AMERICAN PSYCHOLOGICAL ASSOCIATION

Deploying AI Into Workplaces: What Leaders Need to Know

NOVEMBER 2026








Copyright © 2026 by the American Psychological Association. This material may be reproduced and distributed without permission provided that acknowledgment is given to the American Psychological Association. 750 First Street, NE, Washington, DC 20002-4242.

How to cite this report
American Psychological Association. (2026). *Deploying AI into workplaces: What leaders need to know*. https://www.apa.org/pubs/reports/ai-workforce

TABLE OF CONTENTS

How to use this resource 2

What this guidance covers 2

A short checklist for leaders 3

1. Stakes: Match safeguards to the stakes of the deployment 4

2. Trust and reliance: Help people develop and tailor their trust in AI tools over time 5

3. Accountability: Make human oversight meaningful rather than a formality 6

4. Job design: Redesign both work tasks and the job itself 7

5. Expertise: Protect the ways people build deep knowledge 8

6. Fair process: Introduce AI through a fair process 9

7. Monitoring: Respect psychological boundaries and set limits on monitoring 11

About this guidance 12



Organizations are adopting AI tools quickly, but success depends on more than the technology. It also depends on how people use it, how work changes around it, and how organizations manage those changes. Psychological science has something to say about each.

Problems can arise even when the technology is delivering real value to the organization. People may trust an AI tool where it is unreliable, or distrust it where it is reliable. The skills needed to supervise an AI tool can fade or fall out of date. It can become unclear who is responsible for checking the accuracy of AI output, and checking that output can take more time and effort than the tool saves. Psychological and behavioral research helps us anticipate many of these problems.

This guidance does not address whether an organization should adopt AI. It addresses what organizations should consider when AI is introduced into people's work. The evidence comes from research on AI and other automated systems, and from long-standing psychological research on how people make decisions, learn, work in teams, and perform at work.

The guidance is organized around seven objectives. Each identifies something an AI deployment has to get right for the people whose work it changes.

1. **Stakes:** Match safeguards to the stakes of the deployment
2. **Trust and reliance:** Help people develop and tailor their trust in AI tools over time
3. **Accountability:** Make human oversight meaningful rather than a formality
4. **Job design:** Redesign both work tasks and the job itself
5. **Expertise:** Protect the ways people build deep knowledge
6. **Fair process:** Introduce AI through a fair process
7. **Monitoring:** Respect psychological boundaries and set limits on monitoring

HOW TO USE THIS RESOURCE

This guidance is not a compliance standard, an APA professional standard, or an evaluation of any product. Instead, use it to guide decisions and check your reasoning. Specifically:

- Use the checklist in this document before a significant AI deployment and periodically afterward.
- Discuss the questions in this document with people who understand the technology and the work it will affect.
- Revisit the objectives in this document on a regular schedule and after major changes, such as a new system, a change in AI model, a major upgrade, or a reorganization.
- Which objectives matter most will depend on the organization's work, goals, and risks.
- Record the important decisions, who owns them, and when they will be revisited.

WHAT THIS GUIDANCE COVERS

This guidance addresses a central question: From the perspective of psychological science, what should organizations consider when bringing AI tools into the work of their employees, teams, and leaders?

Importantly, this guidance also focuses on people: their judgment, behavior, and well-being.

Two cautions apply throughout this guidance. First, effects will almost certainly differ across organizations, people, tasks, and levels of experience.[1,2,3] Second, AI tools and how they are used will continue to change. Deployment decisions should be revisited as the technology and the work change.

A short checklist for leaders

Consider this checklist when vetting a specific AI tool or reviewing one already in use, to support leaders' judgment and existing decision processes.

STAKES OF THE DEPLOYMENT

- How serious could an error from the AI tool be, and how quickly would someone notice it? Can errors be undone? Who bears the consequences?
- Where does the tool's output go? Could an error affect later decisions or spread across many cases or systems?
- If problems or errors emerge, who can narrow, pause, or stop its use—and how quickly?

TRUST AND RELIANCE

- Is it clear where the AI tool is reliable and where it is not, task by task and case by case?
- Have the people using the tool identified where and why it fails in actual work?
- When will the tool's reliability be checked again as the tool and the work change?

ACCOUNTABILITY

- Who is responsible for outcomes that depend on the AI tool? Can that person actually influence those outcomes?
- Do the people overseeing the tool have the time, expertise, independent information, authority, and support to act on concerns?
- Is there a clear expectation and/or policy for when people should disclose that they used AI, and does everyone know it?

JOB DESIGN

- Before the AI tool arrived, what did the work actually require beyond the formal task list, such as unwritten know-how or informal steps?
- What changes for the work and the workers when the AI tool is introduced?
- Do workloads, schedules, and metrics account for checking AI output and intervening when needed?
- Were expectations for the tool written down before adoption, and have they been compared with what actually happens once the tool is deployed?

EXPERTISE

- Which tasks are deliberately assigned to people instead of an AI tool, and why?
- How will the organization know whether employees are retaining and building the skills it needs?
- If the AI tool takes over a specific task or role, who will have the expertise to monitor it, now and in the future?

FAIR PROCESS

- Can employees explain, in terms of their own work, why the AI tool is being used?
- Were the intended users consulted before the decision to adopt it?
- Do employees see the tool's use and its outcomes as fair?

MONITORING

- Do people know what information is collected or inferred about them from an AI tool, as well as who can see it, how it may be used, and whether they can dispute and correct that information?
- Does the AI tool estimate people's emotions, personality, engagement, or other characteristics or states? If so, what evidence shows those estimates are accurate for the intended use? Either way, has the tool been evaluated across relevant differences among users and situations?
- How does the organization determine whether monitoring is achieving its intended purpose?

The 7 human objectives

The seven objectives below describe the human factors that contribute to whether an AI deployment succeeds. They are not steps in a sequence and don't need to be addressed in order.

Each objective below is a practical goal for the people responsible for deploying AI. Under each heading are the psychological and behavioral concepts it draws on, followed by suggested practices, what success might look like, and warning signs.

1. STAKES: MATCH SAFEGUARDS TO THE STAKES OF THE DEPLOYMENT

Relevant psychological and behavioral concepts: *Judgment under uncertainty, prospective thinking, signal detection, error detection and recovery, common-mode human failure*

Strength of evidence: *Well established for error detection, recovery, and the difficulty of forecasting rare high-consequence events. Emerging for correlated failure across an organization using the same AI system.*

What this means

Different uses of AI tools call for different safeguards. A tool that summarizes meetings needs lighter checks than one that makes consequential decisions or schedules maintenance on safety-critical equipment.[4,5]

Matching safeguards to the stakes of the deployment means thinking through consequences and scenarios that may lie ahead. People struggle with predicting outcomes that are rare, unexpectedly severe, or slow to build up.[6,7,8] So, for higher-stakes uses, write down the possible consequences instead of leaving it to intuition.

"How likely is this?" and "How bad would it be?" remain the basic questions about risk.[9] However, these four questions address considerations that are often overlooked.

1. **Who bears the consequences?** The right safeguards depend partly on who pays for a mistake: the organization, a worker, or someone outside the organization who had no say in using the tool. The stakes are different when the people who are affected are not the ones deciding how the tool is used.
2. **Can the consequence be undone?** Some mistakes can be easily corrected, while others can't. The more difficult a mistake is to fix, the more important it is to prevent the error in the first place. It's also important to consider that preventing one kind of error can make another more likely, and that not acting also carries consequences.
3. **How quickly would anyone know?** An error caught immediately is very different from one that goes unnoticed while it feeds later work. Because AI can operate quickly and at scale, the same problem can recur many times before anyone notices.
4. **How far could the consequence spread?** Output that feeds later decisions can harden into accepted fact that no one re-checks. A tool used the same way across many cases can also repeat the same error in all of those cases. Consider whether one mistake could affect a single case, hundreds of cases, or an entire system or process.

Suggested practices

- Answer the four questions in writing before deployment, with input from the people who do the work.
- Match oversight, training, review, and documentation to the stakes of the deployment.
- Build a quality control system. How will errors be detected, logged, corrected, and contained?
- Do not assume that success in a low-stakes pilot means the tool is ready for high-stakes use.
- Reassess the stakes when the tool, task, users, or volume of work changes significantly.

What effective deployment might look like

- ✔ The safeguards reflect what could go wrong and how difficult it would be to fix.
- ✔ The people responsible for the deployment can explain how an error would be noticed and how quickly.
- ✔ The stakes of the deployment are reassessed when the tool, task, volume, or users change.

Red flags to watch for

- ✖ The same review process is used for low-stakes and high-stakes cases.
- ✖ The same kind of error keeps happening without a change in how the tool is used or checked.
- ✖ Nobody can say how a particular kind of error would be noticed.
- ✖ A successful pilot is used to justify a substantially different use.
- ✖ The tool, task, volume, or users have changed, but the risk assessment has not.



2. TRUST AND RELIANCE: HELP PEOPLE DEVELOP AND TAILOR THEIR TRUST IN AI TOOLS OVER TIME

Relevant psychological and behavioral concepts: *Trust and reliance, automation bias, algorithm aversion, confirmation bias, mental models, metacognition, processing fluency, psychological safety, working memory, and attentional capacity*

Strength of evidence: *Well established for the underlying phenomena. Emerging for how individuals behave with AI systems specifically.*

What this means

Relying on a tool too much or too little (what human–automation researchers call misuse and disuse) carries costs, including safety, legal, ethical, and reputational risk. People who rely on a tool more than is appropriate tend to stop monitoring its output closely and may lose the ability to catch its errors. People who rely on it too little may miss errors the tool would have caught, or waste time re-checking acceptable work.[10,11,12,13,14] People must decide how much of their time and attention to spend checking the tool, and people with different skills and experience may decide differently. How much of that choice to leave to individuals is itself an organizational design decision. Where errors are costly or irreversible, the organization should require specific checks (see Objective 1). Where errors are minor and quickly caught, more individual judgment is reasonable. The goal should be trust that matches what the tool can actually do and that is updated as the tool changes and as people gain experience using it.

There are several behavioral patterns worth considering. First, people can mistake confidence and polish for accuracy.[15] Smooth, well-organized text is easy to read, and what is easy to read tends to be believable.[16,17,18] Second, a tool that does one task well may do another task poorly, and people often fail to adjust their trust from task to task.[19] Third, when a tool is usually right, people pay less attention to checking it, so its occasional errors are easier to miss. A tool that fails once in 100 cases can be harder to oversee than one that fails once in 10, because rare failures are easier to miss.[20] Fourth, chat-based AI tools sometimes agree with whoever is prompting the tool (a tendency called sycophancy).[21] A user can ask a leading question, get agreement, and grow more confident without learning anything new. The quality of the answers also depends on the user, because expertise affects both what people ask and how well they judge the answer. So a tool's reliability depends on the task, the user, and how it is used.[22,23,24]

How much people rely on a tool depends heavily on choices the organization makes. Leaders may not control how the tool is built, but they do decide how it is introduced, what people are told about its weaknesses, what must be checked, and what gets rewarded. Individual differences matter too,[25,26] though there is little evidence yet on how they play out with AI.

Training should show the tool failing as well as working well. Examples should include realistic conditions and both high and low-stakes tasks. Trust also should evolve as the tool, the work, and people's experience change.[27,28]

Suggested practices

- Test and explain where the tool is reliable and where it is not, task by task, context by context in the language used by the people who do the work. Update the explanation as the tool and the work change. Treat vendor accuracy claims as a starting point. They may not reflect your data, your tasks, or your users.
- Show people real examples of the tool failing. Demonstrations of what it can do may not provide a complete understanding. Seeing realistic errors gives people a more accurate picture. Encourage people to report failures they find, but recognize that people new to a task may not yet be able to tell a good answer from a plausible-sounding one. Seeing examples helps them develop that skill.

- Measure the tool against the standards that have traditionally defined good work. Be clear about what you are measuring and comparing: which task, and whether for individuals, teams, or units, and by what standard. Compare error rates and quality in addition to speed. Faster output that needs more correction later may not be an improvement.
- Make clear that the tool agreeing with you is not confirmation of accuracy. A tool that tends to agree with whoever asks is not a valid second opinion. Real checking means consulting a trusted independent source, an independent calculation, or an expert colleague who has not seen the tool's answer.
- Where the stakes warrant it, build checks into the workflow instead of relying on people to remember. Consider task-specific checklists rather than relying only on general reminders to be careful. Start by putting checklists in place where an error would quickly become costly or impossible to fix. Finding and reporting an error should be encouraged, not penalized.
- Schedule regular reviews of how people understand and use the tool rather than waiting for a mistake to force a closer look. The reason is twofold: As confidence grows, people tend to check the tool's output less closely, and the tools themselves keep changing.
- Expect people to vary in their use of AI. Someone who gets good results without the tool is still an effective worker. Their approach may be an example for others.
- Do not make adoption or usage rates the main goal. Usage targets can reward reliance over good judgment about when to use the tool and when to check it.



What effective deployment might look like

- ✔ People can name, without prompting, the tasks where they check the tool's output and the tasks where they don't.
- ✔ Errors are caught by the people using the tool, before the work leaves the team.
- ✔ People rely more on the tool for tasks it handles well and less for tasks it doesn't.

Red flags to watch for

- ✖ Errors are being detected by clients, auditors, or downstream teams.
- ✖ People keep prompting the tool until it provides the answer they want.
- ✖ People cannot name any task where they would not trust the tool.
- ✖ Usage is high, but almost no problems are being reported.
- ✖ People spend less time reviewing the tool's output even though the tool's performance has not changed.

3. ACCOUNTABILITY: MAKE HUMAN OVERSIGHT MEANINGFUL RATHER THAN A FORMALITY

Relevant psychological and behavioral concepts: *Meaningful human agency, diffusion of responsibility, role clarity, moral responsibility gaps*

Strength of evidence: *Well established for role clarity, diffusion of responsibility, and their consequences. Emerging for attribution and moral responsibility in mixed human-AI work.*

What this means

Having a person involved in the process may not equal true oversight. Oversight is more likely to be effective when the reviewer has:

- the time, energy, and attention to do the review;
- the expertise to tell a correct answer from a plausible-sounding but incorrect one;
- information from reliable sources other than the tool;
- the authority to overrule the tool; and
- a workplace that does not penalize them for overruling it.

Without all five, review risks becoming a formality.[29,30,31,32]

Time, energy, and attention are especially important. A reviewer is unlikely to be able to provide meaningful oversight if the workload leaves too little capacity to review the work. Further, when something goes wrong, the person closest to the work or assigned to review the work may be blamed even if they lacked the time, information, or authority to prevent the error.[33,34] This can make oversight feel more like risk and exposure than meaningful responsibility.

Suggested practices

- Decide in advance which decisions require human involvement—however confident the system appears—and explain why each decision is on the list. Update the list as tools and roles change.
- For each important workflow, write down what the AI may do, what a person must do, and who is responsible for the result. If you are unable to do this, the roles are not yet settled. Where workflows change too fast to document, decide instead who may deploy the tool, who may change what it does, and who owns the outcome.
- Check each oversight role against the five conditions above. If any condition is missing, either give the role what it needs or drop the requirement.
- Do not hold people accountable for outcomes they had no power to change. Where responsibility belongs higher up, provide clarity as to where the responsibility resides and how to access the responsible individuals. Applaud people who responsibly report errors and negative outcomes.
- When a system keeps acting after a person's involvement ends, define what it may do, what will stop it, and who is told if it stops. Make sure there's a plan for continuing critical work if the tool becomes unavailable, and that everyone understands it.
- Set a clear rule for when people must disclose that they used AI. Leaving it to individual judgment leads to inconsistent practice.
- Decide in advance if and when people may use AI tools the organization has not approved.

What effective deployment might look like

- ✔ For each important workflow, one named person is responsible for the result and has the authority to change it, with a trained backup.
- ✔ Reviewers sometimes change or reject the tool's output.
- ✔ People openly report when they have used AI, including when its use contributed to an error.

Red flags to watch for

- ✖ Nobody can identify who owns a decision the system is effectively making.
- ✖ The designated reviewer approves nearly everything, and everyone treats that as satisfactory.
- ✖ After an incident, responsibility falls on someone who did not have authority to prevent it.
- ✖ People are unsure when to disclose that they used AI.



4. JOB DESIGN: REDESIGN BOTH WORK TASKS AND THE JOB ITSELF

Relevant psychological and behavioral concepts: *Cognitive load, attention, situational awareness, hidden labor, job and work redesign, job demands-resources model, incentive alignment, change overload, job design*

Strength of evidence: *Well established for cognitive load, attention, and the effects of poorly designed job demand-resource balance. Emerging for the size and distribution of verification work in AI-assisted roles.*

What this means

Organizations often ask which tasks AI will replace. Research on work design suggests a better question: how will AI change the work people do? AI often changes work rather than simply removing it.[35,36,37] Searching, recalling, drafting, and routine thinking may shrink, while checking, combining information, monitoring, and handling the cases the tool does not manage may grow.[38,39] An adjuster who used to spend the morning assembling a claims file now spends it checking a file the tool assembled and handling the cases it couldn't. Whether that trade improves the job depends on what gets easier, what gets harder, and what is new.[40] Organizations tend to count the work that disappears and miss the work that is added, so both should be measured.[41] Not every task should be assigned based on just an evaluation of efficiency. Some activities are important to keep in human hands because they build relationships or responsibilities that are valuable to the organization's mission.

Redesigning existing work is harder than designing a new job. This is because existing work has informal routines,

unwritten steps, and workarounds that the people making deployment decisions often cannot see.[42] Unless there is an ongoing, competent effort to conduct job analyses, there may be no data recorded anywhere that explains the specifics. AI may also make new roles possible or necessary. Either way, plan for the AI-related human judgment, oversight, and checking that the work will require. In a new role, this workflow can be built in from the start. In an existing role, this workflow must be deliberately added.[43]

Suggested practices

- Divide work according to what people and technology each do well. Use technology to keep people's efforts focused on work they do best and work where humans outperform AI, which may include work that needs judgment, creativity, and collaboration. Some tasks should stay with people even when AI could do them, because they sustain relationships, meaning, or commitments central to the organization's mission.
- Before changing the work, map out how the work is done in practice. Include the small steps no one typically writes down, and design the deployment around that reality.
- Treat checking and fixing AI output as real work. Allow time for it, assign it to people with the expertise to catch mistakes, and account for this quality control in people's workloads and schedules.
- Ask the people who are doing the work what has changed for them. Look for changes in workload, mental demands, motivation, and the ability to keep track of what is happening, including differences by role, expertise, and career stage.
- Ensure workloads, incentives, and metrics support the work you expect. If oversight requires sustained attention, provide time for it.



- Check whether metrics reward the behavior you want. If careful review matters, do not reward speed alone or at all. Watch for whether metrics push some people toward disengagement, burnout, or leaving the organization.

What effective deployment might look like

- ✔ Plans, schedules, and staffing account for verification and repair work.
- ✔ People describe the work as different or even more enjoyable, easier, and less time consuming.
- ✔ Productivity gains show up in the metrics and are confirmed by the people doing the work.

Red flags to watch for

- ✖ Reported productivity is up and people describe the work as more tiring, draining, or difficult.
- ✖ AI handles the easier tasks, leaving employees with mostly difficult work.
- ✖ People are doing verification work after hours because there is not enough time for it during the workday.
- ✖ Every rollout is followed by another before people have had time to adapt.

5. EXPERTISE: PROTECT THE WAYS PEOPLE BUILD DEEP KNOWLEDGE

Relevant psychological and behavioral concepts: *Skill acquisition and decay, desirable difficulty, elaborative processing, novice-expert differences, cognitive offloading, deskilling, and out-of-the-loop performance*

Strength of evidence: *Well established for skill acquisition, practice, and decay. Emerging for the specific effects of AI tools on professional skill formation and change.*

What this means

Deploying AI can change how people build expertise.[44,45] The same tool can affect two people doing the same task in different ways.[46,47,48,49] Experienced workers may use the tool to skip effort they no longer need to expend, because they believe they already have the skills to judge the output. Newer workers may use the tool to skip the very effort that builds those skills. Both may look equally productive, but they may not be building the same ability to supervise the tool later.[50,51,52] Which way it goes depends on what people already know, what the task demands, and what the organization rewards. If speed is rewarded, people will likely take the faster path. The gap can widen over time, because getting good answers from these tools requires strong judgment (see Objective 2). Unless the organization makes room for learning, newer workers may get less from the tool today and may build less of the judgment they will need tomorrow.

People build skill by working through tasks that are difficult enough to stretch their current abilities. Researchers sometimes call this "desirable difficulty."[53] Difficulty alone is not enough. Learning requires the right kinds of challenges, such as attempting a problem before seeing the answer and practicing solving problems multiple times.[54,55] If a tool does too much of the work on tasks a person is still learning, that person may not develop the judgment needed to supervise it.

Because building a skill is hard and often uncomfortable, handing the work to an AI tool is tempting, and people face that choice many times a day, often under time pressure. Work design and incentives can make the harder path worth taking. Even people who can supervise AI tools well today will need to keep learning as the tools and their field change.

What makes sense for each individual can still lead to problems for the organization. When questions that once went to a knowledgeable colleague go to an AI tool instead, the answer may be just as good, but three things can be lost:

- colleagues lose the chance to explain things to others, which deepens their own understanding;
- the person asking does not learn who knows what, making it harder to connect people and tasks later; and
- the organization gradually loses track of where its expertise lies.

These effects can take time to appear. Leaders should make sure that handing work to a tool that does it faster, or even better, does not erode the expertise the organization will need later.[56,57] Expertise is an organizational asset that provides value and competitive advantage.[58,59] It is expensive to buy or rebuild, if it can be replaced at all.[60]

Suggested practices

- Decide which tasks people should keep doing themselves to build skill and explain the reason to them. This should be an organizational decision, not left to individual managers' preferences.
- Where people are still learning, have them attempt the work before seeing the tool's version, then compare. Starting from the tool's output may not build the same skill.
- Track whether people are building and maintaining the skills the organization needs.
- Before moving a task from a person to AI, ask what other skills depend on doing that task.
- Watch what happens to collaboration as AI use spreads. If colleagues consult each other less, consider what knowledge may no longer be shared.
- Preserve opportunities for people to approach problems in different ways. Shared tools can lead people toward the same approaches and answers, potentially eroding innovation.

What effective deployment might look like

- ✔ Newer staff can explain the reasoning behind work the tool helped produce.
- ✔ People who are still learning skills try the work themselves before comparing it with the tool's version.
- ✔ Colleagues still bring challenging questions to each other, and people know who to turn to for different types of expertise.

Red flags to watch for

- ✖ Senior staff report spending more time reviewing work that is longer and more polished, but not necessarily more accurate or carefully thought through.
- ✖ Nobody can independently evaluate whether the AI tool's output is correct or comprehensive.
- ✖ New staff get up to speed faster, but develop less skill.

6. FAIR PROCESS: INTRODUCE AI THROUGH A FAIR PROCESS

Relevant psychological and behavioral concepts: *Procedural justice (voice, transparency, consistency, correctability), psychological safety, job insecurity and role ambiguity, psychological contract, organizational change, attribution and contribution*

Strength of evidence: *Well established for procedural justice, psychological safety, and the effects of insecurity and role ambiguity. Emerging for disclosure norms specific to AI use.*

What this means

How a tool is introduced affects how people respond to it, regardless of how good the tool is.[61,62,63] A fair process gives the people affected a meaningful opportunity to raise concerns and influence decisions. That does not mean everyone gets what they want. It means their input is genuinely considered and can lead to changes in how, when, or whether a tool is deployed. It also does not mean using a process designed primarily to win acceptance. Asking for input after the important decisions have already been made is not meaningful consultation, is likely to be detected, and may backfire.[64,65,66,67,68]

People want to know why a tool is being introduced—and what they are told influences their appraisal of the tool.[69,70] "Efficiency" or "competitors are doing it" may not provide much of an explanation. Connect the tool to the organization's mission and the problems it can help solve. Explain what the tool is intended to accomplish and its limits. That gives people something concrete to evaluate and discuss.[71,72]

Job insecurity can also affect how people respond to a deployment.[73,74] People who are unsure about their jobs may take fewer risks, share less information, and watch anxiously for signs about their future. Saying nothing about how a deployment will change work may be heard as bad news. Transparency tends to build more trust than reassurance:

Share what you know, what you don't know yet, and when you expect to know more. Job changes can also affect professional identity, and most people do not see their roles as primarily checking a machine's work. This can matter even for people whose jobs are secure, particularly when they have spent years developing expertise in work that is now being automated.

People closest to the work often see consequences that decision makers may not perceive. Asking for their input can improve how the deployment works as well as how well it is accepted.[75] This means giving people a real chance to contribute before decisions are final, and recognizing the time that takes.

As AI becomes part of shared work, it gets harder to tell who contributed what. Sorting out credit can already be difficult on human teams[76] and can become even more difficult when AI produced part of the work.[77,78,79,80] This matters because performance reviews, credit, and promotions depend on knowing who contributed what.[81,82] Left unaddressed, this can make individual contributions more difficult to evaluate fairly and accurately.

Finally, deployment plans assume problems will be reported and fixed, which happens primarily when people feel safe reporting that a tool did not help, that it made an error, or that their own use of it caused a problem.[83,84] Otherwise the organization finds out too late.[85,86] Leaders should also follow their own rules on disclosing and checking AI use. What leaders do often shows others what is expected. Without safe reporting, accurate trust in the tool is unlikely to develop[87] (see Objective 2).

Suggested practices

- Explain the deployment in terms of the problem it is meant to solve and its limitations. If a new AI system cannot be explained clearly to those expected to use it, it should not be deployed.
- Consult all affected stakeholders before a deployment decision is made. This includes those who will be using the new system, those who will be affected by the outputs of that system, and external monitors and regulators.
- Say what you know about effects on roles, what you don't know, and when you will likely know more. Give people time to learn and adjust.
- Make it safe, and expected, for people to report tool limitations, tool failures, and user errors. Communicate clearly that newly deployed tools and systems are not expected to be perfect.
- Recognize that new work expectations to learn, oversee, and respond to new AI tools and systems are new job tasks. Adding AI can be asking workers to do more with the same existing resources. Communicate this understanding clearly to affected workers.
- Apply the same disclosure rules to leaders as to everyone else.
- Create a simple way to fix problems, and show people when and how something was fixed because someone raised a concern.

What effective deployment might look like

- ✔ People can explain why the tool is being used in terms of their own work.
- ✔ Problems get reported early, including problems caused by a person's own use of the tool.
- ✔ People can point to changes that were made because someone raised a concern.

Red flags to watch for

- ✖ Adoption numbers look fine, but people talk about the tool differently in private than in meetings with leaders.
- ✖ Use is expanding, but few or no problems are being reported.
- ✖ When asked why a tool was deployed, most people can only say that leadership wanted it.
- ✖ People are relying on rumors to understand how their roles may change.



7. MONITORING: RESPECT PSYCHOLOGICAL BOUNDARIES AND SET LIMITS ON MONITORING

Relevant psychological and behavioral concepts: *Privacy and information boundaries, autonomy and dignity, algorithmic control, electronic performance monitoring and surveillance, reactivity to measurement, psychological safety, procedural justice, validity of inferences about internal states*

Strength of evidence: *Well established for the effect of monitoring on behavior, perceived control, well-being, and performance. Emerging for the specific effects and validity of AI-enabled inferences about individuals' internal states.*

What this means

AI can rapidly collect detailed information about how people work. Monitoring can serve real purposes, such as catching a safety problem or showing where someone needs feedback. But more monitoring is not necessarily better, especially when people worry the data could be used to discipline or dismiss them.[88,89,90] Intensive monitoring often falls short of expected performance gains and can increase stress, lower satisfaction and trust, and invite misuse.[91] The key questions are what to monitor and in how much detail, for what purpose, who sees the data, what decisions it informs, and what people are told.[92,93,94]

Monitoring people for safety, such as detecting fatigue, differs from tracking keystrokes or webcam data to judge performance. The two call for different evidence and safeguards, and even safety monitoring can backfire if it is poorly explained.

Asking workers to consent to monitoring is a weak safeguard when the request comes from their employer. Better tests are whether people were clearly told what is collected and why, whether the data are needed for a stated purpose and used only for it, whether the amount of monitoring matches what is at stake, and whether people can see and challenge what is held about them.

Tracking output, errors, and throughput is often ordinary management. AI can greatly increase the amount and detail of the information available. It can also make monitoring more invasive. Systems that are marketed as reading emotion, engagement, personality, stress, or mental health from people's work may get it wrong and predict their behavior poorly, especially because felt emotions rarely predict behavior.[95] These assessments may be less reliable than they appear. Errors matter, especially when the results could affect a person's work reputation or employment status.[96]

Measuring behavior also changes it. Leaders should bear in mind that monitoring may also change how people behave and perform on the job, sometimes in unintended ways.

Suggested practices

- Tell people plainly what information is collected or inferred about them, why, who can see it, and what decisions it may inform.



- Know what information your AI systems store about individuals (summaries, profiles, stored memories, scores), including records created as a side effect of personalization. Collect only the data a stated purpose requires. Do not let tools infer people's emotions, personality, engagement, or mental health without a specific reason, and evidence that the inference is sufficiently accurate for the intended use. Inferences about mental health raise clinical, ethical, and legal questions beyond the scope of this guidance and call for particular caution.
- Keep monitoring for safety and quality separate from evaluating individuals, and tell people which is which.
- Match the intensity of monitoring to the stakes. Do not use monitoring data or AI-generated inferences for purposes beyond those originally stated without reviewing the new use and informing the people affected.

What effective deployment might look like

- ✔ Tracked metrics are clearly connected to important job functions.
- ✔ People can accurately describe what information is being collected about them.
- ✔ Monitoring data gathered for safety or quality improvement stays out of individual performance decisions.
- ✔ People keep reporting problems as monitoring expands.

Red flags to watch for

- ✖ Employees and the leaders who approved the monitoring cannot accurately describe what information is being collected.
- ✖ Data gathered to improve a process shows up in a performance review.
- ✖ A tool rates individuals' engagement, stress, or attitudes without a reliable way to check whether those ratings are accurate.



About this guidance

IN SCOPE

- AI tools that organizations deploy to help people with substantive work, such as drafting, summarizing, analyzing, scheduling, reviewing, supporting decisions, carrying out transactions, and monitoring operations.
- AI built into equipment, robots, and workflows, where it can set the pace of work, filter what a person sees, or prompt or block an action without the person asking it to.
- AI systems that can take actions without continuous human direction, including systems that carry out multiple steps before a person reviews, intervenes, or is notified.

A note on AI making decisions. Some AI systems make, or effectively make, important decisions. This guidance does not assume that a person should always make the final decision. Decades of research show that formula-based judgment can sometimes match or beat unaided human judgment.[97,98,99] For some tasks, research also supports combining human and AI judgment.[100]

Psychological and behavioral science raise questions that accuracy alone cannot answer. How decisions are made and explained can affect whether people view them as fair, whether they trust the process, and how they respond to it.[101] When a person remains responsible for reviewing an AI-supported decision, that person also needs real authority to intervene.

When AI makes decisions without human involvement, this guidance addresses the effects on people and organizations.

OUT OF SCOPE

Personnel selection and other formal human-resources decisions are outside the scope of this advisory and are addressed separately.

- Considerations and Recommendations for the Validation and Use of AI-Based Assessments for Employee Selection, Society for Industrial and Organizational Psychology, APA Div. 14
- Responsible Use of AI in Assessment, APA Committee on Psychological Tests and Assessment (CPTA)

LEVEL OF THE GUIDANCE

This resource sets out broad objectives for organizations deploying AI. It does not rate, endorse, or compare specific products or prescribe a single approach to deployment.

A NOTE ON THE EVIDENCE BEHIND THIS RESOURCE

This guidance rests on two kinds of evidence.

- **Well established.** Some of the guidance draws on decades of research on how people work with automated systems and how organizations affect behavior. Much of this research predates current AI, but the underlying human behavior is well established.[102,103,104] Examples of well-established areas of psychological and behavioral research referenced throughout this guidance include: trust and reliance on machines, automation bias, attention, out-of-the-loop unfamiliarity, cognitive load, procedural justice, psychological safety, models of job performance, team effectiveness, leadership and change implementation, the effects of job insecurity and role ambiguity, and workplace behavior.
- **Emerging.** Other parts of the guidance address effects being observed with current AI systems. The research on these effects is newer, so the guidance treats the findings as provisional, not settled.[105,106,107,108,109]

EXPERT PANEL

Rapid Expert Advisory Panel (REAP)

Tara Behrend, PhD
John Richard Butler II Endowed Professor, Michigan State University

Emily D. Campion, PhD
Associate Professor & Henry B. Tippie Fellow, University of Iowa Tippie College of Business

HeeSun Choi, PhD
Assistant Professor of Psychological Science, Texas Tech University

Joanie B. Connell, PhD
President, Flexible Work Solutions, Inc.

Nancy J. Cooke, PhD
Professor Emerita, Human Systems Engineering, Arizona State University

AIexis Fink, PhD
Director and Senior Research Scientist, Center for Effective Organizations, University of Southern California

Don Grant, PhD
National Advisor of Healthy Device Management, Newport Healthcare

Theodore L. Hayes, PhD
Research Psychologist

Louis Hickman, PhD
Assistant Professor of Management, Virginia Tech

Richard N. Landers, PhD
John P. Campbell Distinguished Professor of Industrial and Organizational Psychology, University of Minnesota Twin Cities

Brandon May, PhD
Assistant Professor of Forensic Psychology, Florida Institute of Technology

Fred Oswald, PhD
Professor, School of Education, University of California, Irvine

Emanuel Robinson, PhD
Director, Board on Human-Systems Integration, The National Academies of Sciences, Engineering, and Medicine

Ericka Rovira, PhD
Professor of Engineering Psychology, United States Military Academy

Mindy Shoss, PhD
Professor of Psychology, University of Central Florida

Sara B. W. Troutman, PhD
Managing Scientist, Human Factors, Exponent

Heng Xu, PhD
Professor of Information Technology, Analytics, and Operations, Mendoza College of Business, University of Notre Dame

APA Staff

Dennis P. Stolle, JD, PhD
Executive Lead Psychologist, Applied Psychology

Mark Chan, PhD
Director, Applied Psychology

Elizabeth Deegan, MLS
Director, Science Programs for Knowledge, Expertise, and Innovation

Corbin Evans, JD
Deputy Chief, Operations & Strategy, Senior Advisor to the Deputy CEO

Ludmila Nunes, PhD
Senior Director, Science Knowledge, Expertise, and Innovation

This guidance was prepared by APA staff with subsequent input from a volunteer Rapid Expert Advisory Panel (REAP) convened by the American Psychological Association. Panel members participated without compensation and were selected to provide relevant expertise and a diversity of perspectives from academia, industry, government, independent practice, and other organizational settings. Prior to appointment, panelists were asked to disclose potential conflicts of interest consistent with APA procedures.

The REAP provided expert input throughout the development and review of this guidance by APA staff, who coordinated the project and prepared successive drafts for panel review, discussion, and revision. The final guidance is based on relevant and current scientific evidence and professional expertise at the time of publication. Individual panel members may hold different views on particular issues or aspects of the guidance.

REAP members are identified with their professional affiliations for purposes of identification only. Their participation does not imply endorsement by, nor do the views expressed in this guidance necessarily represent the views of their employers or affiliated organizations.

No external funding was received to support the development of this guidance.

Use of generative artificial intelligence

Generative AI tools (models including: Claude Opus 5, Gemini 3.5, and Copilot) were used to support document development, including summarizing panel meetings and comments, drafting outlines, revising and condensing text, and assisting with formatting, copyediting, proofreading, literature searches, and reference formatting. Prompts consisted of task-specific instructions to summarize, draft, revise, or organize identified material.

All AI-generated outputs were reviewed, revised, and verified, and all cited sources were checked against the original publications.

SELECT REFERENCES

1. Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior, 123,* Article 106878. https://doi.org/10.1016/j.chb.2021.106878
2. Probst, T. M., Shoss, M. K., Bankins, S., Behrend, T. S., Dragano, N., Jetha, A., Koopman, J., Selenko, E., Tang, P. M., Tay, L., van den Broek, A., & Vietas, J. (2026). Aligning artificial intelligence with worker health, well-being, and safety: An occupational health psychology agenda. *Occupational Health Science, 10*(2), Article 15. https://doi.org/10.1007/s41542-026-00262-5
3. National Academies of Sciences, Engineering, and Medicine. (2022). *Human-AI teaming: State-of-the-art and research needs.* The National Academies Press. https://doi.org/10.17226/26355
4. Ehrhardt, N., & Utz, S. (2026). Trust formation in AI expert systems: The roles of domain and stakes in shaping trustworthiness and risk. *Journal of Media Psychology: Theories, Methods, and Applications.* Advance online publication. https://doi.org/10.1027/1864-1105/a000516
5. National Academies of Sciences, Engineering, and Medicine. (2025). *Human and organizational factors in AI risk management: Proceedings of a workshop*. The National Academies Press. https://doi.org/10.17226/29046
6. Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280–285. https://doi.org/10.1126/science.3563507
7. Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9
8. Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440–463. https://doi.org/10.1037/a0018963
9. National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1
10. Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies, 58*(6), 697–718. https://doi.org/10.1016/S1071-5819(03)00038-7
11. Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033
12. Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253. https://doi.org/10.1518/001872097778543886
13. Karpinsky, N. D., Chancey, E. T., Palmer, D. B., & Yamani, Y. (2018). Automation trust and attention allocation in multitasking workspace. *Applied Ergonomics, 70,* 194–201. https://doi.org/10.1016/j.apergo.2018.03.008
14. Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392
15. Berkowitz, J. S., Patock, J. R., Nawaz, A., Gonzalez-Hernandez, G., & Tatonetti, N. P. (2026). A crisis of overconfidence: Why confidence, not accuracy, is the real risk in clinical AI. *BioData Mining*, *19*(10). https://doi.org/10.1186/s13040-026-00518-4
16. Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition: An International Journal, 8*(3), 338–342. https://doi.org/10.1006/ccog.1999.0386
17. Reber, R., & Unkelbach, C. (2010). The epistemic status of processing fluency as source for judgments of truth. *Review of Philosophy and Psychology, 1*(4), 563–581. https://doi.org/10.1007/s13164-010-0039-7
18. Schwarz, N. (2018). Of fluency, beauty, and truth: Inferences from metacognitive experiences. In J. Proust & M. Fortier (Eds.), *Metacognitive diversity: An interdisciplinary approach* (pp. 25–46). Oxford University Press.
19. Skitka, L. J., Mosier, K. L., Burdick, M., & Rosenblatt, B. (2000). Automation bias and errors: Are crews better than individuals? *The International Journal of Aviation Psychology*, *10*(1), 85–97. https://doi.org/10.1207/S15327108IJAP1001_5
20. Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392
21. Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R., Cheng, N., Durmus, E., Hatfield-Dodds, Z., Johnston, S. R., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., Zhang, M., & Perez, E. (2025). *Towards understanding sycophancy in language models. arXiv*. https://arxiv.org/abs/2310.13548
22. Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153
23. Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors, 50*(2), 194–210. https://doi.org/10.1518/001872008X288574
24. Do Khac, L. T., & Leyer, M. (2026). Towards an integrative model of organizational human–AI collaboration: A semi-systematic review of the current state of the art. *Technology in Society, 84*, Article 103064. https://doi.org/10.1016/j.techsoc.2025.103064
25. Zhang, M. (2021). Assessing two dimensions of interpersonal trust: Other-focused trust and propensity to trust. *Frontiers in Psychology, 12*, Article 654735. https://doi.org/10.3389/fpsyg.2021.654735
26. Rovira, E., Pak, R., & McLaughlin, A. (2017). Effects of individual differences in working memory on performance and trust with various degrees of automation. *Theoretical Issues in Ergonomics Science, 18*(6), 573–591. https://doi.org/10.1080/1463922X.2016.1252806
27. Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*(3), 407–434. https://doi.org/10.1177/0018720814547570
28. Grimm, D., Gorman, J. C., Cooke, N. J., Demir, M., & McNeese, N. (2023). Dynamical measurement of team resilience. *Journal of Cognitive Engineering and Decision Making, 17*(4), 351–382. https://doi.org/10.1177/15553434231199729

29. Gonzalez, C., Donahue, K., Goldstein, D. G., Heidari, H., Jalali, M. S., Schelble, B., Singh, A., & Woolley, A. W. (2026). Toward a science of human-AI teaming for decision making: A complementarity framework. *PNAS Nexus*, *5*(3), pgag030. https://doi.org/10.1093/pnasnexus/pgag030
30. Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems*, *10*(4), Article 26. https://doi.org/10.1145/3419764
31. Langer, M., Baum, K., & Schlicker, N. (2025). Effective human oversight of AI-based systems: A signal detection perspective on the detection of inaccurate and unfair outputs. *Minds and Machines, 35*(1), Article 1. https://doi.org/10.1007/s11023-024-09701-0
32. Edmondson, A. C., & Bransby, D. P. (2023). Psychological safety comes of age: Observed themes in an established literature. *Annual Review of Organizational Psychology and Organizational Behavior, 10*, 55–78. https://doi.org/10.1146/annurev-orgpsych-120920-055217
33. Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886
34. Elish, M. C. (2019). Moral crumple zones: Cautionary tales in human-robot interaction. *Engaging Science, Technology, and Society, 5*, 40–60. https://doi.org/10.17351/ests2019.260
35. Dell'Acqua, F., McFowland, E., III, Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science, 37*(2), 403–423. https://doi.org/10.1287/orsc.2025.21838
36. Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior, 123*, Article 106878. https://doi.org/10.1016/j.chb.2021.106878
37. Seini, A. B., Adam, I. O., & Preko, M. (2026). Human-AI collaboration in information systems research: A systematic literature review and future research directions. *Human-Intelligent Systems Integration*. Advance online publication. https://doi.org/10.1007/s42454-026-00100-7
38. Simkute, A., Tankelevitch, L., Kewenig, V., Scott, A. E., Sellen, A., & Rintel, S. (2024). Ironies of generative AI: Understanding and mitigating productivity loss in human-AI interaction. *International Journal of Human-Computer Interaction, 41*(5), 2898–2919. https://doi.org/10.1080/10447318.2024.2405782
39. Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*(2), 381–394. https://doi.org/10.1518/001872095779064555
40. Behrend, T. S. (Ed.). (2025). *Human-technology partnerships at work*. Cambridge University Press.
41. Watts, L. L., Gray, B. E., & Medeiros, K. E. (2022). Side effects associated with organizational interventions: A perspective. *Industrial and Organizational Psychology, 15*(1), 76–94. https://doi.org/10.1017/iop.2021.93
42. Brown, J. S., & Duguid, P. (1991). Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. *Organization Science*, *2*(1), 40–57. https://doi.org/10.1287/orsc.2.1.40
43. Miton, H., & Jackson, J. C. (2026). Complex technology requires cultural innovations for distributing cognition. *Trends in Cognitive Sciences, 30*(2), 124–135. https://doi.org/10.1016/j.tics.2025.08.003
44. Hickman, L., & Langer, M. (2025). *Integrating generative AI into theories of job performance* [Manuscript accepted for publication]. *Research in Human Resource Management*. ResearchGate. https://doi.org/10.13140/RG.2.2.18113.01128
45. National Academies of Sciences, Engineering, and Medicine. (2025). *Artificial intelligence and the future of work*. National Academies Press. https://doi.org/10.17226/27644
46. Dell'Acqua, F., McFowland, E., III, Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science, 37*(2), 403–423. https://doi.org/10.1287/orsc.2025.21838
47. Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics, 140*(2), 889–942. https://doi.org/10.1093/qje/qjae044
48. Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior, 123,* Article 106878. https://doi.org/10.1016/j.chb.2021.106878
49. Probst, T. M., Shoss, M. K., Bankins, S., Behrend, T. S., Dragano, N., Jetha, A., Koopman, J., Selenko, E., Tang, P. M., Tay, L., van den Broek, A., & Vietas, J. (2026). Aligning artificial intelligence with worker health, well-being, and safety: An occupational health psychology agenda. *Occupational Health Science, 10*(2), Article 15. https://doi.org/10.1007/s41542-026-00262-5
50. Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–217. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x
51. Bainbridge, L. (1983). Ironies of automation. *Automatica, 19*(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8
52. Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*(2), 381–394. https://doi.org/10.1518/001872095779064555
53. Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition, 9*(4), 475–479. https://doi.org/10.1016/j.jarmac.2020.09.003
54. Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). MIT Press. https://doi.org/10.7551/mitpress/4561.003.0011
55. Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

56. Hickman, L., & Langer, M. (2025). *Integrating generative AI into theories of job performance* [Manuscript accepted for publication]. *Research in Human Resource Management*. ResearchGate. https://doi.org/10.13140/RG.2.2.18113.01128
57. National Academies of Sciences, Engineering, and Medicine. (2022). *Human-AI teaming: State-of-the-art and research needs.* The National Academies Press. https://doi.org/10.17226/26355
58. Parker, S. K., & Grote, G. (2022). Automation, algorithms, and beyond: Why work design matters more than ever in a digital world. *Applied Psychology*, *71*(4), 1171–1204. https://doi.org/10.1111/apps.12241
59. Selenko, E., Bankins, S., Shoss, M., Warburton, J., & Restubog, S. L. D. (2022). Artificial intelligence and the future of work: A functional-identity perspective. *Current Directions in Psychological Science*, *31*(3), 272–279. https://doi.org/10.1177/09637214221091823
60. Ployhart, R. E., Nyberg, A. J., Reilly, G., & Maltarich, M. A. (2014). Human capital is dead; long live human capital resources! *Journal of Management*, *40*(2), 371–398. https://doi.org/10.1177/0149206313512152
61. Klein, K. J., & Knight, A. P. (2005). Innovation implementation: Overcoming the challenge. *Current Directions in Psychological Science*, *14*(5), 243–246. https://doi.org/10.1111/j.0963-7214.2005.00373.x
62. Selenko, E., Bankins, S., Shoss, M., Warburton, J., & Restubog, S. L. D. (2022). Artificial intelligence and the future of work: A functional-identity perspective. *Current Directions in Psychological Science*, *31*(3), 272–279. https://doi.org/10.1177/09637214221091823
63. Bankins, S., & Formosa, P. (2023). The ethical implications of artificial intelligence (AI) for meaningful work. *Journal of Business Ethics, 185*(4), 725–740. https://doi.org/10.1007/s10551-023-05339-7
64. Leventhal, G. S. (1980). What should be done with equity theory? New approaches to the study of fairness in social relationships. In K. J. Gergen, M. S. Greenberg, & R. H. Willis (Eds.), *Social exchange: Advances in theory and research* (pp. 27–55). Springer. https://doi.org/10.1007/978-1-4613-3087-5_2
65. Greenberg, J., & Tyler, T. R. (1987). Why procedural justice in organizations? *Social Justice Research, 1*(2), 127–142. https://doi.org/10.1007/BF01048012
66. Greenberg, J. (1990). Organizational justice: Yesterday, today, and tomorrow. *Journal of Management*, *16*(2), 399–432. https://doi.org/10.1177/014920639001600208
67. Colquitt, J. A., & Zipay, K. P. (2015). Justice, fairness, and employee reactions. *Annual Review of Organizational Psychology and Organizational Behavior, 2,* 75–99. https://doi.org/10.1146/annurev-orgpsych-032414-111457
68. Cantarelli, P., Belle, N., & Hall, J. L. (2023). Information use in public administration and policy decision-making: A research synthesis. *Public Administration Review, 83*(6), 1667–1686. https://doi.org/10.1111/puar.13735
69. Gödöllei, A. F., & Beck, J. W. (2023). Insecure or optimistic? Employees' diverging appraisals of automation, and consequences for job attitudes. *Computers in Human Behavior Reports, 12*, Article 100342. https://doi.org/10.1016/j.chbr.2023.100342
70. Zhou, J., Lu, Y., & Chen, Q. (2025). GAI identity threat: When and why do individuals feel threatened? *Information & Management, 62*(2), Article 104093. https://doi.org/10.1016/j.im.2024.104093
71. Gagné, M., Parker, S. K., Griffin, M. A., Dunlop, P. D., Knight, C., Klonek, F. E., & Parent-Rocheleau, X. (2022). Understanding and shaping the future of work with self-determination theory. *Nature Reviews Psychology, 1*(7), 378–392. https://doi.org/10.1038/s44159-022-00056-w
72. Hermann, E., Puntoni, S., & Morewedge, C. K. (2025). GenAI and the psychology of work. *Trends in Cognitive Sciences, 29*(9), 802–813. https://doi.org/10.1016/j.tics.2025.04.009
73. Gagné, M., Parker, S. K., Griffin, M. A., Dunlop, P. D., Knight, C., Klonek, F. E., & Parent-Rocheleau, X. (2022). Understanding and shaping the future of work with self-determination theory. *Nature Reviews Psychology, 1*(7), 378–392. https://doi.org/10.1038/s44159-022-00056-w
74. Sverke, M., Hellgren, J., & Näswall, K. (2002). No security: A meta-analysis and review of job insecurity and its consequences. *Journal of Occupational Health Psychology, 7*(3), 242–264. https://doi.org/10.1037/1076-8998.7.3.242
75. Pasmore, W., Winby, S., Mohrman, S. A., & Vanasse, R. (2019). Reflections: Sociotechnical systems design and organization change. *Journal of Change Management, 19*(2), 67–85. https://doi.org/10.1080/14697017.2018.1553761
76. Mathieu, J. E., Gallagher, P. T., Domingo, M. A., & Klock, E. A. (2019). Embracing complexity: Reviewing the past decade of team effectiveness research. *Annual Review of Organizational Psychology and Organizational Behavior, 6*, 17–46. https://doi.org/10.1146/annurev-orgpsych-012218-015106
77. Berretta, S., Tausch, A., Ontrup, G., Gilles, B., Peifer, C., & Kluge, A. (2023). Defining human-AI teaming the human-centered way: A scoping review and network analysis. *Frontiers in Artificial Intelligence, 6,* 1250725. https://doi.org/10.3389/frai.2023.1250725
78. Puerta-Beldarrain, M., Gómez-Carmona, O., Sánchez-Corcuera, R., Casado-Mansilla, D., López de Ipiña, D., & Chen, L. (2025). A multifaceted vision of the human-AI collaboration: A comprehensive review. *IEEE Access*, 13, 29375–29405. https://doi.org/10.1109/ACCESS.2025.3536095
79. Wang, Y., de Paula, D., & Heavin, C. (2026). Unpacking human–AI collaboration: Conceptualisations and emerging research streams. *Journal of Decision Systems*, *35*(1), 1-19. https://doi.org/10.1080/12460125.2026.2653697
80. Schmutz, J. B., Outland, N., Kerstan, S., Georganta, E., & Ulfert, A.-S. (2024). AI-teaming: Redefining collaboration in the digital era. *Current Opinion in Psychology, 58*, Article 101837. https://doi.org/10.1016/j.copsyc.2024.101837
81. DeNisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology, 102*(3), 421–433. https://doi.org/10.1037/apl0000085
82. Kamoshida, A., Fang, W., & Fujii, T. (2027). Who creates value? Performance evaluation and knowledge attribution mechanisms in the age of AI-human collaboration. In L. Uden & I. Varlamis (Eds.), *Knowledge management in organisations (*pp. 208–219). Communications in Computer and Information Science (Vol. 3105). Springer. https://doi.org/10.1007/978-3-032-34189-1_17

83. Lee, S. E., & Dahinten, V. S. (2021). Psychological safety as a mediator of the relationship between inclusive leadership and nurse voice behaviors and error reporting. *Journal of Nursing Scholarship, 53*(6), 737–745. https://doi.org/10.1111/jnu.12689
84. Edmondson, A. C., & Bransby, D. P. (2023). Psychological safety comes of age: Observed themes in an established literature. *Annual Review of Organizational Psychology and Organizational Behavior, 10*, 55–78. https://doi.org/10.1146/annurev-orgpsych-120920-055217
85. Edmondson, A. C., & Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annual Review of Organizational Psychology and Organizational Behavior, 1*, 23–43. https://doi.org/10.1146/annurev-orgpsych-031413-091305
86. Morrison, E. W., & Milliken, F. J. (2000). Organizational silence: A barrier to change and development in a pluralistic world. *Academy of Management Review*, *25*(4), 706–725. https://doi.org/10.5465/amr.2000.3707697
87. Goodrich, M., McCabe, K., Basford, J., Bambach, K., Kraut, A., & Love, J. N. (2024). Psychological safety and perceived organizational support in emergency medicine residencies. *AEM Education and Training, 8*(2), e10964. https://doi.org/10.1002/aet2.10964
88. Connell, J. B. (2025). Unethical and harmful effects of artificial intelligence on human interactions and well-being: What organizational consultants can do. *Consulting Psychology Journal, 77*(2), 118–130. https://doi.org/10.1037/cpb0000284
89. Wolff, M. S., Ravid, D. M., & Behrend, T. S. (2025). Guidelines for the use of electronic performance monitoring. *Organizational Dynamics, 54*(1), Article 101026. https://doi.org/10.1016/j.orgdyn.2023.101026
90. Shoss, M. K., & Behrend, T. S. (2023). Technology, work, and inequality. *Technology, Mind, and Behavior, 4*(1), 116–120. https://doi.org/10.1037/tmb0000103
91. Ravid, D. M., White, J. C., Tomczak, D. L., Miles, A. F., & Behrend, T. S. (2023). A meta-analysis of the effects of electronic performance monitoring on work outcomes. *Personnel Psychology*, *76*(1), 5–40. https://doi.org/10.1111/peps.12514
92. Probst, T. M., Shoss, M. K., Bankins, S., Behrend, T. S., Dragano, N., Jetha, A., Koopman, J., Selenko, E., Tang, P. M., Tay, L., van den Broek, A., & Vietas, J. (2026). Aligning artificial intelligence with worker health, well-being, and safety: An occupational health psychology agenda. *Occupational Health Science, 10*(2), Article 15. https://doi.org/10.1007/s41542-026-00262-5
93. Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals, 14*(1), 366–410. https://doi.org/10.5465/annals.2018.0174
94. Connell, J. B. (2025). Unethical and harmful effects of artificial intelligence on human interactions and well-being: What organizational consultants can do. *Consulting Psychology Journal, 77*(2), 118–130. https://doi.org/10.1037/cpb0000284
95. DeWall, C. N., Baumeister, R. F., Chester, D. S., & Bushman, B. J. (2016). How often does currently felt emotion predict social behavior and judgment? A meta-analytic test of two theories. *Emotion Review, 8*(2), 136–143. https://doi.org/10.1177/1754073915572690
96. Xu, H. (2026). From the "Real Me" to the "Statistical Me": Why privacy protection must now govern data, representation, and action. *IEEE Security & Privacy, 24*(4), 76–80. https://doi.org/10.1109/MSEC.2026.3700429
97. Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press. https://doi.org/10.1037/11281-000
98. Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*(4899), 1668–1674. https://doi.org/10.1126/science.2648573
99. Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19–30. https://doi.org/10.1037/1040-3590.12.1.19
100. Gonzalez, C., & Heidari, H. (2025). A cognitive approach to human–AI complementarity in dynamic decision-making. *Nature Reviews Psychology*, *4*(12), 808–822. https://doi.org/10.1038/s44159-025-00499-x
101. Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, *86*(3), 386–400. https://doi.org/10.1037/0021-9010.86.3.386
102. Bainbridge, L. (1983). Ironies of automation. *Automatica, 19*(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8
103. Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230–253. https://doi.org/10.1518/001872097778543886
104. Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*(2), 381–394. https://doi.org/10.1518/001872095779064555
105. 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
106. Dell'Acqua, F., McFowland, E., III, Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science, 37*(2), 403–423. https://doi.org/10.1287/orsc.2025.21838
107. Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. *The Quarterly Journal of Economics, 140*(2), 889–942. https://doi.org/10.1093/qje/qjae044
108. Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior, 45*(2), 159–182. https://doi.org/10.1002/job.2735
109. Behrend, T. S. (Ed.). (2025). *Human-technology partnerships at work*. Cambridge University Press.


AMERICAN
PSYCHOLOGICAL
ASSOCIATION