

## RESPONSE

### Critical thinking: Kahneman and policy making

Robin Mansell\*

*London School of Economics and Political Science, London, UK*

*Robin Mansell is professor of new media and the internet at LSE. She is interested in how and why people communicate with each other, especially when their relationships are mediated by the use of information and communication technologies.*

In examining this challenging and innovative work, my concern is with understanding the processes informing or biasing decision making at the institutional level. Kahneman suggests that ‘the idea that our minds are susceptible to systematic errors is now generally accepted’ (2011, p.10), but it is not clear what this means when we consider choices taken by institutional authorities such as regulators and policy makers. He suggests that ‘organizations are better than individuals when it comes to avoiding errors, because they naturally think more slowly and have the power to impose orderly procedures’ (p.418). Thus, his analysis is aimed not only at understanding the biases of individual decision makers, but also at how institutions express collective choices.

Although the unitary decision maker model implied is questionable, Kahneman’s observations about System 1 and 2 thinking open opportunities for understanding bias at the institutional level. This is relevant beyond the disciplines of psychology and economics. In interdisciplinary fields such as science, technology and innovation studies, it is often assumed that investment in evermore sophisticated digital technologies for processing information means that our connectedness and improved access to information will lead to improved decision making.

In many policy making fora, those who rest their arguments on intuition or on less than what others deem to be ‘full’ information are often said to be irrational or to be ignoring the evidence base that should underpin policy prescriptions. Kahneman’s System 1/System 2 model is helpful insofar as he does not dichotomize rational and irrational decision making. He says that ‘humans are not irrational’ (p.411) and that any expectation of logical consistency in human preferences is a ‘hopeless mirage’ (p.335). Commenting that the way economists use the term ‘rationality’ assumes an impossible standard of consistency across different situations, he also argues that the use of intuition and heuristics (System 1) to make decisions is much more common than effortful decision making. This is because time is ‘the

---

\*Email: [r.e.mansell@lse.ac.uk](mailto:r.e.mansell@lse.ac.uk)

ultimate finite resource' (p.409). System 2 decision making requires that time is taken to process information and is filtered through often poorly understood cognitive processes. Learned patterns may lead to perceived coherency especially when challenging information is disregarded. These issues matter in decision situations where the stakes are high (p.192). This arguably is the case in policy making or regulatory settings where stakeholder interests are affected by decisions taken by institutions with a remit to act in ways which affect individual lives or corporate interests.

Kahneman suggests that institutions can try to counter System 1 biases by enabling individuals to be more alert to sources of bias. He says that they can be sensitized to the factors that give rise to bias when they understand cognitive processes and acquire a vocabulary to talk about them (e.g. the halo effect, competitor neglect, and loss aversion). Kahneman *et al.* (2011) offer a checklist for managers who take decisions that can be used to help keep biases in check. In their paper, it is acknowledged that self interest may be associated with interests in financial gain, organizational power, or reputational gain, for instance. Kahneman also suggests that we should be wary of decision making in many organizations when someone 'controls what you see [and] has a vested interest in what you choose' (p.361). Thus, the presence of vested interests is acknowledged at the individual and institutional levels of decision making and for all choice situations. The biases of System 1 thinking are always present.

The use of checklists and other tools does not provide a basis for optimism that this will necessarily improve judgements. Kahneman suggests only that it may be possible to become better at recognizing situations that are prone to errors in judgement and then to use decision quality control procedures to ensure that deliberate efforts are made to invoke effortful System 2 thinking. The trouble with this is that procedures that are intended to strengthen the quality of decision making can themselves be hijacked by those who are in a position to exercise power over others. In addition, when procedural rules start to dominate, this can serve as a break on human creativity, risk taking, and the flexibility normally required for innovation. Kahneman argues that individuals 'will make better choices when they trust their critics to be sophisticated and fair' (p.418), but he does not seem to address what the implications of quality control procedures are when they are introduced in institutional settings characterized by conflict; for example, by mistrust or asymmetrical power relations between managers and workers.

Nevertheless, Kahneman's work should serve as a wake-up call for policy makers and regulators who persist in resisting decision making in the absence of 'full' evidential information. Even if they do accumulate vast amounts of evidence, System 1 biases are likely to be present at the evidence interpretation stage and this will affect decision outcomes regardless of how long the policy maker or regulator waits. Taking the findings of *Thinking, Slow and Fast* seriously suggests that there is good reason to challenge efforts to standardize the collection of information so as to arrive at the best solutions to problems that policy makers and regulators are charged with addressing. Instead, it would be better to acknowledge that, whatever the state of the evidence base, it is more important to apply critical thinking, to invoke System 2 thinking in contexts where choices are very consequential for people's lives and to acknowledge that, even then, interests will be present and that these are likely to bias the choices that are taken.

In addition to promoting decision quality control procedures and the acquisition of a vocabulary for articulating the ways in which cognitive processes are error

prone, Kahneman seems to suggest that better decisions arise when we rely on automated information systems. He notes that ‘hostility to algorithms will probably soften as their role in everyday life continues to expand’ (p.229). Should we take this to mean that increasing numbers of decisions of consequence in our lives should be taken on our behalf by software agents operating behind the computer screen? Perhaps we should since, as Kahneman says, ‘statistical algorithms greatly outdo humans in noisy environments’ (p.241). With the spread of digital networks and the growing sophistication of information processing capabilities, algorithms are indeed acting more frequently on our behalf. Kahneman points to the fact that ‘formulas may do better than humans in some critical decisions in the world of sports’ (p.229). They help us with recommendations about music and books using Amazon and there are many other recommender systems. However, the creep of formulae into decision making is not always a good thing insofar as the design and implementation of these systems are not immune to bias. Somewhat surprisingly, Kahneman says, ‘we take it for granted that decisions about credit limits are made without the direct intervention of any human judgment’ (p.229). Indeed, many of us do take this for granted, but one consequence is that personal bankruptcy rates are soaring. This is testimony to the negative impacts of over-reliance on these unaccountable systems.

It seems likely that Kahneman would acknowledge this, but he does not explicitly ask the crucial question: who will hold institutions that rely upon automation or the designers of the algorithms to account? This is an important consideration. For instance, financial markets have become very reliant on automated computer progress over the past several decades in their efforts to devise bias (intermediary)-free mechanisms for trading exchanges. The result of these developments has introduced a number of perverse effects in the trading system, including the weakening of norms that foster desirable trading behaviours [Beunza, *et al.* (2011); and see Mansell (2012) on accountability and pervasive automation].

Although he does not appear to acknowledge this explicitly in this book, Kahneman’s argument serves as a cautionary tale when it comes to the cyberoptimism that yields the notion that sophisticated information system algorithms will produce rational outcomes in the choices taken by policy making and regulatory institutions. His argument is that such algorithms may mitigate the negative effects of the role of emotion in decision making. If decisions are guided by the calculus of algorithms, however, there is no reason to think that biases of interests and power relations have been eliminated. In some cases, the evidence produced by algorithmic calculations may be linked to a more robust information base, but this does not mean that decisions will be bias free or well aligned with the interests of all parties.

Kahneman insists that effortful decision making takes time. If time is the ultimate scarce resource, then the time saving attributes of automated information systems may free up time for effortful System 2 thinking. However, if the systems themselves are devised in ways that are not transparent and the designers are not accountable, then institutional decision making may become more likely to rest on heuristics than it is today. It is crucial that the lessons of Kahneman’s model of System 1 and System 2 thinking start to be critically reviewed and integrated in fields of study beyond disputes among those working within the disciplines of economics and psychology.

Interdisciplinary fields such as science, technology and innovation studies will also benefit from a careful consideration of far-reaching implications of trends that

are fostering the automation of decision making. Account needs to be taken of features such as time-saving and less error-prone decisions with respect to probabilities. However, the equally important consequences of new areas of life which are becoming hidden from view and thus becoming unaccountable also need to be examined. These developments surely raise questions about democratic decision making and the norms of human behaviour that we want to inform decisions taken by our governing institutions.

## References

- Beunza, D., MacKenzie, D., Millo, Y. and Pardo-Guerra, J. (2011) *Impersonal Efficiency and the Dangers of a Fully Automated Securities Exchange*, report commissioned for the UK Government Foresight Project, *The Future of Computer Trading in Financial Markets*, 7 June, available from <http://www.bis.gov.uk/assets/foresight/docs/computer-trading/11-1230-dr11-impersonal-efficiency-and-dangers-of-a-fully-automated-securities-exchange> [accessed October 2012].
- Kahneman, D. (2011) *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York.
- Kahneman, D., Lovallo, D. and Olivier, O. (2011) 'Before you make that big decision', *Harvard Business Review*, 89, 6, pp.50–60.
- Mansell, R. (2012) *Imagining the Internet: Communication, Innovation and Governance*, Oxford University Press, Oxford.