University of Reading

In defence of individual rationality

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Borg, E. ORCID: https://orcid.org/0000-0003-2725-9568 (2022) In defence of individual rationality. Proceedings of the Aristotelian Society, 22 (3). pp. 195-217. ISSN 1467-9264 doi: 10.1093/arisoc/aoac009 Available at https://centaur.reading.ac.uk/108379/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1093/arisoc/aoac009

Publisher: Wiley-Blackwell

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Meeting of the Aristotelian Society held at Senate House, University of London, on 25 April 2022 at 6:00 p.m.

In Defence of Individual Rationality

Emma Borg

Common-sense (or folk) psychology holds that (generally) we do what we do for the reasons we have. This common-sense approach is embodied in claims like 'I went to the kitchen because I wanted a drink' and 'She took a coat because she thought it might rain and hoped to stay dry'. However, the veracity of these common-sense psychological explanations has been challenged by experimental evidence (primarily from behavioural economics and social psychology) which appears to show that individuals are systematically irrational—that often we do not do what we do because of the reasons we have. Recently, some of the same experimental evidence has also been used to level a somewhat different challenge at the common-sense view, arguing that the overarching aim of reasoning is not to deliver better or more reason-governed decisions for individual reasoners, but to improve group decision making or to protect an individual's sense of self. This paper explores the range of challenges that experimental work has been taken to raise for the common-sense approach and suggests some potential responses. Overall, I argue that the experimental evidence surveyed should not lead us to a rejection of individual rationality.

According to a fairly standard view in philosophy, intentional actions are reasons-responsive: I reach for the jar because I *want a cookie* and *believe there is a cookie in the jar* (see, for example, Davidson 1963). But how widely applicable is this sort of belief–desire understanding of action? Recent work in social psychology and behavioural economics suggests the answer is 'not very widely', for experimental work apparently shows that very often what we do is *not* a rational reflection of our beliefs and desires. Instead, we act out of intuitive gut feelings, using cognitive short cuts or heuristics which result in us being blown this way and that by situational features that do not constitute rational causes for action. The common-sense view of individuals, as generally acting in light of their reasons, turns out not to survive confrontation with the facts about how we actually decide what to do. In this paper I explore the

© 2022 The Aristotelian Society
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

key experimental evidence offered for this claim and suggest that a defence of individual rationality is in order.

I

Two Arguments against Individual Rationality. To appreciate the arguments against individual rationality, we first need to get a somewhat clearer view of the position these arguments oppose: what is it for an individual to count as making a rational decision in the first place? While there is much to be said in answer to this question, for the purposes of this paper I'll simply set out the view that opponents of individual rationality seem to accept, which we might call 'classical rational choice theory'.[1] According to this approach, to count as rational individual decision making needs to:

- aim at maximizing utility;
- consider all the evidence;
- utilize classical logic in deterministic environments and (Bayesian) probability theory in uncertain/vague situations.[2]

For instance, capturing the first two points, Ariely (2008, p. xxx) writes that 'the assumption that we are rational implies that, in everyday life, we compute the value of all possible options we face and then follow the best possible path of action'. This classical model has been very influential, underpinning the view of humans as *homo economicus* in behavioural economics and *homo politicus* in political theory, where agents are viewed as primarily self-interested individuals capable of weighing up all the available evidence in order to arrive at an ideal solution to any problem or decision they face. I'll return below to the question of whether opponents are right to take classical rational choice theory to characterize human rationality but, given the view, the challenge is clear. For there is a wealth of experimental evidence to show that people's actual decisions in a range of cases don't in fact line up with the predictions of classical

[1] For discussion see, for example, Lieder and Griffiths (2020), Kühberger (2002), Stein (1996).

[2] The probability reasoning in question is almost always assumed to be Bayesian (whereby we modulate degrees of credence in a proposition p based on p's prior credence combined with new evidence). This opens up further questions (for instance, concerning objective versus subjective Bayesian probabilities, and the scope of Bayesian explanation, but I'll leave these to one side; see Mandelbaum 2018 for some further discussion).

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

rational choice theory.[3] Individual decision making thus turns out to be frequently non-rational.[4]

It is important to notice, however, that, on the basis of (often the same) experimental evidence, opponents offer two rather different arguments for this conclusion:[5]

(i) *Use of automatic decision-making systems*: agents often *don't try to reason* at all (instead relying on non-logical, automatic decision-making processes).
(ii) *Improper use of a logical system*: agents often *try but fail to reason properly* (although utilizing deliberative processes, agents don't meet the demands of rationality due to some systemic failure).

I'll consider each of these challenges to individual rationality in turn, sampling the evidence offered for each and suggesting some reasons to be sceptical about the conclusions opponents have reached.[6]

II

Automatic (*Intuitive*) *Decision-Making Systems*. As is well known, many theorists have proposed that human decision making is

[3] Doris (2015, pp. 64–6) frames this challenge as follows: 'Where the causes of [an agent's] cognition or behaviour would not be recognized by the actor as reasons for that cognition or behaviour, were she aware of these causes at the time of performance, these causes are *defeaters*. Where defeaters obtain, the exercise of agency does not obtain. If the presence of defeaters cannot be confidently ruled out for a particular behaviour, it is not justified to attribute the actor an exercise of agency.' The problem is that 'The empirical evidence indicates that defeaters occur quite frequently in everyday life' (2015, p. 68).

[4] This paper focuses on the challenge (paradigmatically as it emerges from Kahneman et al.'s 'heuristics and biases programme') that heuristic use is irrational. As is well known, however, alternative views (such as Gerd Gigerenzer and others' 'adaptive rationality' model) characterize heuristic use as rational. Although it is beyond the scope of this discussion, I'd argue that the objection raised in §2.4, which holds that properties like speed and frugality don't define types of cognition, also challenge alternative approaches like Gigerenzer's (see Borg MS).

[5] To be clear, although I think it useful to distinguish these arguments, they are not held apart in the literature. So although I'll write as if the first challenge belongs to Kahneman and Tversky, while the second belongs to theorists like Mercier and Sperber, this is artificial, for sometimes Kahneman and Tversky write as if the real problem is that we systematically misuse processes which attempt to weigh evidence and reasons (the second challenge, of improper use), and sometimes Mercier and Sperber and others seem to suggest that the real problem is that some of our decision-making processes don't look to reasons and evidence at all (the first challenge, of an automatic system). So the focus should be on the kind of objection in play, not on the exegetical question of who actually signs up to which worry.

[6] For a somewhat fuller coverage of the evidence, see Borg (MS).

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

underpinned by two radically different systems or processes.[7] On the one hand, we have a slow, reflective, logical decision-making system—a broadly rational, belief–desire-based system of the kind familiar to philosophers from discussions about 'folk' or 'common-sense' psychology. The outcomes of these processes are actions performed on the basis of reasons. On the other hand, however, we also have a fast, automatic, non-logical decision-making process—a system of intuition and habit which operates on the basis of quick and dirty 'heuristics', sacrificing accuracy for speed—which Kahneman (2011) speaks of as 'a machine for jumping to conclusions'.

To clarify the notion of a cognitive heuristic, Tversky and Kahneman's seminal 1974 *Science* paper opens with discussion of a visual heuristic: we assess object distance in terms of clarity, so we tend to think that things we can see clearly are close, while those that are blurry are further away.[8] In general, this method for assessing object distance works tolerably well—it doesn't yield precise or fine-grained knowledge of distance, but it tends to permit a broadly accurate ranking of objects over distance. However, in conditions where clarity is either particularly good or particularly bad, the heuristic leads to errors—for example, we tend to think that objects seen in poor light or rain are further away than they actually are because we equate being blurry with being further away, forgetting to adjust for lighting conditions. This example highlights three key features of heuristic thinking. First, heuristic thinking is thinking which neglects some information or options, looking only at a subset of what is relevant. Second, and connected to this, heuristic judgements are intuitive (they seem to be the result of quick, automatic triggering rather than slow, logical deliberation). Finally, related to the aforementioned two properties, and as the visual case makes clear, the use of a heuristic, while practically useful, can lead us astray (for example, by leading us to misrepresent object distance in certain lighting conditions). Since the automatic system uses error-inducing heuristics instead of operating in line with classical rational choice theory, actions caused via its operation do not constitute rational actions.

[7] Although there are important distinctions between 'system' and 'process' talk (see Evans 2009), I'll skate over the issue here, using the two terms interchangeably.

[8] Presenting heuristic thinking via comparison to visual judgements is common. For instance, Thaler and Sunstein (2008, pp. 19–21) open their discussion with an optical illusion concerning relative size, commenting that 'your judgement in this task was biased, and predictably so … Not only were you wrong; you were probably confident that you were right'. See also Mercier and Sperber's (2018) introduction to heuristic thinking.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

The suggestion (at least amongst advocates of the 'default-interventionist' approach; see Evans and Stanovich 2013) is that quick and dirty automatic processes (System 1) provide our default response to decision-making problems, with slow, logical processes (System 2) only being wheeled in when someone becomes aware that their automatic processes have made a mistake.[9] So, why should we accept this 'dual process' model, contrasting, as it does, logical and automatic decisions? The evidence comes from two main sources: the Cognitive Reflection Test and work from Kahneman and Tversky (amongst many others) on the range of cognitive heuristics and biases which apparently infect our decision making.

2.1. *The Cognitive Reflection Test (CRT).* In the Cognitive Reflection Test (Frederick 2005) subjects are asked to make a decision or judgement on an apparently simple question. In each case there is an answer that comes quickly, easily and naturally to mind. However, this intuitive answer can subsequently be shown to be, or comes to be appreciated by the agent as, wrong.[10] The original three cases that Frederick presented are now extremely famous:

(a) A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?

_____ cents

(b) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

_____ minutes

(c) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

_____ days

[9] As several authors have pointed out, this take on the relation between systems is not without problems; for instance, we might wonder how a subject becomes aware of error before System 2 kicks in.

[10] As Toplak et al. (2011, pp. 1275–6) note, 'The three problems on the CRT are of interest to researchers working in the heuristics-and-biases tradition because a strong alternative response is initially primed and then must be overridden. As Kahneman and Frederick made clear … this framework of an incorrectly primed initial response that must be overridden fits in nicely with currently popular dual process frameworks'.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

The intuitive answers to these questions are: 10 cents, 100 minutes, and 24 days. However, the correct (or logical) solutions are: 5 cents, 5 minutes, and 47 days. What Frederick found was that the majority of participants give the intuitive answers and not the logical answers to these questions, apparently demonstrating our automatic system in action.[11]

2.2. *Kahneman and Tversky's Heuristics and Biases*. Tversky and Kahneman (1974) introduce their project by focusing on three examples of cognitive heuristics people commonly deploy in decision making, but which lead subjects into error in certain settings. These three original heuristics are:

(a) Representativeness: assessing the likelihood of x being F via x's similarity to other Fs.
(b) Availability: assessing the likelihood of F occurring via the subject's familiarity with other instances of F.
(c) Anchoring and adjustment: assessing the likelihood or value of x from an accepted contextual anchor point, with adjustments made from that point.

In subsequent work, Kahneman and Tversky, and others working in the heuristics and biases programme, went on to add many further heuristics to the list.[12] While surveying every suggested heuristic and bias would take us too far afield, two further influential proposals are worth noting:

(d) Framing effects: 'the large changes of preferences that are sometimes caused by inconsequential variations in the wording of a choice problem' (Kahneman 2011, p. 272).
(e) Overconfidence: 'our excessive confidence in what we believe we know, and our apparent inability to acknowledge the full extent of our ignorance and the uncertainty of the world we live in' (Kahneman 2011, pp. 13–14).

[11] While the original CRT contained just three questions something of a cottage industry has grown up generating CRT-like questions; see, for example, Trouche et al. (2014), Thomson and Oppenheimer (2016).

[12] Lieder and Griffiths (2020, p. 3) note a potential worry with this proliferation: 'Unfortunately, the number of heuristics that have been proposed is so high that it is often difficult to predict which heuristic people will use in a novel situation and what the results will be'.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

To see heuristics in action, Kahneman and Tversky designed experimental probes where logical answers and intuitive answers come apart. As with the CRT, what they found was that a majority deliver the intuitive answer, demonstrating the operations of the automatic system. So, for instance, the use of representativeness was demonstrated via the (now famous) vignette:

> Linda is 31 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 anti-nuclear demonstrations. Which is more probable?

(i) Linda is a bank teller.
(ii) Linda is a bank teller and is active in the feminist movement. (Tversky and Kahneman 1982, p. 92, with slight modifications)

What Tversky and Kahneman found was that a majority of participants answer (ii). Yet clearly this cannot be the logical answer, for a conjunction of claims $A \wedge B$ can never be *more probable* than one conjunct, A, alone. As Thaler and Sunstein write:

> The error stems from the use of the representativeness heuristic: Linda's description seems to match 'bank teller and active in the feminist movement' far better than 'bank teller'. As Stephen Jay Gould … once observed, 'I know [the right answer], yet a little homunculus in my head continues to jump up and down, shouting at me—"But she can't just be a bank teller; read the description!"' Gould's homunculus is the Automatic System in action. (Thaler and Sunstein 2008, p. 29)

Although space prevents full presentation of the experiments run in support of the other heuristics, the general model is the same in each case. For instance, to demonstrate availability in operation, participants can be shown to overstate the likelihood of a given rare event (like a terrorist attack) when a past attack is made particularly salient for them (for example, via newspaper coverage). Or again, in a framing case, if participants are asked whether they would opt to have a hypothetical medical procedure, described in one scenario in terms of having 'an 80% chance of survival' and in another scenario as having 'a 20% chance of death', they reliably opt for the procedure more in the first protocol, described in terms of survival.[13] Yet

[13] The example follows the general shape of the infamous 'Asian flu' example from Tversky and Kahneman (1981).

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

an 80% chance of survival seems equivalent to a 20% chance of death, so the difference in response rates (due solely to the different framing of the same information) seems irrational. For each heuristic, then, experimental evidence supports the idea that we tend to deliver quick, intuitive responses to problems, and that these intuitive responses often fail to match the judgements demanded by classical rational choice theory. Rather than reason our way to logical decisions, people deploy fast but inaccurate heuristics which work well in general but which lead us into predictable errors in a variety of situations (Ariely 2008).

So, should we accept this conclusion, jettisoning the idea that individuals generally make decisions in line with the demands of rationality, instead relying on intuitive gut feels? I think there are reasons to be cautious. First, it's unclear that the experimental evidence in question really supports scepticism about rational decision making. Second, I think there are reasons to reject the overarching dual process framework to which this challenge belongs.

2.3. *Challenging the Experimental Evidence for the Automatic System*. Advocates of individual rationality might seek to challenge the dual process framework by rejecting the experimental evidence used to support it. For a start, as with much work in social psychology, questions can be asked about how well some of the (often most surprising or interesting) experimental results replicate.[14] However, pointing to the replication issue doesn't provide a full response, since many of the key results (such as those discussed above) do, in fact, replicate robustly. An alternative avenue for sceptics about the evidence to explore, then, involves concerns about 'ecological validity'—that is, how closely experimental settings mirror real-world scenarios. For it turns out that improving the ecological validity of experiments (modelling situations more closely on real-life decision making rather than abstract lab-based protocols) often raises the rate at which participants give logical (over intuitive) answers.[15]

[14] For instance, concerns have been raised about replication for Danziger et al.'s (2011) hugely influential claim that hungry judges give stiffer penalties (for example, Weinshall-Margel and Shapard 2011), while John Bargh's finding that people move more slowly after exposure to prompts about old age has failed numerous attempted replications. See the 'Many Labs: Investigating Variation in Replicability' project, https://osf.io/wx7ck/.

[15] See, for instance, Ho and Imai (2008) on the so-called 'ballot order effect'. See also §3.4 below.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

However, the challenge I want to focus on here *accepts* (at least some of) the experimental findings at face value but suggests that these results can be accommodated by advocates of individual rationality.[16] In particular, I want to suggest that a better sensitivity to the pragmatics of experimental prompts reveals that in many cases participants are in fact behaving far more logically than opponents give them credit for.

Take the classic CRT 'bat and ball' case (which I take to be the most compelling of the CRT questions). Here, 5¢ is the correct answer *only if* the experimental prompt is heard as asking a question about the *relative value* of the two items. Yet this is a very unusual kind of query. Another way to hear 'The bat costs $1 more than the ball', I suggest, is as telling you what you have to pay to acquire the bat *after buying the ball*. But if the experimental prompt is heard in that way, then the correct answer to 'How much is the ball?' is in fact '10¢', just as most people say. I contend that what is going on here is not, as advocates of the CRT suggest, mass irrationality as a result of the operations of the automatic system, but rather the simple fact that a majority of people hear the experimental prompt in the more practical way just suggested (while a minority of people are sensitive to the 'relative value' reading and thus give the '5¢' answer).

To reinforce the idea that a better understanding of pragmatics can accommodate the evidence without any appeal to irrationality consider the infamous case of Linda. Here the fact that the speaker bothers to spend time telling participants that Linda fits the stereotype for a feminist is itself informative, and it affects what the speaker is held to convey (Hertwig and Gigerenzer 1999; Mosconi and Macchi 2001). Experimental participants, no less than those engaged in ordinary conversations, are trying to grasp what the speaker is saying to them, and in the representativeness vignette it is clear that a reasonable conclusion is that the speaker intends to imply that Linda is a feminist (after all, why would they bother telling you all the stuff about her activism if not?). Yet if that is how the experimental materials are heard, then answering (ii) is not necessarily irrational, for it displays a highly reasonable sensitivity

[16] This is to agree with Glöckner (2016, p. 610) that 'sometimes there is a nonobvious rational basis for irrational-looking behavior'.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

to the pragmatics of conversation.[17] Finally, consider our example of framing effects, where participants were more likely to choose a procedure described as having 'an 80% chance of survival' over 'a 20% chance of death'.[18] First, note that there is a debate to be had about whether these two descriptions are in fact *semantically* equivalent (something required by any claim of irrationality based on non-equivalent responses) but, second, even waiving that concern, it is clear that the two claims are far from *pragmatically* equivalent.[19] Framing things in terms of the positive '80% survival' pragmatically conveys a hopeful message (that 80% is a good chance of survival in the circumstances), while choosing to focus on the negative (chance of death) pragmatically conveys that the speaker takes the risks to be particularly prominent. A number of theorists working on framing have developed precise mechanisms to explain how these pragmatic readings come about (for discussion, see Fisher and Mandel 2021), and I don't want to take a stand on the precise model to be preferred, but the general claim that framing effects arise due to pragmatic differences in content deserves, I think, serious consideration. In general, then, I suggest that a better appreciation of the fact that experimental participants, just like other conversational participants, respond to what is pragmatically expressed by the speaker, not just what is semantically conveyed, can help deflate allegations of irrationality based on experimental evidence.

[17] There is of course more to say here, since even if the prompt is heard as implying that Linda is a feminist bank teller, the point still stands that a conjunction should not be judged logically more probable than one of its conjuncts alone. However, further moves are certainly available. For instance, it might be argued (as does Gigerenzer 1993, p. 293) that participants are using a frequentist rather than a Bayesian understanding of probability. Alternatively, further pragmatic moves are possible: perhaps participants enrich option (i) ('Linda is bank teller'), hearing it as conveying that she's a bank teller but *not* a feminist, or perhaps they give 'more probable' a loose reading (akin to 'How do you think the world will be?') and operate in line with the Gricean maxim of quantity, giving the most informative description for which they have evidence.

[18] Discussion here has benefited from conversations with Sarah Fisher.

[19] The two descriptions are not semantically equivalent since there may be ways to not-die from the disease that don't depend on treatment. That is to say, some people might survive because they don't contract the disease in the first place or because they turn out to have a genetic mutation which helps them to fight off the illness, and so on (analogously, 'Team *A* won 3 out of 5 games' is not equivalent to 'Team *A* lost 2 out of 5 games', since some not-won games could have been drawn). There is also a related point concerning the interpretation of number claims as stating an exact number versus a threshold interpretation. These kinds of complications show that it is in fact harder than sometimes suggested to find two semantically equivalent phrases.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

2.4. *Challenging the Existence of the Automatic System.* A second worry with arguments based on the operations of the automatic system turns on whether the system really exists (in contrast to a logical system) at all. For it seems, on closer inspection, that there is no property this system could have that is capable of *both* holding apart automatic and logical thinking *and* underpinning a claim of systematic or predictable irrationality.[20] To see this, consider the range of properties appealed to above (and in the literature) when introducing the automatic system, where its operations are held not to be rule-based, and to be fast, unconscious, and effortless.[21] Yet it seems none of these properties are capable of differentiating two different systems.

So, for instance, it seems clear that automatic thinking cannot be characterized as thinking which is not rule-based. For, first, the automatic system is generally held to be computationally tractable, which requires it to be rule-based. Second, the examples of heuristics we have looked at pretty clearly constitute rules. For instance, the visual heuristic 'judge objects which appear blurry to be far away', and the representativeness heuristic 'assess the likelihood of x being F via x's similarity to other Fs', are both perfectly good rules, even if they are approximating rather than exact. So the presence or absence of rules can't be what differentiates the two kinds of system (a point acknowledged by Evans and Stanovich 2013, p. 231, who write, 'Evidence that intuition and deliberation are both rule-based cannot, by any logic, provide a bearing one way or the other on whether they arise from distinct cognitive mechanisms'). Alternatively, then, take the idea that automatic thinking is fast, while logical thinking is slow (an idea which has shaped the entire debate in this area, as evidenced by the title of Kahneman's best-selling *Thinking, Fast and Slow*). Yet it simply seems mistaken to hold that heuristic thinking is always fast or that logical thinking is always slow. Asked the question '15 + 5 = ?', or given a really simple logic puzzle (like 'If Maya had ice cream then she didn't get sweets. Maya had ice cream. Did she get sweets?'), participants are likely (assuming they think the questions are asked seriously) to be able to deliver a logical answer

[20] Samuels et al. (2002) also raise this objection, which they label the 'cross-over problem'. See also Evans (2012), Carruthers (2017).

[21] For example, Kahneman (2011, p. 20): 'System 1 operates automatically and quickly, with little or no effort'.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

extremely quickly.[22] Furthermore, the idea that logical answers can sometimes be as fast as non-logical ones is in fact already embedded in the experimental evidence to hand; for even in CRT cases experimenters found that a minority of people give logically correct answers as their immediate, fast responses. The same goes for the claim that non-logical responses are automatic or effortless, while logical answers are, in contrast, controlled or effortful. Although it is likely that the phenomenology of decision making differs across decision makers, across problems, and across contexts, I suspect that few feel the need to exert more control to arrive at the answer '20' when presented with '15 + 5 = ?' than when giving the answer that 'Linda is a feminist and a bank teller' in the classic representativeness experiment. Instead, I suggest, all these kinds of properties (speed, automaticity, conscious control) are ones that cross-cut any differences in the actual mechanisms of decision making, with both heuristic-based decisions and logical decisions sometimes being fast, sometimes being slow, and so on.

Finally, then, perhaps the advocate of a dual process model could simply appeal directly to the idea that while some decisions are based on the rules of logic, others are based on different, non-logical transitions. In this way, the automatic system would be defined simply through its opposition to the logical system (Sloman 1996). However, while this seems eminently plausible (since it seems highly unlikely that *every* decision we reach runs by way of classical logic or probability theory) it is a further step to the claim that this variability in methods entails irrationality. This takes us back to the assumption of §1 that classical rational choice theory provides the correct model of human rationality. For although the use of heuristics might not count as logical—though see Kruschke (2008) for an argument that even associative thinking should be treated as a form of probabilistic reasoning—this on its own does not show that such thinking is particularly problematic, that is, that it leads us into regular error. As Gigerenzer and colleagues have argued at length, the use

[22] This point has been recognized by advocates of dual process models recently, and it has led some (for example, De Neys 2018) to claim that the right approach is a 'hybrid' one, where the automatic system is capable of delivering *both* intuitive and (simple) logical answers. However, it is very unclear that this kind of model can do the explanatory work the dual process approach was originally introduced to do (for example, explaining the apparent prevalence of non-logical answers in experimental cases), or whether, once this concession is made, much of substance really remains of the dual process approach.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

of heuristics often yields *better* results than the application of logical rules to decision-making problems.[23] Yet if this is right, then it seems perverse to count use of a heuristic as irrational simply because it doesn't fit within the constraints of classical rational choice theory.

Neither logical nor heuristic rules *guarantee* that we get things right, but a guarantee of infallibility would clearly be too high a demand to place on rationality. And if infallibility isn't the bar, then it seems logic and heuristic thinking sit on all fours together, both potentially yielding the kind of *good enough* solutions we ought to demand of rational thinking. What this points to is the need (anticipated in §1) to refine classical rational choice theory in order to make it suitable as an account of human thinking. We should allow that individuals are rational if they:

(i) Aim at *good enough* solutions ('satisficing') rather than maximizing utility (a point recognized long ago in Simon's 1957 notion of 'bounded rationality').
(ii) Look at some but not necessarily all the evidence (Gigerenzer and Gaissmaier 2011).
(iii) Use a reasoning process appropriate to the problem (drawing from classical logic, probability theory *and* associative thinking).

With these kinds of relaxations on the standards for rationality in hand, together with a recognition of the role of pragmatic understanding in experimental contexts, I think we can reject the claim that individuals typically fall short of rational decision making.[24]

However, at this juncture a second line of attack heaves into view, according to which (contra the objection of §11) we do *try* to reason rationally but our attempts are prone to systematic failures. And these failures entail, just as advocates of the automatic system

[23] For instance, take the 'recognition heuristic', where a participant acts on the basis of whether or not a prompt feels familiar. Gigerenzer and Goldstein (1996, p. 41) argue that it can be shown mathematically that an individual operating with recognition alone will perform better on a binary choice test (like 'Which city is bigger, Detroit or Lakewood?') than someone operating with more knowledge (such as selecting city size on the basis of two cues—recognition plus another factor, such as the possession of a city football team). Note, however, that this point pushes further on what exactly is meant by a 'heuristic' (see Borg MS for further discussion).

[24] In line with Evans and Over (1996), Karlan (2021) distinguishes 'locally rational solutions' (good enough solutions for the problem in context) from 'ideally rational solutions' (modally robust results, in line with classical rational choice theory), arguing that both notions of rationality are needed.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

suggest, that individuals often fail to be rational. I turn to this second challenge in §III.

III

The Improper Use Challenge. The second challenge to claims of individual rationality from experimental evidence holds that agents do arrive at decisions or judgements using standard processes of evidence weighing and reasoned response, but their use of these processes fails (in some systematic way) to reach the standards required for rational decision making.[25] The evidence for this stance comes, it seems, from three main sources: the Wason selection task, biased evidence gathering or assimilation ('motivated reasoning'), and what is termed 'belief polarization due to belief disconfirmation'.

3.1. *The Wason Selection Task*.[26] In the classic task designed by Peter Wason, subjects are given a standard piece of conditional reasoning: they are told that there is a deck of cards where each card has a number on one side and a letter on the other. They are then shown a display revealing one side of four different cards, as in figure 1. Subjects are then asked:

- Which card(s) must be turned over to test the rule that if a card has a vowel on one side then it has an even number on the other side?

FIGURE 1. Wason selection task cards.

[25] As Mandelbaum (2018, p. 4) notes, what is needed to show a rationality approach false is 'to find actions that are not just the result of errors in processing. Rather, the irrationality has to result from a system that is set up to properly output the actions we categorize as irrational'.

[26] The Wason task is usually placed alongside the CRT, providing evidence for the automatic system. However, I think it is better viewed as supporting an improper use challenge. For in the Wason case participants do not have an immediate, intuitive feel for an answer; nor are they confident in their responses. Instead, participants think careful and deliberatively but still make mistakes.

© 2022 THE ARISTOTELIAN SOCIETY
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

What Wason (1966) found was that only around 10% of participants deliver the logically correct answer (which is 'E' and '7' in figure 1, as only the reverse of these cards could falsify the rule). Many participants, for instance, incorrectly think the '2' card needs to be turned over. However, turning over the '2' card is irrelevant as, if the '2' card has a vowel on the reverse, that fits with the rule, and if it has a non-vowel on the other side that doesn't matter (as the rule doesn't say 'all *and only* vowels have even numbers on the other side'). As Chater and Oaksford summarize:

> According to Popper's falsificationist view of science, it is never possible to confirm a rule, only to disconfirm it. The rule will be disconfirmed only if there is a logical contradiction between the hypothesis and the data; that is, this requires finding a *p*, *not-q* card, which is explicitly disallowed by the rule. Hence the *p* card should be chosen, because it might have a *not-q* on the back, and the *not-q* card should be chosen, because it might have a *p* on the back, and no other cards should be chosen. This is the 'logical' solution to the task … The fact that people do not follow this pattern has been taken as casting doubt on human rationality. (Chater and Oaksford (1999, p. 62)

3.2. *Myside Bias and Motivated Reasoning.* It turns out that subjects are prone to a range of mistakes around their search for and incorporation of evidence. They (i) hang on to beliefs in light of counter-evidence or don't search properly for counter-evidence ('belief preservation'), (ii) are overconfident (attributing too high a credence to extant beliefs), and (iii) are subject to belief polarization (where groups of like-minded individuals reinforce a shared belief via poor reasons). So, for instance, a subject who has already decided to buy a Honda rather than a Toyota will look longer at Honda ads and skip over Toyota ones (Mandelbaum 2018), revealing that the subject's prior beliefs improperly influence their search for evidence, leading them to seek out evidence that supports what they already believe and ignore or discount counter-evidence. This effect has been widely demonstrated in the experimental literature; for example, subjects who antecedently believe in the efficacy of capital punishment hear equivocal evidence as confirming the truth of this belief, while those who antecedently reject the death penalty hear the same evidence as supporting their belief (Lord et al. 1979), while smokers turn out to be more receptive to arguments denying a link between smoking

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

and lung cancer than non-smokers are (Brock and Balloun 1967). These flaws count as instances of what Mercier and Sperber (2018) call the 'myside bias' (which connects to the confirmation bias noted in §11). The existence of the myside bias spells trouble for claims of individual rationality, since:

> Not only does reasoning fail to fix mistaken intuitions, as this approach claims it should, but it makes people sure that they are right, whether they are right or wrong, and stick to their beliefs for no good reason. Historical examples attest that these are not minor quirks magnified by clever experiments, but real phenomena with tragic consequences. (Mercier and Sperber 2018, p. 244)

While our biased treatment of evidence means, Mercier and Sperber argue, that we must reject the idea that *individuals* are rational, it is nevertheless, they suggest, an adaptive feature, for it underpins the key role that reasoning plays in arriving at better *group* decisions (compare Dutilh Novaes 2018). Since everyone somewhat blindly advocates for their own extant beliefs, the group as a whole is exposed to the strongest possible arguments for each side. Myside bias, Mercier and Sperber (2018, p. 331) argue, thus reveals that reason is fundamentally a social competence and its 'bias and laziness aren't flaws; they are features that help reason fulfil its function'.

3.3. *Protecting an Individual's Sense of Self.* Mandelbaum (2018) is concerned with the peculiar phenomenon whereby subjects sometimes appear to *increase* their level of belief in a given proposition in the face of *accepted* disconfirming evidence ('belief polarization due to belief disconfirmation'). Yet this is the exact opposite of what is predicted by a Bayesian belief-updating procedure. The clearest example of this phenomenon occurs with respect to religious beliefs (Batson 1975), and a particularly noteworthy case involves millennial cults, where members sometimes increase their belief in the truth of the cult's proclamations *after* the date on which the world was previously predicted to end (Dawson 1999). As Mandelbaum stresses, however, the phenomenon also apparently occurs with more mundane beliefs, and he cites experimental work concerning things like an individual's attitude to the safety of nuclear power (Plous 1991) or affirmative action or gun control (Taber and Lodge 2006). These kinds of cases—where, as Mandelbaum argues, subjects seem to ditch Bayesian principles in order to protect their sense of self, hanging on to beliefs that shape their understanding of themselves (for

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

instance, as possessing a particular religious orientation, or as affiliated with certain political or social groupings)—demonstrate that the systems of belief updating we use are systematically designed in a way that can yield non-rational decisions (see also Kahan 2017). That is to say, the system is 'set up to properly output the actions we categorize as irrational' (Mandelbaum 2018, p. 4).

So should we accept that experimental evidence demonstrates that our use of logical decision-making systems is systematically flawed?

3.4. *Challenging the Evidence for Flawed Use of Logical Systems.* A first point to note is that standard answers in the Wason task are only wrong/irrational *if* subjects should be reasoning using Popperian (rather than Bayesian) methods. That is to say, turning over the '2' card in our above example, while it can't falsify the rule, could provide further evidence capable of raising an individual's credence in the rule. So, from a Bayesian perspective, turning over the '2' card might count as acceptable (Oaksford and Chater 1994). Furthermore, as noted in §2.3, individuals tend to do better in reasoning tasks if ecological validity is increased, and as Cosmides and Tooby (1992) have shown, this is certainly the case with the Wason task.[27] So the idea that the Wason task really challenges claims of individual rationality is, I think, moot (Stenning and van Lambalgen 2008).

More importantly, however, I want to take issue with the assumption behind the argument from Mercier and Sperber (and at play in many other places) that motivated reasoning (where what one already believes shapes one's search for or incorporation of evidence) is not, generally speaking, individually rational. Since we have already accepted (§2.4) that rational processes may look at *some* rather than *all* evidence and arrive at *good enough* (rather than *optimal*) decisions, a process which ignores some counter-evidence can nevertheless, it seems, count as entirely rational.[28] Claiming that motivated reasoning is always irrational would ignore the fact that people usually have reasons for believing what they do (for example, evidence from testimony or perception), and that adopting new hypotheses is itself not a cost-free exercise (requiring cognitive

[27] Asked to turn over cards relating to the rule 'You must be over 21 to drink alcohol', participants were far more likely to turn over the potentially falsifying cards showing, for example, 'beer' or '18'.

[28] Mandelbaum (2018) notes that some Bayesian models actually *predict* biased assimilation.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

expenditure in terms of revising previous sets of beliefs, and so on). So the idea that there should be a degree of 'stickiness' about the beliefs one already holds, or that a subject should restrict how much evidence she looks for, far from showing that individual rationality does not hold, seems to be something we might *predict*, given the assumption that individuals are, broadly speaking, rational.[29] Skipping Toyota ads once one has made up one's mind to buy a Honda, contra the claims above, can be rational, as it allows one to shift scarce cognitive reserves to other decision-making issues.

Of course, there is a balance to be struck here: too great a degree of belief preservation (that is, scepticism about presented evidence) *can* slide over into irrational fact-blindness or unwarranted dismissal of counter-evidence, and work needs to be done in any particular case to discover where the line should be drawn (Kahan 2017). Yet in principle, I suggest, motivated reasoning can be rational. And in fact, this point is already recognized within the experimental work cited in support of this second challenge to individual rationality. Thus Taber and Lodge write:

> How we determine the boundary line between rational skepticism and irrational bias is a critical normative question, but one that empirical research may not be able to address. Research can explore the conditions under which persuasion occurs (as social psychologists have for decades), but it cannot establish the conditions under which it should occur. It is, of course, the latter question that needs answering if we are to resolve the controversy over the rationality of motivated reasoning. (Taber and Lodge 2006, p. 768)

This is to recognize that the mere existence of motivated reasoning is not on its own sufficient to challenge claims of individual rationality.

Finally, then, turning to belief polarization due to belief disconfirmation, I suggest that the evidence for this phenomenon with respect to more 'ordinary' beliefs (such as belief in the safety of nuclear power) is in fact fairly weak. For the experimental work in these cases seems better treated as providing examples of motivated cognition. Someone who antecedently believes in the safety of nuclear power may well be more likely than others to interpret an ambiguous or equivocal piece of evidence in a way that supports their extant belief (Plous 1991), but in and of itself, as argued above, this

[29] For 'stickiness' see Kripke (2011, p. 49), Levy (2022).

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

is not an objection to individual rationality. What would need to be shown in addition is that a given individual has strayed across the line from warranted belief preservation into irrational fact blindness. Yet if this can be shown, a supporter of individual rationality, no less than a sceptic, might be happy to treat this individual as failing to be rational about this issue. For the claim that individuals generally act in line with what they believe, desire, and so on, is supposed to be a claim about *typical* intentional action, rather than a universal claim about all decisions. Clearly there will be exceptions. Relatedly, although Mandelbaum may well be right that for certain kinds of beliefs (for instance, religious beliefs) Bayesian reasoning does indeed break down (in favour of some kind of self-protection), still it is unclear whether this is sufficient to show assumptions of individual rationality mistaken (Van Leeuwen 2014). An advocate of individual rationality must already allow that there are a range of cases where a further story needs to be told, for instance, allowing for akrasia, cases of conflicting occurrent and dispositional beliefs, and episodes of mental disorder (see, for example, Egan 2008 on compartmentalized beliefs). That the domain of non-rational thinking may extend to matters of *faith*, such as religious beliefs, seems to me to be a conclusion that a supporter of individual rationality could well live with.

IV

Conclusion. The claim that, typically, our intentional actions are a reflection of our reasons has been challenged in recent years by a range of experimental evidence which seems to show either that we commonly make use of a non-logical/automatic decision-making system or that our use of our logical systems is systematically flawed. However, I have tried to suggest that the rejection of individual rationality on the basis of this experimental evidence may well be too swift. Contra advocates of the automatic system, I think we should reject much of the evidence for irrational decision making (explaining it through greater sensitivity to pragmatics). I also suggested that we should reject the entire dual process model on which this first challenge is based. Contra the second challenge, I argued that belief preservation or motivated reasoning can be (at least in principle) rational. Only belief polarization due to belief disconfirmation is truly problematic, but these are special cases where claims

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

of rationality may indeed be tenuous. I conclude, then, that experimental evidence should not (currently) be taken to refute claims of individual rationality.[30]

Department of Philosophy,
University of Reading,
Whiteknights,
Shinfield Road, Reading RG6 6UR, UK
e.g.n.borg@reading.ac.uk

References

Ariely, Dan 2008: Predictably Irrational: The Hidden Forces that Shape Our Decisions. London: HarperCollins.

Batson, C. Daniel 1975: 'Rational Processing or Rationalization? The Effect of Disconfirming Information on a Stated Religious Belief'. Journal of Personality and Social Psychology, 32(1), pp. 176–84.

Borg, Emma MS: Compelling Reasons. Unpublished manuscript.

Brock, Timothy C., and Joe L. Balloun 1967: 'Behavioral Receptivity to Dissonant Information'. Journal of Personality and Social Psychology, 6(4), pp. 413–28.

Carruthers, Peter 2017: 'Mindreading in Adults: Evaluating Two-Systems Views'. Synthese, 194(3), pp. 673–88.

Chater, Nick, and Mike Oaksford 1999: 'Ten Years of the Rational Analysis of Cognition'. Trends in Cognitive Sciences, 3(2), pp. 57–65.

Cosmides, Leda, and John Tooby 1992: 'Cognitive Adaptions for Social Exchange'. In Jerome H. Barkow, Leda Cosmides, and John Tooby (eds.), The Adapted Mind: Evolutionary Psychology and the Generation of Culture, pp. 163–228. Oxford: Oxford University Press.

Danziger, Shai, Jonathan Levav, and Liora Avnaim-Pesso 2011: 'Extraneous Factors in Judicial Decisions'. Proceedings of the National Academy of Sciences, 108(17), pp. 6889–92.

Davidson, Donald 1963: 'Actions, Reasons, and Causes'. Journal of Philosophy, 60(23), pp. 685–700.

Dawson, Lorne L. 1999: 'When Prophecy Fails and Faith Persists: A Theoretical Overview'. Nova Religio, 3(1), pp. 60–82.

De Neys, Wim 2018: 'Bias, Conflict and Fast Logic'. In W. De Neys (ed.), *Dual Process* 2.0, pp. 47–65. London: Routledge.

[30] Thanks are due to Nat Hansen, Sarah Fisher, Guy Longworth, and audiences at the Aristotelian Society, St Andrews and Reading. This work was supported by a Leverhulme Trust Major Research Fellowship MRF-2019-031.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

Doris, John M. 2015: *Talking to Our Selves: Reflection, Ignorance, and Agency*. Oxford: Oxford University Press.
Dutilh Novaes, Catarina 2018: 'The Enduring Enigma of Reason'. Mind and Language, 33(5), pp. 513–24.
Egan, Andy 2008: 'Seeing and Believing: Perception, Belief Formation and the Divided Mind'. Philosophical Studies, 140(1), pp. 47–63.
Evans, Jonathan St. B. T. 2009: 'How Many Dual-Process Theories Do We Need? One, Two, or Many?' In Jonathan St. B. T. Evans and Keith Frankish (eds.), In Two Minds: Dual Processes and Beyond, pp. 33–54. Oxford: Oxford University Press.
——2012: 'Questions and Challenges for the New Psychology of Reasoning'. Thinking and Reasoning, 18(1), pp. 5–31.
——and David E. Over 1996: Rationality and Reasoning. Hove: Psychology Press.
——and Keith E. Stanovich 2013: 'Dual-Process Theories of Higher Cognition'. Perspectives on Psychological Science, 8(3), pp. 223–41.
Fisher, Sarah A., and David R. Mandel 2021: 'Risky-Choice Framing and Rational Decision-Making'. Philosophy Compass, 16, p. e12763.
Frederick, Shane 2005: 'Cognitive Reflection and Decision Making'. Journal of Economic Perspectives, 19(4), pp. 25–42.
Gigerenzer, Gerd 1993: 'The Bounded Rationality of Probabilistic Mental Models'. In K. I. Manktelow and D. E. Over (eds.), Rationality: Psychological and Philosophical Perspectives, pp. 284–313. London and New York: Routledge.
——and Wolfgang Gaissmaier 2011: 'Heuristic Decision Making'. Annual Review of Psychology, 62, pp. 451–82.
——and Daniel G. Goldstein 1996: 'Reasoning the Fast and Frugal Way: Models of Bounded Rationality'. Psychological Review, 103(4), pp. 650–69.
Glöckner, Andreas 2016: 'The Irrational Hungry Judge Effect Revisited: Simulations Reveal that the Magnitude of the Effect Is Overestimated'. Judgment and Decision Making, 11(6), pp. 601–10.
Hertwig, Ralph, and Gerd Gigerenzer 1999: 'The "Conjunction Fallacy" Revisited: How Intelligent Inferences Look Like Reasoning Errors'. Journal of Behavioral Decision Making, 12(4), pp. 275–305.
Ho, Daniel E., and Kosuke Imai 2008: 'Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment: The California Alphabet Lottery, 1978–2002'. Public Opinion Quarterly, 72(2), pp. 216–40.
Kahan, Dan M. 2017: 'Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition'. Cultural Cognition Project Working Paper Series 164, Yale Law School, Public Law Research Paper 605, Yale Law and Economics Research Paper 575. Available at Social Science Research Network, http://dx.doi.org/10.2139/ssrn.2973067.
Kahneman, Daniel 2011: Thinking, Fast and Slow. London: Penguin.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

Karlan, Brett 2021: 'Reasoning with Heuristics'. Ratio, 34(2), pp. 100–8.
Kripke, Saul A. 2011: Philosophical Troubles: Collected Papers, Volume 1. Oxford: Oxford University Press.
Kruschke, John K. 2008: 'Bayesian Approaches to Associative Learning: From Passive to Active Learning'. Learning and Behavior, 36(3), pp. 210–26.
Kühberger, Anton 2002: 'The Rationality of Risky Decisions'. Theory and Psychology, 12(4), pp. 427–52.
Levy, N. 2022: Bad Beliefs: Why They Happen to Good People. Oxford: Oxford University Press.
Lieder, Falk, and Thomas L. Griffiths 2020: 'Resource-Rational Analysis: Understanding Human Cognition as the Optimal Use of Limited Computational Resources'. Behavioral and Brain Sciences, 43(e1), pp. 1–60.
Lord, Charles G., Lee Ross, and Mark R. Lepper 1979: 'Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence'. Journal of Personality and Social Psychology, 37(11), pp. 2098–109.
Mandelbaum, Eric 2018: 'Troubles with Bayesianism: An Introduction to the Psychological Immune System'. Mind and Language, 34(2), pp. 1–17.
Mercier, Hugo, and Dan Sperber 2018: The Enigma of Reason: A New Theory of Human Understanding. London: Penguin.
Mosconi, Giuseppe, and Laura Macchi 2001: 'The Role of Pragmatic Rules in the Conjunction Fallacy'. Mind and Society, 2(1), pp. 31–57.
Oaksford, Mike, and Nick Chater 1994: 'A Rational Analysis of the Selection Task as Optimal Data Selection'. Psychological Review, 101(4), pp. 608–31.
Plous, S. 1991: 'Biases in the Assimilation of Technological Breakdowns: Do Accidents Make Us Safer?' Journal of Applied Social Psychology, 21(13), pp. 1058–82.
Samuels, Richard, Stephen Stich, and Michael Bishop 2002: 'Ending the Rationality Wars: How to Make Disputes about Human Rationality Disappear'. In Renée Elio (ed.), Common Sense, Reasoning and Rationality, pp. 236–68. New York: Oxford University Press.
Simon, Herbert A. 1957: Models of Man—Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting. New York: John Wiley.
Sloman, Steven A. 1996: 'The Empirical Case for Two Systems of Reasoning'. Psychological Bulletin, 119(1), pp. 3–22.
Stein, Edward 1996: Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science. Oxford: Clarendon Press.
Stenning, Keith, and Michiel van Lambalgen 2008: Human Reasoning and Cognitive Science. Cambridge, ma: mit Press.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022

Taber, Charles S., and Milton Lodge 2006: 'Motivated Skepticism in the Evaluation of Political Beliefs'. American Journal of Political Science, 50(3), pp. 755–69. https://doi.org/10.1111/j.1540-5907.2006.00214.x.

Thaler, Richard H., and Cass R. Sunstein 2008: Nudge: Improving Decisions about Health, Wealth and Happiness. London: Penguin.

Thomson, Keela S., and Daniel M. Oppenheimer 2016: 'Investigating an Alternate Form of the Cognitive Reflection Test'. Judgment and Decision Making, 11(1), pp. 99–113.

Toplak, Maggie E., Richard F. West, and Keith E. Stanovich 2011: 'The Cognitive Reflection Test as a Predictor of Performance on Heuristics-and-Biases Tasks'. Memory and Cognition, 39(7), pp. 1275–89.

Trouche, Emmanuel, Emmanuel Sander, and Hugo Mercier 2014: 'Arguments, More than Confidence, Explain the Good Performance of Reasoning Groups'. Journal of Experimental Psychology: General, 143(5), pp. 1958–71.

Tversky, Amos, and Daniel Kahneman 1974: 'Judgment under Uncertainty: Heuristics and Biases'. Science, 185, pp. 1124–31.

————1981: 'The Framing of Decisions and the Psychology of Choice'. Science, 211, pp. 453–8.

————1982: 'Judgments of and by Representativeness'. In Daniel Kahneman, Paul Slovic and Amos Tversky (eds.), Judgment under Uncertainty: Heuristics and Biases, pp. 84–98. Cambridge: Cambridge University Press.

Van Leeuwen, Neil 2014: 'Religious Credence Is Not Factual Belief'. Cognition, 133(3), pp. 698–715.

Wason, Peter C. 1966: 'Reasoning'. In Brian M. Foss (ed.), New Horizons in Psychology, pp. 135–51. Harmondsworth: Penguin.

Weinshall-Margel, Karen, and John Shapard 2011: 'Overlooked Factors in the Analysis of Parole Decisions'. Proceedings of the National Academy of Sciences, 108(42), p. E833.

© 2022 The Aristotelian Society
Proceedings of the Aristotelian Society, Vol. XX, *Part* XX
https://doi.org/10.1093/arisoc/aoac009

Downloaded from https://academic.oup.com/aristotelian/advance-article/doi/10.1093/arisoc/aoac009/6760361 by University of Reading user on 19 October 2022