PERSPECTIVE
https://doi.org/10.1038/s41562-019-0617-3

nature human behaviour

Corrected: Publisher Correction

# A synthesis of behavioural and mainstream economics

Robert J. Aumann

**Mainstream economic theory is based on the rationality assumption: that people act as best they can to promote their interests. In contrast, behavioural economics holds that people act by behavioural rules of thumb, often with poor results. We propose a synthesis according to which people indeed act by rules, which usually work well, but may work poorly in exceptional or contrived scenarios. The reason is that like physical features, behavioural rules are the product of evolutionary processes; and evolution works on the usual, the common—not the exception, not the contrived scenario.**

The rationality hypothesis—that people act to promote their interests—underlies most of economic theory and indeed economics as a whole. Economic policy revolves largely around the creation of incentives for people to act as the policy maker would like; and to act according to incentives is, of course, to act rationally. Courses in basic economics, in price theory, and in microeconomics revolve around maximizations of utility, first- and second-order conditions, and so on—i.e., rationality. Applications of economic theory to various and sundry areas such as law, criminology, marriage, patents, health, finance, pensions, sports, what have you, are all based on optimization—i.e., rationality.

But for the better part of a century, the rationality hypothesis has been called into question, or modified, in one way or another. Thus Simon[1] suggested the notion of satisficing: that people do not maximize, but only seek an acceptable level of utility; also, that people use heuristics rather than calculating optima[2]. Friedman[3,4] promulgated the 'as if' doctrine: that people do not consciously optimize, but only act as if they do. In experiments such as probability matching[5] and the ultimatum game[6], subjects deviated systematically from utility maximization. And then came behavioural economics (BE): the study of systematic irrationality.

In practice, mainstream economics (ME) uses mathematical models to study how economic agents (consumers, producers, merchants, monopolists, oligopolists, ...) should behave to further their interests; the implicit assumption being that in the real world, 'should' somehow becomes 'do'. In contrast, BE uses little or no mathematics. Rather, it uses surveys and laboratory experiments to study directly not how people should behave, but how they do behave. In the surveys, people are asked how they would act in certain situations or how they respond to certain questions; in the experiments, it is observed how they do act. Dozens of heuristics and 'biases' have been identified: rules of thumb that in the surveys lead to patently incorrect—indeed illogical—responses and in the experiments to patently suboptimal behaviour. In the last two or three decades, BE has become a discipline on its own, with journals, conferences, university courses, summer schools, practical applications and subdisciplines (such as behavioural finance).

Nevertheless, rationality remains the central paradigm of ME. At the same time, the challenge posed by BE—that in fact, people behave irrationally—has not been satisfactorily addressed. ME and BE continue to live uneasily side by side, in spite of the apparent contradiction.

## Rule-rationality

Viewing ME's rationality as a thesis[7], and BE's systematic irrationality as its antithesis, we propose a synthesis: rule-rationality. Ordinary (henceforth 'act-') rationality means that the decision-maker chooses an act that maximizes utility among all acts available in that situation. In contrast, under rule-rationality people do not maximize over acts. Rather, they adopt rules of behaviour that do well in usual, naturally occurring situations; then, when deciding, they choose an act that accords with the adopted rule. Usually these acts maximize utility in the situation at hand, but not always; the decision-maker adopts rules rather than choosing acts.

In general, the rules are not consciously adopted; their adoption results from evolutionary forces, genetic or memetic[8], or from learning, conscious or subconscious (henceforth subsumed under evolution). Specifically, they evolved because they prescribe optimal behaviour in usual, commonly occurring situations; but in exceptional or contrived situations, they may lead to systematically inferior results, because then evolutionary pressures do not apply. Evolution is driven by survival of the fittest, and the fittest are determined by their behaviour in commonly occurring situations (needless to say, situations contrived by an experimental scientist play no role at all in evolution).

An example is the rule 'eat when you have appetite'. For most people, the prescribed behaviour is act-rational, but not for the overweight. The physiological mechanism of appetite evolved to make us eat; our bodies need food. But evolution has not yet had time to take account of the sedentary nature of much of modern life, so the rule may 'misfire', resulting in act-irrational overeating.

Overeating is act-irrational also at the opposite end of the spectrum, for the severely undernourished; there are documented cases of Holocaust survivors who tragically died of overeating upon liberation. Again, evolution did not design the system to deal with that situation, because it is unusual.

Like eating, BE's heuristics are rule-rational: they prescribe act-rational behaviour in usual, commonly occurring situations, because those are the situations to which evolution applies. In unusual or contrived situations, the heuristics may well misfire—prescribe act-irrational behaviour—because evolution does not apply there. It follows that most economic behaviour is indeed act-rational.

BE's founding fathers, Daniel Kahneman and Amos Tversky (KT), themselves remark[9] that "[i]n general, these heuristics are

Department of Mathematics and Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, Jerusalem, Israel.
e-mail: raumann@math.huji.ac.il

quite useful, but sometimes they lead to severe and systematic errors," and again some twenty years later[10], "these heuristics are often useful, but they sometimes lead to characteristic errors or biases." BE traditionally emphasizes the 'sometimes'—i.e., the severe and systematic errors, where the heuristics do not work. But more important is the insight that BE gives us into the 'in general'—where the heuristics do work.

In form—not substance!—the distinction between act- and rule-rationality resembles that between the moral dogmas of act-utilitarianism[11], which prescribes acting to maximize social welfare, and rule-utilitarianism[12], which prescribes acting by rules that generally prescribe act-utilitarian behaviour. Act-utilitarianism justifies Raskolnikov's murder[13] of the old money-lending hag; rule-utilitarianism condemns it. Indeed the term 'rule-rationality' stems[14] from this formal resemblance.

## Examples of rule-rationality

**Linda.** In one of KT's best-known works[15], survey respondents were told of an imaginary Linda, young, single, outspoken and very bright, who as a student was deeply concerned with discrimination and social justice. They were asked whether Linda is more likely to be a bank teller, or a bank teller and an active feminist. The overwhelming response was that 'feminist bank teller' is more likely than 'bank teller', which makes no sense. (Every feminist bank teller is a bank teller.)

Clearly this is a contrived situation, a trick question. The rule (heuristic) involved here is a conversational maxim[16]: a convention of speech, implicit in what is being said, though not explicit. The maxim is that of relevance; the respondents subconsciously take what they are told—indeed what is emphasized—as relevant, which in fact it usually is. When the questioner purposely emphasizes an irrelevant aspect of the situation, they are thrown off; they do not expect to be tricked.

**St. Ives**. Another example of the relevance maxim is the famous nursery rhyme[17]:

As I was going to St. Ives
I met a man with seven wives,
Every wife had seven sacks,
Every sack had seven cats,
Every cat had seven kits.
Kits, cats, sacks and wives,
How many were going to St. Ives?

The kits, cats, sacks and wives, plus the narrator and the man he met, total 2,802. But 2,802 is the wrong answer. The correct answer is 1: the man with the wives, sacks, cats and kits was likely coming from—not going to!—St. Ives. They are irrelevant.

The long and the short of it is that providing relevant information is the norm, the usual, the rule; irrelevant information is exceptional, unusual. So the relevance maxim emerges as a rule that subconsciously governs the responses. People who answer 2,802 are thus rule-rational, though not act-rational.

**Anchoring.** The 'anchoring effect' is the tendency to be influenced by irrelevant numbers. In one study[9], survey respondents observed the spin of a roulette wheel that had been fixed to stop at either 65 or 10. Immediately afterwards, they were asked to estimate the percentage of United Nations countries that are located on the African continent. It turned out that the average estimate of respondents who had seen the wheel stop at 10 was significantly less than that of those who had seen the wheel stop at 65 (25% vs. 45%). Similar effects obtained in many different estimation tasks.

This differs from the previous examples in two ways. First, here the irrelevancy is entirely open and above-board; no attempt is made to trick the respondents into thinking that the roulette wheel is relevant. That it nevertheless significantly influences the responses is eloquent testimony to the deep-seated nature of the relevance rule. Perhaps, indeed, this rule is genetic, has evolved over the millennia; as information that is provided usually is relevant, automatically treating it as such makes for more efficient information-processing.

Second, unlike the 'Linda' and 'St. Ives' effect, the anchoring effect is important in the real world; specifically, in bargaining. A pregnant elephant carved in stone is offered to a Western tourist in India, who has no idea of the market price. The vendor asks for 2,000 rupees; the tourist, who knows he must bargain, offers 300. Thereupon the vendor bargains him up to 600, and the deal is closed at that price; the sale price is largely determined by the asking price. The next customer is Indian and buys a similar elephant from the same vendor for 50 rupees.

But even in the real world, the effect is only moderately important. If you are buying a house or a car in your hometown, rather than a pregnant elephant in a far-off country, you know the market, and the vendor knows that you know. Though anchoring may still kick in, it is a second-order effect.

**Focusing.** The 'focusing bias' is the tendency to place 'too' much importance on one aspect of an event. Respondents in a survey[18] were asked to rank the following outcomes from most to least likely (assuming that Björn Borg reaches the 1981 Wimbledon finals):

Borg will win the match.
Borg will lose the first set.
Borg will lose the first set but win the match.
Borg will win the first set but lose the match.

Most (72%) of the respondents rated 'Borg will lose the first set but win the match' more likely than 'Borg will lose the first set', violating the laws of probability (the former entails the latter).

Both the Linda and Borg scenarios are instances of the 'conjunction fallacy', in which the conjunction of two events is deemed more likely than one of them[18]. But though the errors are similar in form, the underlying rules are different. In Linda, it is a matter of relevance; the respondents are tricked into believing that the information provided is relevant, so they concentrate on that information. In the current example, respondents focus on Borg, whom they know to be outstanding. If they're told only that he loses the first set, he sounds like a loser; if they're told that he loses the first set but wins the match, he sounds like a winner, which seems more likely.

To be sure, placing too much importance on one aspect of an event is not a good idea; that's what 'too' means. But how much is too much? That depends on the circumstances. When the incentives justify devoting considerable time and effort to analyzing the situation, then focusing on one aspect is likely to be suboptimal. But in the Borg example that was not the case; it's safe to assume that the incentives did not justify spending much time to come up with a response. That means that it would be irrational to do so; using a shortcut that focuses on the salient feature—in this case Borg's ability—seems not only reasonable, but optimal.

**Immediacy.** It has long been observed that people act impatiently today but plan to act patiently in the future. Thus, offered a choice between \$10 on the spot and \$11 tomorrow, some experimental subjects choose \$10 on the spot; whereas the same subjects, offered a choice between \$10 in a year and \$11 in a year and a day, choose \$11 in a year and a day, which may be viewed as a violation of act-rationality.

But it does not violate rule-rationality. A well-known proverb, or rule, has it that 'a bird in the hand is worth two in the bush'. If you give me \$10 now, I pocket it, and that's the end of the story. \$11 tomorrow? Maybe yes, maybe no; there's a qualitative difference between now and later. Between 365 and 366 days, there is no such difference.

Particularly interesting in this connection is that separate neural systems are responsible for evaluating immediate and delayed

rewards[19]; specifically, that "short-run impatience is driven by the limbic system, [...] whereas long-run patience is mediated by the lateral prefrontal cortex." Thus for the immediacy bias, which is act-irrational but rule-rational, there is a physiological mechanism, much like appetite for eating.

**Certainty.** The 'certainty effect'[20] is a discontinuity in the evaluation of gambles when passing from near-certainty to certainty. For example, $100,000 with certainty might be preferred to a gamble yielding $150,000 with probability 0.99 and nothing otherwise.

In practice, the discontinuity is well justified. Probability assessments in everyday life are rarely objective (i.e., governed by coin tosses, roulette wheels, or the like). When you invite people to an intimate dinner with a handful of carefully chosen guests, and they say they are 99% certain they'll come, that means that they want to be counted in but reserve the right to opt out. When a contractor tells you that he is 99% certain your house will be ready in eight months, you had better figure at least a year. Like the distinction between now and later, there is a qualitative difference in everyday parlance between certainty and probability 0.99. So when you hear 99%, even if the probabilities are ostensibly objective, your mind fixates on the usual connotation, which is qualitatively different from 100%. Indeed, subjects may suspect, consciously or subconsciously, rightly or wrongly, that the dice are loaded[21].

**The ultimatum game.** In the ultimatum game[6] (UG), two players, the proposer (P) and responder (R), must divide an endowment of $100. If they agree on the division, each gets his or her agreed share. If they do not agree, neither gets anything. The players do not sit face to face and cannot communicate directly. Rather, they sit at computer consoles in separate rooms. P starts by making an offer to R; the offer is numerical only, with no accompanying words. R can respond only by typing 'yes' or 'no' into the computer; no other response is allowed. Once she responds, the game is over. After that, the players get their payoffs (if any) and leave by separate doors. At no stage do they see each other or learn each other's identity. The subjects are students, presumably not particularly long on money.

In this situation, one might expect P to offer R a non-negligible amount—say $10, taking $90 for himself—and for R to accept. That is because there is no rational reason for R to walk away from a non-negligible amount of money; and taking this as given, a rational P should maximize his payoff.

But that is not what happens in the laboratory. Most of the offered splits are around 65–35. And when they are considerably less—say 80–20—they are rejected: R actually walks away from as much as $20.

On the face of it, this seems a clear violation of act-rationality. Not on the part of P, who—perhaps foreseeing the response—is rational in not risking rejection; but on the part of R.

Possible explanations include wounded pride, feeling insulted, self-respect and desiring revenge. Another possible explanation is that R wishes to establish a reputation for rejecting lopsided offers, so that in future negotiations, she will not get such offers. But that explanation does not hold water, because the game is played anonymously; the players' identities are not revealed, so reputations can be neither established nor destroyed.

There are two ways of viewing pride, insults, self-respect and revenge. One is that they themselves are legitimate sources of utility and disutility, so R's rejection of an 80–20 offer is entirely rational; she actually gets positive utility from taking revenge and would get negative utility from accepting an insulting offer. That is a perfectly consistent, logical position.

But conceptually and methodologically, it is not quite satisfactory; one might wish to delve deeper. Rather than taking emotions like the above as given, one might wish to account for them in terms of more fundamental human needs. What purpose—adaptive or otherwise—is served by feeling insulted or taking revenge? What is the function of self-respect?

That's where rule-rationality comes in. Even though it isn't act-rational for R to reject an 80–20 offer, it is rule-rational to do so. As a rule, one should reject lopsided offers, precisely for the reputational reason discussed above: so as to be treated more even-handedly in the future. People use this rule because it is usually act-rational: specifically, it is act-rational in almost all—or all—natural, real-world negotiations, which are not anonymous. The mechanism for executing the rule is a combination of one or more of the emotions discussed above—self-respect, wounded pride, feeling insulted, desiring revenge—which evolved, genetically or memetically, because they usually maximize fitness. In the laboratory, the rule does not lead to act-rationality, which would call for R to accept any positive sum. R here uses the rule that evolved in the natural setting, with face-to-face negotiations; not in the contrived, artificial laboratory setting.

**The dictator game.** Until now, the entire discussion has been about R; it is her behaviour that is rule- but not act-rational. The behaviour of P, who in these experiments usually proposes at least $30, is act-rational, since he may rightly fear rejection by R. But though indeed act-rational, it may also result from applying rules or norms such as fairness, regard for others, appropriate behaviour and decency.

To decide between these alternatives (act- or rule-rationality on P's part), consider what is known as the dictator game[22] (DG). This is like the UG, except that R (here the recipient) cannot reject P's offer; P is in fact not a proposer, but a dictator, who decides by himself on the division of the endowment. In the laboratory, P often grants R a non-negligible proportion of the endowment. As P need not fear rejection, any such amount is act-irrational and must be attributed to rules such as the above (fairness, etc.). So these rules emerge as mechanisms that lead P to behave act-rationally—i.e., make acceptable offers—in the UG and more generally in all negotiations; whereas in the DG, which represents an unusual situation, these same rules lead P to behave act-irrationally. Thus both P and R behave rule-rationally in both the ultimatum and dictator games.

Fairness may be used to explain also R's (as well as P's) behaviour in the UG, by viewing rejections as costly punishments[23] for unfair offers. R then emerges as prosocial, public-spirited—suffering personal loss for the good of Society—'teaching P a lesson' at his own cost. This sounds very different from the self-driven motives (pride, etc.) mentioned above.

To decide between the two, an experiment was conducted[24] in which each subject played the role of R in the UG and also that of P in the DG. Among the subjects who rejected low offers in the UG, some granted considerable sums in the DG, which is consistent with fairness norms; others granted little or nothing in the DG, which is consistent with self-driven norms. Thus the two kinds of norm may coexist.

Our analysis goes beyond identifying the relevant mechanisms—it examines their evolutionary genesis. In most negotiations, reputations are at stake, so rejecting low offers is act-rational; both fairness-type and self-driven mechanisms lead to this outcome, so both kinds evolved. But the DG is not a negotiation; mechanisms that evolved in the context of negotiations may carry over to the DG, or they may not. It is perfectly consistent to grant little or nothing in the DG, while still rejecting low offers in the UG because they're unfair.

Indeed, the UG and DG are completely different scenarios; it is not at all clear why heuristics applying to one should apply to the other, even without touching on their evolution. In the UG, there is an organic relationship between P and R; the outcome depends on both. No such relationship exists in the DG. Why should P grant anything at all to a totally anonymous R? What's 'fair' about that? If he wants to be prosocial, why doesn't he take the entire endowment,

then grant a part to a needy relative or a worthy cause or whatever he deems appropriate?

To summarize: in negotiations, it is for reputational reasons rational as a rule to reject unreasonably low offers. Mechanisms that have evolved to execute this rule include emotions such as self-respect and norms such as fairness. In the contrived and unnatural UG, these mechanisms sometimes make R act irrationally; as for P, he usually acts rationally, either for fear of rejection or because of the fairness mechanism. In spite of glaring differences between the UG and DG, the fairness mechanism may apply to the DG, which also is contrived and unnatural; and then P will act irrationally in the DG. Or it may not apply, and then P will act rationally. In short, the mechanisms work well in usual, commonly occurring situations, but poorly in contrived, unnatural situations.

**Altruism.** In many—perhaps most—human interactions, cooperation is a good idea. Generally, when people help each other, all concerned are better off. Such cooperation is act-rational when the sides enter into an enforceable agreement, like a contract. Or, it is act-rational in a repeated interaction, as when people repeatedly do business with each other[25,26]. In such cases it may take the overt form of altruism: I help you today, ostensibly without any quid pro quo, and you help me tomorrow, also ostensibly without any quid pro quo. Or, we cooperate every day, even though on each day each agent separately would be better off acting selfishly (as in the Prisoner's Dilemma). This is act-rational because people expect others to be generous[27], and if these expectations are not met, may well 'punish'[23] the transgressors. So in repeated interactions, cooperative behaviour is indeed act-rational.

But cooperation occurs also in one-time encounters, even when it is quite clear that the encounter is indeed one-time. What can account for this?

A possible answer is that acting altruistically (within limits)—i.e., truly without a quid pro quo—may be rule-rational. Rather than keeping accounts of who helped whom when, it may be simpler just to be generous, as a rule. Many human interactions are at least potentially repeated or long-term; in such cases, acting generously as a rule will work vis-à-vis others who also are generous as a rule and also vis-à-vis others who do keep accounts. It is not always act-rational, because in an interaction that is one-time for sure—such as tipping in a far-off restaurant that will not be visited again—one could do better by acting selfishly.

What we suggest here is that altruism is a mechanism for achieving cooperation (in the absence of an explicit enforceable agreement), in much the same sense that pride, feeling insulted, self-respect and revenge are mechanisms for getting reasonable offers in the UG, and fairness, decency and so on are mechanisms for making such offers. We intimated in that discussion that such traits evolved—genetically or memetically—because they usually, but not always, maximize fitness. Similarly here, altruism evolved because it promotes cooperation and so usually maximizes fitness.

As between genetic and memetic (i.e., cultural) evolution, the latter may seem more likely to account for altruism. But there is some evidence to suggest that altruism may be to some extent hardwired. In an experiment[28] using the DG, it was found that a common human polymorphism of the arginine vasopressin 1a (AVPR1a) receptor was associated with monetary allocations in the game. Dictators possessing long versions of the AVPR1a RS3 repeat allocated significantly more to the recipient than dictators possessing short versions. So, we have some preliminary evidence for a biological basis for this form of rule-rationality.

As above, one could simply stop there: take altruism as given—a legitimate source of utility—just as revenge, insult, etc., are sometimes taken as legitimate sources of utility and disutility. Indeed, the term 'other-regarding preferences' is sometimes used to 'explain' altruism. But as above, this is conceptually and methodologically unsatisfactory. Rather than taking altruistic preferences as given, one should account for them in terms of more fundamental human needs. What purpose—adaptive or otherwise—do they serve? What is their function?

That is the question addressed in the present treatment. And that question is particularly apt in view of initial evidence that altruism may be at least in part hardwired. And why would it be hardwired? The answer must be because it increases fitness—makes for rule-rational decisions.

We have here discussed only a few categories of systematically act-irrational behaviour, showing how it results from rational rules. Many others remain to be examined.

## How and why

Biologists distinguish[29] between two fundamental questions: how and why. 'How' refers to mechanism; 'why' refers to function. How do we see? With our eyes and their components—lens, retina, optic nerve, and so on—and the relevant parts of our brain. Why do we see? To find our way; recognize people, objects and places; apprehend dangers; grasp opportunities. Indeed, that is the distinction between how and why in general parlance. How do we go to London? By train. Why do we go to London? To meet friends.

Much of the literature on rationality is marred by failure to distinguish between these two fundamental questions. To 'explain' behaviour in the ultimatum and dictator games by other-regarding preferences or preferences for fairness is to answer the question of how, not why. Both questions are important, but they are different. Why does the responder in the UG turn down an offer of $20? Because as a rule, she does not want to get a reputation for accepting a relatively small share of the proceeds from joint ventures. How is this rule expressed? By genes, or memes, for fairness, self-respect, revenge, etc. Why does the dictator in the DG give away money for nothing? Because as a rule, cooperative behaviour is beneficial for all concerned. How is this rule expressed? At least to some extent, through genetics. Why do people eat? As a rule, to get energy and provide materials for the performance of bodily functions. How is this rule expressed? By the physiology of appetite, hunger, food enjoyment. When making quick decisions, why do people take cues from what they are told? Because as a rule, doing so leads to good decisions. How is this rule expressed? By heuristics such as relevance, anchoring, focusing. And so on.

Though BE complements ME, it is not the same. Economics is about incentives, and the incentives created by rule-rationality differ vastly from those created directly by act-rationality. A restaurant serving wholesome but tasteless food will quickly close its doors. International relations are governed largely by rule-rationality[30]. The enormous advertising industry is all about relevance, anchoring, focusing and even immediacy. And so on.

## Previous treatments of BE vis-à-vis ME

Though we know of no previous attempt to reconcile BE with ME, many of the above ideas do occur in the literature[31–45] in one form or another. But the basic idea of rule-rationality—that most irrational behaviour is, after all, accounted for by the rational paradigm—is not explicit in this literature, nor really implicit either. Moreover, the ideas have not heretofore been pulled together; the conclusion has not been drawn that BE and ME complement each other nicely. We here provide a unified conceptual framework: (i) that indeed people do not consciously maximize utility, but act by rules—heuristics or biases—that have evolved, biologically or culturally; (ii) that these rules prescribe behaviour that is generally act-rational and is act-irrational only in exceptional circumstances, not subject to evolutionary pressures; and (iii) that they have evolved for precisely that reason. It follows that economic behaviour is largely rational; so economic theory, based on rationality, is well founded.

This complements BE, which studies the substance of the rules by which people act.

Prominent in the literature is the concept of ecological rationality (ER): that behaviour arises from cultural and biological evolutionary processes[33,39] and so is optimal in the environment. Though related, ER differs from rule-rationality (RR). The central point of RR is the contrast with act-rationality; that of ER, the contrast with—and rejection of—constructivist rationality (CR), which is based on "deductive processes of human reason"[32].

Todd et al.[39] maintain that ER is actually superior to CR. For example, to catch fly balls, CR baseball outfielders need to "estimate the ball's trajectory and run [...] to the spot where the ball will hit the ground." Because of the many factors affecting the trajectory, this cannot be done with sufficient accuracy even consciously in hours, let alone unconsciously in seconds. Rather, outfielders use an ER heuristic that is vastly superior to the CR process: keep the angle between the ground and the line between your eye and the ball (the gaze angle) constant.

Furthermore, they say that as with the gaze heuristic, ER simply yields better results than CR, even ignoring costs of computing and information gathering. Indeed, Gigerenzer[45] goes so far as to say that the Linda respondents who say that 'feminist bank teller' is more likely than 'bank teller' are actually right. And while Smith does not say that ER always trumps CR, he does intimate[33] (footnote 8) that the exceptions may be ignored.

Contrariwise, RR behaviour is not always optimal. The exceptions are significant and systematic; they do not sit out there randomly in the tails of distributions. We know exactly when they occur: namely, when the scenario—the context of the behaviour, the environment—is unusual or contrived. The reason is that evolutionary selection does not apply in such cases. Like ER, RR rests squarely on evolution; but unlike ER, RR accounts for both rational and irrational behaviour. The Linda respondents, though rule-rational, are definitely not act-rational. In plain English, they're wrong: 'feminist bank teller' is less likely than 'bank teller'. It's the rule (the relevance maxim) that is rational; the act definitely is not.

## Conclusion

Mainstream economics studies how people should behave to further their interests; behavioural economics studies how they do behave. As a result of evolution and learning, 'should' and 'do' are effectively the same; they differ only in unusual or contrived scenarios, which have little or no economic impact. BE's heuristics and biases are in fact what makes ME work; ME is the 'why', BE the 'how'. The often-heard assertion that 'people do not behave as economists think' is simply incorrect. Far from contradicting it, BE actually underlies ME.

Received: 22 October 2018; Accepted: 24 April 2019;
Published online: 27 May 2019

## References

1. Simon, H.A. *Administrative Behavior* (Macmillan, 1947).
2. Newell, A., Shaw, C. & Simon, H.A. The process of creative thinking. in *Contemporary Approaches to Creative Thinking* (eds. Gruber, H. E., Terrell, G., & Wertheimer, M.) 63–119 (Atherton, 1962).
3. Hetzel, R. L. *Econ. Q.* **93**, 1–30 (2007).
4. Friedman, M. The methodology of positive economics. in *Essays in Positive Economics* 3–43 (U. Chicago Press, 1953).
5. Siegel, S. & Goldstein, D. A. *J. Exp. Psychol.* **57**, 37–42 (1959).
6. Guth, W., Schmittberger, R. & Schwarze, B. *J. Econ. Behav. Organ.* **3**, 367–388 (1982).
7. Hegel, G.W.F. *Phänomenologie des Geistes* [*Phenomenology of the Spirit*] (Joseph Anton Gebhardt, 1807).
8. Dawkins, R. *The Selfish Gene* (Oxford Univ. Press, 1976).
9. Tversky, A. & Kahneman, D. *Science* **185**, 1124–1131 (1974).
10. Kahneman, D. & Tversky, A. *Psychol. Rev.* **103**(582–591), 592–596 (1996).
11. Bentham, J. *An Introduction to the Principles of Morals and Legislation* (T. Payne, 1789).
12. Harsanyi, J. C. *Theory Decis.* **12**, 115–133 (1980).
13. Dostoyevsky, F. *Prestuplenie i nakazanie* [*Crime and Punishment*]. (The Russian Messenger, 1866).
14. Aumann, R. J. *Games Econ. Behav.* **21**, 2–14 (1997).
15. Tversky, A. & Kahneman, D. Judgments of and by representativeness. in *Judgment under Uncertainty: Heuristics and Biases* (eds. Kahneman, D., Slovic, P. & Tversky, A.) 84–100 (Cambridge U. Press, 1982).
16. Hertwig, R. & Gigerenzer, G. *J. Behav. Decis. Mak.* **12**, 275–305 (1999).
17. *The Real Mother Goose.* (Rand McNally, 1916).
18. Tversky, A. & Kahneman, D. *Psychol. Rev.* **90**, 293–315 (1983).
19. McClure, S. M., Laibson, D. I., Loewenstein, G. & Cohen, J. D. *Science* **306**, 503–507 (2004).
20. Tversky, A. & Kahneman, D. *J. Bus.* **59**, S251–S278 (1986).
21. Hertwig, R. & Ortmann, A. *Ethics Behav.* **18**, 59–92 (2008).
22. Kahneman, D., Knetsch, J. L. & Thaler, R. H. *J. Bus.* **59**, S285–S300 (1986).
23. Espin, A. M., Brañas-Garza, P., Herrman, B. & Gamella, J. F. *Proc. R. Soc. B Biol. Sci.* **279**, 4923–4928 (2012).
24. Brañas-Garza, P., Espin, A. M., Exadaktylos, F. & Herrman, B. *Sci. Rep.* **4**, 6025 (2014).
25. Aumann, R. J. Survey of repeated games. in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern* 11–42 (Wissenschaftsverlag, 1981).
26. Aumann, R. J. *Proc. Natl Acad. Sci. USA* **103**, 17075–17078 (2006).
27. Brañas-Garza, P., Rodriguez-Lara, I. & Sánchez, A. *Sci. Rep.* **7**, 42446 (2017).
28. Knafo, A. et al. *Genes Brain Behav.* **7**, 266–275 (2008).
29. Mayr, E. *The Growth of Biological Thought* (Belknap Press, 1982).
30. O'Neill, B. *Honor, Symbols, and War* (Univ. Michigan Press, 1999).
31. Stigler, G. S. & Becker, G. S. *Am. Econ. Rev.* **67**, 76–90 (1977).
32. Andreoni, J. & Miller, J. *Econometrica* **70**, 737–753 (2002).
33. Smith, V. L. *Am. Econ. Rev.* **93**, 465–508 (2003).
34. Tirole, J. *Am. Econ. Rev.* **99**, 265–294 (2009).
35. Rustagi, D., Engel, S. & Kosfeld, M. *Science* **330**, 961–965 (2010).
36. Fehr, E. & Leibbrandt, A. *J. Public Econ.* **95**, 1144–1155 (2011).
37. Delton, A. W., Krasnow, M. M., Cosmides, L. & Tooby, J. *Proc. Natl Acad. Sci. USA* **108**, 13335–13340 (2011).
38. Kahneman, D. *Thinking, Fast and Slow* (Farrar, Straus and Giroux, 2011).
39. Todd, P.M. & Gigerenzer, G. & ABC Research Group. *Ecological Rationality* (Oxford Univ. Press, 2012).
40. Fawcett, T. W. et al. *Trends Cogn. Sci.* **18**, 153–161 (2014).
41. Winter, E. *Feeling Smart* (Public Affairs, 2014).
42. Heller, Y. & Winter, E. *Int. Econ. Rev.* **57**, 997–1026 (2016).
43. Dye, R. A. *Int. Econ. Rev.* **26**, 233–250 (1985).
44. Chittka, L., Gumbert, A. & Kunze, J. *Behav. Ecol.* **8**, 239–249 (1997).
45. Gigerenzer, G. *Psychol. Rev.* **103**, 592–596 (1996).

## Acknowledgements

We thank A. Rai, a student at the Indian Institute of Technology, Kanpur, and a past intern at the Hebrew University's Federmann Center for the Study of Rationality, for helping us internalize many of the above references. Also, long-standing research support from the Israel Science Foundation is gratefully acknowledged. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Competing Interests

The author declares no competing interests.

## Additional information



**Correspondence** should be addressed to R.J.A.