to reviewer requests as non-sensical as they might be. And yes, people have been citing papers for the wrong reasons forever with apparently little to no understanding of what they actually say (Letrud & Hernes, 2019). But that should not deter us from doing it better in the age of AI.

The future of scholarly writing will not be about rejecting AI, but about writing with it—critically, creatively, and self-consciously. Use AI to lower barriers, spark drafts, and clarify prose, but do not surrender your voice or judgment. In fact, train it on your voice if you already have multiple manuscripts to train it with or, at the very least, be deliberate about the voice with which you want to write with.

Additionally, take full advantage of AI's capacity for structuring, translating, and optimizing—but remember, the most compelling arguments, the sharpest insights, and the greatest leaps in understanding still depend on the human mind's willingness to challenge, surprise, and speak in its own register. More generally, it therefore seems our challenge will become not to master authorship but master copy editing—a worthy challenge we should also train the next generation in. While the role might shift, the ambition should be the same: To accurately get your point across so that it lands with the audience.

Using AI in the journal publication process

Much of the attention around AI in research has focused on authors. But another transformative shift may be happening on the other side—within the journal publication process itself. As submission numbers rise, reviewer fatigue grows, and pressure mounts for faster turnarounds, AI is stepping in not just as a tool, but increasingly as a co-decision-maker. While this shift holds real promise, it also raises foundational questions about judgment, fairness, and the very purpose of peer review (Aczel et al., 2025).

Benefits of working with AI

We can imagine a world where authors submit their manuscript to an *LQ*-specific AI assistant before it even reaches human eyes (Checco et al., 2021). The system could run a quick stats check, flag unclear or inflated claims, and scan for conceptual redundancy against previous *LQ* articles—offering a first-pass assessment of novelty that may motivate the authors to correct or revise aspects before officially submitting (cf. PaperCheck, https://scienceverse.github.io/papercheck/). The same system could highlight problematic framings, identify colonial language or exclusionary sampling, and recommend adjustments that improve the paper's inclusivity before the peer review process begins.

For editors, beyond the above, AI could offer triage support by estimating a submission's likely scholarly impact over time, based on novelty signals and citation trajectory models. In fact, already today's models can distinguish between higher- and lower-quality research closely aligning with established journal prestige measures (Pataranutaporn et al., 2025). Likewise, it could flag deeper methodological issues (cf. Open Science Framework Systematizing Confidence in Open Research and Evidence [OSF SCORE] project that works towards developing and deploying automated tools to provide rapid, scalable, and accurate confidence scores for research claims) or ethical risks (Lehr et al., 2024): Was informed consent truly clear? Are populations used responsibly? Does the study's design mask social or identity-based biases? These capabilities will not replace editorial discretion, but they could provide sharper inputs for early-stage decisions.

In the review process itself, AI could improve both matching and output. It could pair submissions not just with topical experts but with reviewers whose feedback style suits the manuscript—pairing more tentative, early-stage work with more developmental reviewers, and routing technical challenges to methodological specialists. And the reviewers themselves would no longer need to start from a blank page. Indeed, a journal could provide specific reviewer prompts that follow a certain template and can then be idiosyncratically improved upon by the individual reviewers. Consider, for instance, that already 1.5 years ago the overlap of GPT-4 and human reviewers was similar to that of two human reviewers (Liang et al., 2023). Similarly, an AI can help with a rewrite of the human review. A human review draft could, for instance, be adjusted for a more constructive tone and better structure while keeping the content the same. The aim here is not to eliminate human

voice, but to scaffold it—helping reviewers focus their attention on judgment rather than syntax (Van Quaquebeke & Gerpott, 2024).

Second, editors would benefit in the review and revision process as well. With AI assistance, they could synthesize disparate reviewer comments into coherent, calibrated decision letters that balance critique with encouragement. Subsequent revisions could be tracked for substance over style, with AI flagging cases where authors have failed to meaningfully engage with core feedback across multiple rounds.

Third, also authors stand to profit from using AI in the revision process. Here too, the daunting blank page of the response letter becomes a relic of the past. Based on the initially submitted paper and the received reviews, an AI can not only assist with recommendations of how to improve the paper but also provide a first draft of the response letter. At the very least, an AI can also be used to make sense of the reviewer comments and provide perspective on them, or it may offer support on how to best or diplomatically address potentially conflicting reviewer comments.

One day, journals' systems might evolve toward living articles (e.g., https://lifecyclejournal.org/). Pieces could update as new data or debates emerge. AI could be used to signal to readers how the present article may fit into the corpus of more contemporary findings and how to trust its interpretation given developments in methods and theories. Relatedly, further zooming out, AI could offer *meta*-level insight: tracking field-level biases, shifts in citation dynamics, and the replication health of published work. This kind of longitudinal oversight could allow publishers and editors to steer journals with greater transparency and intention.

In the most radical version, humans step out of the review loop entirely. Manuscripts are assessed for soundness, novelty, and responsiveness to critique by an ensemble of independent AI models. In this example, review cycles would be largely automated and revision logs are parsed by machine. The publishing process becomes not only open access—but also open logic. In this futuristic example, human editors will be responsible for the logic as it aligns the AI that does the reviewing and decisioning. The logic (and its updates) could then also be openly stated on the journal website, essentially not much different to what journals already do today with their mission statements and set of editorials but now with a lot more consistency, transparency, and accountability in the process.

Risks of working with AI

The same systems that promise efficiency and consistency also carry serious risks. If trained on historical decisions, AI may reinforce the field's entrenched biases—preferring familiar theories, established authors, and elite institutions over new voices and disruptive ideas (Pataranutaporn et al., 2025). What looks like "fit" may in fact be conformity. Pattern recognition is powerful, but paradigmatic innovation rarely follows a pattern. Indeed, AI risks flattening the nuances that human editors and reviewers bring. An unconventional methodology, a provocative argument, or a culturally grounded insight might not register as valuable within the model's parameters. Worse, we could see reviews start to converge—technically coherent, but lacking the idiosyncratic, the pointed critique, or the provocation that partly also defines meaningful peer dialogue (Naddaf, 2025).

While better understanding what reviewers are trying to say is certainly appreciated—by both editors and authors—we should also be careful not to flatten the messy, creative tension that sometimes drives real breakthroughs. Some of the most meaningful improvements to papers do not come from perfectly clear instructions, but from moments of confusion: when authors wrestle for days with a cryptic reviewer comment, misinterpret its intent, and in the process stumble into an entirely new idea or contribution. Indeed, standard answers by current AIs seem to be focused on doing more rather than smarter ways of doing something (Liang et al., 2023). If AI assists too much in interpreting reviewer feedback or automating responses, we risk losing this productive friction—the chance for human ingenuity to take a reviewer's concern in an unexpected, even radical, direction.

Accountability becomes murkier as well. If a paper is desk-rejected or substantially shaped based on AI input, who takes responsibility? The human editor? The model? The process becomes more efficient, but potentially more opaque. We should take caution from the story of how AI was introduced to the judicial system in the United States. What once was supposed to just be a decision support system to provide probabilities of offenders committing a crime again, quickly turned into the de-facto decision maker (O'Neil, 2017).

There is also a subtler danger: that AI will reward manuscripts optimized for algorithmic preferences—papers that align with citability models, buzzword density, or prior publication trends (all issues that many editors and publishers obsess about). This could encourage authors to write for machines rather than for ideas, reinforcing a shift from contribution to performance. Over time, we risk developing journals that are consistent but bland—monocultures of safe, familiar scholarship, smoothed of friction and surprise.

Synthesis/recommendation

Against the above, our recommendation is to integrate AI not as an arbiter, but as an assistant—to speed the mechanical, flag the questionable, and support the ethical, while preserving space for dissent, depth, and debate. The future of publishing should not be about removing friction; it should be about focusing our effort where it matters most. Where AI can free resources by making everything faster, we should deliberately use these to make the process more intentional.

Using AI in the dissemination process

Publication has traditionally marked the end of the research process. But with AI reshaping how knowledge moves through the world, publishing is no longer a full stop—it is a starting line.

Benefits of working with AI

Already, we are seeing the early stages of AI-driven dissemination. Authors can generate summaries or visual explainers that highlight the practical relevance of their work, with outputs automatically formatted for different platforms—LinkedIn or Instagram posts, blog essays, or media pitches. In a matter of minutes, a dense theoretical article can be reframed into formats that resonate with practitioners, policymakers, or the broader public (Kreps et al., 2022).

More ambitious transformations are becoming feasible. AI can already simulate podcast conversations based on an article's content, turning dense text into dynamic dialogues that feel conversational and accessible (cf. Google's notebookLM). Soon, it may become common to generate personalized videos, where an AI trained on the author's voice and image delivers a narrative presentation—breaking down key findings as if speaking directly to an audience.

But AI's potential does not stop at helping people to understand, it can also help train them. It could help authors turn their research into compelling teaching cases, learning apps, or even games—bringing concepts to life in ways that compellingly bridge research to education. This way, scholars can reach classrooms without needing to become a media or programming pro. For example, the author of a piece on hit production in the music industry (Berg, 2022) developed his own little app so that people can understand the dynamics of what he found (https://justinmberg.com/wp-content/uploads/hitmakergame.html). And naturally those "classrooms" do not need to be situated at a university only. Consider, for instance, the utility of AIs trained on certain scientific findings to be used in general trainings or coachings (Graßmann & Schermuly, 2021). That is, an AI could, for instance, be trained to assist leaders to become more decision resilient, more charismatic in their rhetoric, or generally more effective in their handling of team members

(Weidmann et al., 2025).

Post-publication tracking could evolve too. Imagine a *meta*-science agent assigned to your work: an AI that monitors how your ideas are cited, debated, applied—or misapplied—across the scholarly ecosystem and beyond (Hosseini et al., 2020). It could notify you of unexpected uses or interpretations, suggest clarifications, or even offer automated responses in forums and discussions where your research is being discussed.

Finally, as scholarly articles become increasingly accessible to AI systems, we should expect those same systems to be used by practitioners, policymakers, and journalists to make sense of complex issues (see, for instance, private initiatives such as Skarlinski et al., 2025). Where once only a handful of academics were consulted—and thus shaped public discourse—AI could instead synthesize a broader and more diverse body of knowledge. The result may be a more balanced representation of what the field as a whole has to say.

Risks of working with AI

But as we move into this new terrain, we must tread carefully. The same AI that enables dynamic dissemination could over-customize messages—tweaking interpretations to different audiences in ways that reinforce existing biases or create "filter bubbles". Different stakeholders might hear tailored but subtly divergent versions of the same findings, fragmenting rather than unifying the field's influence.

Metrics, too, become a contested space. AI could be used to optimize social media visibility, boost citation counts, or manipulate altmetrics in ways that distort the fidelity of scholarly impact. The risk is that dissemination becomes a game of optimization for engagement rather than a practice of genuine education—much like we have seen it in the journalistic space as well.

Synthesis/recommendation

The future of disseminating research findings will be characterized not just by reaching more people, faster. It will depend on preserving the integrity of what is shared—and ensuring that the core spirit of inquiry, dialogue, and reflection is not lost in the rush for reach. Used thoughtfully, AI can help scholarship travel further. Used carelessly, it may help it fragment faster, jeopardizing the trust that the public still has in science. This places a dual responsibility on us as researchers—not only to communicate our findings clearly and responsibly, but also to help educate third parties in how to interpret and critically engage with the outputs that AI systems may generate based on our work.

General considerations

While having outlined above specific benefits and risks of working with AI in the research and publication process, below we offer more general points to consider in this debate as well.

General caveats and limitations

As we look to the future of AI in research and publishing, it is worth stepping back to acknowledge the quieter, structural risks that may not show up in any single application but become consequential in aggregate. These are not dramatic misuses or apocalyptic failures. They are the slow erosion of judgment, equity, and understanding. First and most fundamentally: *garbage in, garbage out*. No matter how powerful the AI, it can only work with the data it is fed. If that data reflects past biases, incomplete literatures, or narrow worldviews—as it often does—then the outputs will replicate and reinforce those limitations, just with more polish and speed (Bai et al., 2025). Maybe more importantly, how do we account for the long-term influence of AI providers, i.e. the tech overlords—whose algorithms may subtly shape the direction of scholarly inquiry (Santurkar et al., 2023)?

Access and infrastructure remain significant dividing lines. Well-funded institutions are rapidly advancing with custom-tuned models, large-scale training pipelines, and proprietary data ecosystems, while scholars in less resourced contexts—especially in the Global South—risk being left behind. Yet, there is some cause for optimism. Recent developments in AI have shown that even the most advanced models are becoming markedly more efficient, with lower computational and financial barriers to use. Over time, this trend could enable broader access rather than restrict it. In fact, despite the very real risk of deepening inequalities (Bianchini et al., 2025), AI may ultimately serve as a greater force for democratization, challenging the status quo more than reinforcing it. Indeed, a field experiment conducted amongst BCG consultant, for instance, found that those below the average performance threshold managed to increase their performance on average by 43 % while those above only increased it by 17 % (Dell'Acqua et al., 2023).

There's also the danger of *point-and-click scholarship*: where researchers use AI to generate or interpret results without understanding what is actually happening in the analysis. How was the model trained? What assumptions were made? What do its metrics mean, and where do they break down? AI often speaks with a confidence that far exceeds its actual certainty—and users, especially those newer to research or using AI, may mistake fluency for authority. Relatedly, there is also the deeper issue of *use it or lose it (*Kosmyna et al., 2025*)*. If we outsource too much—writing, reviewing, idea generation, even critique—we risk losing the very practices that define scholarship: reflective thinking, nuanced interpretation, intellectual struggle, and collaborative debate. These are slow, sometimes inefficient processes. But they are also where meaning is made. As reliance on AI grows, we risk a slow atrophy of methodological understanding, creativity, and critical thought—especially for the younger generation that does not know otherwise (Gerlich, 2025; Lee et al., 2025; Robert et al., 2024; Tully et al., 2025).

Broader promise for our discipline

Paradoxically, working with AI might force us to be more precise in how we define, link, and test our ideas. Not because AI tells us what the truth is—but because it forces us to ask more disciplined questions about how we think truth might work. As such, it might push us (finally!!!) toward greater conceptual and methodological clarity (Banks et al., 2018; Cronin et al., 2022; Fischer et al., 2020, 2023; Gallea et al., 2025). There could be a future, for instance, where a specifically trained AI model may expose fuzziness. What exactly do we mean by "transformational leadership" in this study? How does our mechanism differ from adjacent theories? If the logic does not hold up, the system could flag it. Maybe, just maybe, it will allow the field to converge the various concepts and mechanisms into a more finite set—a nomenclature we can then all work with.

Building on the above, imagine subsequently our field as built on keystone datasets of actual empirical data (not just synthetic data as proposed in Wang et al., 2024). Not only could we, with the help of AI, harmonize and converge older data according to a new nomenclature, but such datasets could then be positioned to be open and grow—continuously enriched by scholars across the globe. These datasets would allow for cumulative knowledge building, far beyond the fragmented and siloed structures we often work within today.

Conclusion

We opened this editorial with "the elephant in the room". It is time to name it plainly: at some point, core research processes may no longer require human scholars (Lu et al., 2024; Yamada et al., 2025). Idea generation, study design, data analysis, peer review, even the articulation of practical implications—these are all tasks that AI is beginning to perform, at least in prototype form. The real risk may not be that AI

replaces researchers, but that researchers slowly begin to imitate the logic of machines. The scientific imagination shrinks—not through automation, but through quiet, uncritical alignment. But there is another path. One that takes deliberation. AI could take on much of the procedural weight of research and publishing, freeing up time and cognitive space for us. Let us not squander that time and space trying to produce more, but understand less (Messeri & Crockett, 2024). Rather let us all think more deeply, challenge more boldly, and interpret more finely (Berg et al., 2017). It is an invitation revisit our theories with sharper eyes, clarify our models with greater discipline.

What role remains for humans in a future shaped by AI? The same one that has always mattered most: to ask better questions, to resist easy answers, and to keep alive the reflective, critical, and imaginative spirit that turns information into insight. The opportunity—and the responsibility—rests with us.

Author note

ChatGPT-4o as well as a specifically trained GPT were used to augment ideas of the authors. ChatGPT-4o was additionally used to phrase or paraphrase some of the ideas—as clearly evidenced by the many em dashes—while prompted to retain the tone of the authors. We had ChatGPT o1-pro review sections and provide additional feedback. In doing all of the above, we follow the spirit of what we try to lay out in the present editorial.

The use of AI in research is such a novel and somewhat shunned phenomenon that published sources on some of the related issues are hard to come by. In the following, we therefore, amongst peer-reviewed papers also refer to papers in public repositories, fully acknowledging that their results may not hold up to the scrutiny of peer review (e.g., Lahart, 2025). Additionally, we also rely on what we authors were told by our peers or attempted ourselves. Naturally, it is a snapshot in time, where issues that we currently only speculate about may already be common practice by the time someone will read this paper.

The first author maintains an updated list of AI tools that can be used in the scholarly workflow. Given the amount of tools and speed of development, it will never be complete nor always list the newest tools, but it can be a good starting point for many things we outline in the present paper: https://buff.ly/3zXkFMs.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aczel, B., Barwich, A.-S., Diekman, A. B., Fishbach, A., Goldstone, R. L., Gomez, P., Gundersen, O. E., Von Hippel, P. T., Holcombe, A. O., Lewandowsky, S., Nozari, N., Pestilli, F., & Ioannidis, J. P. A. (2025). The present and future of peer review: Ideas, interventions. In *Proceedings of the National Academy of Sciences, 122*. https://doi.org/10.1073/pnas.2401232121

Algaba, A., Holst, V., Tori, F., Mobini, M., Verbeken, B., Wenmackers, S., & Ginis, V. (2025). *How Deep Do Large Language Models Internalize Scientific Literature and Citation Practices?* (No. arXiv:2504.02767). arXiv. https://doi.org/10.48550/arXiv.2504.02767.

Amabile, T. (2019). Creativity, artificial intelligence, and a world of surprises guidepost letter for academy of management discoveries. *Academy of Management Discoveries*, 351–354. https://doi.org/10.5465/amd.2019.0075

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023). Out of one, many: using language models to simulate human samples. *Political Analysis, 31*(3), 337–351. https://doi.org/10.1017/pan.2023.2

Ashokkumar, A., Hewitt, L., Ghezae, I., & Willer, R. (2024). *Predicting Results of Social Science Experiments Using Large Language Models*.

Bai, X., Wang, A., Sucholutsky, I., & Griffiths, T. L. (2025). Explicitly unbiased large language models still form biased associations. *Proceedings of the National Academy of Sciences, 122*(8), Article e2416228122. https://doi.org/10.1073/pnas.2416228122

Bail, C. A. (2024). Can generative AI improve social science? *Proceedings of the National Academy of Sciences, 121*(21), Article e2314021121. https://doi.org/10.1073/pnas.2314021121

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior, 45*(2), 159–182. https://doi.org/10.1002/job.2735

Banks, G. C., Gooty, J., Ross, R. L., Williams, C. E., & Harrington, N. T. (2018). Construct redundancy in leader behaviors: A review and agenda for the future. *The Leadership Quarterly, 29*(1), 236–251. https://doi.org/10.1016/j.leaqua.2017.12.005

Banks, G. C., Woznyj, H. M., & Mansfield, C. A. (2021). Where is "behavior" in organizational behavior? a call for a revolution in leadership research and beyond. *The Leadership Quarterly*. , Article 101581. https://doi.org/10.1016/j.leaqua.2021.101581

Bechky, B. A., & Davis, G. F. (2025). Resisting the algorithmic management of science: craft and community after generative AI. *Administrative Science Quarterly, 70*(1), 1–22. https://doi.org/10.1177/00018392241304403

Berg, J. M. (2022). One-hit wonders versus hit makers: sustaining success in creative industries. *Administrative Science Quarterly, 67*(3), 630–673. https://doi.org/10.1177/00018392221083650

Berg, M., Seeber, B. K., & Collini, S. (2017). *The slow professor: Challenging the culture of speed in the academy (reprinted in paperback)*. University of Toronto Press.

Bianchini, S., Di Girolamo, V., Ravet, J., & Arranz, D. (2025). *Artificial intelligence in science: Promises or perils for creativity?* Publications Office. https://data.europa.eu/doi/10.2777/6693925.

Bianchini, S., Müller, M., & Pelletier, P. (2022). Artificial intelligence in science: An emerging general method of invention. *Research Policy, 51*(10), Article 104604. https://doi.org/10.1016/j.respol.2022.104604

Binz, M., Alaniz, S., Roskies, A., Aczel, B., Bergstrom, C. T., Allen, C., Schad, D., Wulff, D., West, J. D., Zhang, Q., Shiffrin, R. M., Gershman, S. J., Popov, V., Bender, E. M., Marelli, M., Botvinick, M. M., Akata, Z., & Schulz, E. (2025). How should the advancement of large language models affect the practice of science? *Proceedings of the National Academy of Sciences, 122*(5), Article e2401227121. https://doi.org/10.1073/pnas.2401227121

Bisbee, J., Clinton, J., Dorff, C., Kenkel, B., & Larson, J. (2023). *Synthetic Replacements for Human Survey Data? The Perils of Large Language Models*. SocArXiv. https://doi.org/10.31235/osf.io/5ecfa.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Cao, C., Arora, R., Cento, P., Manta, K., Farahani, E., Cecere, M., Selemon, A., Sang, J., Gong, L. X., Kloosterman, R., Jiang, S., Saleh, R., Margalik, D., Lin, J., Jomy, J., Xie, J., Chen, D., Gorla, J., Lee, S., & Bobrovitz, N. (2025). Automation of systematic reviews with large language models. *Health Informatics*. https://doi.org/10.1101/2025.06.13.25329541

Checco, A., Bracciale, L., Loreti, P., Pinfield, S., & Bianchi, G. (2021). AI-assisted peer review. *Humanities and Social Sciences Communications, 8*(1), 1–11. https://doi.org/10.1057/s41599-020-00703-8

Cornelissen, J., Höllerer, M. A., Boxenbaum, E., Faraj, S., & Gehman, J. (2024). Large language models and the future of organization theory. *Organization Theory, 5*(1). https://doi.org/10.1177/26317877241239056

Cronin, M. A., Stouten, J., & Van Knippenberg, D. (2021). The theory crisis in management research: solving the right problem. *Academy of Management Review, 46*(4), 667–683. https://doi.org/10.5465/amr.2019.0294

Cronin, M. A., Stouten, J., & Van Knippenberg, D. (2022). Why theory on "how theory fits together" benefits management scholarship. *Academy of Management Review, 47*(2), 333–337. https://doi.org/10.5465/amr.2021.0517

Cui, Z., Li, N., & Zhou, H. (2024). *Can AI Replace Human Subjects? A Large-Scale Replication of Psychological Experiments with LLMs* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2409.00128.

Davis, M. S. (1971). That's interesting!: Towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy of the Social Sciences, 1*(2), 309–344. https://doi.org/10.1177/004839317100100211

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the jagged technological frontier: field experimental evidence of the effects of AI on knowledge worker productivity and quality. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4573321

Doshi, A. R., Chai, S., & Troebinger, M. (2025). *How Experience Moderates the Impact of Generative AI Ideas on the Research Process*. SSRN. https://doi.org/10.2139/ssrn.5013086.

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances, 10*(28), Article eadn5290. https://doi.org/10.1126/sciadv.adn5290

Elsevier. (2025). *Generative AI policies for journals* [doshiDos]. https://www.elsevier.com/en-gb/about/policies-and-standards/generative-ai-policies-for-journals.

Eva, N., Howard, J. L., Liden, R. C., Morin, A. J. S., & Schwarz, G. (2024). An Inconvenient Truth: A Comprehensive Examination of the Added Value (or Lack Thereof) of Leadership Measures. *Journal of Management Studies*, joms.13156. https://doi.org/10.1111/joms.13156.

Fan, J., Sun, T., Liu, J., Zhao, T., Zhang, B., Chen, Z., Glorioso, M., & Hack, E. (2023). How well can an AI chatbot infer personality? Examining psychometric properties of machine-inferred personality scores. *The Journal of Applied Psychology, 108*(8), 1277–1299. https://doi.org/10.1037/apl0001082

Feuerriegel, S., Maarouf, A., Bär, D., Geissler, D., Schweisthal, J., Pröllochs, N., Robertson, C. E., Rathje, S., Hartmann, J., Mohammad, S. M., Netzer, O., Siegel, A. A., Plank, B., & Van Bavel, J. J. (2025). Using natural language processing to analyse text data in behavioural science. *Nature Reviews Psychology, 4*(2), 96–111. https://doi.org/10.1038/s44159-024-00392-z

Fischer, T., Hambrick, D. C., Sajons, G. B., & Van Quaquebeke, N. (2020). Beyond the ritualized use of questionnaires: Toward a science of actual behaviors and psychological states. *The Leadership Quarterly, 31*(4), Article 101449. https://doi.org/10.1016/j.leaqua.2023.101752

Fischer, T., Hambrick, D. C., Sajons, G. B., & Van Quaquebeke, N. (2023). Leadership science beyond questionnaires. *The Leadership Quarterly*. , Article 101752. https://doi.org/10.1016/j.leaqua.2023.101752

Fyffe, S., Lee, P., & Kaplan, S. (2024). "Transforming" personality scale development: illustrating the potential of state-of-the-art natural language processing. *Organizational Research Methods, 27*(2), 265–300. https://doi.org/10.1177/10944281231155771

Gallea, Q., Huber, M., & Apostolatos, K. (2025). How 'causal' AI can improve your decision-making. *Artificial Intelligence*. https://www.imd.org/ibyimd/artificial-intelligence/how-causal-ai-can-improve-your-decision-making/.

Gans, J. S. (2025). The efficient market hypothesis when time travel is possible. *Economics Letters, 248*, Article 112209. https://doi.org/10.1016/j.econlet.2025.112209

Gatrell, C., Muzio, D., Post, C., & Wickert, C. (2024). Here, there and everywhere: on the responsible use of Artificial Intelligence (AI) in management research and the peer-review process. *Journal of Management Studies, 61*(3), 739–751. https://doi.org/10.1111/joms.13045

Gerlich, M. (2025). AI tools in society: impacts on cognitive offloading and the future of critical thinking. *Societies, 15*(1), 6. https://doi.org/10.3390/soc15010006

Gibney, E. (2025). AI tools are spotting errors in research papers: Inside a growing movement. *Nature*. , Article d41586–025-00648–5. https://doi.org/10.1038/d41586-025-00648-5

Götz, F. M., Maertens, R., Loomba, S., & Van Der Linden, S. (2024). Let the algorithm speak: How to use neural networks for automatic item generation in psychological scale development. *Psychological Methods, 29*(3), 494–518. https://doi.org/10.1037/met0000540

Graßmann, C., & Schermuly, C. C. (2021). Coaching with artificial intelligence: concepts and capabilities. *Human Resource Development Review, 20*(1), 106–126. https://doi.org/10.1177/1534484320982891

Grimes, M., Von Krogh, G., Feuerriegel, S., Rink, F., & Gruber, M. (2023). From scarcity to abundance: scholars and scholarship in an age of generative artificial intelligence. *Academy of Management Journal, 66*(6), 1617–1624. https://doi.org/10.5465/amj.2023.4006

Güntner, A. V., Meinecke, A. L., & Lüders, Z. E. K. (2023). Interaction coding in leadership research: A critical review and best-practice recommendations to measure behavior. *The Leadership Quarterly, 34*(6), Article 101751. https://doi.org/10.1016/j.leaqua.2023.101751

Hoffmann, M., Boysel, S., Nagle, F., Peng, S., & Xu, K. (2024a). *Generative AI and the nature of work. SSRN*. https://doi.org/10.2139/ssrn.5007084

Hoffmann, S., Lasarov, W., & Dwivedi, Y. K. (2024b). AI-empowered scale development: Testing the potential of ChatGPT. *Technological Forecasting and Social Change, 205*, Article 123488. https://doi.org/10.1016/j.techfore.2024.123488

Holzner, N., Maier, S., & Feuerriegel, S. (2025). *Generative AI and Creativity: A Systematic Literature Review and Meta-Analysis* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2505.17241.

Hosseini, M., Eve, M. P., Gordijn, B., & Neylon, C. (2020). MyCites: a proposal to mark and report inaccurate citations in scholarly publications. *Research Integrity and Peer Review, 5*(13), 1–9.

Jia, N., Luo, X., Fang, Z., & Liao, C. (2024). When and how artificial intelligence augments employee creativity. *Academy of Management Journal, 67*(1), 5–32. https://doi.org/10.5465/amj.2022.0426

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2506.08872.

Krafft, P. M., Macy, M., & Pentland, A. "Sandy." (2017). Bots as Virtual Confederates: Design and Ethics. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 183–190. https://doi.org/10.1145/2998181.2998354.

Kreps, S., McCain, R. M., & Brundage, M. (2022). All the news that's fit to fabricate: AI-generated text as a tool of media misinformation. *Journal of Experimental Political Science, 9*(1), 104–117. https://doi.org/10.1017/XPS.2020.37

Kulkarni, M., Mantere, S., Vaara, E., Van Den Broek, E., Pachidi, S., Glaser, V. L., Gehman, J., Petriglieri, G., Lindebaum, D., Cameron, L. D., Rahman, H. A., Islam, G., & Greenwood, M. (2024). The future of research in an artificial intelligence-driven world. *Journal of Management Inquiry, 33*(3), 207–229. https://doi.org/10.1177/10564926231219622

Lahart, J. (2025, May 16). *MIT Says It No Longer Stands Behind Student's AI Research Paper*. WSJ. https://www.wsj.com/tech/ai/mit-says-it-no-longer-stands-behind-students-ai-research-paper-11434092.

Larson, B. Z., Moser, C., Caza, A., Muehlfeld, K., & Colombo, L. A. (2024). Critical thinking in the age of generative AI. *Academy of Management Learning & Education, 23*(3), 373–378. https://doi.org/10.5465/amle.2024.0338

Lee, B. C., & Chung, J. (2024). An empirical investigation of the impact of ChatGPT on creativity. *Nature Human Behaviour, 8*(10), 1906–1914. https://doi.org/10.1038/s41562-024-01953-1

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative ai on critical thinking: self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–22). https://doi.org/10.1145/3706598.3713778

Lehr, S. A., Caliskan, A., Liyanage, S., & Banaji, M. R. (2024). ChatGPT as research scientist: probing GPT's capabilities as a research librarian, research ethicist, data generator, and data predictor. *Proceedings of the National Academy of Sciences, 121*(35), Article e2404328121. https://doi.org/10.1073/pnas.2404328121

Letrud, K., & Hernes, S. (2019). Affirmative citation bias in scientific myth debunking: a three-in-one case study. *PLoS One, 14*(9), Article e0222213. https://doi.org/10.1371/journal.pone.0222213

Liang, W., Zhang, Y., Cao, H., Wang, B., Ding, D., Yang, X., Vodrahalli, K., He, S., Smith, D., Yin, Y., McFarland, D., & Zou, J. (2023). *Can large language models provide useful feedback on research papers? A large-scale empirical analysis* (No. arXiv:2310.01783). arXiv. https://doi.org/10.48550/arXiv.2310.01783.

Lippert, S., Dreber, A., Johannesson, M., Tierney, W., Cyrus-Lai, W., Uhlmann, E. L., Collaboration, E. E., & Pfeiffer, T. (2024). Can large language models help predict results from a complex behavioural science study? *Royal Society Open Science, 11*(9), Article 240682. https://doi.org/10.1098/rsos.240682

Lu, C., Lu, C., Lange, R. T., Foerster, J., Clune, J., & Ha, D. (2024). *The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery* (No. arXiv:2408.06292). arXiv. https://doi.org/10.48550/arXiv.2408.06292.

Malik, F. S., & Terzidis, O. (2025). A hybrid framework for creating artificial intelligence-augmented systematic literature reviews. *Management Review Quarterly*. https://doi.org/10.1007/s11301-025-00522-8

Mei, Q., Xie, Y., Yuan, W., & Jackson, M. O. (2024). A Turing test of whether AI chatbots are behaviorally similar to humans. *Proceedings of the National Academy of Sciences, 121*(9), Article e2313925121. https://doi.org/10.1073/pnas.2313925121

Meinhart, W. A. (1966). Artificial intelligence, computer simulation of human cognitive and social processes, and management thought. *Academy of Management Journal, 9*(4), 294–307. https://doi.org/10.2307/254948

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature, 627*(8002), 49–58. https://doi.org/10.1038/s41586-024-07146-0

Musslick, S., Bartlett, L. K., Chandramouli, S. H., Dubova, M., Gobet, F., Griffiths, T. L., Hullman, J., King, R. D., Kutz, J. N., Lucas, C. G., Mahesh, S., Pestilli, F., Sloman, S. J., & Holmes, W. R. (2025). Automating the practice of science: Opportunities, challenges, and implications. *Proceedings of the National Academy of Sciences, 122*(5), Article e2401238121. https://doi.org/10.1073/pnas.2401238121

Naddaf, M. (2025). AI is transforming peer review—And many scientists are worried. *Nature, 639*(8056), 852–854. https://doi.org/10.1038/d41586-025-00894-7

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192. https://doi.org/10.1126/science.adh2586

O'Neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy (first paperback edition)*. B/D/W/Y Broadway Books.

Pataranutaporn, P., Powdthavee, N., Achiwaranguprok, C., & Maes, P. (2025). *Can AI Solve the Peer Review Crisis? A Large Scale Cross Model Experiment of LLMs' Performance and Biases in Evaluating over 1000 Economics Papers* (No. arXiv:2502.00070). arXiv. https://doi.org/10.48550/arXiv.2502.00070.

Paulus, P. B., & Kenworthy, J. B. (2019). Effective Brainstorming. In P. B. Paulus, & B. A. Nijstad (Eds.), *The Oxford Handbook of Group Creativity and Innovation* (1st ed., pp. 286–305). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190648077.013.17.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2016). Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. *Organizational Research Methods, 19*(2), 159–203. https://doi.org/10.1177/1094428115624965

Resnik, D. B., & Hosseini, M. (2025). The ethics of using artificial intelligence in scientific research: New guidance needed for a new tool. *AI and Ethics, 5*(2), 1499–1521. https://doi.org/10.1007/s43681-024-00493-8

Robert, S. J., Kadhiravan, S., & McKay, D. (2024). The development and validation of digital amnesia scale. *Current Psychology, 43*(22), 19594–19603. https://doi.org/10.1007/s12144-024-05757-4

Santurkar, S., Durmus, E., Ladhak, F., Lee, C., Liang, P., & Hashimoto, T. (2023). *Whose Opinions Do Language Models Reflect?* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2303.17548.

Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: a guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods, 19*(1), 80–110. https://doi.org/10.1177/1094428115598239

Si, C., Yang, D., & Hashimoto, T. (2024). *Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2409.04109.

Singh Chawla, D. (2024). Is ChatGPT corrupting peer review? Telltale words hint at AI use. *Nature, 628*(8008), 483–484. https://doi.org/10.1038/d41586-024-01051-2

Skarlinski, M., Nadolski, T., Braza, J., Storni, R., Caldas, M., Mitchener, L., Hinks, M., White, A., & Rodriques, S. (2025). Superintelligent AI Agents for Scientific Discovery. *FutureHouse*. https://www.futurehouse.org/research-announcements/launching-futurehouse-platform-ai-agents.

Speer, A. B., Perrotta, J., & Kordsmeyer, T. L. (2024). Taking it easy: off-the-shelf versus fine-tuned supervised modeling of performance appraisal text. *Organizational Research Methods*. https://doi.org/10.1177/10944281241271249

Stollberger, J., Anand, S., & Dick, P. (2025). Capturing a moving target: developing research on and with AI for human relations. *Human Relations*, 499–516. https://doi.org/10.1177/00187267251332075

Thompson, I., Koenig, N., Mracek, D. L., & Tonidandel, S. (2023). Deep learning in employee selection: evaluation of algorithms to automate the scoring of open-ended assessments. *Journal of Business and Psychology, 38*(3), 509–527. https://doi.org/10.1007/s10869-023-09874-y

Tully, S. M., Longoni, C., & Appel, G. (2025). Lower artificial intelligence literacy predicts greater AI receptivity. *Journal of Marketing*. https://doi.org/10.1177/00222429251314491

van Knippenberg, D., & Sitkin, S. B. (2013). A critical assessment of charismatic—transformational leadership research: Back to the drawing board? *The Academy of Management Annals, 7*(1), 1–60. https://doi.org/10.1080/19416520.2013.759433

Van Quaquebeke, N., & Gerpott, F. H. (2023). The now, new, and next of digital leadership: how artificial intelligence (AI) will take over and change leadership as we know it. *Journal of Leadership & Organizational Studies, 30*(3), 265–275. https://doi.org/10.1177/15480518231181731

Van Quaquebeke, N., & Gerpott, F. H. (2024). Artificial intelligence (AI) and workplace communication: promises, perils, and recommended policy. *Journal of Leadership & Organizational Studies*, 375–381. https://doi.org/10.1177/15480518241289644

Van Quaquebeke, N., Salem, M., van Dijke, M., & Wenzel, R. (2022). Conducting organizational survey and experimental research online: From convenient to ambitious in study designs, recruiting, and data quality. *Organizational Psychology Review*, 268–305. https://doi.org/10.1177/20413866221097571

Von Krogh, G., Roberson, Q., & Gruber, M. (2023). Recognizing and utilizing novel research opportunities with artificial intelligence. *Academy of Management Journal, 66*(2), 367–373. https://doi.org/10.5465/amj.2023.4002

Wang, P., Loignon, A. C., Shrestha, S., Banks, G. C., & Oswald, F. L. (2024). Advancing organizational science through synthetic data: a path to enhanced data sharing and collaboration. *Journal of Business and Psychology*. https://doi.org/10.1007/s10869-024-09997-w

Weidmann, B., Xu, Y., & Deming, D. (2025). *Measuring Human Leadership Skills with AI Agents* (No. w33662; p. w33662). National Bureau of Economic Research. https://doi.org/10.3386/w33662.

Wenzel, R., & Van Quaquebeke, N. (2018). The double-edged sword of big data in organizational and management research. *Organizational Research Methods, 21*(3), 548–591. https://doi.org/10.1177/1094428117718627

Yamada, Y., Lange, R. T., Lu, C., Hu, S., Lu, C., Foerster, J., Clune, J., & Ha, D. (2025). *The AI Scientist-v2: Workshop-Level Automated Scientific Discovery via Agentic Tree Search* (No. arXiv:2504.08066). arXiv. https://doi.org/10.48550/arXiv.2504.08066.