A WHITE PAPER

TeamLMI
Focused on Performance

AI as Force Multiplier: Organizational Conditions and the Returns to Artificial Intelligence in the Small and Mid-Sized Enterprise

Kent E. Frese, Ph.D. — TeamLMI · 2026

ABSTRACT

Generative artificial intelligence is diffusing into small and mid-sized enterprises (SMEs) faster than any general-purpose workplace technology in memory, yet early returns are strikingly heterogeneous: firms adopting identical tools report outcomes ranging from substantial productivity gains to expensive disappointment. This paper argues that the variance is not technological but organizational. Drawing on the economics of technology–organization complementarity, the socio-technical tradition, recent labor-market evidence on augmentation and automation, and twenty-five years of practitioner work with small and mid-sized businesses, it advances what I term the *amplifier hypothesis*: artificial intelligence multiplies the organizational conditions into which it is introduced. In firms with a healthy culture, a clear and shared strategy, and capable, committed people, AI functions as a force multiplier; in firms lacking those conditions, it amplifies dysfunction, accelerating misdirected activity rather than correcting it. The hypothesis extends a long-standing practitioner observation (that improving productive practices atop a weak cultural or strategic foundation harms rather than helps) to a technology whose multiplier is unusually large. The paper reviews the 2023–2026 evidence on AI and employment, finding it consistent with conditional rather than unconditional returns. It examines the small-business context, where adoption is accelerating while complementary investment lags. And it identifies a measurement gap with direct practical consequence: no psychometrically validated instrument yet measures *AI-augmented judgment* (the human capacity to decide when to rely on, when to override, and how to direct machine output) in working leaders or SME populations. A research agenda for the construct is proposed. The paper is intended for academic readers and reflective practitioners working at the intersection of industrial-organizational psychology, organization development, and small-enterprise management.

Keywords: artificial intelligence adoption, force multiplier, organizational culture, strategic clarity, technology–organization complementarity, augmentation, automation, judgment, small and mid-sized enterprise, psychometric assessment.

1. Introduction: The Adoption Paradox

Two businesses of similar size, in similar industries, purchase the same artificial-intelligence subscriptions in the same quarter. Eighteen months later, one has measurably expanded its capacity (faster proposals, broader marketing reach, shorter decision cycles) while the other is busier, more fragmented, and no more profitable, with a larger software bill. The pattern is now familiar to anyone who advises small and mid-sized enterprises, and it presents a genuine puzzle: when the technology is held constant, what explains the variance in returns?

The thesis of this paper is that the variance is organizational, and that it is predictable. Artificial intelligence is best understood not as a transformation engine but as a *force multiplier*, and a multiplier operates on whatever quantity it is given. Organizations with sound fundamentals find those fundamentals amplified. Organizations with confused priorities, low trust, or disengaged people find *those* conditions amplified instead, often at considerable speed and expense. The technology is the same in both cases; the multiplicand differs.

The claim is not new in kind. The economics literature has documented for a quarter-century that returns to information technology are conditioned on organizational complements: workplace organization, management practice, human capital, and intangible investment (Brynjolfsson & Hitt, 2000; Bresnahan, Brynjolfsson, & Hitt, 2002; Bloom, Sadun, & Van Reenen, 2012). What is new is the *size* of the multiplier and the *speed* of its diffusion into a population, the small and mid-sized enterprise, that the research literature persistently under-studies (a complaint registered, in the strategic-planning context, in the companion paper to this one; Frese, 2026). When the multiplier was a customer-relationship system or a reporting dashboard, the gap between well-conditioned and poorly-conditioned adopters was meaningful but bounded. Generative AI, which substitutes capital for a broad swath of cognitive *tasks* rather than a narrow band of clerical ones, widens that gap substantially.

The paper proceeds as follows. Section 2 reviews the 2023–2026 evidence on artificial intelligence and work, with attention to what that evidence can and cannot support. Section 3 examines the small-business context specifically. Section 4 states the amplifier hypothesis and grounds it in the complementarity and socio-technical literatures. Section 5 develops the three organizational conditions in detail. Section 6 turns from the organization to the person, arguing that the binding human contribution in AI-augmented work is *judgment*, and documenting a measurement gap of direct practical consequence. Section 7 draws implications for practice, including two illustrative observations from current advisory work. Section 8 states limitations and a research agenda. Section 9 concludes.

A note on register: this paper, like its companion, is written from the position of a practitioner-scholar. Claims supported by peer-reviewed or primary-source evidence are cited as such; claims derived from accumulated consulting experience are identified as practitioner observations and should be weighted accordingly.

2. The Evidence on AI and Work, 2023–2026

The public conversation about artificial intelligence and employment oscillates between displacement alarm and productivity euphoria. The research evidence, read carefully, supports neither pole. Four findings frame the present argument.

First, productivity gains do not automatically become labor demand. Acemoglu and Restrepo's (2019) task-based framework remains the most disciplined account of why. Automation displaces labor from existing tasks; countervailing forces (productivity effects, capital accumulation, and critically the *reinstatement* of labor into newly created tasks) may or may not offset the displacement. The authors' central empirical finding is sobering: over 1987–2017, reinstatement decelerated while displacement accelerated, and changes in the task content of production reduced aggregate labor demand by roughly ten percent. Their category of “so-so technologies,” innovations that displace workers while delivering only modest productivity gains, is a standing caution against assuming that adoption is synonymous with advancement. The presumption that any productivity-raising technology raises demand for labor “simply because it raises productivity,” they write, “is wrong.” Any honest account of AI's organizational promise must begin there.

Second, where AI augments rather than substitutes, a demand-expansion mechanism operates. When the cost of a unit of cognitive output falls, demand for that output can rise enough to increase, rather than reduce, the employment of the people who produce it. The rebound dynamic is popularly associated with Jevons; it was invoked by Microsoft's chief executive in early 2025 (Nadella, 2025) and applied to labor explicitly in recent Goldman Sachs Research analysis of the United States labor market (Goldman Sachs Research, 2026). Bessen's (2018) demand-side analysis supplies the underlying economics: where product demand is price-elastic, productivity growth expands output and employment together. The critical scoping condition, too often dropped in popular treatments, is that the mechanism is verified for *augmented* roles, not as a guarantee of aggregate employment growth. Indeed, the same Goldman Sachs analysis estimates a modest net drag on recent United States payroll growth, concentrated in routine-exposed and early-career segments. Jevons-style optimism, applied to AI and labor, is a conditional mechanism, not a law.

Third, the augmentation–automation boundary is where the labor-market effects sort. The Stanford Digital Economy Lab's analysis of large-scale payroll microdata (Brynjolfsson, Chandar, & Chen, 2025) finds that recent employment declines are concentrated in occupations where AI is positioned to *automate* human labor, while employment among less-exposed workers and *more experienced workers within the same occupations* has remained stable or grown. The pattern is early and subject to revision, and causal attribution remains contested; but its direction is consistent across sources. Experienced people doing judgment-intensive work fare well alongside AI; routine and entry-level task bundles are under genuine pressure. Autor's (2015) older observation, that automation tends to raise the value of the human capabilities it cannot replicate, anticipated precisely this sorting.

Fourth, the macro projections, properly caveated, lean constructive. The World Economic Forum's *Future of Jobs Report 2025*, an employer-expectation survey spanning more than a thousand firms and fourteen million workers, projects 170 million jobs created against 92 million displaced by 2030 (a net gain of 78 million) while ranking analytical thinking, leadership and social influence, and talent management among the fastest-rising skills (World Economic Forum, 2025). The International Labour Organization's refined occupational-exposure index concludes that roughly one in four workers globally is in an occupation with some generative-AI exposure, but that “most jobs will be transformed rather than made redundant” given the continued need for human input (International Labour Organization & NASK, 2025). Both findings require discipline in use. The WEF figures are employer expectations attributed to four concurrent macro trends, not an econometric forecast of AI's isolated effect; and the ILO's “transformation” is not benign by construction, since exposure concentrates in clerical work, falls disproportionately on women, and can mean degraded job quality as readily as enrichment.

The synthesis relevant to this paper is that **the evidence is consistent with conditional returns.** AI rewards configurations of work in which capable people exercise judgment over augmented output, and it pressures configurations built on routine execution. Crucially, and this is the hinge of the argument, *which configuration a given firm presents to the technology is substantially a matter of organizational choice.* The same instrument that one firm deploys to extend its experienced people, another deploys to thin its apprenticeship ranks. The task-level evidence tells us where the gradient lies; it does not relieve any leadership team of the decision about where on that gradient to stand.

3. The Small-Business Context

The research literature on AI adoption, like the strategy literature before it (Frese, 2026), is dominated by large-enterprise studies. The best current evidence on United States small-business adoption comes from the Census Bureau's Business Trends and Outlook Survey as analyzed by the Small Business Administration's Office of Advocacy (Press, 2025). Four findings matter here.

Adoption is accelerating and the size gap is narrowing. Small-firm AI use rose from 6.3 to 8.8 percent in the year ending mid-2025 while large-firm use held near 11 percent; the SBA's read is that small businesses “may only be a year behind.” The figures reflect a strict production-use definition, and the convergence is partly a large-firm plateau, so the claim should be held as trajectory rather than parity. The direction, however, is unambiguous, and more recent survey rounds continue it.

The smallest firms' dominant barrier is perceived irrelevance, not cost or risk. Among firms with fewer than five employees that report no plans to use AI, roughly 82 percent cite “not applicable to my business,” far ahead of knowledge gaps or privacy concerns. For the mid-sized firms that constitute the population this paper chiefly addresses (roughly five to two hundred fifty employees, five to fifty million dollars in revenue), the question has typically moved past relevance to execution: not *whether* but *how well*.

Complementary investment lags adoption badly. Approximately half of small firms already using AI report *no* accompanying investment: no staff training, no integration assistance, no change to data practices. The largest investment gaps relative to large firms are precisely the complements the productivity literature identifies as decisive, namely training staff, engaging integration expertise, and reorganizing data and process. In the vocabulary of this paper, half of small adopters have purchased a multiplier and left the multiplicand unattended.

Small employers are, distinctively, optimistic about AI and headcount. In the BTOS expectations data, the smallest employers are the size class most likely to expect AI to *increase* their employment and least likely to expect a decrease, a pattern the SBA explains through derived-demand economics: price-elastic small firms convert productivity gains into expanded output, which requires more people, not fewer (Press, 2025; cf. Bessen, 2018). The magnitudes are small and the data are expectations rather than realized hiring, but the directional contrast with large-enterprise displacement narratives is notable, and it coheres with the augmentation economics of Section 2.

The small and mid-sized enterprise, in short, is precisely where the amplifier question has the highest stakes: adoption is arriving quickly, complementary investment is thinnest, the potential upside (price-elastic demand, short decision chains, owner-led culture) is unusually large, and the research base is unusually sparse.

4. The Amplifier Hypothesis

The central claim of this paper can be stated in one sentence: **artificial intelligence amplifies the organizational conditions into which it is introduced; it does not create them, and it cannot substitute for them.**

Stated as a falsifiable proposition: *among comparable SMEs adopting comparable AI capability, realized returns will be an increasing function of pre-existing organizational health (specifically, cultural health, strategic clarity, and workforce capability and commitment), and adopters weak on those conditions will show returns at or below zero more often than strong-condition adopters, net of industry and size.*

Three bodies of work converge on this proposition.

The complementarity economics. The finding that information technology pays off only in combination with organizational complements is among the most replicated results in the productivity literature. Brynjolfsson and Hitt (2000) showed that the returns to computing investment were realized through organizational transformation, not hardware; Bresnahan, Brynjolfsson, and Hitt (2002) demonstrated three-way complementarity among IT, workplace organization, and human capital; Bloom, Sadun, and Van Reenen (2012) attributed the United States–Europe productivity divergence of the IT era substantially to *management practices* that allowed identical technology to yield non-identical returns. Brynjolfsson, Rock, and Syverson (2021) generalized the point for general-purpose technologies: measured returns lag adoption because the binding investment is intangible, comprising process redesign, skill-building, and organizational learning. There is no obvious reason to expect generative AI to be the first general-purpose technology exempt from this regularity, and early field evidence suggests it is not. Even where AI confers large individual productivity gains, the gains are uneven, skill-dependent, and bounded by what Dell'Acqua and colleagues (2023) call the technology's “jagged frontier.” That frontier must be learned, and learning it is itself an organizational capability.

The socio-technical tradition. Long before the productivity economists, Trist and Bamforth (1951) documented that introducing technically superior methods into a work system while disregarding its social structure degrades rather than improves performance. The longwall studies' lesson, that technical and social systems require *joint* optimization, is the amplifier hypothesis in mid-century vocabulary. The participative tradition that followed (McGregor, 1960; Likert, 1961) supplies the mechanism: the discretionary effort, candor, and local knowledge that make any new method productive are volunteered by people who are trusted and believed in, and withheld, quietly and rationally, by people who are not. A leadership team's operating assumptions about people (in McGregor's terms, Theory Y or Theory X) thus condition the returns to every tool placed in those people's hands.

The high-performance work systems evidence. The strategic human-resource literature has established with meta-analytic consistency that bundles of people-centered practices (selective hiring, investment in capability, participation, information-sharing) predict firm performance (Huselid, 1995; Combs, Liu, Hall, & Ketchen, 2006; Pfeffer, 1998). The relevance here is compositional: these are the practices that produce the capable, motivated, committed workforce that AI augments most powerfully. An organization that has under-invested in its people has, in amplifier terms, a small multiplicand.

To these three literatures I add a practitioner observation of twenty years' standing, offered with the appropriate epistemic label. In leadership training and advisory work with small and mid-sized businesses since the early 2000s, I have taught a deliberately blunt rule: *improving productive practices inside a broken culture, or in the absence of a coherent strategy, does not save an organization — it wrecks it faster.* Make a confused organization more efficient and you accelerate its progress toward the wrong objective. Streamline a low-trust team and you industrialize its resentment. The rule predates generative AI by two decades; the technology has not repealed it. What the technology has changed is the magnitude. Earlier productive-practice improvements (a better CRM, a cleaner workflow, a sharper dashboard) were modest amplifiers. Generative AI is a large one, and a large amplifier raises the stakes of the foundation in both directions.

It is worth pausing on what the hypothesis does *not* claim. It does not claim that AI is ineffective absent perfect conditions; returns are continuous, not binary. It does not claim that technology choices are irrelevant; fit between tool and task remains real. And it does not claim that conditions are immutable; they are, on the contrary, the proper object of leadership work, which is the practical point of the entire argument.

5. The Conditions, in Order

The framework I use in practice organizes organizational health into three interdependent pillars, each enabling the next, with the ordering itself carrying the load: a **healthy organizational culture** as the foundation; a **clear business strategy, genuinely shared and disciplined in execution**, built upon it; and **productive practices** (efficiency, technology, and personal leadership at every level), which become scalable only when the first two are in place. Artificial intelligence enters this framework without remodeling it. AI is the newest and most powerful occupant of the third pillar, and it obeys the third pillar's standing rule: it pays off in proportion to what sits beneath it.

5.1 Healthy Culture: The Multiplicand's Foundation

Culture, in Schein's formulation, is the pattern of shared assumptions a group has learned as it solved its problems, the thing that determines what people actually do when no one is checking (Schein & Schein, 2017). Two cultural properties bear directly on AI returns. The first is *psychological safety* (Edmondson, 2019): adopting a tool whose output must be questioned, corrected, and overridden requires a climate in which people can say “the machine is wrong” (or, harder, “I was wrong to trust the machine”) without penalty. Low-safety organizations generate a distinctive AI failure mode: confident machine output, unchallenged, flowing into decisions. The second is *trust in leadership intent*. Employees who suspect that efficiency gains will be converted into layoffs rationally resist, slow-walk, or quietly sabotage adoption; employees who believe gains will be converted into growth lean in. The owner-led SME has a structural advantage here, because intent is legible at close range. Legibility, however, cuts both ways.

5.2 Clear Strategy: The Direction of Amplification

A multiplier needs a vector. The strategy condition is not the existence of a planning document but the presence of what the companion paper calls a shared structure of strategic understanding (Frese, 2026; Wallis & Frese, 2017): a leadership team that knows, collectively and concretely, what the organization is trying to become and why. Under that condition, AI capacity is pointed at identified constraints (the bid the firm could never staff, the market it could never reach, the analysis it could never afford). Absent it, AI produces what might be called *accelerated drift*: more proposals, more content, more analysis, more motion, distributed across unexamined priorities. Speed without direction is simply a faster way to be busy, and generative AI supplies speed in unprecedented quantity.

5.3 Capable, Committed People: The Multiplicand Itself

The augmentation evidence of Section 2 is, at bottom, a claim about people. AI extends most powerfully those who know their work deeply, care about outcomes, and can think independently, because they are the ones equipped to direct it, judge it, and catch it failing. A motivated employee who knows the business cold becomes substantially more capable with AI behind them; a disengaged employee becomes a faster version of disengaged. This pattern is the high-performance-work-systems evidence wearing new clothes, and it carries an immediate corollary for selection. The marginal value of hiring for judgment, motivation, and ownership (rather than for routine task proficiency, which is precisely what the technology is absorbing) has risen, and with it the cost of selection error. One capable hire now sits atop more leverage than at any point in the modern history of small business; so does one poor hire.

The ordering claim deserves explicit statement, because it is the most practice-relevant part of the framework: *the conditions are not a menu but a sequence.* Practices scale on strategy; strategy is followed in proportion to culture. An SME leadership team asking “what is our AI strategy?” is, more often than not, asking a third-pillar question while standing on an unexamined first and second pillar. The discipline of asking the pillars in order is, in my experience, worth more than any tool-selection exercise. Is the culture healthy enough that people will actually adopt? Is the strategy clear enough that people know what to point the new capacity at? Are the people capable of being multiplied? That sequence of questions is the practical heart of this paper.

6. From Organization to Person: Judgment and the Measurement Gap

If the organizational analysis above is right, the binding *individual* contribution in AI-augmented work is judgment: the capacity to decide what to delegate to the machine, when to trust its output, when to override it, and how to remain accountable for the result. Likierman's account of judgment, as the combination of what leaders attend to, what they know, who they consult, and how they convert deliberation into commitment, has long positioned it as the defining attribute of leadership (Likierman, 2020). His more recent observation that “the more powerful AI becomes, the more we need human judgement,” with the prescriptive corollary to use AI for efficiency and insight while never outsourcing the decision itself (Likierman, 2024), states the present argument's individual-level half. Likierman's position is authoritative expert commentary rather than settled empirical law, and I cite it as such; but it converges with the augmentation evidence from an independent direction, and it names the construct this section pursues.

Industrial-organizational psychology, the discipline best equipped to measure such a construct, is at present better at *using* the technology than at measuring readiness for it. Survey evidence from the field's professional society indicates that more than ninety percent of responding I-O psychologists now use generative AI at least monthly, and over half weekly, with broadly positive sentiment, while the single most cited concern among users is the *accuracy* of generative output (Zhou & Belwalkar, 2025). The profession, in other words, has adopted the amplifier and is worried about exactly the thing this section is about: the quality control that human judgment must supply.

What does the assessment landscape offer? Less than the demand would suggest. A systematic review of AI-literacy instruments (sixteen scales across twenty-two validation studies, evaluated against the COSMIN measurement-properties framework) found that *no existing scale shows positive evidence across all measurement properties*. Most demonstrate reasonable structural validity and internal consistency; few have been tested for content validity, reliability, or responsiveness; and none, as of the review's mid-2024 cutoff, for cross-cultural validity or measurement error (Lintner, 2024). The most rigorously validated instrument in the family, the AI Literacy Questionnaire, was developed for and validated on secondary-school students (Ng, Wu, Leung, Chiu, & Chu, 2024), a population at some distance from the working leaders and SME workforces with whom the construct matters commercially. At the practitioner end of the market, the most visible recent offering is a major consultancy's “AI-ready leader” profile, organized around six expert-curated behaviors (Ackermann & Visser, 2025). It is directed at the enterprise CHRO and succession pipeline, and its published materials provide no psychometric validation, reliability or validity evidence, or normative data for the profile itself. (The same firm maintains rigorous technical documentation for its established instrument portfolio; the observation here is scoped to the AI-readiness offering, and it describes the category's present state rather than any single vendor's competence.)

Three gaps, then, jointly define the opening. First, a *construct gap*. Existing instruments measure AI literacy (knowledge of, attitudes toward, and self-efficacy with the technology) rather than *AI-augmented judgment*: the evaluative capacity to calibrate trust, delegation, and override in live workflows, which the augmentation evidence and the judgment literature jointly identify as the active ingredient. Knowing what a language model is does not predict knowing when to overrule one. Second, a *population gap*. Validation work concentrates on students and general adult samples, not working leaders, and not the SME populations where, per Section 3, the stakes are presently highest and the research base thinnest. Third, a *rigor gap*. The distance between expert-curated framework and validated, norm-referenced instrument remains uncrossed in this domain, a familiar pattern in the early commercialization of any assessment category, and the gap the discipline exists to close.

Defining the construct with the precision a validation program requires is work beyond this paper's scope, but its outline is visible. AI-augmented judgment would comprise, at minimum, *calibration* (accuracy in estimating when machine output is reliable), *delegation discipline* (matching tasks to the machine's jagged frontier; Dell'Acqua et al., 2023), *override willingness* (the disposition, partly cultural in origin, to challenge confident output), and *accountable integration* (retaining ownership of decisions that incorporate machine input). Each element is plausibly measurable with established I-O methods, with situational-judgment formats an obvious candidate, and each is plausibly trainable, which gives the construct developmental as well as selective utility. Section 8 sketches the validation agenda.

7. Implications for Practice

For the leadership teams of small and mid-sized enterprises, the argument compresses into a sequence of questions that deliberately reverses the usual order of the AI conversation.

Begin with the pillars, not the tools. Before any procurement decision: Is the culture healthy enough that people will adopt openly rather than comply quietly? Is the strategy clear enough (genuinely shared, not merely documented) that new capacity has a target? Are the people, as currently selected and developed, capable of being multiplied? A candid hour on those three questions, in that order, is the highest-yield AI-readiness exercise presently available to an SME leadership team, and it costs nothing. Where the honest answer to the first or second is no, the indicated investment is not software.

Choose a deployment posture deliberately. The augmentation–automation boundary of Section 2 is, within wide limits, a leadership choice. The evidence pattern favors deploying AI to extend capable people into work they could not previously reach. It also counsels particular caution about the quiet temptation to trim the routine and entry-level roles in which future supervisors and managers are grown. A firm that converts AI savings into the elimination of its own apprenticeship pipeline has optimized against its succession plan; the cost arrives later, labeled as a leadership shortage. Honest planning for the routine work AI genuinely pressures (redeployment, role redesign, accelerated development) belongs in the open, not in the annual budget's fine print.

Raise the selection bar where the leverage rose. If judgment, ownership, and the capacity to be multiplied are the appreciating assets, selection and development should be re-weighted accordingly. That shift argues for structured, evidence-based assessment of precisely those qualities, rather than reliance on interview impressions of task proficiency the technology is in the process of absorbing. The selection-error asymmetry of Section 5.3 does the arithmetic.

Use AI to augment collective cognition, not replace it. The companion paper's caution about strategic planning generalizes: early evidence suggests language models usefully accelerate scanning, drafting, and documentation, but degrade the quality of strategic thinking when substituted for the team's own elicitation and sense-making (Frese, 2026; Dell'Acqua et al., 2023). The amplifier should be pointed at the team's understanding, not offered in place of it.

7.1 Two Illustrative Observations

Two recent episodes from advisory practice, suitably anonymized, illustrate the hypothesis operating in both directions.

In the first, an experienced, highly motivated entrepreneur preparing to launch an electrical-contracting firm in a new state used generative tools to produce, within days, a complete brand identity: name treatment, logo, vehicle-wrap designs, and supporting visual assets. Work that would conventionally have consumed weeks of agency time and several thousand dollars was finished in an afternoon. The observation is not that the output was polished, though it was; it is that the *conditions* were already present. The owner knew exactly what he was building, for whom, and why; the tool compressed the distance between a clear intention and its execution. The episode shows the amplifier behaving as advertised: a large multiplier applied to a large multiplicand.

In the second, the principal of an established firm presented, with evident satisfaction, a list of company values generated by a chatbot in roughly thirty seconds. The list was articulate and entirely unobjectionable — and, at the moment of its generation, organizationally inert. Values acquire force when a leadership team believes them, models them, hires by them, and enforces them at cost; none of that work had occurred, and no language model can perform it. The episode is not a failure story (the generated list may yet serve as useful scaffolding if the underlying work is done), but it crystallizes the boundary the amplifier hypothesis draws. The technology can draft the artifacts of organizational health in seconds; it cannot supply the referent that makes the artifacts true. The first pillar cannot be generated.

The two observations bracket the practical message of this paper. The same class of tool, in the same month, in the same advisory practice, produced a genuine force multiplication and an elegant decoration, and the differentiating variable was, in both cases, visible before the first prompt was typed.

8. Limitations and a Research Agenda

The argument advanced here has limits that should be stated directly.

First, the amplifier hypothesis itself, in the conditional form stated in Section 4, has not been directly tested in SME populations adopting generative AI. It is an extrapolation (a disciplined one, I would argue) from the IT-complementarity literature, the socio-technical tradition, and practitioner observation. The complementarity regularity has held across successive general-purpose technologies, but “it has always been so” is an inductive argument, not a demonstration. Direct tests are feasible: panel designs relating pre-adoption measures of cultural health, strategic clarity, and workforce capability to realized AI returns in matched SME samples would convert the hypothesis from proposition to finding, in either direction.

Second, the small-business evidence base of Section 3 rests heavily on a single (high-quality, primary) source family, Census BTOS data as analyzed by the SBA Office of Advocacy, and partly on expectations rather than realized outcomes. The headcount-optimism finding in particular is directional sentiment from a fast-moving period and should be re-examined as realized hiring data accumulate.

Third, the frontier labor-market findings cited in Section 2 (Goldman Sachs Research, 2026; Brynjolfsson, Chandar, & Chen, 2025) are early, contested in their causal attribution, and certain to be revised. The argument here leans on their *pattern*, the sorting along the augmentation–automation boundary, rather than their point estimates, and that pattern is convergent across independent sources; but a reader in 2028 should check.

Fourth, the construct proposal of Section 6 is exactly that — a proposal. The research agenda it implies is conventional in shape and substantial in scale: construct definition and content validation with subject-matter experts; item development (situational-judgment and behavioral-anchor formats are the natural candidates); pilot administration in working-leader and SME samples, where the population gap is most acute; the full COSMIN-style battery of structural validity, reliability, measurement-error, and responsiveness evidence that the existing AI-literacy family lacks (Lintner, 2024); criterion studies against deployment outcomes; and norm development with transparent provenance. The author's applied work is moving in this direction, and the agenda is offered partly as a commitment device. The measurement gap documented in Section 6 should be closed by *someone* to the discipline's standards, and the field should hold whoever attempts it, the present author included, to exactly those standards.

Fifth and finally, this paper inherits the bias it complains of in others only partially corrected: it is written from accumulated experience with North American small and mid-sized enterprises, and its generalization beyond that population is untested.

9. Conclusion

A quarter-century of evidence on technology and organization supports a conclusion that the newest and largest amplifier has made newly urgent: the returns to artificial intelligence are not located in the technology. They are located in the conditions that the technology multiplies — the health of the culture, the clarity and ownership of the strategy, the capability and commitment of the people. Those conditions have always been the proper work of leadership. What has changed is the price of neglecting them and the reward for getting them right, both of which generative AI has raised substantially and, on present evidence, durably.

For the small and mid-sized enterprise, the conclusion is, on balance, encouraging. The conditions are more legible, more directly shapeable, and faster to change in a sixty-person firm than in a sixty-thousand-person one; the demand-side economics of small-firm growth favor converting productivity into expansion rather than contraction; and the technology itself is, for the first time, priced for the small-business budget. The owners best positioned for the coming decade are not those with the most sophisticated tools (the tools are converging toward universal availability) but those who were already serious about their people and their direction, and who become more deliberate about both precisely because the amplifier has grown.

Artificial intelligence is not a reason to believe in people less. It is the strongest reason yet to believe in them more, and to be more rigorous, not less, about who is brought aboard, how their judgment is developed, and whether the conditions around them are worthy of multiplication. The bottleneck was never the technology. It is, as it has always been, leadership.

Author's Note

In keeping with the argument of Section 7, this paper was drafted with substantial assistance from generative AI tools operating under the author's direction, against an evidence base assembled and verified through multi-source research with adversarial fact-checking; every claim, judgment, citation, and error remains the author's responsibility. The process (machine speed under human judgment, with the human accountable for the result) is offered as a small demonstration of the thesis.

References

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives, 33*(2), 3–30.

Ackermann, B., & Visser, K. (2025, December 16). *How to spot an AI-ready leader.* Korn Ferry.

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives, 29*(3), 3–30.

Bessen, J. (2018). *AI and jobs: The role of demand* (NBER Working Paper No. 24235). National Bureau of Economic Research.

Bloom, N., Sadun, R., & Van Reenen, J. (2012). Americans do IT better: US multinationals and the productivity miracle. *American Economic Review, 102*(1), 167–201.

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics, 117*(1), 339–376.

Brynjolfsson, E., Chandar, B., & Chen, R. (2025). *Canaries in the coal mine? Six facts about the recent employment effects of artificial intelligence.* Stanford Digital Economy Lab.

Brynjolfsson, E., & Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives, 14*(4), 23–48.

Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics, 13*(1), 333–372.

Combs, J., Liu, Y., Hall, A., & Ketchen, D. (2006). How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology, 59*(3), 501–528.

Dell'Acqua, F., McFowland, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). *Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality* (Working Paper No. 24-013). Harvard Business School.

Edmondson, A. C. (2019). *The fearless organization: Creating psychological safety in the workplace for learning, innovation, and growth.* Wiley.

Frese, K. E. (2026). *Strategic planning: The next state of the art* [White paper]. TeamLMI.

Goldman Sachs Research. (2026). *The jobs AI is likely to boost — and those it may disrupt.* Goldman Sachs.

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal, 38*(3), 635–672.

International Labour Organization & NASK. (2025). *Generative AI and jobs: A 2025 update.* International Labour Organization.

Likert, R. (1961). *New patterns of management.* McGraw-Hill.

Likierman, A. (2020, January–February). The elements of good judgment. *Harvard Business Review, 98*(1), 102–111.

Likierman, A. (2024). *Human judgement is essential in the age of AI.* London Business School Think.

Lintner, T. (2024). A systematic review of AI literacy scales. *npj Science of Learning, 9.* https://doi.org/10.1038/s41539-024-00264-4

McGregor, D. (1960). *The human side of enterprise.* McGraw-Hill.

Nadella, S. (2025, January 27). *Jevons paradox strikes again!* [Post]. X.

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology.* https://doi.org/10.1111/bjet.13411

Pfeffer, J. (1998). *The human equation: Building profits by putting people first.* Harvard Business School Press.

Press, R. (2025, September 24). *AI in business: Small firms closing in* (Research Spotlight). U.S. Small Business Administration, Office of Advocacy.

Schein, E. H., & Schein, P. A. (2017). *Organizational culture and leadership* (5th ed.). Wiley.

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting. *Human Relations, 4*(1), 3–38.

Wallis, S. E., & Frese, K. (2017). *Strategic planning: A new state of the art* [White paper]. ASK MATT Solutions and TeamLMI.

World Economic Forum. (2025). *The future of jobs report 2025*. World Economic Forum.

Zhou, S., & Belwalkar, B. (2025, Autumn). Use of artificial intelligence in industrial-organizational psychology: Current trends and future outlook. *The Industrial-Organizational Psychologist (TIP)*. Society for Industrial and Organizational Psychology.

About the Author

Kent E. Frese, Ph.D. is the founder and managing partner of TeamLMI and an Industrial-Organizational Psychologist with more than twenty-five years of practical experience in leadership and organization development. He works primarily with small and mid-sized businesses — across manufacturing, technology, professional services, and family-owned enterprises — on strategic planning, leadership development, talent strategy, and long-term succession planning. He is the founder of FactorFactory, a psychometric assessment platform serving small and mid-sized businesses and the advisors who support them. He is a member of the Society for Industrial and Organizational Psychology (SIOP) and mentors doctoral candidates in industrial-organizational psychology. He has held leadership positions from front-line manager through the C-suite and has founded several successful organizations.

For More Information

TeamLMI
3425 Simpson Ferry Road, Suite 204
Camp Hill, PA 17011
(717) 303-3993
inquiry@TeamLMI.com
https://www.teamlmi.com