



Steven Shorrock
Editor in Chief of *HindSight*

WHO ARE WE TO JUDGE? FROM WORK-AS-DONE TO WORK-AS-JUDGED

We all have a habit that we are hardly even aware of; we judge others' work performance, every day, throughout the day. Whether it's the work of people in other organisations, in other parts of our organisation, in our own immediate work environment, when driving home, or at home, we evaluate, appraise and judge others' performance. We don't pay much attention to how we judge, but we ask ourselves all sorts of questions: "Did they do a good job?" "Did they work with due care and attention?" "Would I have done that?" I call this 'work-as-judged', and it has several characteristics that we should bear in mind.

1. We judge in a variety of ways

When it comes to unwanted events, judgement is expressed in various ways, whether formal and planned, or informal and spontaneous. For the most serious unwanted events, work is judged in inquiries, judicial proceedings, court judgements, and media reports. In less serious cases, it may be via investigation reports, audits, or management decisions. But judgements about work are also expressed in private opinions and conversations, now often displayed semi-permanently on social media posts.



2. We judge work on different criteria

Depending on our role and the situation, we emphasise different criteria when judging work. But there are two fundamental criteria, which Erik Hollnagel called the 'efficiency-thoroughness trade-off'. For a safety-related event, the focus is likely to be on thoroughness (e.g., "You were not careful enough") and then perhaps competency (e.g., "You are not skilled or knowledgeable enough"). But where productivity is in question, the focus turns to efficiency (e.g., "You didn't work quickly enough"). To paraphrase Hollnagel, the message is too often that "you should be efficient, unless something goes wrong, in which case you should have been thorough".

3. We judge work indirectly

Once work-as-done is done, it's gone.

The activity cannot be recorded completely; much work is done in the head and so it's not even open to inspection. When making judgements, we therefore use 'proxies' or substitutes for work-as-done. These include work-as-disclosed (e.g., written incident reports, interviews), work-as-observed (e.g., competency checks, video recordings), work-as-measured (e.g., data logs), work-as-analysed (e.g., investigation reports, just culture algorithms), and – crucially – pre-existing or developing ideas about work: work-as-imagined.

We piece these proxies together to form a mental representation of the work in our minds. Essentially, we judge work based on our imagination of what happened, how, and why. But, of course, we need a standard of performance, usually from work-as-prescribed (e.g., procedures) or even normative work-as-imagined (how we think things ought to be done).

Again, these are not real work; they are proxies. But the fidelity of these proxies – how faithfully they really depict work-as-done – strongly affects work-as-judged. The more partial, biased or out-of-context the measures, recordings, or statements, the less just the judgement.

4. We judge work partially

In practice, we form judgements about work from limited fragments of information about work (in recordings, interview notes, etc.). These fragments are usually close in time and space to the outcomes that follow. We tend to see faults in 'sharp end' work as 'causal'. Work at the 'blunt end', and work that was done days, weeks, months or years ago, is not subject to much evaluation. This is partly because there is little that is recorded that can be evaluated, and partly because we can't see the relevance or the influence, let alone prove causation. And much of the crucial context of blunt end work is not recorded. For instance, there are flight deck recorders, but not office recorders, and even voice recorders only record part of the activity.

"In practice, we form judgements about work from limited fragments of information about work."

5. We judge work differently

Work-as-judged changes depending on the contexts of judgement. We judge work differently from one another. We even judge work differently over time. For instance, what we judge as acceptable work performance changes with the shifting personal, social, cultural, and societal contexts (e.g., values, attitudes, and norms) in which we make judgements. The informational and technological contexts also affect how work is judged. For instance, computer logs and other recordings provide information that will affect our imagination of what happened. And the time available affects judgement. Work-as-judged is just as susceptible to time pressure as work-as-done, and work may be judged differently at different points in time as consequences and evidence unfold.

6. We judge work via a range of heuristics and biases

Work-as-judged is affected by a range of heuristics and biases. Let's take just ten biases and heuristics relevant to justice, which seem to have a reasonable evidence base:

1. We tend to judge a decision based on the eventual outcome instead of the quality of the decision at the time it was made (outcome bias).
2. We tend to disregard probability when making a decision under uncertainty (neglect of probability).
3. We tend to judge harmful actions as worse, or less moral, than equally harmful omissions (omission bias).
4. We tend to be overconfident in the accuracy of our judgements (overconfidence effect).
5. We tend to believe things because many others do (bandwagon effect).
6. We tend to search for, interpret, focus on, and remember information in a way that confirms our preconceptions (confirmation bias).
7. We tend to believe that events were predictable at the time that they happened (hindsight bias).
8. We tend to believe previously learned misinformation even after it has been corrected (continued influence effect).
9. We tend to believe that a statement is true if it has been stated multiple times (illusory truth effect).
10. We tend to draw different conclusions from the same information, depending on how that information is presented or 'framed' (framing effect).

In short, we tend to think we are rather objective in our judgement (and certainly more objective than average), but we are not (naïve realism).

7. We judge work in a way that is influenced by language and tools

Related to the framing effect, language and tools have a strong influence on judgement. For instance, the safety literature is awash with negatively framed vocabulary and concepts, such as 'human error', 'unsafe act', 'risk-taking',

and 'violation'. Similarly, safety tools (including taxonomies) for judging work are mostly deficit-based – classifying what went wrong, and not just what went on. And so, our language and our tools guide us to look for the specific ways in which people mess up, while ignoring the overall nature, context and history of work performance.

8. We judge work in a way that is influenced by our profession

Our professions also distort judgement via so-called 'déformation professionnelle' – a sort of job conditioning or occupational acclimatisation. People of different professions (such as safety specialist, competency examiner, prosecutor) attend to, perceive, understand, and judge the same work differently. They have different purposes, pay attention to different things, have different knowledge about work, and use different criteria for judgement.

9. We judge work because we have to

We have to judge work conduct for all sorts of reasons. In organisations, we do this for reasons associated with our own function (ops, engineering, HR, safety, quality, etc.), and in ways that are characteristic of our own function. Regulators, supervisory authorities, investigatory bodies, the media, and the courts judge work...and they must. Front-line staff also judge each other's work and, on a day-to-day basis, this is usually the judgement they fear the most.

"People of different professions attend to, perceive, understand, and judge the same work differently."

But in judging performance, it is important to bear in mind some basic realities about the nature of human performance. ICAO's (2021) Human Performance Principles give us a good

start in this respect.

- Principle 1: People's performance is shaped by their capabilities and limitations.
- Principle 2: People interpret situations differently and perform in ways that make sense to them.
- Principle 3: People adapt to meet the demands of a complex and dynamic work environment.
- Principle 4: People assess risks and make trade-offs.
- Principle 5: People's performance is influenced by working with other people, technology, and the environment.

And now for a crucial final point to bear in mind: each of these principles also applies to the judgement of work. Just as our work performance is variable, so is our judgement of work performance, and for the same sorts of reasons.

The truth is that we can never fully understand work-as-done in a complex situation. We can only construct an understanding. Since this is the basis for judgement, we must remain humble in the knowledge that judgement of work is subject to the same underlying principles of human performance as the work being judged. Mindful of this, we can strive for insight into how and why we judge, in the courtroom, board room, ops room, and living room. ↗

