



---

Digital Article / Developing Employees

## How Do Workers Develop Good Judgment in the AI Era?

AI is simultaneously increasing the need for judgment and destroying the experiences that produce it. *by David S. Duncan*

Published on HBR.org / February 3, 2026 / Reprint [H0929Y](#)



Illustration by Julie Guillem

**A few years ago, when I first** started experimenting with generative AI as a partner at a consulting firm, I noticed something that surprised me: It was helping me a lot more than it was helping my less-experienced colleagues.

In my own work, AI allowed me to move much faster toward familiar kinds of expertise-based outputs such as research synthesis and

analysis. The quality wasn't perfect, but it was directionally right and easy for me to steer.

For newer analysts, the experience using AI was very different. They produced output quickly, but it wasn't meaningfully better than before they employed AI. More strikingly, they often struggled to judge whether the output was any good at all, let alone how to redirect it. In some cases, they weren't even sure how to get started.

This wasn't what I expected. Like many people in 2022, I assumed generative AI would help less-experienced workers punch above their weight, that it would act like a sudden infusion of capability. Instead, I was watching the opposite happen: *The tools amplified existing judgment rather than compensating for its absence.*

I realize now that what I was seeing in miniature helps explain a broader unease that managers across industries are now experiencing when they receive AI-assisted work output: While it is produced quickly and looks polished, it's hard to assess its quality and, therefore, whether to act on it.

This makes sense when you look more closely at how people actually get good results when working with AI tools, which requires constant evaluation of what it's giving you and thoughtful iteration. You prompt AI, get a result, assess, then re-prompt it, making small course corrections as you go.

Historically, that kind of judgment does not come from using AI. It comes from having done work similar to what you're now using AI for and learning from the sometimes poor, slow, imperfect experience of repetition, with real responsibility for the outcomes sharpening your focus.

This leads to a paradox organizations are now confronting, whether they realize it or not: *AI simultaneously increases the need for judgment and erodes the experiences that produce it.*

Harvard Business School's Amy Edmondson and Tomas Chamorro-Premuzic of Russell Reynolds have warned of the risks of using AI to replace entry-level roles entirely. This article focuses on what happens when junior roles remain, but their development value changes fundamentally.

First, let me fully explain what I mean by *judgment*.

### **What Is Judgment?**

Judgment is one of those traits leaders invoke constantly and define rarely. It is often a gating factor for promotion; it becomes increasingly more important the higher someone rises. I've been on both sides of that equation: needing to demonstrate judgment to advance and later looking for it when evaluating others.

Judgment can be defined as *the capacity to act wisely in situations where rules by themselves are insufficient*—by recognizing what matters most in a situation, weighing competing priorities and tradeoffs, anticipating consequences, and deciding when to personally own a choice under uncertainty.

This definition implies that when leaders talk about judgment, they are usually bundling together several distinct forms of it. In practice, it shows up in at least five ways:

1. **Evaluative judgment:** Recognizing whether something is good or bad, strong or weak, appropriate or off base

2. **Contextual judgment:** Knowing when general rules apply and when the situation is different enough to require an exception
3. **Tradeoff judgment:** Weighing competing objectives when no option is clearly right
4. **Anticipatory judgment:** Seeing second-order consequences before they materialize
5. **Ownership judgment:** Deciding when to personally own a decision and its risks rather than deferring or escalating under uncertainty

As a simple illustration, consider the decision of whether or not to greenlight a new product launch. Two leaders might review the same data, understand the same risks, and agree on the likely upside. Yet one greenlights the launch while the other hesitates. The difference lies in judgment: The first leader sees the supply chain risk as manageable and believes a six-month delay would cede too much ground to competitors; the second worries that a quality issue at launch could damage brand trust in ways that take years to repair. Neither is wrong, they're just weighing the same tradeoffs differently based on pattern recognition built over years watching similar decisions play out.

These forms of judgment are not typically taught directly. Rather, they emerge from the structure of work itself.

### **Where Judgment Has Traditionally Come From**

In most organizations, judgment emerges as a byproduct of how work itself is structured. For example, in professional services firms, junior consultants have responsibility for tasks such as conducting research, synthesizing insights from it, building models, and preparing communications. In my experience in that role, early attempts at all of these were flawed and incomplete. But I received feedback and coaching from my managers, improved, and gradually gained judgment as to what high-quality outputs looked like.

This pattern repeats itself at each new level of the consulting firm hierarchy, from analyst to manager to partner. As you rise in the firm, the nature of your responsibilities shifts toward leading projects, managing and developing people, building client relationships, and building the organization. However, judgment develops in similar ways: You're given new responsibilities that stretch you; you try, sometimes fail, get feedback, and learn from the experience.

The pattern is not limited to professional services; it is a common way to develop talent in every type of organization across all functions. To take just two examples (both of which are in areas where AI is already having a big impact):

### **Product Management**

Historically, junior product managers (PMs) learned judgment by writing rough product requirement documents, making tradeoffs explicit, defending prioritization decisions about features to develop, and watching what happens after launch. Today, AI can generate in minutes all of these things: product specs, user stories (descriptions of a product feature, written from the end-users' perspective to define what they want to achieve and why), roadmaps, and more. New PMs now often only review this output instead of originating it.

### **Marketing**

Much of the standard work of marketing teams (and of agencies) like generating ad copy and visual assets, as well as designing campaigns, can now be generated by AI very quickly and at scale (meaning an essentially endless set of options to choose from can be generated). This output gets better as it is trained on previous company campaigns, strategy documents, and competitor activity. This means today's generation of junior marketers don't necessarily have to do all the hard work of creating those things for themselves, thereby missing out on what is a traditional development rite of passage.

Two crucial features of this system are real ownership over decisions and repetition. There's nothing like being accountable for the consequences of a decision to sharpen your judgment. Repetition matters, too, so that by the time you're asked to evaluate the work of others, you've struggled to produce it many times yourself.

None of this is especially efficient, but it works. This system, however, is built on the assumption that humans would continue to perform the core tasks through which judgment is built. In the AI era, that assumption no longer holds.

### **The Paradox: Judgment Without Experience**

This leads to the paradox organizations are now encountering: To use AI effectively, people need judgment about the task at hand, but as AI takes over more of the work, the very experiences that once produced judgment start to disappear.

For individuals, the shift manifests itself in how they experience work and develop their skills. AI-generated output arrives quickly and with a tone of confidence; reviewing and refining that output can feel like meaningful involvement from the human working with it. But because the person is now reacting to a generated answer instead of creating something from a blank page, the cognitive work and related skills developed are quite different.

For organizations as a whole, this effect compounds. Entry-level roles lose the messy, formative tasks that once served as training grounds, leading to mid-level managers being asked to oversee work they never fully learned to do themselves. Senior leaders find that fewer people are capable of stepping into roles that require real judgment under uncertainty.

One downstream consequence of this dynamic was described in an [HBR article](#) on what its authors labeled AI-generated “workslop”: work that looks polished but lacks the substance or context required to move decisions forward. Workslop degrades productivity by causing confusion and rework because fewer people are able to recognize its limits or correct it.

### **Why “Human in the Loop” Isn’t Enough**

Faced with concerns about over-automation, many organizations have converged on a reassuring principle: keep humans in the loop. This takes a variety of forms, including formal reviews for certain types of outputs or decisions, and is an important principle that does reduce risk. However, it does not solve the deeper problem of creating conditions under which judgment can develop.

Other common AI safeguards, while valuable, can inadvertently make this problem worse. Consider escalation protocols, which are rules that require junior employees to pass ambiguous or high-stakes cases up to supervisors to make decisions. These protect quality in the short term, but they also teach newer employees that uncertainty is something to hand off rather than work through. Over time, the organization becomes increasingly dependent on a shrinking group of seasoned decision-makers, while the pipeline of people learning to handle ambiguity narrows.

The fundamental issue is that most “human in the loop” designs optimize for control and for managing known risks rather than the development of those humans who are in the loop.

### **A Concrete Example: Editorial Judgment**

To see what is at stake, consider the work of an expert magazine editor.

When a draft lands on an editor's desk, their task is to evaluate its quality across many dimensions, such as its flow, argument, evidence, tone, audience fit, originality, practical value, and others. They might make all of these assessments simultaneously, while weighing whether the overall quality is promising enough to merit eventual publication. If it is, they'll note how, on each of these dimensions, it needs to be improved to make that possible. They'll also flag what might be missing from the article that needs to be added.

That judgment was built over years of working through messy drafts, providing feedback, going back and forth with aspiring authors on their work, making calls on what to publish, and, importantly, seeing how readers responded. Great editors are often talented writers themselves; seeing their own work go through the editing process has also helped build their judgment. While some of this expertise can be formulated in explicit rules, in other cases it might be tacit or have become so familiar it feels like intuition when it's applied.

Now imagine asking someone without that background to evaluate the same draft. Although the person might be highly intelligent, diligent, and able to identify surface-level issues, they will struggle to assess what truly matters. They won't know which flaws in the submitted draft are fixable or how the piece fits into the broader editorial agenda. Their feedback will tend to be cautious or misdirected, simply because they lack the experience needed to provide it.

This gap illustrates the problem all organizations are now facing with AI. Because AI is so blazingly fast at producing outputs that seem high-quality across so many diverse domains, its human prompters are essentially in the position of an inexperienced editor: They have to review and evaluate output without the lived experience required to form great editorial judgment.

## **The Organizational Risk Leaders Are Missing**

Viewed in isolation, the erosion of judgment can look like a skills issue or a temporary growing pain. Viewed systemically, it presents a leadership and succession risk.

Organizations depend on a steady replenishment of people who can make sound decisions under uncertainty. Historically, that replenishment happened as people moved through roles and accumulated experience with increasingly consequential choices.

As AI automates formative work, fewer people will encounter the situations that once served as training grounds for judgment, with the result that it becomes concentrated in a smaller group of senior leaders whose experience predates widespread automation. Ultimately, succession pipelines will thin out, as many in the next generation of leaders will lack the critical ability to make decisions under the increasingly common conditions of novelty or ambiguity.

There is also a subtler cultural effect. When people are removed from the type of ownership described here early in their careers, the emphasis of their learning will naturally shift toward how to manage upward at the expense of learning how to decide.

These are real risks posed by AI and are distinct from other commonly cited AI-related risks.

## **Getting Started: Rebuilding the Conditions for Judgment**

The challenge organizations now face is how to redesign work in the AI era so that judgment can still develop. A good starting point is to ask diagnostic questions that highlight where, for a given role, decisions get made and the knowledge and skills required to make them. These questions include:

- Who is actually making consequential decisions, and who is merely reviewing work shaped by others or by machines?
- Where do people experience the downstream effects of their choices, including failures?
- Which roles have lost the repetitive, low-stakes tasks that once built judgment over time?
- Where are people being shielded from ambiguity rather than asked to wrestle with it?

These questions provide insight into where the conditions for judgment formation still exist and where AI has removed people from such moments. For the latter, other ways to foster development must be put in place.

While such mechanisms are a source of active research, inspiration can be drawn from domains where gaining real-world experience is too costly or dangerous, such as medicine and the military. Both have long built judgment deliberately through mechanisms like case-based learning, simulation, gradually increasing individuals' responsibilities as they prove their judgment at each level, and structured, post-action reflection.

The defining challenge of the AI era is how to continue to produce people capable of exercising judgment.

*This article was originally published online on February 3, 2026.*



**David S. Duncan** is a partner at the consulting firm [Disruptive Edge](#), where he advises organizations on growth strategy, innovation, and AI adoption. He is the author of [The Secret Lives of Customers](#) and coauthor (with Clayton M. Christensen, Karen Dillon, and Taddy Hall) of [Competing Against Luck](#).