

Cognitive Bias

*Tom Chatfield**

Independent Researcher

*Corresponding author. Email: tom.chatfield@gmail.com

Keywords: behavioural economics; cognitive bias; confirmation bias; evolutionary psychology; heuristics; Daniel Kahneman; overcoming bias

Abstract

Are human beings irredeemably irrational? If so, why? In this article, I suggest that we need a broader appreciation of thought and reasoning to understand why people get things wrong. Although we can never escape cognitive bias, learning to recognize and understand it can help us push back against its dangers – and in particular to do so collectively and collaboratively.

Philosophy is about slow, careful thought: rigorously scrutinizing words, ideas and experiences; seeking good reasons and testing assumptions; trying to address fundamental questions. It's also uniquely human. Other species may possess diverse and remarkable forms of intelligence, but we are alone in our capacity to express and analyse our own minds and, incrementally, to achieve a rich (if imperfect) understanding of our world.

We are also, however, animals. The survivors of countless hostile millennia, we are connected by a common evolutionary heritage to every other living thing. And, like every other living thing, the most fundamental ways in which we deal with our world are defined by those strategies that immemorably supported our survival. In particular, we are neither as reasonable nor as unique as we might wish to believe. And the limits of our reasonableness are inscribed within two distinct ways we think about the world: fast, and slow.

I've borrowed this dichotomy from one of the most famous non-fiction books of the last few decades: *Thinking, Fast and Slow* (2011) by the

behavioural economist and Nobel laureate Daniel Kahneman. Kahneman's research (and that of his late colleague, Amos Tversky) has prompted countless tributes and disputes, but at its root is a relatively uncontroversial and ancient observation. 'Slow' decision-making of the philosophical kind – effortful, reasoned, attentionally engaged, analytical – is a rarity. It has to be, because the intensive use of mental resources is by definition a luxury: something that can bring immense advantages when selectively deployed, but that cannot steer us through the hazards, opportunities and snap decisions of everyday existence. For this, evolution has equipped us with a series of 'fast' decision-making shortcuts known as *heuristics*. These describe the ways in which we constantly make 'good-enough' judgements on the basis of emotion and intuition – as well as suggesting why the business of reasoned analysis is, often, as much about justifying these first impressions as it is about changing our minds.

A heuristic is a cognitive rule of thumb, and is generally experienced as a sense of ease or



comfort. If, for example, something or someone is familiar-looking and associated with safety, I'm likely upon encountering them to feel a sense of 'rightness' that helps me rapidly to respond appropriately. Thanks to the survival and reproduction of thousands of generations of humans who have faced analogous situations, this feeling is a usually a reliable basis for action. It can also encompass plenty of subtleties relating to visual clues, behaviours, habits and experience. If, by contrast, something looks potentially dangerous – perhaps because it's making a strange noise, or has a pungent smell – I'm likely to feel uncomfortable, and to calibrate my actions accordingly. Or, if a situation is novel or ambiguous, I may feel a sense of cautious curiosity upon encountering it; and I may, in due course, subject it to some slow thinking (from a safe distance) alongside other members of my species.

The key to understanding heuristics is their replacement of a complex question with something amenable to a simple, instinctual solution.

As Kahneman puts it: 'This is the essence of intuitive heuristics: when faced with a difficult question, we often answer an easier one instead, usually without noticing the substitution.' Consider the examples in my previous paragraph. The most urgent question underlying all of them goes something like this: 'what are the relative risks and benefits of engaging with this situation in one of a number of different possible ways?' This is a ferociously difficult issue to address exhaustively, and any creature that habitually attempted to do so would be paralysed by prevarication (quite possibly fatally). By contrast, a question like 'how does this novel situation make me feel?' can be answered easily and instinctually. Indeed, the emotions informing this process are in effect a form of biochemical reasoning. And so long as the verdicts they yield are sufficiently reliable – and weighted against lethal outcomes – the efficiency with which they allow us to act is an excellent thing in evolutionary terms.

‘we are neither as reasonable nor as unique as we might wish to believe.’

At this point, before we delve into their difficulties, it's worth emphasizing just how efficacious a guide to decision-making heuristics mostly are. Philosophy, being a business of language and reason, can end up denigrating affective decision-making as inherently irrational or unreasonable. I would suggest, however, that it's more useful to think of such decision-making as perfectly reasonable in a broader sense – not to mention more fundamental to our minds and natures. It is also, however, perilously vulnerable to misapplication; a problem that's particularly acute in contemporary contexts. So long as I'm dealing with something that resembles the circumstances in which my instincts evolved, their affective insights are well worth heeding. When, however, I find myself facing a situation that's markedly *unlike* those humans evolved to assess intuitively – or when someone else is trying actively to exploit the ways in which this history makes me vulnerable – I'm at risk. And, unfortunately, the 'slow' mental processes that should in theory enable me to push back may in practice stay sidelined when I need them most.

The point at which a heuristic leads into error is known as a *cognitive bias*, and describes a flawed judgement that predictably results from our intuitions 'misfiring'. Consider the following question: 'which politician's policies would be best for our country?' Self-evidently, answering this rigorously demands a fair amount of research and consideration. It's neither generic nor the kind of issue that our species has spent the last hundred thousand years repeatedly assessing – but it is something that our slow, deliberative faculties can help us assess. Unfortunately, however, one of the defining features of these faculties is that we have a limited attentional 'budget' when it comes to deploying them. Our slow, thoughtful selves are both indolent and easily exhausted,

while the emotive assessments that shape our fast, intuitive judgements are inextricably interwoven with our perceptions. Thus, Kahneman notes, we are strongly inclined to address a second, easier kind of question in place of the first one – 'which of these politicians looks most pleasing and authoritative?' say, or 'how does this one story you remember hearing about them make you feel?' – while treating the resulting answer as if it adequately addressed both.

One of the most significant cognitive biases at play in the example above is known as the *halo effect*, and describes the fact that – like a halo hovering over a cartoon angel's head – one notably impressive attribute can create a kind of perceptual glow that influences a host of associated judgements. A tall, attractive, confident-looking politician will (unfairly but inexorably) benefit from some unearned credit across other domains thanks to their appearance. They may be perceived as more trustworthy, authoritative or strong-willed than less attractive rivals, even if there's no good reason to believe any of these things. Similarly, a car advert will often feature a beautiful person driving through a beautiful setting, because this image is more likely to create positive associations than footage of the same car in a dark garage with a bucket of fish on its bonnet. Like ripples in a pond, both feelings and their associated judgements tend to permeate our minds on the basis of a few carefully crafted appearances.

What's at work, here, is the human tendency to treat positive (and negative) attributes as correlated: to let our feelings about one feature spill over into other areas, rather than effortfully evaluating every factor individually. It's easy enough to see why, in evolutionary terms, this makes sense. Consider a few obvious attributes, like health and height. Given limited time and information, such features are useful enough proxies for fitness and reproductive desirability. Similarly, it's important for us to be capable of bonding and sustaining cooperative relationships with small groups of peers rather than constantly reassessing every attribute of every individual on a case-by-case basis.

When it comes to the merits of political programmes or products, of course, none of this

applies; or, at least, none of it ought to. It's self-evidently absurd for the symmetry of someone's face or the lighting of a photoshoot to inform my assessment of a complex manufactured object. The fact that an actor has been paid to pretend to drive a car through Italy can't possibly tell me anything worth knowing about its comfort or reliability. Yet – as advertisers and spin doctors know all too well – this won't stop me, or anyone else, from reacting as if it does. To be human is to be obliged to assess technological modernity via a biological apparatus that evolved to help hunter-gathering groups survive on the savannah: to react to faces, bodies, appearances and social norms as if they conveyed fundamental truths rather than tapping into a series of affective short-cuts.

At this point, it's easy to become pessimistic. If you search online for information about cognitive biases, you'll notice that lists of their different forms can run into hundreds of items. As we've seen, misreading a politician's appearance as a meaningful reflection of their policies is an example of the *halo effect* – or, if you don't like the way they look, its opposite, the delightfully named *horn effect*. Yet this is just the beginning. Plenty of further misjudgements are likely to be bound up in such a scenario, including *stereotyping* (which describes our tendency to assess someone or something by how closely it conforms to prejudices about its 'type'), *confirmation bias* (disproportionately seeking out or attending to information that flatters our pre-existing beliefs), *availability bias* (treating information that's immediately available to us as if it were sufficient to resolve any question) and *authority bias* (the tendency to trust any claim that comes from an authority figure). And that's just for starters. Even knowing about these biases won't prevent you from feeling them. Indeed, this knowledge may actually produce further vulnerabilities, such as *expert over-confidence* (the tendency of those who possess expertise in one area wrongly to overestimate the quality of their judgements in other areas) or *the illusion of understanding* (the tendency to believe we understand the world's complexities when all we've actually grasped is a simplified model).

What do these biases have in common? Above all, they are all kinds of *category error*. That is, they all treat feelings and first impressions as decisive facts about the world, rather than as facets of our inner lives. As the word 'decisive' suggests, this means that they prompt us towards definitive judgements rather than open, exploratory processes. Not only are many of our judgements deeply flawed; they're also inherently resistant to the kind of incremental error-correction that defines improvement. Cognitive category errors are, it seems, our irredeemable lot in life. Yet it's here, where the evidence for our inadequacy seems strongest, that some of my own best hopes for human decision-making come in. Just so long as we don't succumb to high-minded revulsion at our own animal natures – or the philosophical denigration of those attributes that have borne our species to this point.

In particular, I would suggest that focusing too narrowly on individual biases (and individualistic accounts of bias) can blind us to the kind of collective curiosity that underpins our most remarkable achievements as a species; and that the right kinds of structures, habits and collaborations can help us become far more than the sum of our cognitive parts.

This, in a sense, is what science and philosophy are about: the pooling of many minds' observations; the testing of theories and explanations in the light of experience and evidence. It's equally clear, however, that none of us can undertake this kind of collaboration most of the time – and that the key question, when it comes to bias, is thus how far the decision-making opportunities surrounding us are or aren't supportive of valid intuitions. What does it mean for us to take control of our cognitive environments, and to have at least *some* faith in the reliability of the heuristics that guide us from moment to moment? Above all, I would suggest, answering this hopefully means abandoning the fantasy that such a thing as perfectly rational or well-informed decision-making can ever exist; and embracing, instead, the fundamentally embodied, subjective and pragmatic business of doing the best we can.

‘At this point, it’s easy to become pessimistic. If you search online for information about cognitive biases, you’ll notice that lists of their different forms can run into hundreds of items.’

As soon as we’ve done this, a few fundamental rules – or, at least, helpful heuristics – come to mind. Our limited budget of attention is a bodily fact: one closely related to energy and mood. Mental experiences are always also physical experiences, and this means that flawed decision-making is most likely to catch us out if we are depleted, bewildered, overwhelmed, or operating in a field where neither instinct nor experience can offer sure guidance. At this point, the simplest advice is also the best: slow down and seek reinforcements. These reinforcements may take the form of others’ experiences; reliable, relevant information; or a rest, a snack or a moment’s reflection. Sometimes, none of these will be available, in which case seeking broader analogies may be the best idea. But the key point remains: valid judgements *cannot* emerge in the absence of reliable expertise, advice or information – and the very best place to seek these is among other people we have good reasons to trust.

A second point relates to cognitive ease. To borrow another line from Kahneman, ‘a reliable way to make people believe in falsehoods is frequent repetition, because familiarity is not easily distinguished from truth’. As advertisers and marketers have long known, the mere fact that something *feels* familiar is often enough to make us trust it – something that’s especially

true if we don’t notice this is going on. Comfort, familiarity and ease disarm critical engagement. By contrast, unfamiliarity and cognitive strain tend to mobilize our ‘slow’ selves. This in turn suggests a strategy for pushing back against the seductions of repetition. If a question or decision is significant, it pays dividends to *defamiliarize* it; to rephrase and reframe it in different ways; to seek out diverse or dissenting views. When we are uncomfortable, we become less intuitive and more hesitant. This is extremely undesirable if we’re playing in a tennis tournament and eager to channel the benefits of a thousand hours’ practice. But it’s a vital corrective if the path of minimal resistance is paved towards an undesirable outcome; and, once again, it entails a fruitful form of friction with other minds.

This brings me to my final point, which speaks to perhaps the most fundamental bias listed earlier in this essay: *confirmation bias*, and the ways in which we tend to pay disproportionate attention to our own views. In a sense, to describe confirmation in these terms is to do little more than formalize the fact that all of us view the world through the lens of our own experiences: that we inevitably bring with us particular perspectives, ideas and expertise, and can no more step outside of these than we can become someone else. What we can do, however, is take as close an interest as possible in what it is *like* to be someone else – and, in particular, in how and why other people might have an utterly different view of the world from our own, for reasons that are as compelling to them as ours are to us.

Seen through the lens of an individual life, it’s easy to see confirmation bias as a jet black mark against human understanding. In the collective context, however, something remarkable can happen when it comes to the strength with which we’re able to argue our side in a debate. People aren’t just prone to seeing the world in one particular way. They are also brilliantly, maddeningly adept at deploying evidence and reason in the service of a particular worldview; of defending and justifying cherished claims. At its worst, this can lead to the despairing entrenchment of differences. At its best, however, the exchange of rivalrous, rigorous points of view is

Tom Chatfield

a fundamental form of truth-seeking. And all it requires to become more than a zero-sum conflict is a commitment on both sides to *listening* as well as to speaking; and to an overarching principle of clarification and illumination.

This is an idealization, of course. Yet it's one that our species achieves, albeit imperfectly, more often than you might think; and certainly more often than you might fear when listing the frailties of a lone mind. We are, first and foremost, animals, our apprehensions of the world saturated with sentiment and partiality. But we are

also a deeply empathetic, cooperative and curious species. From language and culture to technology and child-raising, our achievements rest upon the entwining of many minds and perspectives; upon a remarkable relationship with our own vulnerabilities. Individually, in the information-saturated realm of the twenty-first century, it's all too easy to see ourselves as irredeemably irrational: pawns in an algorithmic game. Collectively, however, we are and have always been something else: self-knowing creatures in the constant business of becoming.

Tom Chatfield

Tom Chatfield is a British author and philosopher of technology. His most recent books are *How to Think* (SAGE) and *Critical Thinking* (SAGE).