



A PRACTITIONER REPORT FOR L&D LEADERS

THE CAPABILITY- INSTITUTION GAP

When AI changes what people can do
faster than organizations can adapt

Better output today does not
guarantee stronger capability
tomorrow.

DR. RAVINDER TULSIANI

Enterprise Capability Executive

rtulsiani.com

SEPTEMBER 2026

EXECUTIVE SUMMARY

When capability changes, the organization has to respond.

The Capability-Institution Gap is the mismatch between what people can do with AI and what their jobs, managers, processes and development systems enable them to do.

An employee may now produce an analysis in an afternoon that once required several colleagues and a week of work. That changes what the employee can attempt. It does not establish that the analysis is sound, that the employee understands it, or that the organization can act on it.

The practical decision for L&D leaders is where to change the work and where to build human capability. This is a work-design and capability-strategy problem. An AI course may help, but it cannot resolve an obsolete approval process, an unclear decision right or a manager who continues rewarding the old task mix.

What matters	What L&D should do
Performance and learning can diverge	Assess the quality of assisted work and the judgment people retain. A polished output is insufficient evidence of competence.
The surrounding system determines value	Define the outcome, task mix, authority and manager support before choosing a learning intervention.
Automation can remove developmental work	Identify which experiences build essential judgment. Replace those experiences deliberately when tasks disappear.

Recommended first move: Run a 90-day pilot across up to three AI-exposed roles. Select one consequential workflow per role, establish a baseline, agree the human-AI responsibilities, and test a small change. The business sponsor owns the result; L&D owns development and capability evidence.

Scale only when the workflow produces an agreed improvement, controls work in practice, and people demonstrate the capability required of them. Report capacity released separately from cash savings. Include review, rework and development time in the calculation.

Evidence boundary: Research shows gains in some settings and weaker results in others. The gap is the author's practitioner framework, not a validated scale or a universal finding. Evidence and limits appear on pages 5 and 20; sources on page 22.

ABOUT THIS REPORT

A practical guide to work, capability and responsibility

This report is written for L&D leaders deciding how to respond when AI changes the work their organizations need people to perform. It can also support discussions with operations, HR, technology, finance and risk colleagues.

Its starting point is the work itself: what must be achieved, who should make the decisions, and what people need to understand to carry that responsibility. The proposed tools help translate those questions into a small, testable operating change.

Read for	Pages
The central argument and evidence	4-5
Implications for work, expertise, experience and management	6-9
Learning, compliance, performance and governance	10-13
Diagnostic, operating model and task boundaries	14-16
90-day pilot, scorecard and decision worksheet	17-19
Evidence limits, leadership decision and references	20-22

How to use it

For a leadership meeting, begin with the executive summary and the worksheet on page 19. Use the diagnostic on page 14 to locate the actual constraint. Assign an owner to each unresolved decision before commissioning development work.

What the examples represent

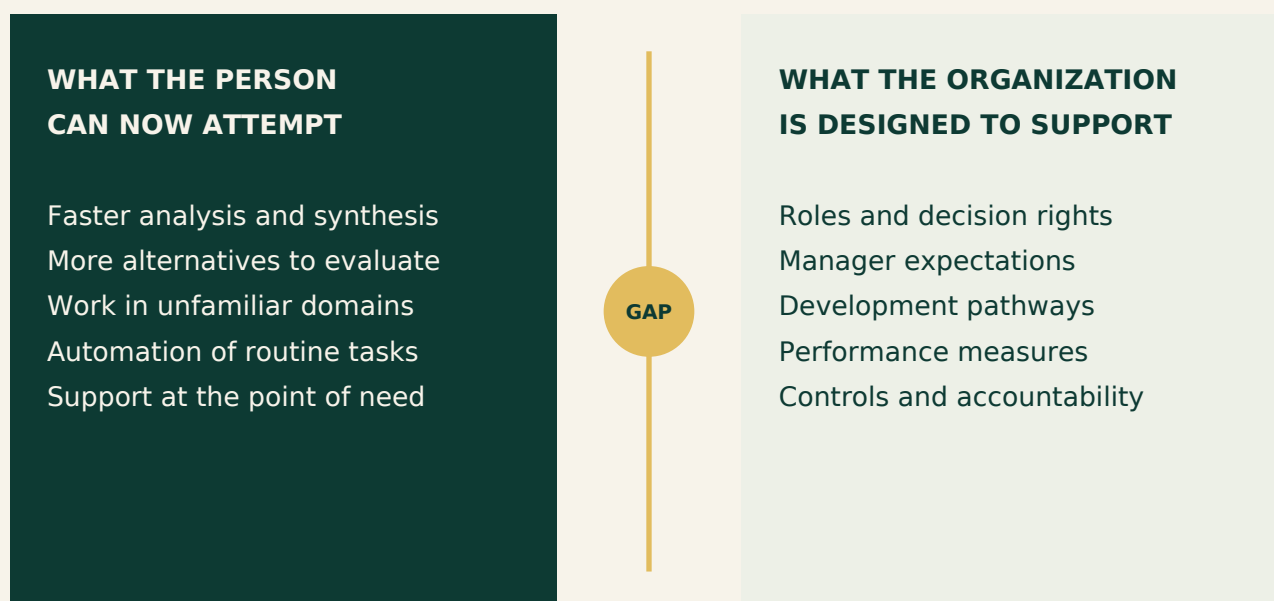
The analyst, management-reporting and compliance scenarios are illustrative composites. Their numbers describe hypothetical situations, not client results. The diagnostic, operating model and pilot gates are the author's recommendations; they are not research instruments or universal thresholds.

Prepared by Dr. Ravinder Tulsiani. Research checked 2 September 2026. This report complements *What Must Humans Remain Responsible For?* by examining the organizational conditions around AI-enabled capability.

Where capability meets the institution

Organizations encode assumptions about human capability in job descriptions, approval routes, staffing models and career ladders. Those assumptions need review when AI changes what one person can reasonably accomplish.

EXHIBIT 1 / THE CAPABILITY-INSTITUTION GAP



Conceptual model. The gap is a mismatch to investigate, not a measured distance or a claim that every organization is behind.

A gap exists when a changed capability encounters a rule, practice or expectation that no longer serves its purpose. Faster production may meet the same approval delay. Better access to information may meet a role that is still limited to gathering it. Automated junior tasks may leave no route to developing senior judgment.

THE L&D RESPONSE

Locate the constraint around the employee before prescribing more learning.

Some constraints should remain. Expert approval, separation of duties and practical experience can protect the organization. Review the reason for a control before changing it; speed alone is not a sufficient reason to remove it.

AI can help. The task and design still matter.

The research supports testing AI in real work. It does not support assuming that every user, task or organization will benefit equally.

Research finding	Boundary for interpretation
Customer support: A study of 5,172 agents found 15% more issues resolved per hour with AI assistance. Less-experienced and lower-skilled workers gained most. [1]	A deployment in one firm. The study also found evidence of learning, including during AI outages. It does not establish effects across occupations.
Occupational exposure: The ILO-NASK index estimates that roughly one in four workers globally is in an occupation with some GenAI exposure. [2]	Exposure is technical potential at task level. It is not a job-loss forecast. The authors judge transformation more likely than full replacement.
Learning: A high-school mathematics experiment found that unguarded AI assistance could improve practice results while weakening later unaided performance. Tutor safeguards mitigated that harm. [3]	This is education evidence. It motivates a workplace test; it does not prove that AI causes employee deskilling.
Software development: In a 2025 trial, 16 experienced developers took 19% longer on familiar repositories when AI was allowed. [4]	A narrow, dated setting. METR's 2026 follow-up could not reliably estimate the current effect because of selection and measurement problems. [5]
Skills: The OECD identifies skills shortages as an adoption barrier and highlights data interpretation, problem-solving and managerial capability. [6]	Survey associations and evidence synthesis do not establish a single causal effect or a uniform skill shift.

For an employer, the useful response is a baseline and a controlled local test. Measure the complete workflow, including verification and rework, and assess the person's required judgment separately from the output.

Numbered references link to the source list on page 22. Findings from different settings should not be combined into a single estimate of AI value.

Reconsider the job before building skills against it.

Adding AI to an existing task can save time while leaving the role unchanged. Work redesign begins with the next question: what should happen to the capacity that has been released?

A team might use that capacity to resolve a backlog, improve customer service, investigate exceptions or take on more complex decisions. It might remove an activity altogether. These are business choices. If nobody makes them, the organization may get faster production of work it no longer needs.

Decision	Example of a useful change
Stop	Retire a recurring report that no longer informs a decision. Do not automate it simply because it is easy to automate.
Automate	Generate a routine status summary from approved data, with proportionate checks and a named owner.
Assist	Use AI to develop options while an employee tests assumptions and recommends an action.
Retain human performance	Keep independent diagnosis or exception handling where the person must demonstrate that capability.
Add responsibility	Use released capacity for customer follow-up or harder cases, with suitable authority and support.

Begin the needs analysis one step earlier

Before accepting a request for AI training, examine the outcome, current task mix, bottlenecks and decision rights. Observe actual work alongside the job description. Ask employees where AI helps, where it creates review effort and which tasks still teach them something important.

L&D should bring the capability questions into this discussion. The business owner should approve the work change; HR should address role and workload implications. Development can then follow an agreed performance requirement.

Access to an answer does not establish competence.

AI lowers the cost of attempting unfamiliar work. Employees can ask for an explanation, draft a model or request a critique without first becoming specialists. The risk appears when plausible output is mistaken for a sound decision.

Domain knowledge still matters because it helps a person notice missing context, weak evidence and unusual cases. A generic instruction to “check the answer” is inadequate unless the reviewer knows what a material error looks like and has a credible way to test it.

That does not mean every employee must become an expert in every task AI assists. The required depth should follow the responsibility. Someone retrieving routine information needs a different standard from someone approving a consequential recommendation.

Responsibility	Capability to demonstrate
Use a routine answer	Locate the approved source, check applicability and recognize when the case falls outside the guidance.
Recommend an action	Explain assumptions, examine alternatives, check evidence and describe the likely consequences.
Approve consequential work	Identify material errors, challenge the model or recommendation, and justify approval against an agreed standard.
Handle an exception	Recognize when the usual approach is unreliable and escalate or intervene within the person's authority.

Assess judgment in context

Give employees an output containing a plausible mistake, incomplete evidence or a conflicting assumption. Ask what they would verify and why. Score the reasoning and the corrective action, not the fluency of the explanation. Match difficulty to the actual responsibility.

Prompting can be part of that development. It becomes more useful when paired with domain practice, credible reference material, feedback from a qualified reviewer and clear standards for acceptable work.

If AI does the early work, how will people become experts?

Many professional roles develop judgment through ordinary work: preparing an analysis, encountering an exception, receiving corrections and trying again. Automation can remove repetitive effort and the learning embedded in it.

The long-term risk is plausible, but its size is uncertain. The practical response is to identify which experiences build capability the organization still needs. Preserve the learning mechanism rather than keeping every manual task unchanged.

Illustrative scenario: the junior analyst

A new analyst uses AI to produce a market assessment, financial model and executive summary in two hours. Her manager likes the result. During review, however, she cannot explain the demand assumptions or identify which figures are estimates.

The polished document answers one question: can she produce a usable-looking output with assistance? It leaves a different question unresolved: can she judge whether the recommendation is dependable?

Development move	Evidence to look for
Frame the problem first	Before using AI, the analyst states the decision, key assumptions and evidence needed.
Compare, challenge and revise	She checks a material calculation, challenges an assumption and explains why she accepted or rejected an AI suggestion.
Practise a changed case	She responds to a new constraint or deliberately flawed output with proportionate support.
Review with an expert	The manager records what improved, what remains unreliable and which responsibility she can take on next.

Use independent tasks only where unaided capability is part of the role requirement. For other responsibilities, test performance with the approved tools and references. The objective is credible competence under the conditions in which the person will work.

Managers decide whether faster work becomes better work.

AI can reduce the time managers spend collecting information and preparing outputs. They still need to decide priorities, set quality standards, develop people and take responsibility for the team's results.

Illustrative scenario: six weekly reports

Six employees each produce a weekly operational report. AI reduces preparation time from several hours to minutes, but the manager still requests six documents in the old format. The team measures success by punctual submission.

The preparation saving may be real. Yet the organization has not established whether all six reports remain useful, whether review time has increased or what the team will do with the released capacity. The next improvement depends on the manager's decisions.

Manager action	Practical question
Clarify the decision	Which decision uses this report? Would a consolidated exception review serve it better?
Set the standard	What must be accurate, traceable and complete before the output is accepted?
Reallocate capacity	Which backlog, customer issue or improvement now receives attention? What will stop?
Build capability	Which employee should investigate an exception and explain the recommendation?
Address uneven support	Do all team members have suitable access, time, coaching and permission to use the agreed tools?

Develop managers through live work

Have each manager bring one workflow to a facilitated review. Require a specific change, a quality standard, a coaching plan and a follow-up measure. Review whether those actions occurred. A manager's course completion is weak evidence that the team's conditions changed.

Decide what people must know, find, do and judge.

A request for information does not always justify a course. Equally, an answer available through AI does not remove the need to understand or practise. The intervention should follow the performance requirement.

Need	Design response	Proof of capability
Know it	Build understanding and recall where the person needs them to recognize a situation or act without delay.	Explain or retrieve the essential knowledge under realistic conditions.
Find it	Provide approved, current information at the point of work, with a route for ambiguity and exceptions.	Find the right source, check its applicability and use it correctly.
Do it	Use practice, feedback and progression for tasks the employee must perform.	Complete a representative task to the required quality with the permitted support.
Judge it	Use difficult cases, competing evidence and expert feedback. Require a defensible decision.	Identify uncertainty, test assumptions, choose an action and explain the consequences.

A single role may require all four. An employee might retrieve a policy, remember a critical prohibition, practise a difficult conversation and judge when an exception needs escalation. Treating these as one content requirement produces an unfocused intervention.

Be selective about independent performance

Some knowledge must be retained because it supports error detection, urgent action or fallback operation. Other information can be reliably accessed. Do not require memorization merely because a course has always assessed it; do not remove it merely because a system can retrieve it.

AI-based support also needs ownership. Someone must maintain the approved source, test the answer against the work context and manage changes. L&D can help design and assess the experience, but an unmaintained information source remains an operational problem.

A completion rate cannot show whether people will act.

Illustrative scenario: an organization reports 98% completion of a 45-minute conduct course. Employees can also ask an approved AI tool to explain the policy. Neither fact shows whether they will act appropriately under pressure.

Consider a manager who asks an employee to bypass a control to meet a deadline. The employee understands the policy but expects consequences for delaying the work. More policy explanation may leave the performance problem intact.

Design around the moment that matters

What to examine	A practical response
Recognition	Use a realistic case with incomplete information. Ask the employee to identify the issue before presenting the rule.
Action	Practise a response to the manager, the documentation needed and the correct escalation route.
Environment	Confirm that the escalation route works and that managers support its use under operational pressure.
Evidence	Assess a new scenario and sample how actual cases are handled. Review quality and response time with the control owner.

The course may still be required. Keep any mandated elements and make them usable. Add practice where employees need judgment, and fix the operational barrier where knowledge is already sufficient. The control owner should decide the policy and evidence requirements.

Interpret behaviour carefully

A rise in reported concerns can indicate better recognition or greater confidence in reporting. It can also reflect more underlying problems. Read reporting volumes alongside case quality, severity, resolution and employee feedback. A lower number is not automatically a better result.

Time saved is capacity. Value depends on what happens next.

AI makes many outputs easier to produce. That weakens their usefulness as evidence of individual competence and makes output volume a poor substitute for business contribution.

Measure the complete workflow. Include prompt preparation, source checking, expert review, corrections and downstream rework. A fast draft that creates a slower approval process may move effort rather than remove it.

Illustrative weekly calculation	Hours
Preparation time released across six employees	12
Additional verification and rework	(3)
Protected practice and expert feedback	(2)
Capacity available for redeployment	7

Hypothetical example. All figures are assumptions for illustration, not a forecast or a reported result.

The seven hours become valuable when used for something the business needs: resolving more customer issues, clearing an important backlog or reducing delay. They become cash savings only when actual spending falls. An hourly cost applied to released time is a capacity valuation, not evidence of reduced expenditure.

Separate four kinds of evidence

Workflow: cycle time, accepted outputs, error severity and rework.

Business: the customer, operating or financial outcome the workflow serves.

Human capability: the judgment employees can demonstrate.

Adoption: access and use, interpreted as implementation information.

Finance and the business sponsor should agree how value will be recognized before the pilot. Include tool and integration costs, support, assessment and ongoing control effort. Avoid counting the same benefit once as time saved and again as the output produced with that time.

Give responsibility an owner and a usable control.

An employee may use AI to draft a recommendation, access sensitive data or initiate an action. These are different permissions. Organizations need to decide what is allowed before expecting employees to apply the rules.

NIST's AI Risk Management Framework calls for defined roles and responsibilities for human-AI arrangements and oversight. [7] That principle becomes useful when translated into specific authority, review requirements and escalation routes for the workflow.

Decision to resolve	Accountable owner to name
Outcome and task allocation	Business or process owner, with HR where roles or workload change.
Tool access and data use	The designated technology and data owners, with relevant privacy or security authority.
Approval, exceptions and stop conditions	The accountable business or control owner, with relevant risk and professional expertise.
Required human competence	The business owner sets the standard; L&D designs development and gathers capability evidence.
Ongoing supervision and practice	The line manager, with time and resources agreed by the sponsor.

Suggested ownership pattern. Apply the organization's actual mandates and name individuals for the pilot. L&D should not inherit decisions outside its authority.

Test whether oversight can work

A reviewer needs competence, evidence, authority, time and a route to intervene. If any is missing, a sign-off can create the appearance of control without its substance. Test the reviewer with a realistic error and observe what they can actually do.

For an agent that can act, define the permitted action, scope, approval trigger, activity record and stop or recovery procedure. A demonstration that the agent can complete a task does not establish permission to execute it in live operations.

Find the constraint before choosing the intervention.

Choose one workflow. Ask for observable evidence, not a general rating of AI readiness. Record an owner and a next action wherever the answer is unclear.

Area	Question	Evidence to bring
Work	What has AI made faster, possible or unnecessary?	Before-and-after work samples; complete cycle times.
Tasks	What should AI perform, assist or leave to people?	Task map; quality and consequence requirements.
Role	Does the responsibility still fit the changed task mix?	Job expectations; handoffs; new decisions.
Expertise	What must people understand to judge the result?	Material errors; reviewer criteria; required knowledge.
Experience	Which developmental tasks may disappear?	Early-career assignments; feedback and progression.
Management	What must the manager change?	Quality standards; coaching; capacity allocation.
Performance	What result matters beyond output volume?	Baseline; accepted work; customer or operating outcome.
Governance	Who may act, approve, stop and escalate?	Named authority; permissions; tested controls.

Three checks across the whole workflow

Access: Are differences in tools, language, accessibility, confidence or manager support limiting who benefits?

Evidence: What would demonstrate an improvement, and what result would make us stop?

Dependency: What happens when the tool, source or expert reviewer is unavailable?

Move from the request to the performance requirement.

L&D can keep strong instructional design while changing where it starts. The function becomes more useful when it helps establish the work requirement before it accepts a learning solution.

Step	Decision	Minimum useful output
1 / Work	What is changing?	A workflow with a specific problem and business owner.
2 / Performance	What result is required?	A baseline, quality standard and desired outcome.
3 / Capability	What must people know, find, do and judge?	A responsibility-based capability requirement.
4 / Experience	How will people develop it?	Practice, feedback and progression in realistic work.
5 / Support	What must change around the person?	Tools, guidance, manager action and clear authority.
6 / Evidence	Has the change worked?	Workflow results and demonstrated competence.

Use the smallest intervention that addresses the cause

A missing skill may need practice. A retrieval problem may need reliable guidance. An unclear permission may need a decision. A manager who blocks the new method may need a different expectation and follow-up. These causes can coexist; one course rarely resolves all of them.

Keep the operating model proportionate. A low-consequence task may need a short review and a simple check. A consequential workflow may need specialist review, stronger evidence and a tested fallback. The framework should help people decide, not become an additional approval queue.

Specify what AI may do and what the person must prove.

Illustrative workflow: a junior analyst prepares a recommendation for a business investment. Authority follows the organization's existing approval requirements.

Activity	AI role	Human responsibility and evidence
Frame the decision	Help explore possible questions.	Analyst states the business decision, constraints and success criteria before accepting a frame.
Gather and synthesize	Summarize approved sources.	Analyst traces material claims to current sources and identifies missing evidence.
Build the analysis	Assist calculations, comparisons and scenarios.	Analyst validates material assumptions and calculations; a qualified reviewer checks consequential judgments.
Recommend an option	Generate alternatives and challenge reasoning.	Analyst compares trade-offs and explains why the recommendation fits the situation.
Approve and act	Prepare a decision pack within permitted access.	Authorized owner accepts or rejects the recommendation. No autonomous commitment without explicit permission.

Make the capability test specific

Introduce a plausible source error or change an important assumption. Can the analyst spot the issue, quantify its effect and revise the recommendation? Can the approver explain what they relied on? Those are stronger tests than asking whether either person feels confident using AI.

Retain an activity record sufficient to explain the decision. For training or assessment, use approved, sanitized cases where appropriate. The amount of review should follow consequence and uncertainty, with clear triggers for expert involvement.

This task map illustrates the approach. It is not a financial recommendation or a universal allocation of authority.

Run a bounded pilot with three decision gates.

Choose up to three roles where AI already affects consequential work. Select one workflow per role and appoint a business sponsor. Start with one role if reviewer or development capacity is limited.

When	Work and ownership	Gate to continue
Days 1-30 Diagnose	Business owner and employees map the work. L&D identifies capability and experience requirements. Finance helps establish the value baseline.	Named owner; credible baseline; permitted tools and data; defined responsibilities; agreed outcome and stop conditions.
Days 31-60 Test	Manager implements a small workflow change. L&D provides practice and assessment. Relevant control owners test review, escalation and fallback.	Acceptable quality in representative cases; employees meet the capability standard; controls can be exercised.
Days 61-90 Decide	Business owner reviews full workflow results. L&D reports capability evidence. Finance reviews realized value; control owners report exceptions.	Scale, revise or stop using the criteria agreed at the start. Record the decision, owner and next review date.

Use a comparison that can answer the question

Where feasible, compare similar work with and without the changed approach, or introduce the change in stages. Account for case difficulty, worker experience and review standards. A before-and-after comparison is useful, but it cannot isolate the cause when several things change together.

Do not spend the same hour twice

Give employees time for practice and managers time for review. Budget those hours explicitly. A pilot that relies on unpaid effort or assumes all released time is immediately available will misstate both feasibility and value.

Stop or narrow the test if material errors go undetected, decision authority is unclear, required capability deteriorates or the workflow creates disproportionate review effort. Fix the specific cause before increasing scope.

Measure assisted results and the capability behind them.

Agree the threshold, sample and owner before the pilot begins. Use the same quality standard for the baseline and the changed workflow. Avoid a universal pass score for work with different consequences.

Dimension	Practical measure	Interpretation
Workflow performance	Accepted work per unit of time; total cycle time; rework and error severity.	Include preparation, review and correction. Faster drafting alone is incomplete.
Human judgment	Performance on a new case, changed assumption or seeded material error.	Use the tools normally permitted. Add an unaided check only where independent capability is required.
Business value	Backlog reduced, customer outcome improved or spending actually avoided.	Separate realized results from forecasts and released capacity.
Control reliability	Errors detected; escalation accuracy; intervention and recovery time.	Check that a qualified person can stop or correct the work.
Development and access	Progression to harder work; feedback received; results by access and support conditions.	Investigate unequal conditions before attributing differences to motivation or ability.

Read mixed results as a design signal

Assisted work improves; judgment does not

Keep the supported use within its approved boundary. Strengthen experience and feedback before assigning more independent responsibility.

Judgment improves; business results do not

Inspect handoffs, demand, approval delay and capacity allocation. Additional learning may not be the next useful investment.

Record the tool version and workflow conditions. Recheck after a material change, an incident or a shift in task difficulty. A successful pilot does not establish permanent readiness.

Make the next decision concrete.

Complete this with the business owner, a manager, employees and the relevant support functions. Take one workflow through the discussion.

WORKFLOW / SPONSOR

Name the work and the person accountable for the outcome.

PROBLEM / BASELINE

What result is inadequate today? What evidence supports that view?

PROPOSED CHANGE

What stops, becomes AI-assisted, is automated or gains responsibility?

AUTHORITY / CONTROLS

Who may act, approve, stop and escalate? What must be checked?

HUMAN CAPABILITY

What must people know, find, do and judge? What proves it?

EXPERIENCE / SUPPORT

What practice, manager action, access and guidance must be provided?

VALUE / COST

Where will capacity go? What result and full costs will be measured?

DECISION / OWNER / DATE

Pilot, revise or stop? Name the next action, owner and review date.

Before leaving the meeting: agree a success threshold, a stop condition and who will supply the evidence. An unresolved authority or control decision needs an owner before live use expands.

Treat the framework as a hypothesis to test.

What the evidence establishes

The studies cited here show that AI can improve some forms of work, that effects vary by setting, and that learning depends on how assistance is designed. Occupational exposure and skills research indicate substantial scope for change. They do not establish one trajectory for every worker or organization. [1-6]

What this report infers

The Capability-Institution Gap connects these findings to a practical concern: work expectations and supporting systems may lag behind what people can attempt with AI. This is the author's synthesis. It has not been validated as a distinct construct, diagnostic scale or causal explanation.

The strongest objection

An apparent institutional gap may be a sensible control or a response to weak technology, poor data or uncertain demand. A process that looks slow may protect a consequential decision. The evidence does not justify redesigning every role or assuming that more autonomy will improve results.

The response is to examine the constraint in context. If a control remains useful, improve its execution. If tool reliability is the problem, address that before changing responsibility. If learning is the limiting factor, provide it. Use the framework to test explanations, not to confirm a preferred AI strategy.

What could change the recommendation

Stronger tool reliability may justify changing the task allocation. Evidence of declining human judgment may require more protected practice. A rise in review cost may narrow the business case. A role may need deeper expertise even when its production work becomes easier.

Research limits worth keeping visible

The evidence combines workplace studies, occupational analysis, education research and policy synthesis. These methods answer different questions. Some findings are dated snapshots; some rely on self-report. The school experiment does not establish workplace deskilling, and the case scenarios in this report are not empirical evidence.

The long-term effects on employment, expertise and career progression remain uncertain. The recommended pilot is designed to produce local evidence without relying on a precise forecast.

Give L&D a work question before giving it a course request.

AI changes what employees can attempt. The organization still has to decide what work is useful, which responsibilities people should carry and how they will develop the judgment those responsibilities require.

For L&D, that is a practical mandate. Bring evidence about capability into decisions about work. Identify where learning will help and where another owner must change the conditions. Design experience where automation removes it. Assess whether people can handle the decisions they are being asked to make.

THE NEXT MOVE

Choose one workflow. Agree the human responsibility. Test the capability and the result.

Take three decisions into the next leadership meeting

- 1. Select the work.** Choose an AI-exposed workflow with a material outcome and an accountable sponsor.
- 2. Agree the change.** Specify task allocation, decision authority, manager support and the human capability that must remain dependable.
- 3. Require evidence.** Set a 90-day decision date with measures for workflow quality, human judgment and realized value.

The durable objective is an organization that can notice when work changes and adjust development around it. More AI use is useful only to the extent that it helps people and the business perform better under conditions they can sustain.

Dr. Ravinder Tulsiani

Enterprise Capability Executive
rtulsiani.com

SOURCE NOTES

Research and further reading

Links open the original publication or the research organization's report. Source numbers correspond to the evidence statements in this report.

[1] Brynjolfsson, E., Li, D., & Raymond, L. (2025). [Generative AI at Work.](#)

The Quarterly Journal of Economics, 140(2), 889-942. Workplace field study.

[2] Gmyrek, P., et al. (2025). [Generative AI and Jobs: A Refined Global Index of Occupational Exposure.](#)

ILO Working Paper 140; ILO-NASK. Occupational exposure analysis.

[3] Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakçı, Ö., & Mariman, R. (2025). [Generative AI without guardrails can harm learning: Evidence from high school mathematics.](#)

PNAS, 122(26), e2422633122. Field experiment in education; see the publisher's corrected record.

[4] Becker, J., Rush, N., Barnes, E., & Rein, D. (2025). [Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity.](#)

METR research report; randomized trial of 16 developers and 246 tasks.

[5] Becker, J., Rush, N., Cunningham, T., Rein, D., & Mahamud, K. (2026). [We are Changing our Developer Productivity Experiment Design.](#)

METR, 24 February. Follow-up explaining selection and measurement limitations.

[6] OECD (2026). [AI and skills: What we know so far.](#)

OECD Publishing, 5 June. Policy brief synthesizing research on skills and adoption.

[7] National Institute of Standards and Technology (2023). [Artificial Intelligence Risk Management Framework \(AI RMF 1.0\).](#)

NIST AI 100-1. See GOVERN 3.2 on roles and responsibilities for human-AI arrangements and oversight. Voluntary guidance.

How sources were used

Research findings are attributed where they appear. The gap model, diagnostic, work examples, operating sequence and pilot tools are practitioner synthesis. Illustrative examples are labelled. No combined effect size or universal return on investment is claimed.

Research checked 2 September 2026. The report uses the evidence appropriate to each claim rather than assuming that the newest available finding is the most reliable.