Baruch Fischhoff

The microbiomes of gut-level decisions

The greater the variety of microbiota in our gut, the greater the variety of food we can digest. The role of any one organism is clearest when it is missing. Vegans may find that they have lost the organisms needed to digest meat, if they go off the wagon. Fast-food junkies may find whole grain foods a challenge. Children who did not have peanuts in infancy may have allergies later on. Patients who take broad-spectrum antibiotics may be overwhelmed by normally docile *Clostridium difficile* after competing bacteria have been knocked back.

By analogy, the greater the variety of perspectives in our minds, the greater the variety of evidence we can digest. The role of a perspective, too, is clearest when it is missing. Students denied music, art, or environmental education may get less from the world around them. Members of isolated communities may empathise less with people with different backgrounds. People afraid of maths or science may not even try to parse statistical evidence.

The mind and the gut must cope with whatever comes their way, trying to extract benefits and minimise risks. Comparing those processes suggests some of the complementary abilities and limits, as the sciences and humanities treat risky decisions. It also suggests ways that they might inform one another's disciplinary work and collaborate on projects that need them all.

Two strategies for managing complexity

Herbert Simon described two generic approaches to making complex decisions. One, he called 'bounded rationality'. It entails deliberately ignoring enough elements of a decision to be able to identify the optimal choice for the elements that remain. The second, he called 'satisficing'. It entails ignoring nothing, but being satisfied with a choice that is good enough, rather than optimal.[1]

Both strategies depend on heuristics, generally useful rules that can backfire when applied clumsily or ill-advisedly. Heuristics that guide bounded rationality specify which issues to ignore. They fail when the

bounds are set too narrowly, so that critical issues are missed, or too broadly, so that decision-makers are overwhelmed by what is still left. Satisficing heuristics identify potential solutions and ways to evaluate their adequacy. They fail when imagination cannot produce viable options or anticipate the conditions under which they will be exercised.

A boundedly rational heuristic for healthy eating might be a vegan diet. It could fail people whose work, social, or living conditions (e.g., company cafeteria, peer pressure, food deserts) preclude consuming complementary proteins. A satisficing heuristic for healthy eating might be a Mediterranean diet. It could fail people who lack essential genes, microbiota, or trace elements in their food and water. Heuristics for tasty eating might be more robust, except perhaps with finicky kids and cats.

Strategic academic decisions

Scholars face analogous choices when dealing with the complex domains that they study. Generally speaking, those in the sciences opt for bounded rationality, whereas those in the humanities opt for satisficing. They might profitably join forces in thinking about when broad, imperfect, satisficing gut-level decisions can inform, and perhaps should override, narrow, precise boundedly rational ones.

Each scientific discipline sets bounds on the issues that it considers. For example, psychologists typically assume that *how* people respond to particular situations is universal enough for their studies to rely on convenience samples (often WEIRD, Western Educated Industrialized Rich and Democratic),[2] while acknowledging that *what* situations people face can vary widely. Scientists value colleagues who pose interesting questions within their discipline's bounds. They can be baffled or threatened by scientists who work outside those bounds.

Each humanities discipline, as viewed by this scientist, has a form of discourse that satisfices for its members, capturing some of their perspectives on complex concerns. Its scholars value colleagues whose accounts are thought-provoking enough to advance their thinking and incomplete enough to invite comment. They can be exasperated by the reductionism of scientists, especially those who purport to say something meaningful about human affairs or the natural world.

Interdisciplinary science is rare, in part, because it requires scientists to accept the compromises of living within other disciplines' bounds. As a result, scientists sometimes find it easier to borrow ideas from the humanities, even when they cannot fully appreciate the discourse that produced them. The scientific study of risk has done just that, at pivotal

junctions, drawing on the humanities in ways that might alarm those whose ideas it has garbled in the process.

The sciences of risk

Each scientific approach to risk represents a way of knowing that is suited to digesting some experiences, while producing indigestion (or worse) with others. The bully on the block is the rational actor model.[3] Although it emerged from rich philosophical traditions, their nuances have been lost as various disciplines have shaped themselves around the model, led by neoclassical economics. One bound that the model imposes is 'consequentialism', considering only the outcomes of a choice and not the process leading to it, unless that process is also an outcome (e.g., being the decider). A second bound is that decision-making is all in the head, and not the heart or gut.

The rational actor model is founded on philosophical demonstrations of the powerful inferences that follow, when preferences are bounded to ones that follow the utility theory axioms.[4] One axiom requires all options to be comparable, such that a decision-maker is never unable or unwilling to prefer one or be indifferent between them (as with a Sophie's Choice). A second axiom bounds rationality to options that have a price, such that none is sacred. A third axiom holds that the relative attractiveness of two options not depend on what else is on offer. Some decisions fall within those bounds. Some do not, including ones where people need to feel good about a choice or care about how it is made, or by whom.[5]

The rational actor model has echoes in many sciences. Its seemingly clear terms have revealed structures that can be observed and studied in disparate domains (e.g., the costs and benefits of gathering more information, the difference between single-play and multiple-play competitions, the value of writing off sunk costs). In adept hands, the model is sometimes fairly successful at predicting choices, even ones made in non-rational ways (e.g., by trial and error, mimicry, surrogacy).[6]

Bounded models

However, the hegemony of the rational actor model has come at a price. Disciplines in its sway privilege descriptions that place decisions within its bounds. They promote researchers who are proficient at model formalisms, at the expense of other skills. They also make broad claims about the model's applicability. Romantic love? Reading? Exercise? Curiosity? Self-control? Whatever the topic, someone has created a rational actor model for it, with perhaps some ability to inform or be informed by

the human experience, while ignoring, and sometimes dismissing, complicating concerns.

One response to these excesses is to reject rational actor models altogether. By some accounts, demonstrations of biased judgements and poor choices do just that. However, without the rational actor model's conceptual structure, such demonstrations can become undisciplined litanies, allowing observers to criticise any decision that doesn't look quite right or suffers misfortune. Such inventories of irrationality are brute force reactions to rational actor models run amok, powerful, but chaotic attempts to restore order, akin to the fecal transplants used to cure *Clostridium difficile*, by restoring normal gut biota.

In social discourse, both rationality and irrationality can be corrosive memes. The former can promote unwarranted faith in market mechanisms, leaving people to fend for themselves and denying them needed protections. The latter can undermine faith in the public, as unable to manage its affairs. Just as the rational actor model is indifferent to how decisions are made, so is the reaction to it. Neither worries much about empowerment, self-efficacy, learning, or civil society.

And their limits

Concepts drawn from the humanities have been instrumental in helping the sciences study risk in ways that escape the bounds of the rational actor model without losing the order and insight that it can bring. That borrowing has required domesticating rich satisficing perspectives from the humanities to the ascetic forms of bounded sciences. Examining the fidelity of that translation process might benefit all parties, suggesting distinctions that the sciences have missed and concerns that the humanities might productively worry.

In scientific terms, studying any decision entails three interdependent forms of research: analysis (identifying options and assessing where they might lead), description (characterising decision-makers' relevant beliefs and values), and intervention (closing the gap between the analytical ideal and the descriptive reality). The process is iterative. If an intervention fails, then the analysis or the description may be flawed.[7]

The examples that follow sketch how the humanities have informed each of these three pursuits: the analysis of risks, by uncovering ethical assumptions in conventional practices; the description of risk behaviours, by situating them in individuals' lives; and interventions, by clarifying the social contracts that they entail. They suggest work that boundedly rational sciences cannot do alone.

What are the decisions?

Analytical research breaks down ('decomposes') complex problems into more comprehensible parts. Such analyses were initially created to compare designs of new technologies, notably nuclear power, chemical plants, and space travel. Soon, though, they were redirected to estimate absolute risk levels, often hoping to reassure sceptics, by demonstrating that a technology had 'acceptable risk'. Those estimates sometimes find their way into the elaborate accounting schemes of cost-benefit analyses.[8]

Although ostensibly transparent, these calculations are often impenetrable, even to technical reviewers. Moreover, their complexity supports a patina of objectivity that obscures the ethical assumptions that any analysis must make. Once exposed, those assumptions can be understood by anyone and illuminated by scholars with broader perspectives. Indeed, ethicists have dogged risk analysis since its inception, asking questions about what is estimated (e.g., just mortality or also morbidity?), how it is estimated (e.g., does equity matter?), and how to accommodate unquantified concerns (e.g., feelings of dread).[9]

Translating these broader, satisficing perspectives into the boundedly rational world of risk analysis has been a struggle. It takes analysts out of their comfort zone, by giving standing to concerns that are not easily quantified. It limits their sovereignty, by requiring skills from other disciplines. It undermines the promise of authority and objectivity that their clients may demand.[10]

How do people view them?

Descriptive research asks how people view their choices, in terms that can be compared to the analytical account. Revealed preference analysis, the predominant descriptive approach in economics, assumes that people are informed rational actors, and then infers what matters to them from their observed actions. It also assumes 'stable preferences', whereby people apply the same overarching values (or multi-attribute utility function) to all their decisions.[11] These studies have no place for satisficing heuristics, even when they can do as well as elaborate calculations.[12]

The sciences often use pathology to understand normally healthy processes, by showing when they can go astray. With risk, that study has focused on the biases that can arise from generally useful heuristics.[13] However, those biases reflect emergent properties of complex environments, which boundedly rational approaches have no orderly way to address. For example, the availability heuristic treats hard-to-imagine

risks as unlikely, leaving people surprised when unimagined events happen. Scientists have just begun to tap the work of scholars who study what makes likely and unlikely narratives compelling.[14]

The humanities have long informed the study of decisions that violate some utility theory axioms. Those include decisions involving 'sacred values', held so strongly that no trade-offs are possible. They also include decisions involving preferences held so weakly that people do not know what they want, as with difficult medical choices or conflicts among sacred values.[15] Anthropology, which straddles the humanities and social sciences, sometimes uncomfortably, has been instrumental in making these links.[16]

What can be done about them?

Interventions try to align analysis and behaviour, by refining one or both. Their success depends on whether the critical issues are within their disciplinary bounds and whether there is clear feedback on how well interventions are working. Both conditions are met, for example, when people need to know the size of a well-studied risk (e.g., annual homicides, probability of an auto accident) and alternative presentation formats can be tested in controlled settings that resemble their real-world counterparts. For example, researchers have developed a Drug Fact Box that allows most people to learn the risks and benefits of pharmaceuticals, expressed in the quantitative terms needed for informed decisions.[17] Although the US Food and Drug Administration (FDA) has not adopted the Fact Box, it now summarises its approval decisions in roughly those terms.[18]

FDA's experience shows some of the challenges of designing interventions. It has fought pressure for rational-actor model quantification, so that it can consult with patients about satisficing approval rules. However, it has also been criticised for murky decision rules regarding opioids, which appear to have ignored predictable public health issues (e.g., effects on families, illegal prescribing practices).[19] The skill of the sciences of risk lies in evaluating options. The skill of designing options worth evaluating may lie in the arts and humanities.

With interventions intended to inform choices, ethical issues often lurk in seemingly technical issues, such as how terms are defined and uncertainty is disclosed. With interventions intended to change behaviour, the enterprise itself requires ethical treatment. For example, nudges,[20] currently popular in some circles, manipulate people 'for their own good' to pursue rational actor solutions, as determined by someone else. In research tests of these interventions, Institutional Review

Boards provide input from the humanities (and the community). In actual interventions, such protections are rare.

Interacting ways of knowing

Each science of risk has its own way of knowing, suited to digesting experiences that fall within its bounds. Each science leaves other problems only partially digested. Each expands its range slowly, by exploring its boundary conditions. That expansion reflects ad hoc responses to possible insights or novel problems. As a result, scientific disciplines cannot readily situate themselves in the space of issues that they treat incompletely or ignore entirely.

In each area, analysis, description, and intervention, the humanities have made valuable contributions to expanding the sciences' awareness of their bounds, by offering integrative satisficing perspectives. Sometimes, that insight has come from scientists reading something that got them thinking, or remembering something from a class long ago. Sometimes, it has come through humanities scholars wondering aloud about a science's practices and proposing alternative approaches. There could be more such borrowing, if scientific research were more accessible, so that humanities scholars could be fuller, more forceful commentators. Providing such access was the goal of this article, as was making the case that scientists could use the help.

How might scientists return the favour? One possible way is to provide context for scientific research that might catch the eye of humanities scholars. J.H. Hexter once cautioned fellow historians against 'jumping aboard intellectually sinking ships', picking up fads whose time had passed in their home disciplines.[21] Outsiders cannot know what insiders think about work in a discipline or when members have gone beyond its bounds, and their own competency, in positioning themselves as public intellectuals. We might tell them. Even better, we might share the deliberations that embody the wisdom of any discipline. With a little decoding, we might benefit from sitting in on one another's research meetings and seminars.

When individuals face decisions involving risks, they often must weigh their gut feelings against scientists' boundedly rational advice. Knowing the bounds of those scientists' discipline might help them to decide whether to follow that advice or whether satisficing analyses from the humanities will be more useful in digesting what is on their plates. Collaboration across the sciences and humanities might be satisfying for both, and for those they hope to serve.

Notes

I am grateful for the support of the Carnegie Corporation of New York and the Swedish Foundation for the Humanities and Social Science. The essay is based on a talk accompanying receipt of an honorary Doctorate of Humanities, University of Lund. The views expressed are my own.

1 Herbert A. Simon, *Models of Man* (New York: Wiley, 1957).
2 Douglas Medin, Bethany Ojalehto, Ananda Marin and Megan Bang, 'Systems of (non-)diversity', *Nature Human Behavior*, 1:0088 (2017).
3 Gary S. Becker, *The Economic Approach to Human Behavior* (Chicago: University of Chicago Press, 1976).
4 Nils-Eric Sahlin, *The Philosophy of F.P. Ramsey* (Cambridge: Cambridge University Press, 1990).
5 Jonathan Baron, *Thinking and Deciding* (4th edn), (New York: Cambridge University Press, 2008).
6 Robyn Dawes, *Rational Choice in an Uncertain World* (New York: Harcourt Brace Jovanovich, 1988).
7 Baruch Fischhoff and John Kadvany, *Risk: A Very Short Introduction* (Oxford: Oxford University Press, 2011).
8 Baruch Fischhoff, 'The realities of risk-cost-benefit analysis', *Science*, 350:6260 (2015), 527; https://doi.org/10.1126/science.aaa6516.
9 Ibid.
10 Heather Douglas, *Science, Policy, and the Value-Free Ideal* (Pittsburgh: Pittsburgh University Press, 2009).
11 Becker, *The Economic Approach to Human Behavior*.
12 Dawes, *Rational Choice in an Uncertain World*.
13 Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011).
14 Caroline A. Bartel and Rahul Garud, 'The role of narratives in sustaining organizational innovation', *Organizational Science*, 20 (2009), 107–17.
15 Sarah Lichtenstein and Paul Slovic (eds), *Construction of Preferences* (London, Earthscan, 2006).
16 Scott Atran. 'The devoted actor: unconditional commitment and intractable conflict across cultures', *Current Anthropology,* 57:suppl. 13 (2016), S192–S203.
17 Lisa M. Schwartz and Steven Woloshin, 'The Drug Facts Box: Improving the communication of prescription drug information', *Proceedings of the National Academy of Sciences USA* 110:Suppl. 3 (2014), 14069–14074.
18 Food and Drug Administration. *Benefit-risk assessment in drug regulatory decision making. Draft PDUFA VI implementation plan. FY2018–2022* (Washington DC, 2018).
19 National Academy of Medicine, *Pain Management and the Opioid Epidemic: Balancing Societal and Individual Benefits and Risks of Prescription Opioid Use* (National Academy Press, 2017).
20 Richard Thaler and Cass Sunstein, *Nudge: Improving Decisions about Health, Wealth and Happiness* (New Haven CT: Yale University Press, 2008).
21 J.H. Hexter, *The History Primer* (New York: Basic Books, 1971).