The Myth of Rationality **Bob Stone**

Introduction

If you were at this event 11 years ago, and you remember what you heard there, first you have a better memory than mine, but secondly you may faintly recognise some of the things I'm about to say.

Rationality is a very important concept to philosophers. Plato thinks intelligent reasoning gives us all the answers to everything; Kant claims that no one does anything genuinely moral unless she does it because her reason tells her that it is the one moral choice.

Yet there is also a tradition of scepticism about our ability to reason. Plato wants to restrict political power to the handful of men, and women, who alone are governed by their reason rather than by some lower mental capacity – which just happens to be philosophers like him. And Kant is so pessimistic about the prospect of anyone being able to listen to his reason rather than to other inclinations that he places reason outside the perceptible world of natural cause and effect and locates it in the transcendent world of things-in-themselves. Bertrand Russell complained: "It has been said that man is a rational animal. All my life I have been searching for evidence which could support this."

I don't think that the pessimistic philosophers are worried about people's IQ. They have noticed what we all notice about other people: even the cleverest of them do stupid things and think stupid thoughts. It is as though our minds are overwhelmed by inclinations other than the desire to reason clearly, and so any ability we may have to reason is negated by all that baggage.

What I aim to do in this talk – which of course will be an entirely rational exception to the general rule I have just mentioned – is, first, to show how what we believe is rational thinking is often nothing of the sort, and, second, to look closely at what remains – what seems like straightforward reasoning – and argue that it merits the name not through the nature of what goes on in the head but purely from the outcome.

The now classic book by Daniel Kahneman, *Thinking, Fast and Slow*, gives the conclusions of a mass of research carried out over a few decades into the way people think. It has the merit of being very clear, very thorough, and very entertaining.

Most of the research is in the form of administering tests – usually questions – to large numbers of people, often students at American universities and so, we presume, moderately intelligent. (Ha ha!) The tests can be, and often have been, easily replicated by other psychologists who are only too keen to show up methodological failings; and so there are some controversies, at least about the implications of some results. But the general findings seem pretty well established. Furthermore, when you read about them, you quickly recognise your own limitations, or at least those of others.

1. Cognitive biases

Kahneman distinguishes two 'systems' in our way of thinking, which he calls 'System 1' and 'System 2'. He does stress that these are not meant to be a literal description of two parts of the brain or mind, just a device for distinguishing two different ways in which we think. System 2 is the genuinely rational one, which thinks clearly and logically and, for that reason, often slowly.

But it is System 1 that is dominant, even when we believe we are thinking rationally. The result is that, although it gets it right most of the time, we do tend to believe things not because there is some reason to suppose they are true, but because of certain 'cognitive biases'.

Here are a few examples.

- We accept an idea because it is familiar to us, or it is presented to us in a nice way, or it fits what we already believe.
- We are also happy to jump to conclusions on the basis of minimal evidence, as if 'what you see is all there is' (WYSIATI).
- We are influenced by words or pictures we see just before we are told something – whether by accident, or placed there by people trying to influence us – including things we don't even know we've seen, as in subliminal advertising. If a small number has just come up on the roulette table, we will be likely to make a lower estimate of some totally unrelated fact – like the price of a champagne bottle – than if it had been a high number.
- We are particularly hopeless in dealing with statistics, and I want to dwell on one example to illustrate.

Here is the famous 'Linda Problem', where this scenario was presented to students (in the 1980s in America).

Linda is thirty-one years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.

The students were asked to rank these scenarios in order of probability, which you see on the screen:

a) Linda is a teacher in elementary school
b) Linda works in a bookstore and takes yoga classes
c) Linda is active in the feminist movement
d) Linda is a psychiatric social worker
e) Linda is a member of the League of Women Voters
f) Linda is a bank teller
g) Linda is an insurance salesperson
h) Linda is a bank teller and is active in the feminist movement

I'm not going to read all those out, because the only feature I and the researchers are interested in is scenarios (f) and (h). Nearly all ranked (h) above (f). And, even when only those two alternatives were presented to groups of students in several universities, between 85% and 90% put (h) above (f). Of course, I don't have to explain to you that, regardless of what qualities Linda has, it is simply more likely that she, or anyone at all, is a bank teller than that she is a bank teller with feminist tendencies: the latter category is a sub-category of the former, and therefore smaller.

Just to check that this wasn't a cognitive flaw prevalent only in the young, or in Americans, I gave the 'test', or an equivalent one, to various people nearer my age that I met shortly after reading the book. The result is the same: nearly everyone puts (h) above (f). They all saw the point immediately when I explained it, and felt rather sheepish; I had to admit that I'd made the same mistake when I first read the Linda problem.

There is some dispute about the implications of such tests. I suspect that, when people read 'bank teller' along with 'bank teller . . . active in the feminist movement', they imagine the mere bank teller as one who is not active in the feminist movement, even though that's not what it says; so they are comparing a non-feminist bank teller with a feminist bank teller. They do not consciously think to themselves, " 'bank teller' means 'bank teller who is not active in the feminist movement' ". But it forms part of their thinking.

And that is a crucial aspect of thinking: much of what is included in the process is not consciously present. In the case of subliminal suggestions, we don't *ever* know that they are, at least until someone tells us. In the case of tests like these, we realise afterwards what we must have been thinking unconsciously, but were not aware at the time.

The great mass of psychological evidence gathered by Kahneman demonstrates that the intuition that shapes most of our thinking, though essential for getting through life without getting bogged down every few seconds in systematic reasoning, is full of cognitive biases which makes much of it plain wrong.

In a way this is not surprising. When I engage in thinking, I am not using some special faculty of the mind or switching on some dedicated rational neurons in the brain, like switching on a calculator. When philosophers discuss the nature of the 'conscious being' – i.e. the nature of 'me' or 'I' – it is always assumed that, whatever it is, it is something unified, or at least something that gives the illusion of being unified. It is the same 'I' that sees, hears, feels, thinks and remembers, continuous with the 'I' that saw, heard, felt, thought and remembered yesterday and ten years ago. The 'I' that engages in thinking at a particular moment is the 'I' that I have become over a number of decades, with all kinds of ingrained attitudes, memories, tastes, feelings that are a background to any experience I have whatever, not to mention more temporary accompaniments such as mood, tension, discomfort in the leg, and what I have just heard or seen.

So, if I have to do some simple piece of reasoning such as multiplying 15 by 17 in my head, there will be a context in which I am doing it. Maybe someone has asked me, "What is 15 by 17?" My reaction to the question may be a desire to help someone I like, irritation that I have to take my mind off the TV, suspicion that the inquirer is testing me. I contend that it is simply impossible to have nothing in my mind except the process of reasoning – unless I'm a computer.

In the case of something fairly simple, like multiplication, it is unlikely that any of those things will affect the answer, though they may slow me down, or speed me up, or possibly lead me to think too quickly and make a silly mistake.

How much more susceptible to distortion by acquired personality and surrounding feelings is a more complicated piece of reasoning, especially where inductive or practical judgment is involved rather than merely following mathematical or logical rules.

To enlarge on something I mentioned in my introductory remarks earlier today, we've all noticed that just about everybody in the world holds more or less exactly the same beliefs about important matters as their parents, or, if not their parents, the kind of people among whom they live. If someone is asked why he holds that belief, he does not usually say "because my parents brought me up to believe it", but gives some kind of rationale. From this it is hard to avoid the conclusion that almost all reasons given to justify beliefs about anything important are not the actual *reason for*, in the sense of *cause of*, the belief's being held; they are thought up in order to justify a belief that is going to be held regardless of the relevant evidence and reasons. Not quite all reasons, however; each one of us, in our own minds, is an exception to that rule!

2. Genuine 'reasoning'

Now some might think that I'm exaggerating. You may agree that an awful lot of what people believe to be rational thought is in fact not rational at all, but hopelessly skewed by forms of cognitive bias, such as wishful thinking, laziness, love of the familiar, loyalty to one's earlier beliefs or those of the family or community. Yet, at the end of the day, you still believe, it is possible to transcend (in a not necessarily Kantian sense) all those anti-rational influences and to think rationally, according to Kahneman's 'System 2'.

To express the contrast in a different way, Tim Bayne – a professor of psychology who wrote the OUP *Very Short Guide to Thought* – wrote in the *New Scientist* (21 Sept 2013) that there are two types of train of thought: the associative one, where one thought leads naturally to another, like a game of word association, and the systematic variety which uses evidence and logic.

My contention is that this distinction is illusory, and that the apparently systematic type of thought is merely a subset of the associative one. So let us see what actually happens in the so-called systematic kind of thinking.

You are set a problem to solve by sheer thinking. The first thing to note is that different people – even those who come up with the same answer – have quite different things going on in their minds. Some may talk to themselves out loud, others may have their minds full of pictures, or images of words, or imagined sounds of words. Different things will occur to them at different times, such as earlier problems they've heard about, maybe Linda the bank teller, which warn them against certain ideas, or for that matter quite extraneous memories of something associated with a word or name in the problem. Some might think systematically along a certain route, others might keep being sidetracked by memories and ideas from their experience – including ideas that make the solution suddenly easier, but which don't occur to the more single-minded.

How do we judge whether all this mish-mash of stuff going on in your mind is 'rational' or not? Simply by the result that you give at the end of it. If you wanted to convince others that you had thought rationally, you would not try to reconstruct all the things that flitted through your mind, but would explain how the answer is a correct answer to the question. It is the analysis of the *outcome* of the thinking, as expressed in words or other symbols, that decides whether the thinking that produced it, and indeed the thinker, are rational.

That may remind you of Wittgenstein's *Philosophical Investigations* (which I've read since writing the original version of this talk): words which apparently pick out mental processes, such as 'understanding' or 'intending', do no such thing, he argues. It is the *result* of the process that enables us to characterize whatever was going on in the mind as 'understanding', 'intending', or – in our case – 'reasoning'.

An analogy is winning a race: what characterizes a winning performance – whether Usain Bolt's in the Olympics or mine, at the age of 5, in an egg-and-spoon race (I'm told) – is not some common feature of our performances (there is none!), but the result: no one got to the finishing line before us.

Let us look at two types of problem-solving, deductive and practical. First, a simple problem of arithmetic which has one right answer. 7 x 2. Like most people, I simply remember from learning it years ago that the answer is 14. No reasoning needed there, just memory. Less automatic would be 7 x 15, as our tables only went up to 12. It would occur to me, perhaps, that 7 x 15 is really 7 x 10 plus 7 x 5; I can do those two sums automatically, and the addition of 100 and 35 is automatic. There are other ways of doing it, too, e.g. multiplying 7 by 30 and dividing the result by 2 – equally valid and rational, as long as they make it 105.

That entire operation, however it actually works, consists of memories – whether of learnt facts or of learnt methods – each prompted by the one before, in the manner of association, according to the way that I've been trained.

Another problem of the same sort is this, which I remember hearing in this very building. There is a life-threatening medical condition which one person in a million has. A man goes to see the doctor and asks if he has it. The doctor gives him a test which he says is 99% reliable, i.e. it gives the right answer on 99% of occasions. The test is positive: it says the man does have the condition. How worried should he be?

Again, I followed the inexorable logic: if the test is 99% likely to be right, and it says the man has the condition, then the man is 99% likely to have the condition. My thinking was as systematic as you can get, so I thought, pure logic, pure rationality. Except that, as I'm sure you all know, such a deduction is totally wrong.

The correct answer, as was explained to us, is that the man's chances of having the condition have in fact risen from 1 in a million to about 1 in 10,000. From now on, when given a similar problem, I shall remember the method and quite likely get it right. It's all a matter of having learnt, and being prompted by the situation to remember, the correct steps to take. [Whether I get the answer to such questions right or wrong depends entirely on whether I have learnt the correct facts and procedures, and whether – if I have learnt them – they occur to me when prompted.] (By the way, explanation of that counter-intuitive example will be available in the bar!)

My second type of problem is the much more interesting, and difficult, kind where there is no one right answer waiting to be worked out. You need inductive, practical reasoning. For example, how should I reduce my carbon footprint? How can we bring peace to the Middle East? I shall use a rather more manageable, and trivial, example, as I have only 30 minutes.

The case of the misbehaving padlock

On Monday, I went to the University of Worcester's gym class for oldies like me. I've only just started going, and I needed to buy a padlock to keep my clothes and valuables safe while I was torturing myself on the dark Satanic machines. It's the sort of padlock where you need to choose a 3-digit number as your code, and then you can always unlock it. That worked like a dream last week, but this week, the same number I'd chosen, 149, failed to unlock the padlock.

Right, practical problem requiring rational thought. As physical things aren't really my forte, the only alternatives that occurred to me were to panic, to seek help from some university member, to call the police, to drive into town and buy a hacksaw (except that my car keys were in the locker), and – finally – to ask if any of my fellow-gymnasts happened to be a locksmith.

One of them offered to help, and within a minute or two he'd unlocked the padlock. It occurred to him, what hadn't occurred to me, that a likely cause of the trouble (leaving any memory lapse on my part out of it) was that *one* of the three rotating wheels was faulty. So he tried keeping two of the numbers constant while trying a different number in the other wheel. He soon found that the number 139 (rather than my chosen 149) was the number that the padlock believed was the code; I was able to reclaim my wallet and my trousers.

Now what is the point of this story? The way I've told it might suggest that the way to solve my problem was to engage in a purely rational process of thinking which would lead to the right answer. But surely I could have done that myself? Yet I didn't think of that possible solution, it simply didn't occur to me, and so I couldn't just decide to employ that rational method. The helpful fellow-gymnast may have been more familiar with padlocks than I was, but in any case, whatever the reason, it DID occur to him and not to me. Of course, it *might* not have worked; the padlock might have been more seriously damaged. But he could see that, even if that method didn't work, it would take less time to realise that than testing any other method he might have thought of, such as trying all 1000 possible combinations until one worked.

Association

My contention is that *all* apparently rational thinking, of any type, is just like any other conscious activity: a sequence of thoughts that prompt each other, by association, with the sequence determined by the existing personality and experience of the thinker and any input from the senses or from other people. A small amount of thinking may be fairly systematic – but even then, it has to be a correct system, which occurs to us because we remember learning it, not just any old method. Basically, thinking consists of ideas or memories just occurring to us.

That last phrase might suggest that our thinking is not deliberate but in a sense out of our control. Quite right. Think about a memory. It is either triggered by the situation, or it isn't. Whether it is triggered depends partly on the situation itself – e.g. the question that is posed – but even more on the past history of the person concerned. You cannot deliberately remember something, in the sense of deliberately bringing it to mind at a particular moment. (I find that people are stubbornly resistant to that idea!) If the memory is already in your mind, and if you happen to have a desire to keep it there, and if there is nothing to distract you, then you can keep remembering it for a time.

Similarly, you cannot *avoid* remembering something if the memory is triggered. When I was about 7, my father told me that, if I went down to the kitchen at 3 in the morning, I would find there, on the table, all for me to spend as I pleased, one million pounds . . . just so long as I didn't think about crocodiles. I got the point immediately, and I've spent the last 70 years (nearly) trying to think of some way round it – like getting so old that I really do forget what it is I'm not allowed to think about. The point is that what comes into our minds is triggered by a combination of what is there already and stimuli from outside. This is not some wacky point, but sheer logic: you can't *decide* to think about something unless you are already thinking about it. Think about it! We are at the mercy of what occurs to us.

Not only is thinking, even rational thinking, not something we deliberately do. It is often not even done consciously. Consider what happens when we talk. Most of us, unless we are drunk or asleep or mad, talk rationally, in the sense that what we say makes sense, is understood by the speaker, conveys what we mean, and is understood by the listener. At least some of the things that are said by us are intelligent, thoughtful, logical. But an occasion like this, where I am saying words that I have already thought up, is quite exceptional. In the vast majority of cases where people talk, as in the question session tomorrow, or over dinner, we talk impromptu.

That means that, although we know roughly what point we want to make, or what impression to convey, the words just come tumbling out unplanned – unless we are trying to speak in a foreign language we don't know very well. We don't know the end of the sentence we have started, how many clauses it will have, and so on; our choice of words is not the result of deliberate selection but an automatic product of our past history, our personality and the requirements of the moment. And yet the words make sense. Ditto driving a car without crashing, walking without falling over.

Conclusion

First of all, much of our thinking is quick and superficial, System 1, and that is very often perfectly adequate. But we are prone to all kinds of cognitive biases, even when we think we are being intelligent and rational.

Secondly, even when we are being as rational as we can be, System 2, the rational solutions we come up with are *either* triggered automatically by remembered cues, e.g. 7+5=12, whether we like it or not – a kind of Pavlovian system of stimulus and response, occurring in very basic calculations; *or* they occur to us in a variety of non-systematic ways, a kind of thought association, each thought triggered by the previous thought in a way that is determined by our experience, our memories, our mood, in particular the people we are talking or working with.

My contention is that the phrase 'rational thinking' is a transferred epithet: 'rational' properly applies not to the anarchic constituents of our conscious minds – the images, the imagined words, the extraneous interruptions, the occasional strings of systematic (correct or incorrect) ideas – but to the outcome. The thinking itself is not done in some deliberately chosen sequence; it consists of just one thought after another. Thoughts, like everything else in the world, just happen.