Short read for "The accuracy-coherence tradeoff in cognition"

David Thorstad

By many measures, human cognition is remarkably successful. Our beliefs are often accurate and our decisions often good. We make quick and efficient judgments and decisions when needed. Humans learn well from many different types of information. And we pose and solve cognitive problems that no other species on earth can even imagine.

These cognitive successes help humans to achieve our goals. On the basis of cognition rather than brute force, we have become the dominant species on the planet. We are the only species to have explored outer space, cured a disease, or studied philosophy.

But there is one measure on which human cognition performs less well: we are not especially coherent creatures. On occasion, humans have been observed to violate nearly every axiom of logic, probability theory and decision theory ever proposed. It is important not to exaggerate the extent of human incoherence. But at the same time, it is widely accepted that humans are by far the least coherent creatures on the planet. An incoherent rat is a noteworthy scientific finding. By contrast, we humans are wont to believe six incoherent things before breakfast.

This finding raises a puzzle. Why is it that humans, the most cognitively sophisticated creatures on earth, would also be the least coherent? This finding is especially puzzling if we identify rationality with coherence. Then, to say that humans are the least coherent creatures on the planet is also to say that we are the last rational. This is surprising. Why would the most sophisticated creatures on the planet fall the furthest short of rationality?

We can make progress on this puzzle by appealing to two factors: bounded rationality and tradeoffs. Bounded rationality describes the situation of cognitively limited agents who must ration scarce cognitive resources to achieve the best results possible in their environment. Bounded rationality is a form of decisionmaking under scarcity. A defining feature of decisionmaking under scarcity is the existence of tradeoffs: bounded agents

cannot achieve all of the cognitive outcomes that they would like to achieve, so they must allocate cognitive resources in a way that best balances a variety of competing cognitive goals. Some of the best-known tradeoffs in bounded rationality are widely taken as motivations for adopting more complex and multifaceted normative standards for bounded agents.

Most famously, there is often an *accuracy-effort tradeoff* in cognition. In many situations, the accuracy of an agent's judgments and the quality of her decisions trades off against the effort of making them. The accuracy-effort tradeoff is standardly taken to show that the rationality of cognitive processes cannot only be a matter of their accuracy. It is sometimes rational for agents to use less-accurate processes in order to save cognitive and noncognitive resources for other inquiries, as well as for the many other activities that make up a full human life.

In this paper, I argue that there is a comparably general *accuracy-coherence tradeoff* in cognition. In many situations, the accuracy of an agent's judgments and the quality of her decisions trades off against their coherence. After a point, modifying cognitive processes such as reasoning to increase the expected accuracy of our judgments or the quality of our decisions often increases the expected incoherence of our judgments and decisions. If an accuracy-coherence tradeoff in cognition could be found, what would the normative consequences be?

First, an accuracy-coherence tradeoff would put some pressure against coherence-based theories of bounded rationality. Just as the accuracy-effort tradeoff suggests that effort must be balanced against accuracy in selecting cognitive policies, the accuracy-coherence tradeoff suggests that coherence needs to be balanced against accuracy in selecting cognitive policies. At the very least, this suggests that coherence needs to be weighed against other goals in assessing human rationality. To assess the rationality of bounded agents, we need to ask not only how coherent they are, but also how well their cognition conduces to other goals such as accuracy.

Second, an accuracy-coherence tradeoff would complicate the inference from experi-

mental observations of incoherence to allegations of irrationality. Striking a good balance between accuracy and coherence in cognition often requires agents to accept a nontrivial risk of incoherence, even when they could reduce their risk of incoherence at the cost of accuracy. But this means that observed instances of incoherence may result from fully rational cognitive processes which strike the best balance between cognitive goals such as accuracy and coherence. If that is right, then we cannot infer straightforwardly from incoherence to irrationality, because it may be that this incoherence could only have been avoided by making irrational sacrifices along other important cognitive dimensions.

What evidence is there for an accuracy-coherence tradeoff in cognition? It may help to begin with an example from decisionmaking. Suppose you are buying a car. You might decide using lexicographic choice. You would order features of cars by their importance. Perhaps the most important feature is that is has an automatic transmission; next most important is price; then other features such as safety rating and comfort follow. You would compare the available cars by their most important feature, choosing the car which scores best on this feature. If several cars score just as well on this feature, for example because they all have an automatic transmission, you would compare cars along the second-most-important feature, price, continuing in this way until a decision was reached.

Lexicographic choice is quite a silly way to buy a car. Our lexicographic chooser will always buy the cheapest automatic car unless two automatic cars are tied in price. A traditional and cognitively efficient way to improve upon lexicographic choice is semilexicographic choice. Semilexicographic choice fixes, for each cue, a small difference in values which will be ignored. For example, we might ignore price differences under $1,000 and safety differences no greater than one star. The turn to semilexicographic choice is widely held to yield improvements in decision quality. A semilexicographic chooser will not buy the cheapest automatic car if a comparably-priced automatic is significantly safer.

But semilexicographic choice is less coherent than lexicographic choice. To see the problem, suppose that three automatic cars have the following costs and safety ratings. If given the pairwise choice between Car A and Car B, our semilexicographic chooser will

	Car A	Car B	Car C
Cost (Thousands of dollars)	19	18.6	17.8
Safety Rating (Stars)	4	2.5	1

pick car A. Between Car B and Car C, she will choose Car B. And between Car C and Car A, she will choose Car C. It is natural to interpret this result as a collection of intransitive preferences.

Lexicographic choice does not have this problem. The pairwise choices made by a lexicographic chooser are always transitive. In this way, the move from lexicographic to semilexicographic choice is an example of the accuracy-coherence tradeoff: going semilexicographic decreases the coherence of an agent's decisions, but increases their quality. Similar examples can be constructed for judgment rather than choice tasks.

But one example does not demonstrate a systematic tradeoff, nor does it help us to understand why accuracy and coherence can trade off during cognition. What drives the accuracy-coherence tradeoff? By way of illustration, I identify two factors which often generate an accuracy-coherence tradeoff.

The first factor is *strategic variety*. It is often held that boundedly rational agents should adopt a robust toolbox of cognitive strategies in order to ensure that appropriate strategies are available to confront cognitive challenges. Up to a point, increasing the variety of strategies used increases the expected accuracy of an agent's judgments and the expected quality of her decisions by providing her with strategies suitable to a wider range of challenges. But increasing the variety of an agent's cognitive strategies also decreases the expected coherence of her judgments and decisions by creating the possibility that different strategies will be applied to similar problems, with very different results.

Another factor driving the accuracy-coherence tradeoff is *cognitive complexity*. In many cases, agents can increase the expected accuracy of their judgments or the expected quality of their decisions by adopting more complex forms of processing. But complex processes also create the opportunity for new forms of incoherence, such as framing effects, which are more difficult to produce using simpler rules. This fact explains our original puzzle

of why humans are less coherent than nonhuman animals, and more generally why incoherence is more frequently documented in the most sophisticated creatures. Increasing cognitive complexity tends to decrease coherence. Why, then, would sophisticated creatures choose to use complex cognitive processes when they could instead choose simpler ones? Plausibly, because increasing complexity often increases accuracy and decision quality. In this way, increases in cognitive complexity often drive an accuracy-coherence tradeoff between increased accuracy and decreased coherence.

All of these arguments rely essentially on considerations of bounded rationality. And that is as it should be. In the unbounded case, disputes about the rational status of coherence are not, primarily, extensional disputes. Defenders of coherence requirements hold that coherence is a fundamental requirement of rationality. Detractors hold that coherence requirements are true, or mostly true, but not fundamental. Rather, they are derivative on other requirements such as believing what your evidence supports.

But matters are otherwise for bounded agents. Defenders of coherence requirements hold that bounded agents should be coherent, or at least strive to be as coherent as they can be. Detractors think that rationality may require bounded agents to sacrifice some amount of expected coherence for the sake of other cognitive goods, such as accuracy. And now we have arrived at an extensional dispute. What is at issue for bounded agents is not merely whether coherence requirements are derivative or fundamental, but also whether they are true at all. Insofar as coherence trades off with accuracy for bounded agents, we may have reason to relax coherence requirements on bounded agents to recognize the importance of accuracy and other cognitive goods.

Acknowledgments

I am grateful to Tomi Francis and Elliott Thornley for their comments on this short read.