THE American Journal of Humanities and Social Sciences Research (THE AJHSSR) 2026

E-ISSN: 2581-8868

Volume-09, Issue-06, pp-22-41

www.theajhssr.com Crossref DOI: https://doi.org/10.56805/ajhssr

Research Paper Open Access

AI-Assisted Strategic Decision-Making: A Framework for Accountable Experimentation

[1]Dr Muhammad Emdadul Haque, [2]Dr Shreya Chakravarty

ABSTRACT

Purpose: This article offers a governance-focused approach to strategic decision-making with AI assistance in situations characterised by high strategic uncertainty. In contrast to the view of strategy that primarily relies on plans, forecasts, and causal reasoning, the analysis centers on how managers are able to convert strategic beliefs into evidence-based, revisable, and accountable commitments in an environment where generative artificial intelligence (GenAI) can generate analyses, scenarios, and recommendations within a few minutes. Design, methodology, and approach: Conceptual synthesis based on a 56-source literature review focusing on the theory-based approach to strategy, strategic experimentation, behavioral strategy, human-AI collaboration, and AI governance. The review is explicitly focused on concept development and is not meta-analytic. The sources have been selected based on their direct relevance to strategic problem formulation, causality, experimentation, delegation, verification, and/or accountability. Findings: The conceptual synthesis reveals five recurring trade-offs between: prediction and intervention; breadth of the option set and validity of the causal reasoning; speed and verification effort; augmentation and diffusion of responsibility; data exploitation and strategic exploration. In order to overcome those trade-offs, the paper proposes a five-stage Accountable Strategic Experimentation Cycle: (i) strategic problem formulation; (ii) generation of focal and rival hypotheses; (iii) assumptions, evidence and decision rights audit; (iv) bounded strategic experiment; (v) evaluation of evidence and updates of confidence/commitment. A Strategic Experiment Ability Matrix further distinguishes the roles of AI depending on the degree of causal uncertainty and the level of commitment reversibility. Practical implications: The proposed framework provides senior managers and board members with a useful tool for implementing AI-based approaches without equating fluent output with strategic evidence. It integrates AI use with commitments in stages, independent challenge, evidence threshold, stop/scale criteria, and human accountability. Originality/value: The value of the research is in developing a coherent process for AI-assisted strategic learning drawing on well-established mechanisms of causal strategy, experimentation, human-AI collaboration, and AI governance. It does not aim at introducing any new mechanism but rather outlines its contingent integration and governance before material commitment is involved.

KEYWORDS: AI-assisted strategic decision-making; accountable experimentation; generative artificial intelligence; theory-based view; human-AI collaboration; decision governance; uncertainty; accountability

1. INTRODUCTION

Actions preceeding full knowledge about future are at the very core of strategic management. Managers make resource allocations, target markets, redesign institutions and bet on technology relying on data which is incomplete, controversial and self-generated. Artificial intelligence changes the pace and scale of this activity. There is experimental evidence that Generative AI can improve the speed and output quality of structured professional tasks (Noy and Zhang, 2023; Brynjolfsson, Li, and Raymond, 2025) whereas discipline-specific studies show that large language models can generate and evaluate strategic options but demonstrate inconsistent performance in the presence of uncertainty, path-dependency and delayed feedback (Csaszar, Ketkar, and Kim, 2024; Allen and McDonald, 2026). Both of these capabilities make AI relevant for the work on strategy, however, they also introduce the new risk for management - the feeling of analytical completeness can come before a strategic statement is causally justified or empirically tested.

Current literature on strategy has advanced the idea that strategy is not just a set of goals or collection of standard frameworks. Theory-based view states that managers develop problem-beliefs, causation-beliefs and solutions-beliefs specific to their firm (Felin and Zenger, 2017). Correspondingly, there is literature emphasising the need of strategic statements' coherence and transparency of its underlying assumptions (Sørensen and Carroll, 2021). Lazzarini, Zenger, and Felin (2026) extend this view developing a concept of strategy as a theory of change whose

causal structure can be mapped and analysed. This contribution is important as it transforms strategic thinking from simple description to explicit causal hypotheses.

The current article represents a continuation of this academic track but with another perspective. It avoids the discussion of the structure of causal map and its properties such as complexity or circularity. The article focuses on what happens after managers and AI systems develop a strategic theory. In the modern business environment, the practical question is increasingly related to managed translation of the strategic idea - how to transform an interesting narrative into an evidence-based experiment and then into an organisational commitment, while keeping control over the underlying assumptions and consequences. This question becomes important since GenAI can make weak strategic ideas compelling, generate alternative ideas in faster pace than managerial investigation and bring premature convergence around unverifiable outputs. The recent studies on agentic AI design and strategic AI lock-in make the difference between productive search and accountable commitment absolutely necessary (Camuffo et al., 2026; Neshenko and Ryall, 2026).

Literature on the interaction of artificial intelligence and strategy demonstrates a clear divergence of views on the border between machine contribution and human judgement. Csaszar et al. (2026) distinguish the strategic task capability of the AI and organisational willingness to delegate. This controversy also exists in Strategy Science as there is much disagreement on whether AI is able to perform high-level strategising or only to be a supporting technology (Chatterji et al., 2026). Felin and Holweg (2024) and Sako and Felin (2025) emphasise the role of forward-looking causal reasoning and experimentation. Human-AI research shows that good augmentation depends on task structure, metaknowledge, expertise and the way how human and machine search are integrated (Fügener et al., 2022; Raisch and Fomina, 2025). All of these literatures create a fruitful opportunity to build a unified framework focused on strategic testing rather than general AI capability.

In this way, the article addresses three research questions. RQ1 is about the failure modes arising when GenAI is used for generating strategic theories under high strategic uncertainty. RQ2 is about the process of converting an AI-assisted strategic hypothesis into the credible evidence, while keeping human accountability. RQ3 is about how the use of AI should depend on uncertainty of the underlying causal mechanism and reversibility of the strategic commitment. The dichotomy between capability and delegation in the current literature on strategy and commitment-flexibility in real options theory suggest that these contingencies are important (Csaszar et al., 2026; Trigeorgis and Reuer, 2017). Answering these questions, the paper develops Accountable Strategic Experimentation Cycle and Strategic Experimentability Matrix. Together, these constructs define strategy as an accountable learning system, where theories guide tests and tests discipline resource commitments.

The contribution is fourfold. First, the paper connects the theory-based view with the literature on disciplined experimentation showing that the value of a strategic theory is realised through designing the disconfirmable test (Andries, Debackere and van Looy, 2013; Camuffo et al., 2024). Second, it extends the human-AI collaboration literature into strategic context distinguishing analytical assistance from authority to commit organisation's resources (Berente et al., 2021; Raisch and Fomina, 2025). Third, it integrates the governance principles like provenance, monitoring and human oversight into the strategy process rather than treat them as downstream compliance. Fourth, it provides propositions and criteria for implementation. The contribution is integrative as it doesn't claim to invent problem framing, competing hypotheses, experimentation, staged commitment or human oversight as separate mechanisms.

Figure 1 summarises the conceptual departure from causal strategy representation towards a governed learning process in which strategic claims must generate evidence before larger commitments are made.

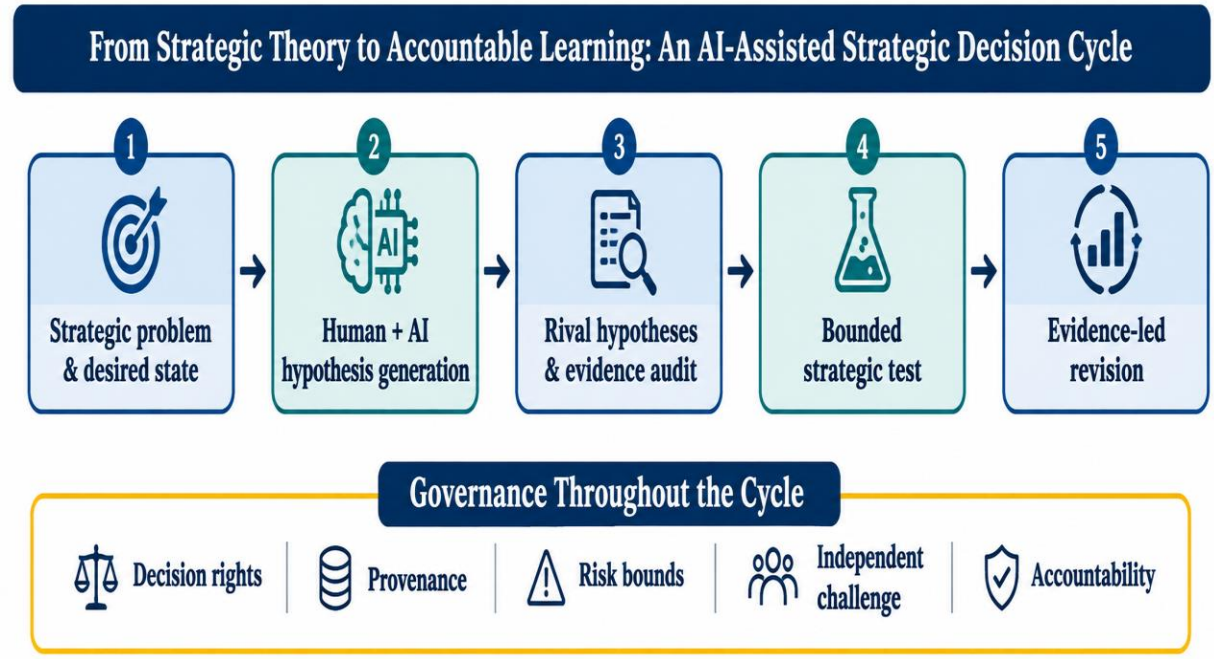


Figure 1. Conceptual departure of this article: from causal strategy representation to an accountable learning process. Author-developed conceptual synthesis based on the reviewed literature.

2. CONCEPTUAL POSITIONING AND RESEARCH GAP

2.1 From strategic representation to strategic verification

The interaction between cognition and actions has been an ongoing issue in the study of strategy. Behavioral strategy focuses on how managerial decision making is affected by bounded rationality, heuristics, biases, and social processes (Powell, Lovallo, and Fox, 2011; Gavetti, 2012). Another line of work takes problem framing as an essential strategic activity: the way the problem is formulated restricts the set of possible solution and knowledge needed (Nickerson and Zenger, 2004). The theory-based approach extends this perspective by viewing managers as theorists creating their own theory about sources of value (Felin and Zenger, 2017).

Yet a theory is not made strategically relevant just because it is coherent internally. Strategic environments are dynamic and interactive; competing firms, customers, government agencies, and technologies can react to whatever intervention you consider. The question, then, is not whether a story can be told, but rather whether the series of decisions will yield a pattern of data that changes your confidence in the story. In other words, the attention switches from representation to verification. Similarily, Sorenson (2024) argues that theory affects search and learning in that it specifies the places to look and the combinations to test. Additionally, the real-options literature tells us that uncertainty may make staged, reversible investments valuable in order to preserve learning before commitment (Trigeorgis and Reuer, 2017). This paper interprets all this discussion in terms of a governance problem, which says that the strategic theory should not only dictate what managers should believe, but also what evidence will change their mind.

2.2 Why GenAI changes the verification problem

Generative AI expands the range of analyses that can be done before a decision is made. It is able to create competitive scenarios, summarise market information, formulate assumptions, create synthetic customer profiles, draft business cases, and confront company narratives. Empirical strategy research has already shown that large language models are capable of generating and evaluating entrepreneurial strategies on par with human evaluators (Csaszar, Ketkar & Kim, 2024). Nevertheless, the ability does not equate to knowledge about whether a certain new approach will work in a given company. Felin and Holweg (2024) distinguish between prediction based on data and causation based on theory, while Sako and Felin (2025) stress that many consequential decisions involve actions that generate new data instead of analysing the existing data more deeply. Sen, Workiewicz, and Puranam (2026) also explain why analogical reasoning is fallible in case of inappropriate analogies chosen for addressing the strategic question at hand.

The risk associated with strategy-making therefore becomes asymmetric. While generative AI makes the cost of creating a compelling recommendation dramatically cheaper, the cost of validating the causal mechanisms behind it stays the same. If the output is fluent, detailed, and apparently balanced, managers may confuse the quality of articulation with the quality of the evidence provided. At the same time, algorithm aversion makes managers ignore valuable machine recommendations after observing errors while algorithm appreciation may make some users overvalue algorithmic recommendations (Dietvorst, Simmons & Massey, 2015; Logg, Minson & Moore, 2019). Finally, delegation experiments prove that lack of sufficient metaknowledge hinders people from delegating tasks effectively even if they want to delegate to AI (Fügener et al., 2022).

2.3 How the Present Study Extends Causal Strategy Research

Table 1 clarifies the distinction between representing a causal strategy and governing the experimentation required to test and revise it.

Table 1. Conceptual distinction between causal strategy representation and accountable experimentation.

Dimension	Strategy as theory of change	Present study: accountable experimentation
Primary question	How can a strategic theory specify causal paths from action to outcomes?	How should an AI-assisted strategic theory be challenged, tested and governed before resource commitment?
Unit of analysis	Structure of the strategic theory or causal model.	Decision process linking hypothesis, experiment, evidence and commitment.
Central risk	Poorly specified, complex or difficult-to-evaluate causal structure.	Plausible but unverified AI-assisted narratives, weak challenge and responsibility diffusion.
Core mechanism	Explicit causal representation and theory attributes.	Rival hypotheses, evidence thresholds, bounded experiments, staged decision rights and revision.
Main output	A clearer strategic theory of change.	An auditable learning cycle that calibrates confidence and commitment.

The difference is more than just skin deep. Lazzarini, Zenger, and Felin (2026) provide an intellectual beginning, with the assertion that strategy can be conceived as a set of causal theories. While this study agrees with the above statement, it pushes the discussion forward in terms of the governance of strategic learning. The key idea of the study is that within an artificial intelligence abundant environment, the advantage of theorising depends on whether an organisation has the ability to create good tests, tolerate dissent, articulate assumptions, and change commitments in light of new information.

3. METHODOLOGY: PROBLEM-DRIVEN CONCEPTUAL SYNTHESIS

For this study, a problem-driven conceptual synthesis is chosen in place of a systematic review or meta-analysis. Such a methodology is suitable since the research problem involves partially separate literatures using different constructs and approaches: strategic cognition and theory, entrepreneurial experimentation, behavioral decision making, human-AI cooperation, and AI governance. In order to build a new framework of explanations and prescriptions, the aim is to identify the mechanisms that appear repeatedly in these literatures, rather than estimate the pooled effect size.

Evidence collection was carried out in October and on 1 November 2026 through a targeted search on the publishers' and repository websites: INFORMS, Academy of Management, Wiley, Springer, SAGE, Elsevier, Oxford Academic, AAAS/Science, MIS Quarterly/AIS, and institutional repositories; in addition, official AI governance documents from NIST, ISO, and EUR-Lex were used. The search strings were comprised of keywords like "theory-based view" and strategy; "strategic decision" and artificial intelligence; "generative AI" and strategy; experimentation and strategic decision making; human-AI cooperation; algorithm aversion; causal reasoning; AI governance; human oversight; decision rights. Backward and forward citation analysis was used to identify the seminal works. Bibliographic information and DOIs were cross-verified with the official records before inclusion.

Only those documents have been selected that could contribute to the achievement of any of the following analytical tasks: (1) problem formulation and causal strategic reasoning; (2) disciplined experimentation and learning; (3) human-AI task division, cooperation, or decision-making process; and (4) governance, monitoring, and accountability. Recent evidence for the years of 2021 to 2026 was preferred for AI topics, but older studies have been included where they offered the durable theory or well-substantiated behavioral evidence. The selected corpus includes 56 sources: 49 journal articles and scholarly journals, three scholarly or management books, and four authoritative AI governance or standard-setting sources. The corpus is intended to be selective and concept-building, and not an exhaustive census of AI-and-strategy research.

The analysis of the literature was performed in three stages. At first, the roles assigned to humans, AI, and the organisational process in each source in terms of strategic function were coded. Secondly, the tensions arising in these literatures were compared with respect to the situations where the AI advantage in one aspect has created vulnerabilities in another; contradictory evidence concerning AI strengths, weaknesses, algorithm aversion, algorithm appreciation, and opacity was collected rather than discarded. Thirdly, the mechanisms were recombined into the process requirements and put together into five stages of the Accountable Strategic Experimentation Cycle. The Strategic Experimentability Matrix was designed subsequently via the questioning of the situations where strategic testing is possible and commitment is hard to undo. The developed architecture was conceptually validated with regard to the literature on automation, augmentation, delegation, and governance, rather than considered as empirical validation. Table 2 summarises the four evidence streams used to construct and challenge the conceptual architecture.

Table 2. Evidence streams used in the conceptual synthesis.

Evidence stream	Representative sources	Role in framework development
Strategic theory and problem framing	Nickerson and Zenger (2004); Felin and Zenger (2017); Sorenson (2024); Lazzarini et al. (2026)	Defines strategy as problem formulation, causal belief and coherent argument rather than a static plan.
Experimentation and learning	Andries et al. (2013); Camuffo et al. (2020, 2024); Agarwal et al. (2025); Trigeorgis and Reuer (2017)	Shows why hypotheses, bounded tests, flexibility and disciplined learning matter under uncertainty.
Human-AI decision making	Shrestha et al. (2019); Fügener et al. (2022); Raisch and Fomina (2025); Hillebrand et al. (2025); Li and Tian (2026)	Identifies complementarity, delegation, opacity, metaknowledge, trust and organisation-design contingencies.
AI and strategic reasoning	Csaszar, Ketkar and Kim (2024); Kanis et al. (2026); Allen and McDonald (2026); Sen et al. (2026); Camuffo et al. (2026); Neshenko and Ryall (2026)	Provides recent evidence on AI strategy generation, foresight, analogical reasoning, benchmarking, agentic design and strategic lock-in.
Governance and assurance	Berente et al. (2021); Tabassi (2023); Autio et al. (2024); ISO/IEC (2023); European Parliament and Council (2024)	Adds monitoring, documentation, risk proportionality, human oversight and accountability requirements.

3.1 Methodological boundaries

This synthesis is limited by four factors. The first limitation is that the fast pace of the technological development of AI makes performance claims obsolete because, according to recent benchmarking results, the model's performance varies significantly across different strategic uses (Allen and McDonald, 2026; Camuffo et al., 2026). As a result, the framework incorporates the use of the governance concepts which can retain their value as models evolve. The second limiting factor is that some of the strategy papers are either purely theoretical or contain debates. Therefore, the evidence presented here does not mean that a specific governance design causes the improvement in performance of the company. The third limitation is related to experimentation since some of the strategic decisions are illegal, ethical, or irreversible. Finally, the proposed framework operates on the organisational level and does not substitute other industry-specific measures aimed at mitigating risks.

3.2 Conceptual provenance and contribution boundary

The outlined cycle is a configuration. Problem formulation, competing theories, experiments, gradual commitment, provenance, human supervision, and decision authority all have antecedents in the literature on strategy, organization, and governance (Nickerson and Zenger, 2004; Camuffo et al., 2024; Trigeorgis and Reuer, 2017; Berente et al., 2021; Tabassi, 2023). As such, the contribution is configurational - it is about how these components can be configured and scaled to achieve AI-assisted strategic decision-making, linking the governance burden with causal ambiguity and reversibility. This paper does not claim novelty for the individual mechanisms and/or empirical support for the cycle as a whole. It provides an empirical framework that could be tested, invalidated, and improved in further studies.

4. THEORETICAL FOUNDATIONS

4.1 Strategy begins with problem framing, not answer generation

A recurring insight across strategy research is that the quality of the problem definition constrains the quality of the For Nickerson and Zenger (2004), problem framing is the unit of analysis when searching for knowledge, while, for Felin and Zenger (2017), it is the ability of economic agents to generate new questions and theories of value. The experimental research of AI-assisted strategic foresight suggests that, when there is time pressure, the quality of the mental representation is important (Kanis, Mann, and Stumpf-Wollersheim, 2026). In the case of GenAI, a problem framing is especially important, as an AI system is capable of giving a great answer to a badly-formulated strategic question. If a company asks “How to cut cost by 15%?”, it is likely to receive many useful recommendations without asking whether the cost cutting is a right strategy in the first place, does not interfere with the company's differentiating capability or the value proposition has to be redefined.

At the same time, problem framing sets the boundary of evidence. Strategic questions must state the outcome of interest, relevant stakeholders, the time horizon, the restrictions, and the mechanisms through which actions influence the value creation. It eliminates one of the main errors of using AI for strategic analysis, namely, the uncontrolled scope expansion, in which the model suggests many factors but cannot prioritise them. According to recent research on AI-augmented strategic decision-making, such process is characterised by changes in search, representation, and aggregation, which highlights the importance of framing before the evaluation (Csaszar, Ketkar, and Kim, 2024). As such, in the suggested cycle, the first step is the establishment of a decision statement and the outcome criteria.

4.2 Strategic theories create value when they direct learning

What makes a strategic theory valuable is its ability to influence the organisation’s perceptions, test and resource allocation. From a theory-based perspective, idiosyncratic beliefs are considered a source of diversity because different managers have unique perceptions of the environment and create value in a different way (Felin and Zenger, 2017). The idea is further developed by Lazzarini (2025) and Lazzarini, Zenger, and Felin (2026) and applied to explicit causal mechanisms. In addition, the current article adds an additional criterion for evaluating a useful theory: it must be discriminative. Testing two competing theories that predict the same outcome has little value from the learning point of view. In turn, a theory that cannot provide criteria for counterevidence is likely to survive through post-hoc explanations.

The reasoning is very consistent with scientific entrepreneurship research. For example, Camuffo et al. (2020) show how training entrepreneurs to develop predictions and test hypotheses could improve decision-making precision. The wider replication effect was found by Camuffo et al. (2024) who revealed the impacts on idea termination and pivoting of 759 firms. Agarwal et al. (2025) provide an experimental proof that developing a theory of value improves outcomes of Tanzanian entrepreneurs. Finally, earlier longitudinal research shows how experimentation can preserve variety and learning under uncertain conditions (Andries, Debackere, and van Looy, 2013). It does not mean that large corporations should follow the science lab in their strategies, but there is some evidence that discipline in doubts and tests can improve search under uncertainty.

4.3 AI changes the economics of search but not the logic of accountability

AI lowers the cost involved with scanning expansive conceptual spaces. It is able to produce several configurations of business models, competitive reactions to these models, and scenarios, as well as criticising managerial beliefs. This is especially useful during exploratory stages, when companies need variety before choosing. March's (1991) distinction of exploration and exploitation still holds: companies that specialise in their routine activities excessively will become efficient but brittle strategically. GenAI can complement individual creativity; however, there is evidence that AI-supported creative outputs tend to converge as a collective (Doshi and Hauser, 2024). Therefore, whether AI increases the strategic variety or not depends on how AI is stimulated, challenged, and controlled.

Scholarly literature on human-AI collaboration warns against the understanding of augmentation as a division of labor between machine analysis and human decision-making. Previous research on human-AI symbiosis and hybrid intelligence treated performance as a characteristic of the coupled system rather than a transfer from one side to another (Jarrahi, 2018; Dellermann et al., 2019a). As shown by Shrestha, Ben-Menahem, and von Krogh (2019), decision structures depend on a number of things, including search-space specificity, interpretability, speed, and replicability, whereas Raisch and Fomina (2025) differentiate alternative types of hybrid searches. AI usage changes the sources of competitive advantage through the creation of new capabilities of the human-AI interface, and organisational design literature stresses principled approaches to coupling human and AI (Krakowski, Luger, and Raisch, 2023; Kolbjørnsrud, 2024). As far as strategy is concerned, accountability can be displaced gradually even when the formal responsibility remains with people.

Therefore, the problem of opacity is both organisational and technological. Lebovitz, Lifshitz-Assaf, and Levina (2022) have demonstrated that professionals who are faced with opaque outputs from an AI system behave differently depending on whether they develop the practices to question the system. Related research shows that AI-supported decision-making allows for a more reflective behavior among users, when they engage with and resolve contradictions in machine advice (Abdel-Karim et al., 2023), and the research on hybrid decision-support design highlights careful allocation of human and machine capabilities (Dellermann et al., 2019b). In strategy, the analogous practices include tracing the origin of claims, separating the factual information from model-generated assumptions, testing different prompts, recording refuted hypotheses, and asking a named executive to justify the causal chain of reasoning without using the model's language.

4.4 Governance must enter before the decision, not after it

Governance systems nowadays give more weight to risk management throughout the life cycle. The NIST AI Risk Management Framework and the GenAI Profile focus on context, risk assessment, control management, and responsible governance throughout the life cycle of AI (Tabassi, 2023; Autio et al., 2024). The ISO/IEC 42001 standard views AI management as a complex organisational system requiring policy formulation, role assignment, monitoring, and improvement (ISO/IEC, 2023). Finally, the EU AI Act stipulates that a high-risk AI system should undergo human oversight and that there should be management of overreliance on AI results (European Parliament and Council, 2024).

Even though those sources do not present any theory of strategy, they have a lot to do with the concept of governance, which is very relevant. A strategic decision might produce great impact even if the tool itself is not defined as a high-risk AI system by the legislation. Actions like screening acquisitions, leaving the market, corporate reorganisation, price setting, redesigning of supply chain, and capital allocation are some examples of actions that will produce impact. Research on algorithmic control shows that technology may shift control over processes in the organisation and therefore turn the question of governance into the problem of organisational design (Kellogg, Valentine, and Christin, 2020; Berente et al., 2021).

5. FIVE TENSIONS IN AI-AUGMENTED STRATEGY

5.1 Prediction versus intervention

AI performs exceptionally well where past or synthetic data can be leveraged to identify patterns and predict outcomes. However, strategy involves redesigning the system whose outcome is to be predicted. For example, a pricing strategy will change how customers act; entry into the market will affect competitor's behavior; and a platform strategy will impact on complementor behavior. According to Sako & Felin (2025), prediction cannot be easily extrapolated to strategic decisions because intervening variables might alter state variables that do not have any relevant data available at all. Nonetheless, prediction could still offer a form of strategic foresight when done appropriately for the right context, as exemplified in recent studies in acquisitions (Qu, Kumar, and Tong,

2026). Implication in governance is not to shy away from prediction but to set the line between where historical associations become irrelevant to commitment.

5.2 Option breadth versus causal validity

GenAI can create a wide range of possibilities quickly. While exploring in the initial stage, this diversity comes in handy since the manager usually conducts a narrow search process and draws upon analogies. Nevertheless, the quantity of options presents an additional choice dilemma for the manager who lacks a good theory of value and is likely to pick one on the basis of fluency, managerial fashion, and confirmation of past preferences. Behavioral Strategy holds that cognitive and social biases affect the process of strategic choice (Powell, Lovallo, and Fox, 2011). The AI technology makes the individuals novel while at the same time constraining the group diversity and can be prone to making mistakes due to poor analogy even if it is quick (Doshi and Hauser, 2024; Sen, Workiewicz, and Puranam, 2026).

5.3 Speed versus verification burden

The productivity benefit from GenAI is one-sided: the creation of outputs is low-cost, while the process of validation can remain high-cost. Experiments confirm the productivity boost in structured knowledge jobs, but these gains lack strategic validity (Brynjolfsson, Li, and Raymond, 2025; Noy and Zhang, 2023). Strategic benchmarking shows systematic weaknesses of even the most advanced models when dealing with long-term strategic uncertainty (Allen and McDonald, 2026). Kögel, Meile, and Canos-Daros (2026) find systematic disparities in the AI and senior managers' answers to projects' decision-making situations. The solution, therefore, lies in limiting the number and significance of the claims made by AI. The suggested cycle uses the concept of evidence thresholds and materiality: it is only necessary to validate claims that impact the decision, while lower-priority support information is proportionally validated.

5.4 Augmentation versus responsibility diffusion

In this regard, human-in-the-loop systems can create a perceived accountability if the human's only task is approval. If the set of options is provided by AI, if AI chooses the relevant evidence and AI writes the justification, then formal approval by humans would not mean any substantial assessment. Research on AI governance and algorithmic control reveals that responsibility can be transferred due to different factors such as autonomy, opacity, and certain configurations of control (Berente et al., 2021; Kellogg, Valentine, and Christin, 2020). Raisch and Krakowski (2021) note that automation and augmentation processes are related, whereas Hillebrand, Raisch, and Schad (2025) show how AI changes management's agency and control. Therefore, at the audit step, decision rights are formally established, and there must be a decision owner who is a human and explains strategically significant points.

5.5 Data exploitation versus strategic exploration

AI algorithms built around historical patterns are necessarily good at capitalising on existing regularities. Nevertheless, the strategic thinking process often calls for a deviation from the norm. As emphasised by Felin and Holweg (2024), there is a difference between backward-looking predictive regularity and forward-looking causal theorising. In accordance with findings of hybrid problem-solving research, various combinations of human and AI-based searches lead to different results, whereas formal strategy models warn about the possibility of autonomous AI search converging on self-confirming but not optimal equilibriums (Raisch & Fomina, 2025; Neshenko & Ryall, 2026). Thus, the task of the manager is to use AI not only as a search accelerator but also as a tool to discover less obvious alternatives. This is done through the suggested cycle of hypothesis generation and selection which entails consideration of at least one alternative explanation or scenario.

6. ACCOUNTABLE STRATEGIC EXPERIMENTATION: FROM THEORY TO GOVERNED TESTING

The Accountable Strategic Experimentation Process cycle turns these five tensions into a repeatable strategy-process model. This model is designed not to burden individuals with too much bureaucracy regarding their decisions. Rather, it provides a scalable architecture whereby the process becomes lighter when decisions are reversible and have low stakes, and the process becomes more demanding as the degree of uncertainty, the degree of irrevocability, and exposure of stakeholders increases. This idea of contingency makes sense in light of the current distinction made between AI capabilities and delegation and the findings that different types of AI architectures lead to different forms of strategic thinking (Csaszar et al., 2026; Camuffo et al., 2026).



Figure 2. Accountable Strategic Experimentation Cycle for AI-assisted strategy. The arrows show a continuous sequence from framing through evidence-based updating and back to reframing. Author-developed conceptual synthesis based on the reviewed literature.

6.1 Frame the strategic problem

The procedure starts with clearly defining the strategic problem before posing any solutions. Managers define what decision needs to be made, what goal it should achieve, what time frame it has, what key limitations exist, and what are the consequences for the relevant stakeholders. Research on problem framing suggests that the way a problem is framed dictates the type of knowledge and searching that will be required to solve it (Nickerson & Zenger, 2004). Recent experiments have shown that artificial intelligence can affect strategic foresight through changing decision-makers' mental framing (Kanis, Mann, & Stumpf-Wollersheim, 2026). Thus, the framing statement should be narrow enough to delimit the search but broad enough to allow reframing. At this point, artificial intelligence functions well as a framing critic but not as a framer, since the task of framing belongs to managers.

6.2 Develop focal and rival hypotheses

Having delineated the problem space, managers develop a hypothesis as well as at least one alternative hypothesis. Specifically, the central hypothesis is expected to identify the relationship between the intervention and certain intermediary mechanisms and outcomes. The alternative hypothesis should be able to explain the problem in a different way and/or predict different outcomes in reaction to the planned intervention. Such a process helps to minimise the risk of confirmation bias and a conversational AI's inclination to stick to the assumptions stated in the prompt. While theoretical considerations may help in focusing on relevant scenarios, recent findings on LLM-driven analogical reasoning suggest that candidate scenarios still need to be aligned with the core problem (Sorenson, 2024; Sen, Workiewicz, and Puranam, 2026). The key objective is not to generate the maximum number of scenarios, but to derive a short list of distinct explanations.

6.3 Audit assumptions, evidence and decision rights

The audit stage transforms a persuasive case into a decision object that can be audited and held accountable. Each substantive assertion in the case is classified into one of five categories: a verified fact, evidence from an outside source, an estimation derived from the modeling process, a managerial assumption, and a normative judgment. The origin of each verifiable claim that is not made by the manager is recorded. The managers specify which of the assumptions are invalid in such a way that it would negate the case. They also determine the threshold of evidence required before the scaling of the case. In AI in organisations literature, autonomy, learning, and inscrutability are also managerial problems, requiring control decisions (Berente et al., 2021). Risk management frameworks such as those of NIST, ISO/IEC 42001, and EU AI Act also emphasise that control should be well-documented and role-based (Tabassi, 2023; ISO/IEC, 2023; European Parliament and Council, 2024).

6.4 Run bounded strategic tests

A strategic experiment is a limited experiment carried out to reduce uncertainty before making a major commitment by an organisation. Such an experiment may take the form of a prototype, geographical pilot, shadow process, split testing, staged contract, reversible price testing, limited partnership, or operational simulation. The design of a strategic experiment should clearly articulate the hypothesis, unit of analysis, lead and lag indicators, length, guardrails, stopping rules, and decision rule. Experiences drawn from experiments on business models, randomised entrepreneurial interventions, and research on real options confirm the value of staged, bounded commitments for learning amid uncertainty (Andries, Debackere & Van Looy, 2013; Camuffo et al., 2020, 2024; Trigeorgis & Reuer, 2017). In cases where randomisation is possible, it may make for stronger causal inferences; nonetheless, the proposed process does not have to involve RCTs for all strategic tests. The fundamental concept is disciplined exposure, whereby organisations are able to generate enough variation to learn but limit downside risks. Thomke (2020) also sees experimentation as a capability of organisations rather than an analytical technique.

6.5 Evaluate evidence and update commitment

In the final stage, the process is not about evaluating the project on the basis of its success or failure but rather the effect of the evidence on the level of belief in the causal theory and organisational commitment. Management compares the evidence collected with the predictions made by both the central and alternative theories and evaluates unintended consequences, distinguishing between execution failure and theory failure. Literature on organisational learning emphasises the need for lessons learned in terms of decision rules (Bingham and Eisenhardt, 2011; Sorenson, 2024), while theory-guided search highlights the significance of evidence that leads to shifting the direction of future inquiry. The outcome might involve stopping the project, scaling it, changing the design, obtaining more evidence, or reframing the problem. Whenever possible, the assessment should be based on predetermined criteria.Table 3 translates each stage of the cycle into a managerial question, an appropriate AI contribution and the evidence or control required before progression.

Table 3. Operational requirements for the five-stage Accountable Strategic Experimentation Cycle.

Stage	**Core managerial question**	**AI contribution**	**Required evidence / control**
Frame	What problem are we actually solving, for whom, over what horizon?	Critique framing; surface constraints and stakeholders.	Decision statement; outcomes; boundaries; named decision owner.
Develop hypotheses	What causal explanation links action to value, and what rival could also explain the problem?	Generate alternatives, counterfactuals and second-order effects.	Focal and rival hypotheses; discriminating predictions.
Audit	Which claims are facts, estimates or assumptions, and who is accountable?	Trace sources; flag inconsistencies; stress-test assumptions.	Provenance log; material assumptions; independent challenger; decision rights.
Run test	What is the smallest credible intervention that can change our confidence?	Design scenarios, metrics and monitoring support.	Bounded exposure; guardrails; stop/scale criteria; protected data.
Evaluate / update	What did we learn, and how should commitment change?	Compare results to predictions; identify anomalies and patterns.	Pre-specified review; explanation of deviation; decision record; revised causal theory.

7. THE STRATEGIC EXPERIMENTABILITY MATRIX

Application should not be done consistently with uniform rigor. Particularly insightful are two dimensions that allow one to judge the role of AI and the amount of evidence required: causal uncertainty and commitment reversibility. Current academic research on AI-and-strategy distinguishes between what the technology can do and what the organisation should delegate to it, while real options theory explains why flexibility becomes more valuable when there is uncertainty and irreversibility involved (Csaszar et al., 2026; Trigeorgis and Reuer, 2017). While causal uncertainty deals with certainty of cause-effect relationship in question, reversibility relates to the

possibility and cost of uncommitting after receiving additional evidence. These two dimensions define four decision-making environments.

First quadrants consist of situations where causal uncertainty is low, and reversibility is high. Examples are bounded price experiments, media allocations under constraints or operational tuning in general. AI can be trusted with greater degree of analysis and execution in such situations because feedback loop is fast, and risks are low. Research on the use of prediction by machines in acquisition shows the benefits of machine learning in providing foresight, especially when outcome is measurable (Qu, Kumar and Tong, 2026). Second quadrants relate to low causal uncertainty and low reversibility - for instance, long-term contracts or big system migrations on solid technologies. In this case, AI can assist in analysis, but independent verification is still necessary as the presence of a known mechanism does not preclude commitment risk.

Third quadrant consists of cases where causal uncertainty is high, and reversibility is also high. It represents the inherent territory of experimental explorations - novel customer offerings, prototyping of processes, market trial experiments, and so on. Machine learning allows to expand the search field in this quadrant, while humans need to define problems and conduct rapid tests to get the missing evidence. Special agentic machines might help in strategic thinking when architecture is aligned with user and task, but current state of strategic benchmarks still shows significant vulnerabilities in uncertain environment (Camuffo et al., 2026; Allen and McDonald, 2026). Finally, fourth quadrant consists of situations where both causal uncertainty and commitment irreversibility are high. Examples here include major acquisitions, exiting markets, reputation-related restructuring, platform bets and similar actions. Modeling of AI-driven strategy warns against self-reinforcing biases in learning systems, further supporting the need for independent testing and human decision-making (Neshenko and Ryall, 2026).



Figure 3. Strategic Experimentability Matrix: matching AI role and governance response to causal uncertainty and commitment reversibility. Author-developed conceptual synthesis based on the reviewed literature.

Table 4. Recommended AI and human roles by strategic decision environment.

Environment	Typical AI role	Human role	Governance emphasis
Bounded optimisation	Recommend or execute only within pre-approved, reversible limits.	Set objectives, constraints and exception rules.	Performance monitoring, data quality, rollback capability.
Validated commitment	Synthesise evidence, scenarios and implementation risks.	Validate mechanism and authorise commitment.	Independent validation, sensitivity testing, staged commitment and documented sign-off.
Exploratory testing	Generate options, counterfactuals and experiment variants.	Frame novelty, choose tests and interpret ambiguous evidence.	Learning metrics, rapid stop rules, preserved dissent.
Strategic deliberation	Support synthesis and red-teaming; no autonomous resource commitment.	Own causal theory, stakeholder trade-offs and final authority.	Multiple hypotheses, provenance, board challenge, staged commitment, accountability record.

8. RESEARCH PROPOSITIONS

Six propositions for empirical evaluation emerge from the conceptual synthesis of research streams. These propositions have been deliberately constructed in such a way as to make operationalisation possible without reference to any particular kind of AI system or vendor.

Proposition 1 - Option-generation boundary. In regard to strategically novel issues, GenAI-supported option generation will lead to a wider set of alternatives to be considered; but the positive effect on the quality of decision-making will emerge only if managers formulate alternative causal explanations of strategic problems and criteria for discriminating evidence beforehand. Proposition 1 is based on evidence about the ability of AI systems to widen strategy generation and risks of convergence in generating content (Csaszar, Ketkar, and Kim, 2024; Doshi and Hauser, 2024).

Proposition 2 - Reversibility contingency. The benefit of more delegated responsibility to AI will be stronger when the strategic commitment is reversible and the outcomes are measurable and weaker when the commitment is difficult to be reversed and causal feedback is rare. Proposition 2 follows the distinction between AI capabilities and delegation as well as the importance of flexibility in uncertain circumstances (Csaszar et al., 2026; Trigeorgis and Reuer, 2017).

Proposition 3 - Auditability and calibrated trust. The separation of evidence, assumptions, model-based estimates, and managerial judgment is expected to mitigate automation bias and algorithm aversion as a result of improving the calibration of trust in AI-supported strategic decision-making. Prior studies showed both the phenomenon of algorithm aversion and appreciation; while the research on delegation demonstrated the effect of metaknowledge on productive collaboration with AI (Dietvorst, Simmons, and Massey, 2015; Logg, Minson, and Moore, 2019; Fünger et al., 2022).

Proposition 4 - Experimentation as a mediator. The link between the quality of strategic theory formation and subsequent resource allocation will be mediated by the organisation's capacity to design bounded experiments which provide evidence that can confirm, discriminate between, or disconfirm rival hypotheses. Proposition 4 extends evidence that both simultaneous and hypothesis-driven experimentation increases learning under uncertainty (Andries, Debackere, and van Looy, 2013; Camuffo et al., 2024).

Proposition 5 - Decision-right separation. AI-enhanced strategy processes will result in higher accountability when hypothesis generation, critical analysis, and the decision commitment right are separated among different roles. Studies of decision structure in organisations and AI management indicate that task allocation and authority should be based on the degree of interpretability, autonomy, and risk (Shrestha, Ben-Menahem, and von Krogh, 2019; Berente et al., 2021).

Proposition 6 - Disconfirming review and adaptation. Organisations which use pre-established criteria for disconfirming evidence and stop/scale thresholds in evaluation of AI-assisted strategic initiatives adapt better than organisations which primarily justify initiatives through post-factum narratives. Proposition 6 is consistent with findings in process-learning and recent concerns regarding the risk of AI-driven strategy being stuck in self-reinforcing equilibria (Bingham and Eisenhardt, 2011; Neshenko and Ryall, 2026).

9. MANAGERIAL IMPLICATIONS

9.1 Treat AI output as a hypothesis object, not a decision object

Most important behavioral changes involve semantics. An improved version of the AI-generated recommendation should be referred to as a hypothesis package before the substance of the recommendations and the assumptions behind it are audited. This is relevant as it places the burden of proof on the managers. Rather than asking whether the output looks plausible, the managers are encouraged to think about what additional information will make them more or less confident in the advice, and what action could help gather such evidence. According to recent findings in the strategy literature, LLMs are capable of delivering reliable results for some generation and evaluation problems while having some difficulties with more sophisticated strategic simulation tasks (Csaszar, Ketkar, and Kim, 2024; Allen and McDonald, 2026).

9.2 Require visible reasoning at the point of commitment

In case of substantial decisions, the decision owner must be able to articulate the causal reasoning, assumptions, alternative explanations, downside limits, and evidence thresholds independently of the AI-generated narrative. There is no need to reveal internal weights of the models here. What is needed is organisational interpretability - a humanly defensible explanation of why the selected action is expected to work and what would cause the reconsideration of the decision. Interrogation routines prove to be crucial for experts who use opaque systems, and the ability of the users to disagree with the results of AI can motivate reflection (Lebovitz, Lifshitz-Assaf, and Levina, 2022; Abdel-Karim et al., 2023). This is true for strategic reasoning as well.

9.3 Separate exploration productivity from commitment authority

It is legitimate for AI to be more autonomous in exploration than in commitment. Multiple models or agents can be used by the teams for exploring evidence, developing scenarios and counterarguments, while keeping the ultimate control over the allocation of resources to humans. Both experimental research on delegation and hybrid problem solving theory prove that the performance depends on the allocation of tasks and search routines between humans and machines, and not on the AI usage per se (Fünger et al., 2022; Raisch and Fomina, 2025). This design helps to avoid the false dichotomy of 'AI decides vs AI is just a writing tool'. This produces the multi-layered design, in which machine autonomy is large in the analytical sandbox while organisational commitment stays commensurate with stakes and reversibility.

9.4 Build a strategic evidence ledger

The lightweight strategic evidence ledger can track each material hypothesis, sources, assumptions, experiments, outcomes and decisions changes. This allows tracing the results across strategy cycles and prevents organisational amnesia after turnover of the management teams or leadership. AI-management and risk governance literature highlights documentation, monitoring and accountability as the main organisational controls (Berente et al., 2021; Autio et al., 2024). The ledger allows distinguishing reasonable decisions made on the basis of the evidence available from badly managed decisions. It can be implemented within the existing systems of project, risk or knowledge management; what is required is the structured continuity rather than new software implementation.

10. IMPLEMENTATION AND GOVERNANCE DESIGN

Implementation is possible using three layers of governance instead of one company-wide process. Simple decisions cause little friction, whereas important and irreparable decisions require stronger challenge. This proportionality is consistent with strategic management literature, which distinguishes capability from delegation, and formal AI governance schemes, which adjust governance according to situation and risk (Csaszar et al., 2026; ISO/IEC, 2023). The following table provides an example of such a scheme.

Table 5. Proportionate governance levels for AI-assisted strategic decisions.

Level	Decision characteristics	Minimum controls	Illustrative examples
Level 1: bounded	Reversible, measurable, limited stakeholder impact.	Named owner; documented objective; automated monitoring; exception and rollback rules.	Campaign allocation, routine pricing test, internal workflow optimisation.
Level 2: material	Moderate commitment or stakeholder impact; material assumptions.	strategic evidence record; independent challenge; provenance check; pilot or staged commitment; executive sign-off.	New service launch, major supplier change, regional entry, AI-enabled process redesign.
Level 3: strategic	High novelty, low reversibility, major capital/reputation/workforce consequences.	Multiple rival hypotheses; board-level challenge; external evidence where appropriate; scenario stress test; staged options; explicit accountability and post-decision review.	Acquisition, major market exit, platform transformation, restructuring, long-duration technology bet.

The governance framework must tackle boundaries for data and tools. An external AI service may reveal confidential information in its own right, and generated content may have included unverifiable third-party content. NIST's GenAI profile identifies specific risks associated with generative systems, including confabulation and issues related to the integrity of information (Autio et al., 2024). Organisations should define allowed tools, data processing rules, retention requirements, and escalation procedures. These controls protect not only the confidentiality of the information but also the integrity of the decision process - if the analysts cannot replicate the production of a claim, strategic learning is jeopardised.

The second implementation challenge relates to competence. Human supervision implies competent people who can question the output generated by the AI. In case managers lack domain knowledge, causal reasoning skills or authority to overrule the output, the so-called "human in the loop" becomes formalistic rather than substantive. Research on ML-assisted causal inference warns about the necessity of human competence in the analysis of results even when machine learning improves analytical processes (Rathje, Katila, and Reineke, 2024). The same is stated in the EU AI Act in regard to high-risk systems (European Parliament and Council, 2024). This means that in addition to financial, marketing, operational, organisational and evidence assessment skills, strategy teams need to develop their skills in the use of AI technologies.

Finally, organisations should measure the quality of the process of strategy making and not just the performance of individual AI tools. Relevant indicators might be: the share of material decisions supported by a rival hypothesis; the share of AI-assisted claims for which the source is verifiable; the share of major commitments following a bounded test or staged option; the frequency with which contradictory evidence results in changing a decision; and the delay between experimental results and revisions of the strategy. Strategy-specific benchmarking and agentic-AI experiments show the usefulness of evaluation mechanisms compared to apparent fluency or confidence (Allen and McDonald, 2026; Camuffo et al., 2026).

11. DISCUSSION

11.1 Why accountable experimentation is a different strategic capability

The suggested framework redefines AI advantage. When the models are widely available, the competitive advantage cannot be based only on access to fluent analytical capability. More robust competitive advantages can arise because of an organisation’s ability to ask better questions, develop better hypotheses, generate its own evidence by actions, and make decisions fast without undermining judgment. From empirical strategy research, it follows that adoption of AI leads to the creation of new human-AI capabilities instead of substitution of existing human capabilities (Krakowski, Luger, and Raisch, 2023). Machine learning strategy research points out that automation does not substitute expertise (Rathje, Katila, and Reineke, 2024). It is consistent with the heterogeneity of beliefs idea (Felin and Zenger, 2017) from the theory-based approach and logic of dynamic capabilities (Teece, Pisano, and Shuen, 1997).

The suggested framework also helps to understand that "human judgment" is not about everything that AI still cannot do. The role of humans is narrower: creating an organisational mission, determining the relevant uncertainties, authorising interventions, making trade-offs among the interests of stakeholders, interpreting ambiguous evidence, and taking responsibility for the commitments. The research in hybrid problem-solving shows that different combinations of human and AI search result in different solution characteristics, and the contemporary literature review evidence emphasises the role of careful allocation of tasks, trust, and control in this area (Raisch and Fomina, 2025; Li and Tian, 2026). The AI can help humans to perform all these activities, but it does not substitute the need of governance to assign authority.

11.2 Productive disagreement as a design feature

The underestimated threat linked to GenAI is premature consensus. A conversational agent is programmed to respond coherently to the framing provided by the user and thus may be used to reinforce the prevailing narrative inside a person. Research findings show that the use of AI in assisting individuals can increase their creativity while making team output less diverse. This illustrates the necessity of independent challenge in team strategy work (Doshi and Hauser, 2024). The discussion about the potential of AI in doing strategy emphasises the role of disagreement, problem framing, and verification (Chatterji et al., 2026). The cycle suggested above institutionalises disagreement in order to have rival hypotheses and independent challenge. The goal here is not endless dispute, but the commitment made after the clash of competing narratives.

11.3 High strategic uncertainty cannot be reduced to more data

Additional data can help improve the estimates of the stable representation of the problem, but high strategic uncertainty often involves the representation of the problem itself: which mechanism is significant, how competitors will react, which customer behavior will be generated by an intervention, or what future state would be preferable. Pearl and Mackenzie (2018) define three types of reasoning - association, intervention, and counterfactual - involved in causal inference. Strategy-related research on using AI for foresight also considers the prediction of outcomes to be input for, not substitute of, strategic judgment (Qu, Kumar and Tong, 2026). That explains the focus on experiments and staged commitments in the suggested framework because, in many cases, the necessary data does not yet exist and managers need to create them.

12. LIMITATIONS AND FUTURE RESEARCH

As a conceptual paper, the article does not prove that the suggested cycle is validated as a better process of strategising. Literature review is a work of single author; it was non-exhaustive and purposive; no analysis of inter-rater reliability, PRISMA completeness or evidence weighting by quality is offered. Hence, the main goal of future research is testing comparative experiments. Field experiments can compare regular strategy sessions assisted by AI with strategy sessions based on the proposed cycle and measure the diversity of hypotheses generated, the quality of evidence, calibration of confidence, reversal of decisions and their performance outcomes. Recently, experimental and benchmarking studies offer good methodological examples for exploring strategic foresight and performance under AI assistance (Kanis, Mann, and Stumpf-Wollersheim, 2026; Allen and McDonald, 2026). Such experiments can manipulate the problem novelty and reversibility to explore the Strategic Experimentability Matrix.

First, future research should explore conditions under which competing hypotheses creation helps and when it causes unnecessary complexity. Strategic teams have limited time; too many options can cause delays of the actions. Researchers can investigate if one independent competing hypothesis is enough, if AI generated competitors are as helpful as human generated ones, and whether the quality of disagreement depends on domain expertise. There is work in progress on analogical reasoning and hybrid problem solving to explore such issues (Sen, Workiewicz, and Puranam, 2026; Raisch and Fomina, 2025).

Second, the framework assumes that experimenting is possible. Some decisions cannot be experimented ethically or economically. Thus, researchers need to find the substitutes evidentiary tools for the situations with low reversibility, such as natural experiments, digital twins, pre-mortems, prediction markets, expert elicitation, and real options structures. Main research question here is how much uncertainty can be reduced before the commitment in the cases where direct experimenting is impossible. Prior research on simultaneous experimenting and real options offers theoretical background for such comparison (Andries, Debackere, and van Looy, 2013; Trigeorgis and Reuer, 2017).

Third, AI systems will continue to develop. With the future advances in AI ability of making causal inferences, running autonomous experiments or performing agentic actions, some of the proposed boundaries can change.

However, increased capability might increase rather than decrease governance needs because the wider autonomy increases the possible errors. Current research on agentic AI and formal models of AI-based strategy suggests performance opportunities but also the risks of lock-in (Camuffo et al., 2026; Neshenko and Ryall, 2026). Thus, future longitudinal research should trace both the technical capabilities and delegations of the decisions to AI as separate variables.

Fourth, organisational politics need more investigation. Hypotheses are being examined in the organisational context, not in lab environment: business units compete for the resources, executives support initiatives and performance measures can be strategically chosen. There is already research suggesting that the technological changes reshape the control and power balance along with decision effectiveness (Kellogg, Valentine, and Christin, 2020; Hillebrand, Raisch, and Schad, 2025). Future research can expand the suggested cycle by modeling how power, incentives and identity affect the choice of hypotheses tested, credibility of evidence and time when wrong theories are abandoned.

13. CONCLUSION

The impact of Generative AI on the production of strategic analysis happens faster than the epistemological foundations of strategy are revised. Organisations can create scenarios, recommendations, and persuasive stories at an incredible speed and there is experimental evidence of high productivity gains for professional knowledge work (Noy and Zhang, 2023; Brynjolfsson, Li, and Raymond, 2025). Yet, organisations keep encountering the same old challenge of how to act in situations of uncertainty without a guarantee that one's assumed causal relationship is right. The danger here is that cheaper articulation might be taken for cheaper truth.

In this article, we address this issue by reconceptualising strategy as an accountable experimentation process. With the help of theory-based approach and the emerging understanding of strategy as a theory of change, we claim that a good strategic theory should build commitment through a well-governed challenge-evidence sequence (Felin and Zenger, 2017; Lazzarini, Zenger, and Felin, 2026). The Accountable Strategic Experimentation Cycle explicates the mentioned sequence through five steps: framing of the strategic problem, formulation of focal and rival theories, auditing of assumptions and decision rights, a bounded test, and calibration of confidence and commitment. The Strategic Experimentability Matrix further governs the use of AI depending on causal uncertainty and reversibility, and not based on the technologic enthusiasm, consistent with the recent evidence of importance of AI system design and problem fit for strategy (Camuffo et al., 2026).

The key principle in such an approach is clear and powerful - AI is supposed to facilitate strategic learning, but not avoid it. The value comes from the fact that machines help with exploration of options, interrogation of evidence, and monitoring, while accountable humans are responsible for framing, intervention, stakeholder choices, and irreversible commitments. Recent studies suggest that sustainable competitive advantage can increasingly be built by creating capabilities through human-AI interactions and design of hybrid problem solving, and not by having access to generic models (Krakowski, Luger, and Raisch, 2023; Raisch and Fomina, 2025). In the world where analysis appears to be competent, it might be strategic advantage coming from one's own experiments, evidence, and judgment.

REFERENCES

1. Abdel-Karim, B.M., Pfeuffer, N., Carl, K.V. and Hinz, O. (2023) 'How AI-based systems can induce reflections: the case of AI-augmented diagnostic work', MIS Quarterly, 47(4), pp. 1395-1424. https://doi.org/10.25300/MISQ/2022/16773.
2. Agarwal, R., Bacco, F., Camuffo, A., Coali, A., Gambardella, A., Msangi, H., Sonka, S., Temu, A., Waized, B. and Wormald, A. (2025) 'Does a theory-of-value add value? Evidence from a randomised control trial with Tanzanian entrepreneurs', Organization Science, 36(2), pp. 601-625. https://doi.org/10.1287/orsc.2023.17590.
3. Allen, R.T. and McDonald, R.M. (2026) 'How well can AI do strategy? Empirical benchmarking using strategy simulations', Strategy Science, 11(1), pp. 93-117. https://doi.org/10.1287/stsc.2025.0444.
4. Andries, P., Debackere, K. and van Looy, B. (2013) 'Simultaneous experimentation as a learning strategy: business model development under uncertainty', Strategic Entrepreneurship Journal, 7(4), pp. 288-310. https://doi.org/10.1002/sej.1170.
5. Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P. and Roberts, K. (2024) Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-

1. Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.600-1.
6. Berente, N., Gu, B., Recker, J. and Santhanam, R. (2021) 'Managing artificial intelligence', MIS Quarterly, 45(3), pp. 1433-1450. https://doi.org/10.25300/MISQ/2021/16274.
7. Bingham, C.B. and Eisenhardt, K.M. (2011) 'Rational heuristics: the simple rules that strategists learn from process experience', Strategic Management Journal, 32(13), pp. 1437-1464. https://doi.org/10.1002/smj.965.
8. Brynjolfsson, E., Li, D. and Raymond, L. (2025) 'Generative AI at work', Quarterly Journal of Economics, 140(2), pp. 889-942. https://doi.org/10.1093/qje/qjae044.
9. Camuffo, A., Cordova, A., Gambardella, A. and Spina, C. (2020) 'A scientific approach to entrepreneurial decision making: evidence from a randomized control trial', Management Science, 66(2), pp. 564-586. https://doi.org/10.1287/mnsc.2018.3249.
10. Camuffo, A., Gambardella, A., Messinese, D., Novelli, E., Paolucci, E. and Spina, C. (2024) 'A scientific approach to entrepreneurial decision-making: large-scale replication and extension', Strategic Management Journal, 45(6), pp. 1209-1237. https://doi.org/10.1002/smj.3580.
11. Camuffo, A., Gambardella, A., Kazemi, S. and Pandey, A. (2026) 'Beyond black boxes: designing and testing agentic AI systems for strategy', Strategy Science, 11(1), pp. 137-156. https://doi.org/10.1287/stsc.2025.0432.
12. Chatterji, A., Csaszar, F.A., Evans, J., Felin, T., Hullman, J., Lakhani, K.R., Sako, M. and Zenger, T. (2026) 'Can AI do strategy? A dialogue and debate', Strategy Science, 11(1), pp. 16-30. https://doi.org/10.1287/stsc.2026.ed.v11.n1.
13. Csaszar, F.A., Ketkar, H. and Kim, H. (2024) 'Artificial intelligence and strategic decision-making: evidence from entrepreneurs and investors', Strategy Science, 9(4), pp. 322-345. https://doi.org/10.1287/stsc.2024.0190.
14. Csaszar, F.A., Lee, G., Zemsky, P. and Zenger, T. (2026) 'Can AI do strategy?', Strategy Science, 11(1), pp. 1-15. https://doi.org/10.1287/stsc.2026.intro.v11.n1.
15. Dellermann, D., Ebel, P., Söllner, M. and Leimeister, J.M. (2019a) 'Hybrid intelligence', Business & Information Systems Engineering, 61(5), pp. 637-643. https://doi.org/10.1007/s12599-019-00595-2.
16. Dellermann, D., Lipusch, N., Ebel, P. and Leimeister, J.M. (2019b) 'Design principles for a hybrid intelligence decision support system for business model validation', Electronic Markets, 29, pp. 423-441. https://doi.org/10.1007/s12525-018-0309-2.
17. Dietvorst, B.J., Simmons, J.P. and Massey, C. (2015) 'Algorithm aversion: people erroneously avoid algorithms after seeing them err', Journal of Experimental Psychology: General, 144(1), pp. 114-126. https://doi.org/10.1037/xge0000033.
18. Doshi, A.R. and Hauser, O.P. (2024) 'Generative AI enhances individual creativity but reduces the collective diversity of novel content', Science Advances, 10(28), eadn5290. https://doi.org/10.1126/sciadv.adn5290.
19. European Parliament and Council (2024) Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), as amended by Regulation (EU) 2026/1744. Consolidated text current to 27 July 2026. ELI: http://data.europa.eu/eli/reg/2024/1689/2026-07-27 (Accessed: 1 November 2026).
20. Felin, T. and Holweg, M. (2024) 'Theory is all you need: AI, human cognition, and causal reasoning', Strategy Science, 9(4), pp. 346-371. https://doi.org/10.1287/stsc.2024.0189.
21. Felin, T. and Zenger, T.R. (2017) 'The theory-based view: economic actors as theorists', Strategy Science, 2(4), pp. 258-271. https://doi.org/10.1287/stsc.2017.0048.
22. Felin, T., Sako, M. and Hullman, J. (2026) 'AI and actor-specific decisions', AI & Society. https://doi.org/10.1007/s00146-026-03202-7.
23. Fügener, A., Grahl, J., Gupta, A. and Ketter, W. (2022) 'Cognitive challenges in human-artificial intelligence collaboration: investigating the path toward productive delegation', Information Systems Research, 33(2), pp. 678-696. https://doi.org/10.1287/isre.2021.1079.
24. Gavetti, G. (2012) 'Toward a behavioral theory of strategy', Organization Science, 23(1), pp. 267-285. https://doi.org/10.1287/orsc.1110.0644.
25. Hillebrand, L., Raisch, S. and Schad, J. (2025) 'Managing with artificial intelligence: an integrative framework', Academy of Management Annals, 19(1), pp. 343-375. https://doi.org/10.5465/annals.2022.0072.
26. ISO/IEC (2023) ISO/IEC 42001:2023 Information technology - Artificial intelligence - Management system. Geneva: International Organization for Standardisation and International Electrotechnical Commission. Available at: https://www.iso.org/standard/81230.html (Accessed: 1 November 2026).
27. Jarrahi, M.H. (2018) 'Artificial intelligence and the future of work: human-AI symbiosis in organisational decision making', Business Horizons, 61(4), pp. 577-586. https://doi.org/10.1016/j.bushor.2018.03.007.

28. Kanis, T., Mann, J.E. and Stumpf-Wollersheim, J. (2026) 'AI-augmented strategic decision-making under time constraints: an experimental study on mental representations and strategic foresight', Strategy Science, 11(1), pp. 75-92. https://doi.org/10.1287/stsc.2025.0442.
29. Kellogg, K.C., Valentine, M.A. and Christin, A. (2020) 'Algorithms at work: the new contested terrain of control', Academy of Management Annals, 14(1), pp. 366-410. https://doi.org/10.5465/annals.2018.0174.
30. Kögel, D., Meile, S. and Canos-Daros, L. (2026) 'Who's steering whom and in what direction? An experiment on AI versus top management decision-making in projects', Project Management Journal, 57(2), pp. 205-221. https://doi.org/10.1177/87569728251379338.
31. Kolbjørnsrud, V. (2024) 'Designing the intelligent organization: six principles for human-AI collaboration', California Management Review, 66(2), pp. 44-64. https://doi.org/10.1177/00081256231211020.
32. Krakowski, S., Luger, J. and Raisch, S. (2023) 'Artificial intelligence and the changing sources of competitive advantage', Strategic Management Journal, 44(6), pp. 1425-1452. https://doi.org/10.1002/smj.3387.
33. Lazzarini, S.G. (2025) 'Theory-based view in strategy: recent contributions and opportunities for development', Revista de Administração de Empresas, 65(5), e2025-0464. https://doi.org/10.1590/S0034-759020250507.
34. Lazzarini, S.G., Zenger, T.R. and Felin, T. (2026) 'Strategy as a theory of change', Strategic Management Journal, advance online publication. https://doi.org/10.1002/smj.70139.
35. Lebovitz, S., Lifshitz-Assaf, H. and Levina, N. (2022) 'To engage or not to engage with AI for critical judgments: how professionals deal with opacity when using AI for medical diagnosis', Organization Science, 33(1), pp. 126-148. https://doi.org/10.1287/orsc.2021.1549.
36. Li, H. and Tian, F. (2026) 'Advancing decision-making through AI-human collaboration: a systematic review and conceptual framework', Group Decision and Negotiation, 35, article 26. https://doi.org/10.1007/s10726-026-09980-1.
37. Logg, J.M., Minson, J.A. and Moore, D.A. (2019) 'Algorithm appreciation: people prefer algorithmic to human judgment', Organizational Behavior and Human Decision Processes, 151, pp. 90-103. https://doi.org/10.1016/j.obhdp.2018.12.005.
38. March, J.G. (1991) 'Exploration and exploitation in organizational learning', Organization Science, 2(1), pp. 71-87. https://doi.org/10.1287/orsc.2.1.71.
39. Neshenko, N. and Ryall, M.D. (2026) 'When artificial intelligence does strategy: learning, good times, lock-in, and human-driven strategic renewal', Strategy Science, 11(1), pp. 157-179. https://doi.org/10.1287/stsc.2025.0448.
40. Nickerson, J.A. and Zenger, T.R. (2004) 'A knowledge-based theory of the firm - the problem-solving perspective', Organization Science, 15(6), pp. 617-632. https://doi.org/10.1287/orsc.1040.0093.
41. Noy, S. and Zhang, W. (2023) 'Experimental evidence on the productivity effects of generative artificial intelligence', Science, 381(6654), pp. 187-192. https://doi.org/10.1126/science.adh2586.
42. Pearl, J. and Mackenzie, D. (2018) The Book of Why: The New Science of Cause and Effect. New York: Basic Books.
43. Powell, T.C., Lovallo, D. and Fox, C.R. (2011) 'Behavioral strategy', Strategic Management Journal, 32(13), pp. 1369-1386. https://doi.org/10.1002/smj.968.
44. Qu, X., Kumar, M.V.S. and Tong, T.W. (2026) 'The role of predictions in acquisition decision making: the strategic value of AI-driven foresight', Strategy Science, 11(1), pp. 55-74. https://doi.org/10.1287/stsc.2025.0418.
45. Raisch, S. and Fomina, K. (2025) 'Combining human and artificial intelligence: hybrid problem-solving in organizations', Academy of Management Review, 50(2), pp. 441-464. https://doi.org/10.5465/amr.2021.0421.
46. Raisch, S. and Krakowski, S. (2021) 'Artificial intelligence and management: the automation-augmentation paradox', Academy of Management Review, 46(1), pp. 192-210. https://doi.org/10.5465/amr.2018.0072.
47. Rathje, J., Katila, R. and Reineke, P. (2024) 'Making the most of AI and machine learning in organizations and strategy research: supervised machine learning, causal inference, and matching models', Strategic Management Journal, 45(10), pp. 1926-1953. https://doi.org/10.1002/smj.3604.
48. Sako, M. and Felin, T. (2025) 'Does AI prediction scale to decision making?', Communications of the ACM, 68(4), pp. 18-21. https://doi.org/10.1145/3722138.
49. Sen, P., Workiewicz, M. and Puranam, P. (2026) 'Can LLMs aid analogical reasoning for strategic decisions? A comparative study', Strategy Science, 11(1), pp. 118-136. https://doi.org/10.1287/stsc.2025.0426.
50. Shrestha, Y.R., Ben-Menahem, S.M. and von Krogh, G. (2019) 'Organizational decision-making structures in the age of artificial intelligence', California Management Review, 61(4), pp. 66-83. https://doi.org/10.1177/0008125619862257.
51. Sorenson, O. (2024) 'Theory, search, and learning', Strategy Science, 9(4), pp. 372-381. https://doi.org/10.1287/stsc.2024.0179.

52. Sørensen, J.B. and Carroll, G.R. (2021) Making Great Strategy: Arguing for Organizational Advantage. New York: Columbia University Press.
53. Tabassi, E. (2023) Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1.
54. Teece, D.J., Pisano, G. and Shuen, A. (1997) 'Dynamic capabilities and strategic management', Strategic Management Journal, 18(7), pp. 509-533. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z.
55. Thomke, S. (2020) Experimentation Works: The Surprising Power of Business Experiments. Boston, MA: Harvard Business Review Press.
56. Trigeorgis, L. and Reuer, J.J. (2017) 'Real options theory in strategic management', Strategic Management Journal, 38(1), pp. 42-63. https://doi.org/10.1002/smj.2593.

APPENDIX A. STRATEGIC EVIDENCE AND COMMITMENT RECORD

The following one-page record can be adapted for board papers, innovation gates, investment committees or transformation programmes. It operationalises the Accountable Strategic Experimentation Cycle by making the reasoning chain, evidence status, challenge process and decision update visible without requiring disclosure of confidential model internals.

Appendix Table A1. Illustrative strategic evidence and commitment record.

Field	Required entry
Frame	Decision statement; desired outcome; time horizon; constraints; stakeholders; decision owner.
Focal hypothesis	Intervention; mechanism; expected intermediate result; strategic outcome.
Rival hypothesis	Credible alternative mechanism or explanation; observation that would distinguish it from the focal theory.
Evidence status	For each material claim: fact; external evidence; estimate; assumption; judgement; source and date.
AI use	Tool/model and version/date; strategic task supported; material outputs retained; human checks performed.
Critical assumptions	Top assumptions ranked by decision sensitivity and uncertainty.
Experiment / staged action	Scope; duration; sample or unit; metrics; guardrails; stop/scale thresholds; reversibility plan.
Challenge and decision rights	Independent challenger; accountable executive; escalation authority; approval body.
Evaluation	Observed results versus focal and rival predictions; unintended effects; confidence update.
Decision update	Stop; scale; redesign; collect more evidence; reformulate; rationale and next review point.