Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Fluent but unfounded: generative AI, cue–evidence decoupling and founder judgement

Abbas Mohammed Sulaiman

abbas.mohammed@ukh.edu.krd

University of Kurdistan Hewler

Zina Saad Hussein

University of Kurdistan Hewler

Article

Keywords:

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11354871/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Generative artificial intelligence (AI) can now produce, within minutes, a venture risk assessment that reads like the work of an experienced analyst. Whether founders, and the investors who read their documents, can still tell how much such an assessment rests on is an open question with consequences for entrepreneurship, finance and the governance of AI. This article develops a theory of cue–evidence decoupling. People judge the strength of evidence from cues such as fluency, structure and agreement, and generative AI supplies these cues whether or not the evidence is there. Because founders act when the uncertainty they perceive falls to a level they can accept, decoupling lowers perceived uncertainty without lowering actual uncertainty. It also weakens written venture materials as costly signals to investors. Because feedback in entrepreneurship is slow and noisy, founders should learn to correct for decoupling only slowly. To make the theory testable before outcomes are known, the article draws on risk science to define the knowledge-strength gap: the number of strength-of-knowledge criteria on which a founder rates an AI-assisted risk assessment as better supported than blind raters judge a normalised version of it to be. Eight propositions specify when the gap widens (novel ventures; estimating rather than listing risks), when it narrows (an AI critic; feedback on past estimates), how it raises resource commitment and how disclosure of AI use shifts investors' errors. Each proposition is paired with a randomised test and a result that would refute it. The theory implies that AI may mislead founders most where they need it most. It recasts the debate on AI and Knightian uncertainty as a question about founders' beliefs, and it suggests that safeguards for AI-assisted decisions, including transparency rules, should target the grounds of confidence and not only the accuracy of AI output.

Introduction

Generative artificial intelligence (AI) is changing how people form beliefs about matters they cannot check for themselves (Bentley, 2025; Kidd & Birhane, 2023). Entrepreneurship is a consequential case. Founders now ask AI assistants to list what could go wrong with their ventures, to estimate how likely each risk is and to write the result up for investors (Otis et al., 2026; Shepherd & Majchrzak, 2022). Generative AI is spreading quickly through new and small firms as a tool for analysis, planning and business model design (Giuggioli & Pellegrini, 2023; Kanbach et al., 2024).

Consider a founder who plans a drone-based crop-monitoring service for smallholder farmers in a region where no such service exists. The founder asks an AI assistant to assess the venture's risks. Within a minute it returns a well-organised register of regulatory, technical and demand risks, each with a likelihood and a mitigation. Yet the assistant has no data on what these farmers would pay and no record of comparable ventures, so it cannot know whether its estimates are sound. The founder must decide whether the assessment is good enough to act on. This article is about that decision, and the example recurs throughout.

The example raises a problem that current theory does not address. Theories of entrepreneurial action hold that founders act when the uncertainty they perceive falls to a level they are willing to bear

(McMullen & Shepherd, 2006). Perceived uncertainty depends on how founders judge their evidence. People judge evidence largely from cues, such as how fluent, detailed and confident it looks (Alter & Oppenheimer, 2009; Koriat, 1997). Before generative AI, these cues were costly to produce without knowing the domain, so they carried information. Generative AI now supplies them at almost no cost, whatever the state of the underlying knowledge. If founders keep reading the cues as before, AI will lower the uncertainty they perceive without lowering the uncertainty they face.

Entrepreneurship research has long documented that founders perceive less risk than others in the same situation and are prone to overconfidence (Hayward et al., 2006; Simon et al., 2000). It has also begun to ask how AI changes entrepreneurial judgement (Chalmers et al., 2021; Obschonka & Audretsch, 2020; Rady et al., 2026). Research on human–AI interaction shows, in turn, that people often over-rely on fluent AI output (Klingbeil et al., 2024; Steyvers et al., 2025). Two things are still missing. The first is a theory of how AI changes the link between the evidence founders hold and the confidence they place in it. The second is a way to observe that link when founders decide. Overconfidence is usually detected by comparing beliefs with outcomes. For a new venture, outcomes may take years to arrive, and comparable cases may not exist.

This article addresses both gaps. Its approach is conceptual: it integrates research on metacognition, risk science, entrepreneurial action and signalling into a mechanism-based theory, derives testable propositions and specifies for each a randomised test and a condition under which it would be refuted. The core claim is that generative AI separates the cues people use to judge evidence from the evidence itself, a process termed cue–evidence decoupling. For founders, decoupling lowers perceived uncertainty; for investors, it erodes the value of written venture materials as signals; and correction through learning should be slow. To make the theory testable before outcomes are known, it borrows a standard from risk analysis, the strength of the knowledge behind a judgement (Flage & Aven, 2009), and uses it to define the knowledge-strength gap: the number of criteria on which a founder rates an AI-assisted risk assessment as better supported than blind raters do.

The article makes three contributions. First, it extends theories of entrepreneurial action (McMullen & Shepherd, 2006) and the metacognitive view of entrepreneurial cognition (Haynie et al., 2012). It explains how a tool, rather than a disposition, can move founders across the threshold of action without new evidence. Because tools can be redesigned, the explanation also points to remedies. Second, it gives the debate on AI and Knightian uncertainty (Ramoglou et al., 2025; Townsend et al., 2025a) a consequence that can be observed: AI should mislead founders most for the ventures it knows least about, because its fluency does not fall when its knowledge does. Third, it links founder cognition to entrepreneurial finance and to the governance of AI. If AI makes polish costless, business plans and pitch decks can no longer serve as costly signals of preparation (Chen et al., 2009; Connelly et al., 2011), and disclosure of AI use, the remedy favoured by transparency rules (European Union, 2024), may trade one error for another. These contributions speak beyond entrepreneurship to a broader social science question: how societies should judge knowledge claims when the appearance of expertise has become cheap.

Theoretical development

Entrepreneurial action and perceived uncertainty

Entrepreneurial action is judgement under uncertainty (Knight, 1921; Packard et al., 2017). Founders must decide whether to commit resources before the information that would settle the question exists. Much of their work consists of forming, testing and revising beliefs about a venture's prospects (Camuffo et al., 2020; Zellweger & Zenger, 2023). Venture risk assessment, which means listing what could go wrong, judging how likely and how severe each risk is and deciding whether to proceed, is where these beliefs are written down.

McMullen and Shepherd (2006) explain when beliefs lead to action. Action depends on two things: the knowledge a person brings to an opportunity and the motivation to bear the uncertainty that remains. A person acts when perceived uncertainty falls to a level they are willing to bear. The trigger is perceived, not actual, uncertainty, and knowledge enters through the person's judgement of it: what moves a founder is the belief that they know enough. Anything that raises this belief without raising the knowledge behind it moves the founder towards action without good grounds. AI could also act on the second element, motivation: an assistant that agrees with a founder may encourage them to bear more uncertainty. The theory concerns the knowledge element, but Proposition 6 is designed to separate the two.

Founders' judgements of risk are known to be favourable. People who start ventures tend to perceive less risk, rather than to accept more of it, and lower perceived risk goes with the decision to start a venture (Simon et al., 2000). This pattern is linked to overconfidence (Hayward et al., 2006) and to an inside view that ignores comparable cases (Kahneman & Lovallo, 1993). These findings are the baseline against which any effect of AI must be shown.

Founders rarely assess risk only in their heads. They write business plans, pitch decks and investor memos, both to think a venture through and to justify it to others. For investors, the quality of such documents has worked as a costly signal: a detailed and coherent plan was hard to write without effort and some command of the domain (Connelly et al., 2011). Venture capitalists respond to signs of a founder's preparation (Chen et al., 2009). Venture risk assessment therefore has two audiences: the founder who acts on it and the investors who fund the action.

Current debates about AI in entrepreneurship and strategy turn on what the technology does to judgement. One view holds that AI can take over parts of prediction and evaluation, leaving judgement to humans (Csaszar et al., 2024; Doshi et al., 2025). Theory-based and scientific views add that founders create value by forming and testing theories about unproven ventures, which data-driven prediction cannot replace (Camuffo et al., 2020; Felin & Holweg, 2024). A third line asks whether AI can say anything useful about futures without precedent (Ramoglou et al., 2025; Townsend et al., 2025a). All three views assume that founders can tell how much their beliefs rest on. The theory developed here questions this assumption.

Judging evidence from cues

People have no direct access to how much they know. According to the cue-utilisation account of metacognition, they infer it from cues available at the moment of judgement (Koriat, 1997). Fluency is one such cue: information that is easy to process feels more certain and is more likely to be believed (Alter & Oppenheimer, 2009). Advice given with confidence is judged more accurate and carries more weight (Sniezek & Van Swol, 2001). These judgements then steer behaviour: people stop searching, and act, when they feel they know (Nelson & Narens, 1990).

Entrepreneurship scholars have recognised that such monitoring matters. Metacognitive ability, the capacity to monitor and adapt one's thinking, is a source of cognitive adaptability in entrepreneurial tasks and shapes how entrepreneurs use feedback (Haynie et al., 2012). This work treats monitoring as a capability and assumes that the cues available to it carry information. The question here is what happens when they do not.

Cues carry information when they vary with actual knowledge. For a founder working alone, a detailed and coherent risk assessment is hard to produce without knowing the domain. Having produced one is therefore a partly valid sign of knowledge, for the founder and for an investor alike.

Cue–evidence decoupling

Generative AI breaks this link. It produces fluent, structured and confident text on any topic, and its fluency does not fall when its knowledge does (Messeri & Crockett, 2024; Steyvers et al., 2025). Its output is generated without regard to whether it is true (Hicks et al., 2024), and it rarely signals uncertainty even when it is wrong (Zhou et al., 2024). It can also fabricate content outright (Sun et al., 2024), and it tends to agree with the views users express (Sharma et al., 2024). Because people treat confident, fluent answers as a sign of knowledge, AI can distort beliefs in ways users cannot easily detect (Kidd & Birhane, 2023). The result is cue–evidence decoupling: the cues by which people judge evidence vary independently of the evidence. In lens-model terms, decoupling lowers the ecological validity of the cues, that is, how closely they track what they are used to judge, while people's reliance on the cues stays the same. Accurate judgement needs both valid cues and appropriate reliance on them (Karelaia & Hogarth, 2008).

The idea that cues can mislead is old. What is new is that a single technology now lowers the validity of the same cues for all founders and investors at once, at the point where entrepreneurial theory locates the decision to act. Decoupling is a property of the information environment, not of the founder: sound metacognitive habits formed when the cues carried information can still mislead. Trained professionals are not immune. Attorneys in a United States federal court were sanctioned for filing a brief that cited court decisions a chatbot had invented (Mata v. Avianca, Inc., 2023). Unlike most founders, they worked in a field where every citation can be checked.

Decoupling has consequences at two levels. For the founder, it moves perceived uncertainty towards the threshold of action (McMullen & Shepherd, 2006) without supplying the knowledge the threshold is meant to reflect. For investors, it erodes the signalling value of written venture materials. When polish costs little, it no longer separates prepared founders from unprepared ones, and a separating signal becomes a pooling one (Connelly et al., 2011).

Two points define the scope of the theory. First, the founder and the AI produce an assessment together, but not as equals. The founder decides and the AI advises, as in a judge–advisor system (Sniezek & Van Swol, 2001). The term founder–AI team is used only in this sense. Second, the mechanism is not trust in automation (Lee & See, 2004). Trust concerns a user's attitude towards a system. Decoupling concerns the user's judgement of the evidence behind an assessment. Fluent output can inflate that judgement even in a user who distrusts AI. Organisational research agrees that how professionals engage with AI output matters more than their general attitude towards it (Lebovitz et al., 2022; Raisch & Krakowski, 2021).

Is decoupling temporary?

People can learn how valid a cue is, but only through feedback, and entrepreneurship offers poor conditions for such learning. Feedback on a venture's risks is slow, noisy and confounded by luck. Entrepreneurs also update slowly: in one study, self-employed people revised their expectations by only about 16% of the new information they received and gave most weight to their prior beliefs (Parker, 2006). Reliance can also sustain itself: agreeable AI is trusted and preferred even as it distorts judgement (Cheng et al., 2026). Decoupling should therefore last longer for founders than in settings with fast and clear feedback.

Learning should also differ between founders and investors. A founder sees the outcome of only one or a few ventures, whereas an investor sees many and can learn sooner that polished documents predict little. For founders, business failure can prompt learning and more realistic expectations, although responses to failure vary (Ucbasaran et al., 2013). Investors face the opposite risk: if they learn to discount polish, they may discount sound work along with weak work. Propositions 7 and 8 state these predictions.

A standard that applies before outcomes

To test whether decoupling inflates founders' judgements, a standard is needed that applies at the time of assessment. Risk analysis offers one. In its uncertainty-based view, a probability is a judgement that rests on background knowledge, and the same number can rest on strong or on weak knowledge (Aven, 2016). A risk description is therefore incomplete unless it also says how strong that knowledge is. Flage and Aven (2009) proposed four criteria. Knowledge is strong when the phenomena are well understood, reliable data are available, the assumptions are reasonable and experts broadly agree. It is weak when

one or more of these conditions clearly fails. Because such ratings depend on interpretation, raters may disagree (Cox, 2008), a concern built into the measure below.

The criteria must be linked carefully to entrepreneurship's own concepts. Knightian uncertainty has two readings. On an ontological reading, the future of a novel venture is not yet determined, so no one can know it. On an epistemic reading, the future is knowable in principle but unknown to the founder (Packard et al., 2017; Townsend et al., 2024). The strength-of-knowledge criteria address the epistemic side: they ask how well the available knowledge supports a judgement, not whether the future is knowable. Under ontological uncertainty they still give a clear answer, that knowledge is weak, because no reliable data or expert consensus about an undetermined future can exist. The standard is therefore strictest where the theory predicts the most distortion.

Evidence on AI and judgement

Does AI improve judgement under uncertainty?

A meta-analysis of 106 experiments found that human–AI teams performed, on average, below the better of the human or the AI alone; teams lost ground in decision tasks and gained in content creation (Vaccaro et al., 2024). Against this, generative AI raised the speed and quality of professional writing (Noy & Zhang, 2023) and the productivity of customer-support agents, especially less experienced ones (Brynjolfsson et al., 2025). Assistance from large language models (LLMs) improved forecasting accuracy by roughly a quarter in a preregistered study of 991 forecasters (Schoenegger et al., 2025), and LLM evaluations of business strategies resembled those of experts and investors (Csaszar et al., 2024; Doshi et al., 2025). The gains have a boundary. In a field experiment with management consultants, AI raised performance on tasks within its capabilities but made consultants less likely to reach correct answers on a task just outside them (Dell'Acqua et al., 2026). The positive results come mostly from tasks with reference classes or checkable answers. Whether they extend to the risks of an unprecedented venture is not known.

Do the surface features of AI output inflate confidence?

Users overestimated the accuracy of LLM answers, and longer explanations raised their confidence without raising accuracy (Steyvers et al., 2025). People who used AI on reasoning problems scored higher but overestimated their scores (Fernandes et al., 2026), and knowledge workers who were more confident in generative AI reported less critical thinking when using it (Lee et al., 2025). Participants followed AI advice even when it contradicted available information and was costly to follow (Klingbeil et al., 2024), and LLM-based search made decisions faster, but users over-relied on its answers when it erred (Spatharioti et al., 2025). Agreeable responses made users more convinced that they were right (Cheng et al., 2026). Yet people also discount what they know to be machine-made. Headlines labelled as AI-generated were rated less accurate, whether or not they were (Altay & Gilardi, 2024), and people

abandon algorithms after seeing them err (Dietvorst et al., 2015). The net effect of fluent AI text on judgements of evidence remains open.

Can AI address Knightian uncertainty?

Townsend et al. (2025a) argue that the problems of Knightian uncertainty limit what AI can predict. Ramoglou et al. (2025) reply that AI forces a rethinking of the concept, and the original authors respond (Townsend et al., 2025b). Madjdi and Wurth (2026) argue that AI produces plausibility signals from historical data, which may inflate confidence and penalise ventures with little precedent. Rady et al. (2026) describe an ideator's dilemma: founders often cannot tell whether an AI suggestion is a breakthrough or a fabrication. This debate has not yet asked how AI changes a founder's judgement of how much they know.

Do remedies work, and is confidence harmful?

With AI, an LLM devil's advocate improved reliance in group decisions (Chiang et al., 2024), and among business professionals, moderate and questioning use of generative AI went with less anchoring, confirmation and overconfidence (Cristofaro et al., 2026). When LLMs expressed uncertainty in the first person, users relied less on incorrect answers (Kim et al., 2024). Remedies can also trade one error for another. Interventions that force users to think reduce over-reliance, but users dislike them and they can lower reliance on correct advice (Buçinca et al., 2021). Inflated confidence is also not clearly harmful. Lower perceived risk goes with the decision to start a venture (Simon et al., 2000), and optimism may explain why people start ventures despite low expected returns (Åstebro et al., 2014). On the other side, overconfidence is argued to raise venture failure (Hayward et al., 2006).

Evidence from founders establishes that they judge their prospects favourably (Simon et al., 2000) and that the same AI advice helps some founders and harms others: in a field experiment in Kenya, an AI business assistant had no detectable average effect, but its effect for entrepreneurs performing poorly at baseline was more than 0.20 standard deviations lower than for high performers, a gap traced to the advice they chose to implement (Otis et al., 2026). The remaining link, from the cues in AI output to a founder's judgement of evidence, has been observed only in general samples. The propositions are designed to test it with founders.

Framework and construct

The founder–AI team and its division of labour

An assessment has distinct tasks, and the AI can take a different role in each. As generator, it lists risk events. As estimator, it supplies probabilities and magnitudes. As critic, it challenges an assessment the founder has drafted. These roles follow the tasks of building a risk description (Stødle et al., 2025) and the advisory roles studied in AI-assisted decision making (Chiang et al., 2024). They matter because the evidence on AI differs by task (Proposition 4). Figure 1 shows the framework and the propositions.

The knowledge-strength gap

When a founder works with an AI assistant, two things can change. The knowledge behind the assessment may improve, because the assistant supplies information the founder lacked. The founder's judgement of that knowledge may also rise, because the assistant supplies fluent, structured and agreeable text. Decoupling implies that the second change does not depend on the first. The resulting discrepancy is termed the knowledge-strength gap.

For a given assessment, the gap is the number of the four strength-of-knowledge criteria on which the founder rates the assessment as better supported than independent raters do. The founder rates each criterion as weak, moderate or strong. A panel of raters with domain and venture-evaluation experience rates the same assessment on the same criteria without knowing how it was produced. On each criterion, the two ratings are compared by order only, and the result is over-rated, matched or under-rated. The gap is the number of criteria over-rated, from 0 to 4; under-rating is reported separately. Table 1 gives a worked example for the drone venture. In the original scheme, the four criteria jointly yield one overall rating (Flage & Aven, 2009), so whether the founder's overall rating is higher than the panel's is also reported.

The standard must not be exposed to the bias it is meant to detect. If polish inflates founders' ratings, it may inflate raters' ratings too, and the gap would then shrink exactly when AI is used. Raters therefore judge a normalised version of each assessment. Its risks, estimates, reasons and sources are copied into a plain template with a fixed structure, so that raters see the evidence without the cues. A second coder checks that nothing was lost. The fourth criterion is also adapted. Raters judge whether the assessment's assumptions reflect agreement among experts in the field, as shown by the sources it cites or omits, not whether they agree with one another. Their own agreement is a separate test of the measure.

Normalisation may, however, strip out structure that genuinely aids understanding, so that part of the gap reflects lost information rather than over-rating. Founders therefore also rate the normalised version, after rating the original. The difference between the two ratings is a presentation effect: it measures how much presentation alone moves the founder and needs no panel. It also helps with novel ventures, where raters will judge knowledge as weak and the panel-based gap therefore has a low floor.

Three further features are deliberate. First, the gap is a count, not a difference of scores. Weak, moderate and strong are ordered labels, and arithmetic on such ratings is a known source of error (Cox, 2008). Second, the panel's rating is an assessed standard, not a true one, because expert judgement is fallible. A precondition is therefore set: panel ratings must reach a pre-registered level of agreement, such as a weighted kappa of 0.60 (Cohen, 1968). If trained raters cannot agree, the construct fails, and this is its first test. Third, the gap can be scored when the assessment is made, which is when founders decide. The Supplementary Information sets out the full protocol, including the rating rubric and analysis plan.

Table 1
Worked example of the knowledge-strength gap for one assessment (hypothetical ratings).

Criterion	Founder's rating	Panel's rating	Comparison
Phenomena understood	Strong	Moderate	Over-rated
Data available and reliable	Moderate	Weak	Over-rated
Assumptions reasonable	Moderate	Moderate	Matched
Experts agree	Moderate	Weak	Over-rated
Knowledge-strength gap			3 of 4

Distinguishing the gap from neighbouring constructs

The gap is related to several constructs but identical to none (Table 2). Overconfidence compares beliefs with outcomes (Moore & Healy, 2008), so it requires outcomes. The calibration gap compares users' confidence in an LLM's answer with the model's own confidence (Steyvers et al., 2025); it concerns trust in the model, not the evidence behind a joint assessment. Metacognitive accuracy compares self-rated with actual performance on tasks that have right answers (Fernandes et al., 2026), which venture risk assessments lack. The account of inflated evaluative confidence (Madjdi & Wurth, 2026) names the phenomenon but offers no standard for measuring it. The gap can be scored before outcomes are known, which is when it matters for entrepreneurship.

Table 2
The knowledge-strength gap and its nearest neighbours.

Construct	What it compares	Why the gap differs
Overconfidence (Moore & Healy, 2008)	Belief with realised outcomes	The gap needs no outcomes and is scored at the time of assessment
Calibration gap (Steyvers et al., 2025)	User's confidence in a model's answer with the model's confidence	The gap concerns the evidence behind a human–AI assessment, not trust in the model
Metacognitive accuracy (Fernandes et al., 2026)	Self-rated with actual performance on tasks with right answers	A venture risk assessment has no answer key, so the gap uses a risk-science standard
Inflated evaluative confidence (Madjdi & Wurth, 2026)	Plausibility signals with epistemic grounding, at ecosystem level	The gap specifies criteria for grounding at the level of one assessment
Strength of knowledge (Flage & Aven, 2009)	An assessment with stated criteria	The gap treats the assessor's own rating as open to bias

Propositions

Eight propositions follow. Each rests on a condition that a researcher can assign at random, such as whether a founder works with AI or which role the AI plays, so none depends on inferring causes from correlations. Table 3 lists the test for each proposition and the result that would refute it. Each test fixes in advance the smallest effect that would matter in practice, written δ (Lakens, 2017). A proposition is refuted when the observed effect is reliably smaller than δ or runs in the opposite direction. With δ set at ten percentage points in the share of criteria over-rated, a two-group comparison needs about 170 founders per group; the Supplementary Information gives the calculation and shows how within-founder designs keep the studies feasible.

AI assistance and the gap

The core claim follows from decoupling. An AI assistant adds two things to a founder's assessment: content and cues. Content may raise the assessed strength of knowledge. Cues raise the perceived strength of knowledge, whether or not the content does. Because AI produces fluent and confident text on any topic, the cues arrive in full even when the content is thin. The founder's rating should therefore rise more than the raters' rating. Existing evidence fits this reading without testing it. AI users overestimated both their own performance (Fernandes et al., 2026) and the accuracy of LLM answers (Steyvers et al., 2025); in each case, perceived quality ran ahead of a measurable standard.

Proposition 1

Founders who produce a venture risk assessment with a generative AI assistant over-rate its strength of knowledge on more criteria than founders who produce one unaided.

AI may, however, improve the knowledge itself. LLM assistance improved forecasts (Schoenegger et al., 2025), and LLMs evaluated strategies about as well as experienced investors (Csaszar et al., 2024). If AI improves an assessment as much as it improves the founder's opinion of it, the raters' ratings rise in step and the gap does not grow. Proposition 1 does not deny that AI can help; it predicts that perceived support rises more than assessed support. The test assigns founders at random to work alone or with an assistant, on a common case or on their own ventures.

Cues as the mechanism

Proposition 1

could hold for reasons other than cues, such as trust in AI. Proposition 2 isolates the mechanism by changing the cues while holding the content fixed. Two cues stand out. Polish is the fluency, structure, length and confident tone of the output. It excludes specificity, which adds information and so belongs to content. Agreement is the degree to which the output endorses the view the user has expressed.

Longer explanations raised users' confidence in LLM answers without raising accuracy (Steyvers et al., 2025). Assistants tend to conform to users' views (Sharma et al., 2024), and agreement made users more convinced they were right (Cheng et al., 2026). Founders often hold a favourable view of their venture (Simon et al., 2000), so agreement is likely to be frequent. Both cues are visible outside the

laboratory. In April 2025 a leading developer withdrew an update to its chatbot after users reported that it had become overly flattering and agreeable (OpenAI, 2025). A risk register like the one in the drone example once took days to write; it now takes a prompt.

Proposition 2

Holding the content of an assessment constant, (a) more polished output and (b) output that agrees with the founder's stated view each raise the founder's strength-of-knowledge ratings.

Polish may not work this way. People discount what they know to be machine-made (Altay & Gilardi, 2024; Dietvorst et al., 2015), and founders may read polish as a mark of the machine rather than of knowledge. Agreement is a design choice that developers can reduce, so part (b) may not hold for later systems. The test presents the same information units (each risk, estimate, reason and source) in plain or polished form and in neutral or agreeing wording. Because content is identical, any rise in the founder's rating is over-rating, and no panel is needed.

Novelty as a boundary condition

The gap should not be the same for all ventures. AI output draws on patterns in its training material, which for a venture in an established category include comparable cases, base rates and known failure modes. For a venture with no reference class, the patterns are remote from the case (Townsend et al., 2025a), and the knowledge is weak on every criterion. The cues, however, do not weaken: output for a novel venture is as fluent as output for a familiar one. Madjdi and Wurth (2026) make a related argument at the level of the ecosystem.

Evidence on language models supports this asymmetry. A model's accuracy depends on how common a task and its answer are in its training material; accuracy falls for rare versions of otherwise identical tasks, while the output stays just as fluent (McCoy et al., 2024). Professionals who used AI on a task beyond its capabilities did worse than those who worked alone (Dell'Acqua et al., 2026). The remedy for an optimistic inside view is to consult a reference class of similar cases (Kahneman & Lovallo, 1993). A novel venture, such as the drone service, has none. The founder of a new restaurant can compare AI estimates with the known fate of other restaurants; the founder of the drone service cannot.

Proposition 3

The increase in over-rating produced by AI assistance is larger for a venture with no reference class than for a venture in an established category.

Two arguments point the other way. AI may help most where users know least: generative AI raised the output of less creative writers more than that of others (Doshi & Hauser, 2024), and of less experienced workers more than that of experienced ones (Brynjolfsson et al., 2025). If AI fills a novel venture's knowledge gap, assessed strength rises most there, and the gap narrows. Ramoglou et al. (2025) also dispute that Knightian uncertainty limits AI as Townsend et al. (2025a) claim. Because raters will judge

the knowledge behind a novel venture as weak, over-rating may be more common for novel ventures with or without AI. The proposition is therefore stated as an interaction that compares AI-assisted and unaided founders facing the same case, which removes baseline differences between cases; the presentation effect gives a second test that does not depend on the low floor.

Identifying versus estimating risks

Risk assessment is not one task. A listed risk can be judged on sight, because the founder can see whether it applies. A probability cannot be inspected for correctness, so judgement of an estimate falls back on the cues. The evidence on AI differs in the same way. An LLM did better at brainstorming failure modes than at giving expert-level guidance (Collier et al., 2025), generative AI produced solutions to business problems that were rated as valuable though somewhat less novel than human ones (Boussioux et al., 2024), and human–AI teams gained in creation tasks and lost in decision tasks (Vaccaro et al., 2024).

Proposition 4

AI assistance (a) increases the number of panel-validated risks a founder identifies and (b) raises over-rating more for the founder's probability estimates than for the founder's judgement of which listed risks apply.

Neither part is assured. LLM help improved general forecasting (Schoenegger et al., 2025), which counts against part (b). AI lists may also be generic: product safety professionals found LLM guidance too generic (Collier et al., 2025), and AI can make outputs more alike across users (Doshi & Hauser, 2024). Part (a) therefore counts only risks the panel validates as relevant. The two halves of part (b) are measured differently, because the strength-of-knowledge criteria apply to probabilities, not to lists. For estimates, over-rating is the knowledge-strength gap. For the list, it is the difference between the share of listed risks the founder judges relevant and the share the panel validates. If the proposition holds, founders should use AI to list risks and keep estimates tied to outside evidence.

A critic role as a remedy

If cues of knowledge widen the gap, cues of its absence should narrow it. An assistant told to challenge the founder's assessment names missing data, weak assumptions and reasons the venture could fail. A devil's-advocate role improved reliance in AI-assisted group decisions (Chiang et al., 2024), questioning AI output went with less bias (Cristofaro et al., 2026), and expressions of uncertainty reduced reliance on incorrect LLM answers (Kim et al., 2024). A critic role also fits the scientific approach to entrepreneurship, which asks founders to state what would prove their beliefs wrong (Camuffo et al., 2020).

Proposition 5

An AI assistant instructed to challenge the founder's assessment produces less over-rating than either unaided work or an assistant instructed to supply estimates, and more rejection of sound AI input than an

estimator.

The second clause reflects a cost. Interventions that make users think harder reduce over-reliance but can also lower reliance on sound advice, and users dislike them (Buçinca et al., 2021). There is also a cost specific to ventures. Scrutiny based on historical patterns can penalise ventures with little precedent (Madjdi & Wurth, 2026), so a critic may talk founders out of ventures worth trying. The prediction is therefore a trade-off, not a net benefit. The test assigns founders to work alone, with an estimator, with a critic or with a generic critic that asks only fixed questions, such as "What evidence supports this estimate?" If over-rating falls under the generic critic, the effect runs through cues of absence; if it falls only when the critic names specific gaps, it runs through content.

From the gap to commitment

The gap matters for entrepreneurship only if it changes what founders do. In the cue-utilisation account, people who feel they know stop searching and act (Nelson & Narens, 1990), and lower perceived risk goes with starting a venture (Simon et al., 2000). A founder who believes an assessment is well supported should perceive less risk and commit more.

Inflated commitment matters because early commitments are hard to reverse. People escalate their commitment to a course of action they feel responsible for, even after negative feedback (Staw, 1976). Effectual reasoning offers a guard: it ties commitment to what a founder can afford to lose rather than to expected returns (Sarasvathy, 2001). A founder who believes their knowledge is strong has less reason to use this guard, so the gap's effect on commitment should be weaker among founders who reason in terms of affordable loss.

The proposition claims that the gap, and not polish as such, drives commitment. A correlation between the gap and commitment cannot show this, because polish could raise commitment in other ways. The proposed mediator is therefore manipulated (Imai et al., 2013). Founders see the polished or plain output from Proposition 2, and half then see an independent rating card showing the panel's ratings on the four criteria, which closes the gap. A placebo card from the same experts, rating only the clarity of the writing, controls for deference to expert authority. If polish raises commitment through the gap, its effect should shrink with the rating card but not with the placebo card. If agreement still raises commitment after the gap is closed, it works through motivation rather than knowledge.

Proposition 6

Polished AI output raises founders' commitment of resources to the venture, and this effect is smaller when the knowledge-strength gap is closed by an independent rating than when a placebo rating is shown.

More commitment is not necessarily worse. Optimism may be what gets ventures started (Åstebro et al., 2014), and the biases that lower perceived risk may have benefits as well as costs (Simon et al., 2000). The proposition predicts the direction of the effect and leaves the judgement of welfare to the

Discussion. If Proposition 2(a) fails, the cue was not manipulated, and Proposition 6 is untested rather than refuted. Commitment is measured with real stakes, such as the share of a budget given to launching rather than to further testing.

Recalibration

Founders can learn to discount fluent cues, but slowly, and feedback is the mechanism. A founder who learns that an AI-assisted estimate was wrong has evidence that the cues misled them. Repeated feedback of this kind should lower their reliance on the cues and narrow the gap in later assessments. Evidence from education points the same way: an AI tutor that gave hints rather than answers mitigated the learning losses that unrestricted use produced (Bastani et al., 2025). The effect should be stronger for founders who have experienced venture failure, which can make people readier to believe they can be wrong (Ucbasaran et al., 2013).

Proposition 7

Feedback on the accuracy of earlier AI-assisted estimates reduces the knowledge-strength gap in later assessments, and the reduction is larger for founders who have experienced venture failure.

This proposition may fail. Entrepreneurs give heavy weight to prior beliefs and update slowly (Parker, 2006), and the same study found no difference in updating by experience. Feedback may therefore have little effect, or none that differs with failure experience. The test gives founders a sequence of cases, with accuracy feedback for a random half, and measures the gap on the final case.

From founders to investors

Investors and accelerator selectors also infer the strength of evidence partly from presentation (Chen et al., 2009). An obvious remedy is disclosure: if founders state that AI helped produce an assessment, investors can discount its polish. Disclosure, however, is a blunt cue. People who learn that content was produced with AI tend to judge it less accurate, whether or not it is (Altay & Gilardi, 2024; Longoni et al., 2022), and news labelled as AI-generated was judged less trustworthy even when it was not judged less accurate, unless its sources were listed (Toff & Simon, 2025). People are also poor at telling AI-generated from human-written language (Jakesch et al., 2023). An investor who cannot see the evidence directly may discount sound assessments along with weak ones. Disclosure should therefore trade one error for another, as the critic role does.

This prediction runs against a common view. Transparency is often treated as a signal that builds warranted trust in AI (Park & Yoon, 2025), and transparency rules treat disclosure of AI use as helping readers judge content (European Union, 2024). The theory implies that disclosure helps only where readers can separate cue from evidence; otherwise, it shifts errors from over-rating to under-rating. The test shows investors the same assessments, varies only whether AI use is disclosed and compares their ratings with the panel's.

Proposition 8

Disclosing that a venture risk assessment was produced with AI reduces investors' over-rating of weakly supported assessments, but also lowers their ratings of assessments that the panel judges well supported.

Table 3
Tests and refutation conditions for the eight propositions. δ is the pre-registered smallest effect of interest, and "< δ" means reliably smaller than δ.

Proposition	Test	Supported if	Refuted if
P1. Gap	Founders work alone or with AI; panel rates normalised versions	Over-rating at least δ higher with AI	Difference < δ, or reversed
P2. Cues	Same information units: plain vs. polished; neutral vs. agreeing	Ratings at least δ higher under each cue	Difference < δ, or ratings fall
P3. Novelty	Alone vs. AI; each founder rates a novel and an established case	AI effect at least δ larger for the novel case	Interaction < δ, or reversed
P4. Identify vs. estimate	Panel validates listed risks; gap scored for estimates	(a) More validated risks with AI; (b) AI effect at least δ larger for estimates	Either part fails
P5. Critic	Alone, estimator, critic or generic critic	Over-rating at least δ lower with the critic; more rejection of sound input	Over-rating not reduced by δ
P6. Commitment	Polish crossed with rating card vs. placebo card; choice with real stakes	Polish raises commitment; effect smaller with rating card than with placebo	No polish effect, or no difference between cards
P7. Recalibration	Sequence of cases with or without accuracy feedback	Final gap at least δ smaller with feedback; larger drop after failure experience	Feedback effect < δ
P8. Disclosure	Investors rate the same assessments with or without AI disclosure	Over-rating of weak assessments and ratings of sound ones both fall by at least δ	Either effect < δ

Discussion

When the framework would be wrong

The framework has a clear rival. On this account, generative AI improves the knowledge behind an assessment at least as fast as it improves the founder's opinion of it, as the positive findings reviewed above suggest. This is a real possibility, especially as models improve, and three results would decide between the accounts. If over-rating is no more common with AI than without it, the gap does not exist as described (Proposition 1). If ratings do not move when cues change and content is fixed, decoupling

is not the mechanism (Proposition 2), and any gap would need another explanation, such as trust (Lee & See, 2004) or advice-taking in general (Sniezek & Van Swol, 2001). If the gap is no wider for novel ventures, decoupling has no special force under Knightian uncertainty (Proposition 3). The construct can also fail before any proposition is tested, if raters cannot agree. Founder characteristics may condition all of these results: domain expertise gives an independent check on AI output (Otis et al., 2026), whereas familiarity with AI may not protect, as higher AI literacy has gone with less accurate self-assessment (Fernandes et al., 2026).

Is the gap harmful?

A gap between perceived and assessed knowledge is an error of judgement, but whether it is harmful is a separate question. For the individual founder, over-rating implies commitments the evidence does not support, and overconfidence is argued to raise venture failure (Hayward et al., 2006). For the economy, optimistic entry produces experiments that would not otherwise be run (Åstebro et al., 2014). A critic that narrows the gap may also suppress ventures with no precedent (Madjdi & Wurth, 2026).

Entrepreneurship theory offers a way through that does not require settling this question. When knowledge is weak, founders can still act boldly if they limit commitments to what they can afford to lose and learn from small experiments first (Camuffo et al., 2020; Sarasvathy, 2001). These strategies require not strong knowledge but recognition that knowledge is weak, and decoupling undermines exactly this recognition. Does AI help founders act boldly on weak knowledge they recognise as weak? Or does it lead them to act as if the knowledge were strong?

Implications for theory

The framework speaks to three conversations. The first concerns entrepreneurial action and cognition. Research on overconfidence explains why founders perceive less uncertainty than they face by pointing to dispositions and heuristics (Hayward et al., 2006; Simon et al., 2000). The present account adds a third source: the information environment in which founders judge evidence. The same founder can be well calibrated in one environment and overconfident in another, so overconfidence is partly a property of the founder's tools. Monitoring that relies on cues (Haynie et al., 2012) should lose value as AI spreads, while monitoring that relies on independent checks, such as experiments and outside ratings, should gain value, which gives the scientific approach to entrepreneurship (Camuffo et al., 2020) a new rationale. In aggregate, if AI lowers perceived uncertainty for many potential founders at once, it should raise entry. Experiments show that overconfidence draws in more entrants than a market can profitably support (Camerer & Lovallo, 1999). Whether AI-assisted entry is excess entry is an empirical question that the gap can help answer.

The second conversation concerns AI and Knightian uncertainty. The exchange between Townsend et al. (2025a) and Ramoglou et al. (2025) asks what AI can compute about the future. The present theory shifts the question from computation to belief. Whatever AI can compute, its fluency does not fall when its knowledge does. It should therefore mislead founders most for the ventures it can compute least about (Proposition 3). If so, the limits of AI under Knightian uncertainty matter less for what AI predicts

than for what founders infer from how it writes. This consequence can be observed without first settling whether such uncertainty is ontological or epistemic.

The third conversation concerns entrepreneurial finance. Signalling accounts assume that some signals are costly enough to separate types (Connelly et al., 2011). Research on venture finance shows that investors respond to signs of preparation (Chen et al., 2009), and an algorithm has outperformed business angels other than highly experienced ones (Blohm et al., 2022). Generative AI removes the cost of one class of signal, written polish, for all founders at once. Investors will either be misled by polish or learn to discount it, and Proposition 8 predicts that such learning will penalise sound work along with weak work. Either way, screening should shift towards evidence that AI cannot easily produce, such as live interaction and the results of experiments, which would change which founders obtain resources.

Beyond entrepreneurship, the account adds to wider social science debates on AI-mediated knowledge. Concerns that generative AI reduces many voices to one (Bentley, 2025), distorts beliefs (Kidd & Birhane, 2023) and fosters dependence (Cheng et al., 2026) have mostly been framed around information quality and learning. Cue–evidence decoupling identifies a more general mechanism, a fall in the validity of the cues by which people judge knowledge claims, that should operate wherever polished documents have served as evidence of competence, from grant applications and policy briefs to student essays. The knowledge-strength gap offers a measure that can be carried to these settings.

Implications for practice and policy

For founders, the main implication is a division of labour. AI is useful for listing risks, because it widens the set of things a founder considers (Proposition 4). Founders can mark which parts of an assessment came from AI, ask someone who did not help produce it to rate how well supported it is, and run a second pass with the AI in a critic role (Proposition 5). Entrepreneurship scholars have also begun to recommend that founders check AI output against independent sources (Rady et al., 2026).

For investors and accelerators, which for most new ventures are the closest thing to a risk function, the implication is to ask for evidence that polish cannot supply. A venture risk assessment can be required to state the strength of knowledge behind its main estimates, to show the result of a critic pass and to report the experiments behind its key assumptions. Simple disclosure of AI use is not enough and may penalise sound work (Proposition 8). Accelerators can also give founders feedback on the accuracy of their past estimates (Proposition 7). Developers of AI tools for founders can build in a critic pass, flag claims that rest on no cited source, express uncertainty and limit agreement with the user.

For policymakers, the implication is that transparency about AI use and transparency about evidence are different things. Public programmes that fund early-stage ventures can ask applicants for a short statement of the strength of knowledge behind their main claims. Such a statement would complement transparency rules, which address whether content was generated by AI but not how well supported it is (European Union, 2024).

Research agenda and limitations

Four studies would test the eight propositions: Study 1 (Propositions 2 and 6) needs no panel and should be run first; Study 2 covers Propositions 1, 3 and 4; Study 3 covers Proposition 5; and Study 4 covers feedback and investor disclosure (Propositions 7 and 8). Founders can be recruited through accelerators and screened online panels in more than one country, so that the role of entrepreneurial and regulatory context can be examined. Each study should fix and report the AI model, its version and the prompts used, and repeat key conditions when models change. The Supplementary Information sets out the full protocol.

Field evidence would complement the experiments. The public release of ChatGPT in November 2022 offers a natural break. Researchers can ask whether the link between the polish of business plan competition entries, accelerator applications and crowdfunding campaigns and later venture outcomes weakened after that date, as the signalling argument predicts. An accelerator could also give the critic routine to some cohorts and not others and follow later pivots, closures and funding. A further question is whether AI makes founders' risk assessments more alike (Doshi & Hauser, 2024).

The framework has four limits. First, the mechanism has so far been observed among students, online participants and professionals without a stake in their tasks, whereas founders are invested in the answer and assess ventures that may have no precedent. The propositions are therefore framed as tests, not findings. Second, generative AI changes quickly, and agreement and the effects of cues may change as users learn and models improve (Sharma et al., 2024), so the propositions are stated so that they can be retested. Third, the strength-of-knowledge criteria were developed outside entrepreneurship, so their fit to ventures must be shown rather than assumed. The count also ignores how far apart two ratings are, which is the price of avoiding arithmetic on ordered labels. Fourth, the theory considers one founder and one assistant. Founding teams, whose members may check one another, and AI agents that act on their own raise questions that are left open.

Conclusion

Generative AI has made it easy to produce a venture risk assessment that looks well supported. Whether it is well supported is a different question, and for most founders no one is placed to ask it. This article has argued that AI decouples the cues by which founders and investors judge evidence from the evidence itself, lowering perceived uncertainty without lowering actual uncertainty and eroding the signalling value of venture documents. The knowledge-strength gap makes these claims observable at the moment of decision, and each of the eight propositions states how it could be shown to be wrong. If the propositions hold, founders, those who fund them and those who regulate AI need safeguards aimed at confidence, not only at the accuracy of AI. If they fail, the view of AI as a reliable aid to entrepreneurial judgement will rest on firmer ground.

Declarations

Data availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study. The protocol for the proposed studies is provided in the Supplementary Information.

Competing interests

The authors declare no competing interests.

Ethical approval

Ethical approval was not required as this conceptual article does not involve human participants or their data. The studies proposed in the article will require institutional ethics approval before any data are collected.

Informed consent

Informed consent was not required as this article does not involve human participants.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Use of generative AI

During the preparation of this work the authors used ChatGPT to support literature discovery and language refinement. Every source identified with the tool was retrieved, read and checked by the authors. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the article.

Author Contribution

Abbas Mohammed Sulaiman: Conceptualization, Methodology, Investigation, Visualization, Writing – original draft, Writing – review & editing, Supervision.Zina Saad Hussein: Conceptualization, Project administration, Writing – original draft, Writing – review & editingBoth authors read and approved the final manuscript.

References

1. Altay S, Gilardi F (2024) People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation. PNAS Nexus 3(10):pgae403. https://doi.org/10.1093/pnasnexus/pgae403

2. Alter AL, Oppenheimer DM (2009) Uniting the tribes of fluency to form a metacognitive nation. Personality Social Psychol Rev 13(3):219–235. https://doi.org/10.1177/1088868309341564
3. Åstebro T, Herz H, Nanda R, Weber RA (2014) Seeking the roots of entrepreneurship: insights from behavioral economics. J Economic Perspect 28(3):49–70. https://doi.org/10.1257/jep.28.3.49
4. Aven T (2016) Risk assessment and risk management: review of recent advances on their foundation. Eur J Oper Res 253(1):1–13. https://doi.org/10.1016/j.ejor.2015.12.023
5. Bastani H, Bastani O, Sungu A, Ge H, Kabakcı Ö, Mariman R (2025) Generative AI without guardrails can harm learning: evidence from high school mathematics. Proceedings of the National Academy of Sciences 122(26):e2422633122. https://doi.org/10.1073/pnas.2422633122
6. Bentley SV (2025) Knowing you know nothing in the age of generative AI. Humanit Social Sci Commun 12:409. https://doi.org/10.1057/s41599-025-04731-0
7. Blohm I, Antretter T, Sirén C, Grichnik D, Wincent J (2022) It's a peoples game, isn't it?! A comparison between the investment returns of business angels and machine learning algorithms. Entrepreneurship Theory Pract 46(4):1054–1091. https://doi.org/10.1177/1042258720945206
8. Boussioux L, Lane JN, Zhang M, Jacimovic V, Lakhani KR (2024) The crowdless future? Generative AI and creative problem-solving. Organ Sci 35(5):1589–1607. https://doi.org/10.1287/orsc.2023.18430
9. Brynjolfsson E, Li D, Raymond L (2025) Generative AI at work. Quart J Econ 140(2):889–942. https://doi.org/10.1093/qje/qjae044
10. Buçinca Z, Malaya MB, Gajos KZ (2021) To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. Proceedings of the ACM on Human-Computer Interaction 5(CSCW1):1–21. https://doi.org/10.1145/3449287
11. Camerer C, Lovallo D (1999) Overconfidence and excess entry: an experimental approach. Am Econ Rev 89(1):306–318. https://doi.org/10.1257/aer.89.1.306
12. Camuffo A, Cordova A, Gambardella A, Spina C (2020) A scientific approach to entrepreneurial decision making: evidence from a randomized control trial. Manage Sci 66(2):564–586. https://doi.org/10.1287/mnsc.2018.3249
13. Chalmers D, MacKenzie NG, Carter S (2021) Artificial intelligence and entrepreneurship: implications for venture creation in the fourth industrial revolution. Entrepreneurship Theory Pract 45(5):1028–1053. https://doi.org/10.1177/1042258720934581
14. Chen X-P, Yao X, Kotha S (2009) Entrepreneur passion and preparedness in business plan presentations: a persuasion analysis of venture capitalists' funding decisions. Acad Manag J 52(1):199–214. https://doi.org/10.5465/amj.2009.36462018
15. Cheng M, Lee C, Khadpe P, Yu S, Han D, Jurafsky D (2026) Sycophantic AI decreases prosocial intentions and promotes dependence. Science 391(6792):eaec8352. https://doi.org/10.1126/science.aec8352
16. Chiang C-W, Lu Z, Li Z, Yin M (2024) Enhancing AI-assisted group decision making through LLM-powered devil's advocate. In: Proceedings of the 29th International Conference on Intelligent User

Interfaces. ACM, New York, p 103–119. https://doi.org/10.1145/3640543.3645199

17. Cohen J (1968) Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull 70(4):213–220. https://doi.org/10.1037/h0026256
18. Collier ZA, Gruss RJ, Abrahams AS (2025) How good are large language models at product risk assessment? Risk Anal 45(4):766–789. https://doi.org/10.1111/risa.14351
19. Connelly BL, Certo ST, Ireland RD, Reutzel CR (2011) Signaling theory: a review and assessment. J Manag 37(1):39–67. https://doi.org/10.1177/0149206310388419
20. Cox LA Jr (2008) What's wrong with risk matrices? Risk Anal 28(2):497–512. https://doi.org/10.1111/j.1539-6924.2008.01030.x
21. Cristofaro M, Giardino PL, Augier M-S (2026) Generative AI and cognitive biases in strategic decision-making. Strategic Organ. https://doi.org/10.1177/14761270261457350
22. Csaszar FA, Ketkar H, Kim H (2024) Artificial intelligence and strategic decision-making: evidence from entrepreneurs and investors. Strategy Sci 9(4):322–345. https://doi.org/10.1287/stsc.2024.0190
23. Dell'Acqua F, McFowland E, Mollick E, Lifshitz H, Kellogg KC, Rajendran S, Krayer L, Candelon F, Lakhani KR (2026) Navigating the jagged technological frontier: field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. Organ Sci 37(2):403–423. https://doi.org/10.1287/orsc.2025.21838
24. Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: people erroneously avoid algorithms after seeing them err. J Exp Psychol Gen 144(1):114–126. https://doi.org/10.1037/xge0000033
25. Doshi AR, Bell JJ, Mirzayev E, Vanneste BS (2025) Generative artificial intelligence and evaluating strategic decisions. Strateg Manag J 46(3):583–610. https://doi.org/10.1002/smj.3677
26. Doshi AR, Hauser OP (2024) Generative AI enhances individual creativity but reduces the collective diversity of novel content. Sci Adv 10(28):eadn5290. https://doi.org/10.1126/sciadv.adn5290
27. European Union (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official J Eur Union L 2024/1689. http://data.europa.eu/eli/reg/2024/1689/oj
28. Felin T, Holweg M (2024) Theory is all you need: AI, human cognition, and causal reasoning. Strategy Sci 9(4):346–371. https://doi.org/10.1287/stsc.2024.0189
29. Fernandes D, Villa S, Nicholls S, Haavisto O, Buschek D, Schmidt A, Kosch T, Shen C, Welsch R (2026) AI makes you smarter but none the wiser: the disconnect between performance and metacognition. Comput Hum Behav 175:108779. https://doi.org/10.1016/j.chb.2025.108779
30. Flage R, Aven T (2009) Expressing and communicating uncertainty in relation to quantitative risk analysis. Reliability & Risk Analysis: Theory & Applications 2(13, Part 1):9–18
31. Giuggioli G, Pellegrini MM (2023) Artificial intelligence as an enabler for entrepreneurs: a systematic literature review and an agenda for future research. Int J Entrepreneurial Behav Res 29(4):816–837.

https://doi.org/10.1108/IJEBR-05-2021-0426

32. Haynie JM, Shepherd DA, Patzelt H (2012) Cognitive adaptability and an entrepreneurial task: the role of metacognitive ability and feedback. Entrepreneurship Theory Pract 36(2):237–265. https://doi.org/10.1111/j.1540-6520.2010.00410.x
33. Hayward MLA, Shepherd DA, Griffin D (2006) A hubris theory of entrepreneurship. Manage Sci 52(2):160–172. https://doi.org/10.1287/mnsc.1050.0483
34. Hicks MT, Humphries J, Slater J (2024) ChatGPT is bullshit. Ethics Inf Technol 26(2):38. https://doi.org/10.1007/s10676-024-09775-5
35. Imai K, Tingley D, Yamamoto T (2013) Experimental designs for identifying causal mechanisms. J Royal Stat Society: Ser (Statistics Society) 176(1):5–51. https://doi.org/10.1111/j.1467-985X.2012.01032.x
36. Jakesch M, Hancock JT, Naaman M (2023) Human heuristics for AI-generated language are flawed. Proceedings of the National Academy of Sciences 120(11):e2208839120. https://doi.org/10.1073/pnas.2208839120
37. Kahneman D, Lovallo D (1993) Timid choices and bold forecasts: a cognitive perspective on risk taking. Manage Sci 39(1):17–31. https://doi.org/10.1287/mnsc.39.1.17
38. Kanbach DK, Heiduk L, Blueher G, Schreiter M, Lahmann A (2024) The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective. RMS 18(4):1189–1220. https://doi.org/10.1007/s11846-023-00696-z
39. Karelaia N, Hogarth RM (2008) Determinants of linear judgment: a meta-analysis of lens model studies. Psychol Bull 134(3):404–426. https://doi.org/10.1037/0033-2909.134.3.404
40. Kidd C, Birhane A (2023) How AI can distort human beliefs. Science 380(6651):1222–1223. https://doi.org/10.1126/science.adi0248
41. Kim SSY, Liao QV, Vorvoreanu M, Ballard S, Vaughan JW (2024) I'm not sure, but..: examining the impact of large language models' uncertainty expression on user reliance and trust. In: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency. ACM, New York, p 822–835. https://doi.org/10.1145/3630106.3658941
42. Klingbeil A, Grützner C, Schreck P (2024) Trust and reliance on AI—an experimental study on the extent and costs of overreliance on AI. Comput Hum Behav 160:108352. https://doi.org/10.1016/j.chb.2024.108352
43. Knight FH (1921) Risk, uncertainty and profit. Houghton Mifflin, Boston
44. Koriat A (1997) Monitoring one's own knowledge during study: a cue-utilization approach to judgments of learning. J Exp Psychol Gen 126(4):349–370. https://doi.org/10.1037/0096-3445.126.4.349
45. Lakens D (2017) Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. Social Psychol Personality Sci 8(4):355–362. https://doi.org/10.1177/1948550617697177

46. Lebovitz S, Lifshitz-Assaf H, Levina N (2022) To engage or not to engage with AI for critical judgments: how professionals deal with opacity when using AI for medical diagnosis. Organ Sci 33(1):126–148. https://doi.org/10.1287/orsc.2021.1549
47. 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. ACM, New York, p 1–22. https://doi.org/10.1145/3706598.3713778
48. Lee JD, See KA (2004) Trust in automation: designing for appropriate reliance. Hum Factors 46(1):50–80. https://doi.org/10.1518/hfes.46.1.50_30392
49. Longoni C, Fradkin A, Cian L, Pennycook G (2022) News from generative artificial intelligence is believed less. In: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency. ACM, New York, p 97–106. https://doi.org/10.1145/3531146.3533077
50. Madjdi F, Wurth B (2026) AI-mediated plausibility regimes: entrepreneurial judgment, epistemic risk, and the distribution of entrepreneurial futures. J Bus Venturing Insights 26:e00644. https://doi.org/10.1016/j.jbvi.2026.e00644
51. Mata v, Avianca (2023) Inc. 678 F. Supp. 3d 443 (S.D.N.Y. 2023)
52. McCoy RT, Yao S, Friedman D, Hardy MD, Griffiths TL (2024) Embers of autoregression show how large language models are shaped by the problem they are trained to solve. Proceedings of the National Academy of Sciences 121(41):e2322420121. https://doi.org/10.1073/pnas.2322420121
53. McMullen JS, Shepherd DA (2006) Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. Acad Manage Rev 31(1):132–152. https://doi.org/10.5465/amr.2006.19379628
54. Messeri L, Crockett MJ (2024) Artificial intelligence and illusions of understanding in scientific research. Nature 627(8002):49–58. https://doi.org/10.1038/s41586-024-07146-0
55. Moore DA, Healy PJ (2008) The trouble with overconfidence. Psychol Rev 115(2):502–517. https://doi.org/10.1037/0033-295X.115.2.502
56. Nelson TO, Narens L (1990) Metamemory: a theoretical framework and new findings. In: Bower GH (ed) The psychology of learning and motivation, vol 26. Academic, San Diego, pp 125–173
57. 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
58. Obschonka M, Audretsch DB (2020) Artificial intelligence and big data in entrepreneurship: a new era has begun. Small Bus Econ 55(3):529–539. https://doi.org/10.1007/s11187-019-00202-4
59. OpenAI (2025) Sycophancy in GPT-4o: what happened and what we're doing about it. https://openai.com/index/sycophancy-in-gpt-4o/. Accessed 12 Oct 2026
60. Otis NG, Clarke R, Delecourt S, Holtz D, Koning R (2026) The uneven impact of generative artificial intelligence on entrepreneurial performance: evidence from a field experiment in Kenya. Manage Sci. https://doi.org/10.1287/mnsc.2024.06909

61. Packard MD, Clark BB, Klein PG (2017) Uncertainty types and transitions in the entrepreneurial process. Organ Sci 28(5):840–856. https://doi.org/10.1287/orsc.2017.1143
62. Park K, Yoon HY (2025) AI algorithm transparency, pipelines for trust not prisms: mitigating general negative attitudes and enhancing trust toward AI. Humanit Social Sci Commun 12:1160. https://doi.org/10.1057/s41599-025-05116-z
63. Parker SC (2006) Learning about the unknown: how fast do entrepreneurs adjust their beliefs? J Bus Ventur 21(1):1–26. https://doi.org/10.1016/j.jbusvent.2004.07.005
64. Rady J, Townsend D, Hunt R (2026) From algorithmic hallucinations to alien minds: addressing the ideator's dilemma through entrepreneurial work. J Bus Ventur 41(1):106550. https://doi.org/10.1016/j.jbusvent.2025.106550
65. Raisch S, Krakowski S (2021) Artificial intelligence and management: the automation–augmentation paradox. Acad Manage Rev 46(1):192–210. https://doi.org/10.5465/amr.2018.0072
66. Ramoglou S, Schaefer R, Chandra Y, McMullen JS (2025) Artificial intelligence forces us to rethink Knightian uncertainty: a commentary on Townsend et al.'s Are the futures computable? Acad Manage Rev 50(2):471–473. https://doi.org/10.5465/amr.2024.0127
67. Sarasvathy SD (2001) Causation and effectuation: toward a theoretical shift from economic inevitability to entrepreneurial contingency. Acad Manage Rev 26(2):243–263. https://doi.org/10.5465/amr.2001.4378020
68. Schoenegger P, Park PS, Karger E, Trott S, Tetlock PE (2025) AI-augmented predictions: LLM assistants improve human forecasting accuracy. ACM Trans Interact Intell Syst 15(1):1–25. https://doi.org/10.1145/3707649
69. Sharma M, Tong M, Korbak T et al (2024) Towards understanding sycophancy in language models. In: Proceedings of the 12th International Conference on Learning Representations (ICLR 2024), Vienna, 7–11 May 2024. https://proceedings.iclr.cc/paper_files/paper/2024/hash/0105f7972202c1d4fb817da9f21a9663-Abstract-Conference.html
70. Shepherd DA, Majchrzak A (2022) Machines augmenting entrepreneurs: opportunities (and threats) at the nexus of artificial intelligence and entrepreneurship. J Bus Ventur 37(4):106227. https://doi.org/10.1016/j.jbusvent.2022.106227
71. Simon M, Houghton SM, Aquino K (2000) Cognitive biases, risk perception, and venture formation: how individuals decide to start companies. J Bus Ventur 15(2):113–134. https://doi.org/10.1016/S0883-9026(98)00003-2
72. Sniezek JA, Van Swol LM (2001) Trust, confidence, and expertise in a judge-advisor system. Organ Behav Hum Decis Process 84(2):288–307. https://doi.org/10.1006/obhd.2000.2926
73. Spatharioti SE, Rothschild D, Goldstein DG, Hofman JM (2025) Effects of LLM-based search on decision making: speed, accuracy, and overreliance. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. ACM, New York, p 1–15. https://doi.org/10.1145/3706598.3714082

74. Staw BM (1976) Knee-deep in the big muddy: a study of escalating commitment to a chosen course of action. Organizational Behav Hum Perform 16(1):27–44. https://doi.org/10.1016/0030-5073(76)90005-2
75. Steyvers M, Tejeda H, Kumar A, Belem C, Karny S, Hu X, Mayer LW, Smyth P (2025) What large language models know and what people think they know. Nat Mach Intell 7(2):221–231. https://doi.org/10.1038/s42256-024-00976-7
76. Stødle K, Flage R, Guikema S, Aven T (2025) Artificial intelligence for risk analysis—a risk characterization perspective on advances, opportunities, and limitations. Risk Anal 45(4):738–751. https://doi.org/10.1111/risa.14307
77. Sun Y, Sheng D, Zhou Z, Wu Y (2024) AI hallucination: towards a comprehensive classification of distorted information in artificial intelligence-generated content. Humanit Social Sci Commun 11:1278. https://doi.org/10.1057/s41599-024-03811-x
78. Toff B, Simon FM (2025) Or they could just not use it? the dilemma of AI disclosure for audience trust in news. Int J Press/Politics 30(4):881–903. https://doi.org/10.1177/19401612241308697
79. Townsend DM, Hunt RA, Rady J (2024) Chance, probability, and uncertainty at the edge of human reasoning: what is Knightian uncertainty? Strateg Entrepreneurship J 18(3):451–474. https://doi.org/10.1002/sej.1516
80. Townsend DM, Hunt RA, Rady J, Manocha P, Jin JH (2025a) Are the futures computable? Knightian uncertainty and artificial intelligence. Acad Manage Rev 50(2):415–440. https://doi.org/10.5465/amr.2022.0237
81. Townsend DM, Hunt RA, Rady J, Manocha P, Jin JH (2025b) Do androids dream of entrepreneurial possibilities? A reply to Ramoglou et al.'s Artificial intelligence forces us to rethink Knightian uncertainty. Acad Manage Rev 50(2):474–476. https://doi.org/10.5465/amr.2024.0299
82. Ucbasaran D, Shepherd DA, Lockett A, Lyon SJ (2013) Life after business failure: the process and consequences of business failure for entrepreneurs. J Manag 39(1):163–202. https://doi.org/10.1177/0149206312457823
83. Vaccaro M, Almaatouq A, Malone T (2024) When combinations of humans and AI are useful: a systematic review and meta-analysis. Nat Hum Behav 8(12):2293–2303. https://doi.org/10.1038/s41562-024-02024-1
84. Zellweger T, Zenger T (2023) Entrepreneurs as scientists: a pragmatist approach to producing value out of uncertainty. Acad Manage Rev 48(3):379–408. https://doi.org/10.5465/amr.2020.0503
85. Zhou K, Hwang J, Ren X, Sap M (2024) Relying on the unreliable: the impact of language models' reluctance to express uncertainty. In: Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, Bangkok, p 3623–3643. https://doi.org/10.18653/v1/2024.acl-long.198

Figures

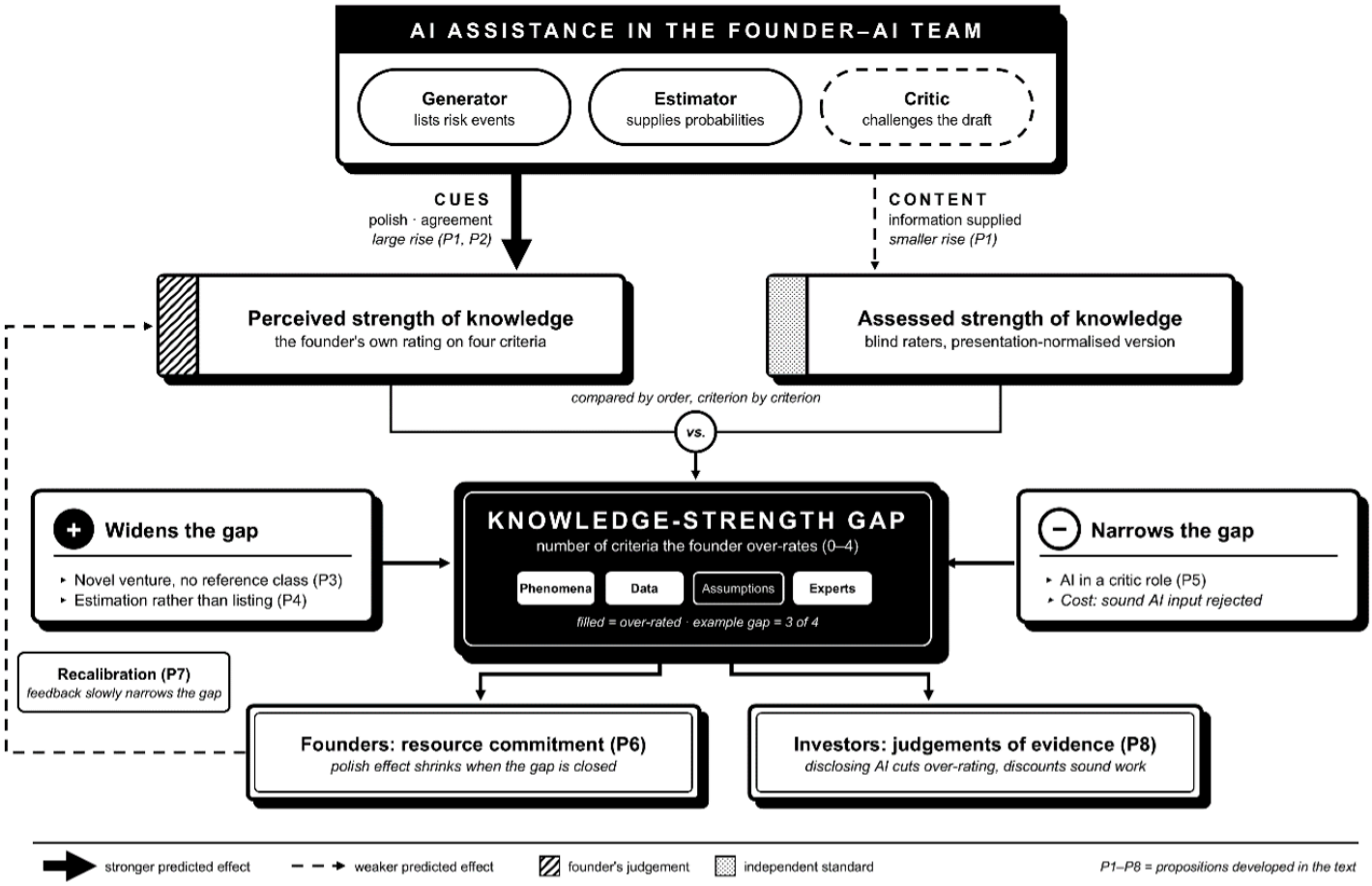


Figure 1

The founder–AI team, the knowledge-strength gap and the eight propositions (P1–P8). AI assistance raises the founder's perceived strength of knowledge mainly through cues (polish and agreement), and it raises the strength of knowledge assessed by blind raters, who see a normalised version, mainly through content. Comparing the two, criterion by criterion, yields the knowledge-strength gap. Venture novelty and estimation tasks widen the gap, and an AI critic narrows it. The gap raises founders' commitment and, through disclosure of AI use, shapes investors' judgements. Feedback on past estimates slowly narrows the gap. Solid arrows indicate stronger predicted effects and dashed arrows weaker predicted effects.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.docx