

Interview

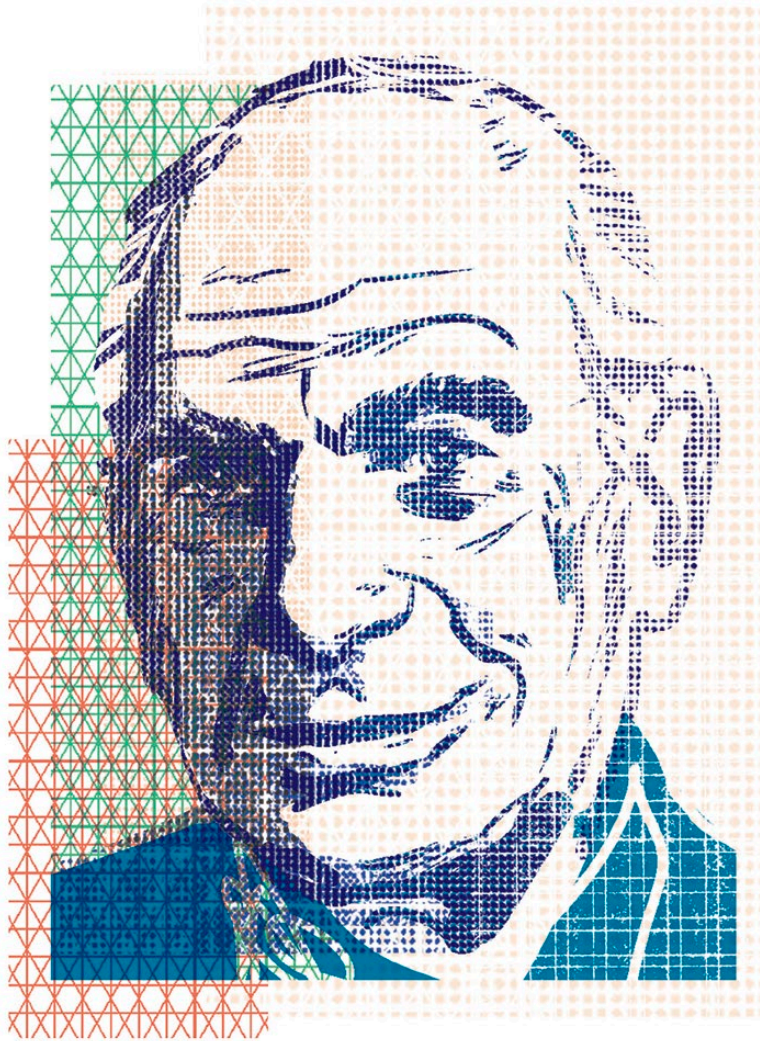


Illustration by Shonagh Rae

“Try to design an approach to making a judgment; don’t just go into it trusting your intuition.”

Nobel Prize-winning psychologist Daniel Kahneman discusses the stubbornness of cognitive biases, the “noise” that besets human decisions, and how institutions can learn to make fairer judgments.

Cognitive psychologist Daniel Kahneman has spent his career studying the ways humans think, including the cognitive shortcuts and biases that shape—and sometimes misshape—our decisions. He was awarded the Nobel Prize in Economics in 2002 for integrating psychological research on how we make decisions under uncertainty into economics, and he is a member of the National Academy of Sciences.

Kahneman's 2011 bestseller, *Thinking, Fast and Slow*, won the National Academies Best Book Award and the Los Angeles Times Book Prize. His most recent book, coauthored with Olivier Sibony and Cass Sunstein and published last year, is *Noise: A Flaw in Human Judgment*. *Issues in Science and Technology* editor Sara Frueh spoke with Kahneman to get his insights on how we make decisions, the nature of noise and the problems it causes, and whether algorithms could improve decisionmaking.

During the pandemic, people have had to make elaborate risk assessments to decide whether to visit loved ones, or send their kids to school, or sometimes just leave the house. How did you see the phenomena that you've explored in your work—our intuitive and effortful ways of thinking and our mental shortcuts and biases—operating in the context of the pandemic?

Kahneman: Well, the first thing that was very salient at the beginning of the pandemic was that people really find it difficult to deal with exponential growth. I recognized this in myself. I was about to take a flight to France when there were just a hundred cases in France; that didn't look like much, except it was doubling every couple of days. And that was really quite powerful.

What we've seen since is that people think about risk a lot, but it doesn't look as if we have a very explicit idea of what those risks are. And I've been struck by the role of emotion, which I hadn't emphasized in my previous work. Some people are very afraid and other people are much less afraid, and it's the level of fear that seems to dominate behavior. You have people who have barely gone out over the two years and other people who have exploited every opening and every opportunity. And that looks more like a different emotional response than a different risk calculation, because we haven't had much material to really make calculations.

In Thinking, Fast and Slow, you wrote that you weren't generally optimistic about the potential for individuals to control the cognitive biases that send our thinking off track. In the decade since that was published, have you seen any evidence or intervention that has convinced you otherwise?

Kahneman: No, not really. I mean, there have been some published successes, but they were fairly minor. Not all that

much has happened. Again, my optimism with respect to the ability of individuals to improve their thinking is limited. As I think I said in *Thinking, Fast and Slow*, I have more confidence in the ability of institutions to improve their thinking than in the ability of individuals to improve their thinking.

That brings us to your more recent work on noise, which you define as unwanted variability in judgments that should be identical—for example, when different judges hand down dramatically different sentences for the same crime, or doctors make different diagnoses for the same patient. How did you get interested in noise? What convinced you that this is a big enough problem that it needed public attention?

Kahneman: Well, I had an experience with an insurance company where I ran an experiment—the kind of experiment we now call a noise audit—where we presented the same cases to a large number of underwriters and we asked them to set a premium for these cases. Now, nobody would expect two underwriters looking at a complex case to arrive at exactly the same number. Underwriting is a matter of judgment, so you'd expect some disagreement.

I tried to identify how much disagreement people would tolerate by asking the executives: If you took two underwriters, by how much would you expect them to differ? And there is a number that comes up more frequently than any other, which is 10%. This is not only for underwriters; it seems to be a general number, that where judgment matters, 10% variability seems tolerable.

But in fact, among the underwriters the variability was closer to 50%—five times as much expected, and that is qualitatively different. I mean, if there is that much variability in the underwriting system, you should go back to the drawing board, because clearly they're adding a lot of noise.

And the other thing that is important is that this was completely new to the people I talked to. The organization had a very large noise problem and was completely unaware of it. It's really that combination—of there being noise, and people not being aware of it and not recognizing it as a problem—that made it tempting to do something about it.

In what systems did you and your coauthors find high levels of noise? Where do you think noise showed up in the worst way?

Kahneman: Well, we found noise wherever we looked for it. Where I personally found it the most shocking is in the justice system. And that is an extraordinarily interesting case because there is huge variability among judges in terms of the sentences they are imposing for the same offenses. And yet judges really do not want to be made to be uniform. It seems to strike very deeply—the possibility of enforcing or even suggesting that uniformity is desirable is already quite threatening.

What do you think is driving that reaction?

Kahneman: In part, there really is a tradition that justice is something determined by an experienced individual with high ethical standards and understanding of the norms of society. Justice is whatever that individual decides with full, complete knowledge of the circumstances, and no other individual who doesn't have that information can make judgments about it. The judge is like an instrument for determining what is just. And once you threaten that—and the idea of noise really threatens that—then it becomes very difficult, I think, for judges to reconcile themselves to the situation.

It's also the case that if the extent of noise in the judicial system were something that people talked about a lot, then people would lose respect for the justice system. But it's not talked about a lot. And it's quite remarkable, there are few efforts to do anything about this in the justice system.

And then, of course, there are things that are not part of the justice system but operate in similar ways—so, asylum judges, patent officers, reviewers of grants, and even teachers who grade students in ways that determine their future. Noise in all those systems seems to potentially be the source of unfairness.

In other institutions, in insurance companies for example, it's clear that noise in underwriting is costly, and it leads to decisions that are not good for the business. I'm more hopeful about reducing noise in business than reducing noise in the judicial system. But you asked me what shocks me most—it's there.

With regard to the justice system, one of the objections people have to making sentencing more uniform and less noisy is that judges really need to be allowed to take into account the particulars of a case. Are there ways to lessen noise in the justice system while still allowing for that to happen?

Kahneman: There are circumstances where very clearly you have a critical piece of information that's a deal-breaker, and clearly deal-breakers should be allowed. You don't want a system that doesn't allow for those. What is insidious are aspects of the situation that are not by themselves deal-breakers but that people have intuitions about. Then we see that the weight that people give to the information is really not optimal, and that's where noise comes in.

The quality of people's decisions in many cases doesn't increase consistently with the amount of information that they have. There are some items of information that help, and then it reaches a point where more information is actually more likely to make you go astray than to add to the quality of your decisionmaking.

It turns out that people really do best with a small

amount of information, and that when they begin to consider the details and the complexities of the individual case—except if it's a deal-breaker—they're likely to give improper weight to insignificant matters.

What are some examples of practices that organizations can use to reduce noise?

Kahneman: I think the book is, in an important respect, premature. That is, in general with an idea of this size, there should be at least 20 years before you publish a book, because there's a lot of research to be done. Now as it happened, I was 80 when I had that idea, and so I didn't have 20 years. When we speak of what we call "decision hygiene," which are procedures that hopefully will reduce noise, a fair amount of that is speculative—that is, it hasn't been tested through research. It's backed up indirectly—we didn't completely invent things out of our head—but there is a lot of work that needs to be done to establish those things. So, that's sort of a confession.

"I have more confidence in the ability of institutions to improve their thinking than in the ability of individuals to improve their thinking."

Given that, we do have ideas about procedures that are better than others, and the main example in my mind was a contrast between structured and unstructured hiring interviews. Unstructured interviews are when interviewers do what comes naturally. The structured interview breaks up the problems into dimensions, gets separate judgments on each dimension, and delays global evaluation until the end of the process, when all the information available can be considered at once.

We know that neither structured nor unstructured interviews are very good predictors of success on the job, which is extremely difficult to predict. But within those limits, the structured interview is clearly better than the unstructured one.

If you think of decisions, then decisions involve options, and you can think of the options as similar to job candidates. That means that each option has attributes, and you want to assess those attributes separately. And we expect that approach to have the same kind of advantages that structured interviews have over unstructured interviews.

So, the most important recommendation of decision hygiene is structuring. Try to design an approach to making a judgment or solving a problem, and don't just go into it trusting your intuition to give you the right answer.

One method you've considered to reduce noise is through the use of algorithms. But there have been a lot of concerns raised about their use in decisionmaking, in particular that they might amplify racial and gender biases. How should we weigh the benefits and risks of using algorithms?

Kahneman: Well, I think that there is widespread antipathy to algorithms, and it's a special case of people's preference for the natural over the artificial. In general we prefer something that is authentic over something that is fabricated, and we prefer something that's human over something that is mechanical. And so we are strongly biased against algorithms. I think that's true for all of us. Other things being equal, we would prefer a diagnosis to be made or a sentence to be passed by a human rather than by an algorithm. That's an emotional thing.

But that feeling has to be weighed against the fact that algorithms, when they're feasible, have major advantages over human judgment—one of them being that they are noise-free. That is, when you present the same problem to an algorithm on two occasions, you are going to get the same answer. So, that's one big advantage of algorithms. The other is that they're improvable. So, if you detect a bias or you detect something that is wrong, you can improve the algorithm much more easily than you can improve human judgment.

And the third is that humans are biased and noisy. It's not as if we're talking of humans not being biased. The biases of humans are hidden by the noise in their judgment, whereas when there is a bias in an algorithm, you can see it because there is no noise to hide it. But the idea that only algorithms are biased is ridiculous; to the extent they have their biases, they learn them from people.

A famous example is, I think, an attempt to measure and predict crime in different areas. If the measure of crime is arrests, then you're going to end up with something that is grossly racially biased because arrests are grossly racially biased. So typically, the biases are introduced into algorithms by human decisions about how to define the problem. But if you take care to define the problem properly, the algorithm is not going to invent biases.

What about the concern that rules intended to reduce noise might also reduce creativity and ingenuity? How might that happen? Are there ways to reduce noise that won't have that effect?

Kahneman: Well, I think there is a real risk that when you produce procedures that guide judgment and decisionmaking, and that makes it more homogeneous and more uniform, the risk is demoralization and bureaucratization. No question, that risk exists. And so reducing noise without demoralizing people—that's a skill

that has to be acquired, and clearly that's a constraint on the implementation of noise reduction techniques.

Whether noise reduction will impair creativity or not, I'm really not sure. And that is because you really want to create a distinction between the final decision and the process of creating that decision. And in the process of creating a decision, diversity is a very good thing. When you're constructing a committee to make decisions—whether of hiring or of strategy—you do not want people to come from exactly the same background and to have the same inclinations. You want diversity. You want different points of view represented, and you want different sources of knowledge represented. In some occasions increasing diversity in the making of the decision could reduce noise in the decision itself.

The real deep principle of what we call decision hygiene is independence. That is, you want items of information to be as independent of each other as possible. For example, you want witnesses who don't talk to each other, and preferably who saw the same event from different perspectives. You do not want all your information to be redundant. So, good decisions are decisions that are made on the basis of diverse information.

You mentioned that a lot of research still needs to be done about noise. What are some key questions that you would most like to see answered?

Kahneman: Well, I think the most urgent questions are about mitigation, and they're about really verifying decision hygiene and improving our recommendations and testing them. That would be the first thing I would hope would happen.

Then, it would be very interesting to study, I think, individual differences in judgment, in different kinds of judgment. Where do the differences come from, and can they be anticipated?

And what is the real value of experience? Experience always increases confidence, and experienced people have more confidence in the quality of their judgment. But this is true even when they get absolutely no feedback from the environment about the quality of their decisions. So, underwriters never know whether they set the right premium or not, and yet they become more confident. Now, how do people become more confident? Well, it's when they begin agreeing with themselves. That's basically the criterion that, "Oh, I had a similar problem and that's what I decided then and I feel like deciding the same thing now." And that gives people confidence that they're doing the right thing—with absolutely no objective justification.

Also, studying what we call "respect-experts"—that is, what distinguishes those people who become well-recognized experts in the absence of objective feedback? Trying to understand that phenomenon is interesting.