ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

HEURISTICS AND PUBLIC POLICY: DECISION MAKING UNDER BOUNDED RATIONALITY

Sanjit Dhami
Ali al-Nowaihi
Cass R. Sunstein

Discussion Paper No. 963

06/2018

Harvard Law School
Cambridge, MA 02138

Heuristics and Public Policy: Decision Making Under Bounded Rationality

Sanjit Dhami* **Ali al-Nowaihi†** **Cass R. Sunstein‡**

18 June 2018

Abstract

How do human beings make decisions when, as the evidence indicates, the assumptions of the Bayesian rationality approach in economics do not hold? Do human beings optimize, or can they? Several decades of research have shown that people possess a toolkit of heuristics to make decisions under certainty, risk, subjective uncertainty, and true uncertainty (or Knightian uncertainty). We outline recent advances in knowledge about the use of heuristics and departures from Bayesian rationality, with particular emphasis on growing formalization of those departures, which add necessary precision. We also explore the relationship between bounded rationality and libertarian paternalism, or nudges, and show that some recent objections, founded on psychological work on the usefulness of certain heuristics, are based on serious misunderstandings.

Keywords: Heuristics; biases; Bayesian rationality; ecological rationality; true uncertainty; libertarian paternalism.

*Professor of Economics, University of Leicester, United Kingdom.
†Professor Emeritus of Economics, University of Leicester, United Kingdom.
‡Robert Walmsley University Professor, Harvard University, USA.

1. Introduction

How do human beings make decisions under certainty, risk, subjective uncertainty, and true uncertainty?[1] Neoclassical economics, the dominant paradigm in economics, does not always offer predictions for true uncertainty, but it does give a precise answer in the remaining cases. This is encapsulated in the *Bayesian rationality approach* (BRA), which lies at the heart of modern economics. In the most extreme version of the BRA, decision makers (firms, governments, and the person on the street) have complete, transitive, and continuous preferences; possess unlimited attention, computation power, and memory; are not influenced by frame-dependence of problems if the frames are informationally equivalent; make cold, calculated decisions in which emotions play no role; effortlessly follow all the laws of statistics and mathematics including all the latest research in these areas; engage in instantaneous mathematical optimization to static and dynamic problems; and update their prior beliefs using Bayes' law. Furthermore, they conform to the axioms of expected utility theory under risk; subjective expected utility under subjective uncertainty; and exponential discounting when making decisions over time.

Aiming to clarify debates about both rationality and public policy, we have three goals here. The first is to offer a disciplined, contemporary overview of departures from BRA in human behavior, with special emphasis on the role of heuristics. As we shall show, recent advances have allowed far more precision and formalization. The second is to demonstrate that although many advances have been made, and far more remains to be learned, the fundamental claims of the original work on heuristics, undertaken by Daniel Kahneman and Amos Tversky, remain largely intact. Many objections to that work are rooted in fundamental misunderstandings of its purposes (and also its central claims). The third is to demonstrate that for purposes of law and policy, libertarian paternalism, or nudging, does not depend on controversial psychological claims. A GPS device is helpful to human beings – no matter how we think about heuristics, and even if we agree that heuristics generally work well, in the sense they are helpful in the contexts in which most people use them.

Our starting point is that for many years the BRA was almost an article of faith in economics. It was the dominant approach taught in most economics departments, with relatively little discussion of the empirical validity of the assumptions that lie behind it or, indeed, any discussions of an alternative. Even now, some of the leading textbooks in microeconomics and game theory, which lay the foundation for the subject, continue to offer little or no empirical motivation for the theoretical models they use.

[1] We beileve that most readers will have heard of these terms. In any case, these are defined in Section 2. For the moment, we clarify that true uncertainty, a term used by Frank Knight, refers to a situation where one cannot define or even imagine the set of all possible states and their associated probabilities.

This state of affairs is captured in a colorful comment by Gintis (2009, p. xvi): "Economic theory has been particularly compromised by its neglect of the facts concerning human behavior... I happened to be reading a popular introductory graduate text on quantum mechanics, as well as a leading graduate text in microeconomics. The physics text began with the anomaly of blackbody radiation,...The text continued, page after page, with new anomalies...and new, partially successful models explaining the anomalies. In about 1925, this culminated with Heisenberg's wave mechanics and Schrödinger's equation, which fully unified the field. By contrast, the microeconomics text, despite its beauty, did not contain a single fact in the whole thousand-page volume. Rather the authors built economic theory in axiomatic fashion, making assumptions on the basis of their intuitive plausibility, their incorporation of the 'stylized facts' of everyday life, or their appeal to the principles of rational thought....We will see that empirical evidence challenges some of the core assumptions in classical game theory and neoclassical economics."

The behavioral economics revolution in economics has made significant progress in incorporating a more accurate understanding of human behavior (Kahneman and Tversky, 2000; Gintis, 2009; Thaler, 2015; Dhami, 2016). To an increasing degree, economics courses give attention to behavioral models and findings. It is important to see that much (not all) of behavioral economics still uses the optimization framework – but relaxes almost everything else. What if individuals do not optimize, or are simply not able to optimize? In this case, behavioral economics draws upon the influential work of Tversky and Kahneman (1971, 1974) to illustrate a range of simple rules of thumb (heuristics) that are fast and frugal (in terms of the time and information required) to solve economic problems (Kahneman et al., 1982; Kahneman, 2011; Dhami 2016, Part 7).

The work of Kahneman, Tversky, and other researchers (abbreviated by KT&O) on heuristics shows that the behavior of people is biased relative to the BRA; hence the name, *heuristics and biases*. That work has demonstrated that the BRA in economics is not tenable, not even in an 'as if' sense. Consistent with our first goal, we aim to show that the heuristics used in the KT&O program are not merely labels. Attempting to go beyond the first decades of work, we explain that most heuristics can be given precise mathematical definitions and are consistent with the bulk of the evidence (Section 3). Many of the objections to the KT&O program can be successfully addressed with suitable modifications and a fuller consideration of the empirical evidence (Section 5). One of the most common objections is that the biases in the KT&O program are washed away if one uses a frequency format rather than a probability format. We show that this distinction is not relevant to many of the leading heuristics. When it is relevant, a frequency format reduces the biases in some cases, notably the conjunction fallacy. However, in most cases, if no framing confounds are introduced, a majority of the subjects still exhibit the claimed biases (Section 5.3).

We consider a range of other objections to the KT&O program, as it is best understood, and find that the criticisms are overstated or simply do not stand up to a fuller scrutiny. This includes the nature of probability and subject errors (Section 5.1); the "we cannot be that dumb" critique (Section 5.2); lack of specification of empirical counterparts of the proposed heuristics (Section 5.4); the use of the recognition heuristic to explain an event (the gambler's fallacy) and its negation (the hot hands fallacy) (Section 5.5); the question of an appropriate statistical norm (Section 5.6). Finally, in Section 5.7, we briefly outline the System 1 and System 2 distinction proposed by Kahneman (2011) and suggest that it is useful in understanding the nature of heuristics in human life.

The work of Gerd Gigerenzer and others (abbreviated by G&O) on fast and frugal heuristics draws from and overlaps with the KT&O program, especially insofar as it shows that heuristics generally work well, as Kahneman and Tversky repeatedly emphasized. It claims to find its motivation in the original work of Herbert Simon (1955) that stressed the *procedural rationality* of solutions to problems (see Section 7). The stated domain of the G&O program is *large worlds* (or true uncertainty), although in some cases, notably the *priority heuristic*, it can also deal with *small worlds* (or risk and subjective uncertainty). We consider the foundational elements of the G&O program in Section 7.

G&O begin with the plausible idea, also found in KT&O, that people may have an adaptive toolbox of heuristics from which they draw heuristics depending on the context and frame of the problem (*ecological rationality*; see Section 3.2). In principle, this is eminently plausible, and consistent with the stated objective of procedural rationality. But theories cannot be judged on plausibility. We attempt to shed more light on the points of disagreement with the KT&O approach, which has come to be known as the *great rationality debate* in psychology (Stanovich and West, 2000; Dhami, 2016, Section 19.15).

The analysis of the problem of decision making under true uncertainty is arguably more challenging than under any of the other cases (certainty, risk, subjective uncertainty). We believe, based on the evidence, that the G&O program has not yet given us a persuasive account of decision making under true uncertainty. Economics has no optimization benchmark to offer in the case of true uncertainty, which makes it difficult to evaluate the performance of any candidate heuristic (Section 8.1); Gigerenzer (2008) is aware of this issue. The G&O program has compared the performance of its proposed heuristics against benchmarks claimed to be optimization benchmarks, typically logistic regression or weighted tallying. However, none of these benchmarks under true uncertainty is persuasive, and certainly not recommended by any optimization theory in economics.

We consider two further issues in Section 8. Is it a good option to train people in the use of statistics (Section 8.3)? We agree that in principle, training in statistics is an excellent idea. But which policy problems would it solve? Should a prime minister or a president -

focused, say, on clean air, highway safety, obesity, cigarette smoking, diabetes, savings, or poverty reduction – emphasize statistical training as a top priority? Behaviorally informed approaches have at least dented such problems, and many others (Sunstein, 2013; Halpern, 2015; Benartzi et al., 2017). Gigerenzer (2008) reports an improvement in the ability of doctors and children to use Bayes' Law when trained in a frequency format. That is good news. At the same time, these training problems involve simple conditional probability calculations, such as $P(A \mid B)$, but in actual practice doctors really need to compute complex conditional probabilities such as $P(A \mid B, C, D, E, \ldots)$, where $A, B, C, \ldots$ are events in some sample space. For instance, what is the probability that someone will die of a cardiac arrest, conditional on multiple contributing factors such as smoking, drinking alcohol, history of heart attacks in the family, cholesterol test scores, ethnic background, etc? These complex calculations are virtually impossible to do whether one has been taught in a frequency format or a probability format. It is probably a more productive use of resources to allow doctors to use software to compute these complex conditional probabilities (as in the National Health Service in the UK). Of far greater importance is to decide on the relevant thresholds of these conditional probabilities, over which pro-active medical treatment is needed.

Section 11 argues that the domains of choice in the KT&O and the G&O programs are often non-overlapping. KT&O largely considered situations of risk and uncertainty to which the BRA applies; their aim was to test the BRA. On the other hand, G&O are typically interested in the large worlds situation (true uncertainty), where one cannot even list or imagine the possible outcomes and/or objective/subjective probabilities. As such, a great deal of the debate that pits the two positions as adversarial is, in our view, unfortunate and misleading.[2] The main raison d'être of the G&O program is its attempt to answer the question of how people make decisions under true uncertainty; we do not believe that objective has been accomplished yet. We also raise other potential avenues of exploration for the quest to answer this fundamental question, such as social norms and mental models.

The G&O program often differentiates itself from the KT&O program on the following grounds in common and in published discourse (Gigerenzer, 2008, 2014; Gigerenzer et al., 1999) that we paraphrase as follows: (1) The KT&O program suggests that people are fallible, hardwired with defective mental software; and prone to errors; that heuristics are bad; and that the appropriate normative norm of human behavior is BRA (Gigerenzer, 1996; Gigerenzer and Todd, 1999; Gigerenzer, 2014). (2) In contrast, the G&O program is

[2] We fully realize that some of the G&O heuristics apply to the domain of risk and subjective uncertainty (e.g., the priority heuristic). We are also fully aware of the criticisms by the G&O program of the KT&O program about the frequency versus probability format, ecological rationality, and one event probabilities; we consider these in Section 5.

designed to show that heuristics are good, and do better than optimization methods once ecological rationality is taken into account.

This distinction is unhelpful and inaccurate. KT&O did not mean to argue that heuristics are good or bad. To the extent that they addressed that topic, it was to say that the heuristics they identified generally worked well, but also led to severe and systematic errors (which is demonstrably true). Their main goal was to test if human behavior was consistent with the BRA; and they found it was not. But they went further by identifying various classes of heuristics that explain human behavior in different contexts and frames. The following two passages from Kahneman (2000, p. 682) show just how close the two sides in the debate really are on core issues: (1) "Contrary to a common perception, researchers working in the heuristics and biases (HB) mode are less interested in demonstrating human irrationality than in understanding the psychology of intuitive judgment and choice." (2) "All heuristics make us smart, more often than not..."

We also offer some claims about law and public policy. Welfare economics often assumes that people follow the BRA and make informed choices that are in their best interests; such preferences are termed as *normative preferences*. However, the evidence from behavioral economics suggests that individuals sometimes do not make choices that are in their best interests (Thaler and Sunstein, 2009; Dhami, 2016; Dhami and al-Nowaihi, 2018). This may be the case when individuals lack information, have limited attention, misperceive risks, or face self-control problems arising from various forms of present-biased preferences (and have imperfect awareness of such problems). Thus, individuals might undersave for retirement, not enroll in pension plans, consume various goods (such as cigarettes) that harm the quality of life, make decisions in an emotional hot state that they regret later, and procrastinate in making choices.

Libertarian paternalism (LP), or more simply "nudging," has been an effort to help human beings to avoid errors while also preserving freedom of choice. A GPS device is an example. It allows people to go their own way, but helps them to arrive at their preferred destination. Most broadly, LP is an effort to increase *navigability*. It can be useful when people use heuristics that produce errors. Policies that are consistent with LP do not distort the choices of those who follow the BRA (or do so minimally), but significantly improve the welfare of those who do not. Consider warnings and reminders, which may overcome the problem of limited attention. Use of default options, an example of an LP policy, has been extremely effective in a large number of domains, and can be more cost-effective than traditional tax/subsidy and direct regulation methods advocated in classical welfare economics (Thaler and Sunstein, 2009; Thaler, 2015; Dhami, 2016, Part 8; Sunstein, 2016, 2017).

As we shall show, endorsement of LP does not depend on a commitment to psychological claims that may be controversial. Most puzzlingly, the perceived adversarial position between the KT&O and G&O approaches, described above, has nonetheless given rise to a critique of the LP approach (Gigerenzer, 2015). Section 12 considers three main elements of this critique. (1) Choice architects (those who enact the LP policy) may not be benevolent; rather, they may be self-interested or malicious. (2) The rationale for nudges lies in people's alleged irrationality. (3) Nudges ignore other policy interventions that might help. None of these criticisms has merit.

The 'lack of a benevolent policymaker' criticism applies to any economic policy, not just nudges; this is analyzed in political economy, a well established field in economics. If policymakers are not benevolent, the strongest objections should be to mandates and bans, not against nudges, which maintain freedom of choice (in part because of an insistence that policymakers may err). A primary reason for nudging, as opposed to mandates and bans, is precisely the possibility that policymakers are not benevolent (or adequately informed).

Those who embrace nudges do not speak of irrationality. The rationale for many successful nudges lies in a lack of information, in limited attention, and in self-control problems (and imperfect awareness of such problems). None of these should be controversial, certainly not in the abstract. Objections to LP, based on psychological claims, do not grapple with what LP means in practice (Sunstein, 2013; Halpern, 2015)– that is, with the particular policies that advocates or practitioners of LP have embraced in actual policymaking roles.

Section 13 concludes.

2. A simple taxonomy of situations

This section introduces a basic taxonomy of situations that are of potential interest to the study of heuristics. Its purpose is to ensure a common understanding of some of the terms that we shall use.

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a fixed, finite, set of real numbers such that $x_1 < x_2 < \ldots < x_n$, possibly outcomes or wealth levels arising from some choices exercised by a decision maker. A *lottery*, or *gamble*, is written as

$$L = (x_1, p_1; x_2, p_2; \ldots; x_n, p_n), \quad (2.1)$$

where $p_1, p_2, \ldots, p_n$, are the respective probabilities corresponding to the outcomes $x_1, x_2, \ldots, x_n$, such that $p_i \in [0, 1]$ and $\sum_{i=1}^{n} p_i = 1$.

When examining individual choice, economists are typically interested in four kinds of situations.

1. *Certainty*: These are choices among outcomes received with certainty and take the form: Choose $A = (x_i, 1)$ or $B = (x_j, 1)$.

2. *Risk*: These are choices among lotteries of the form given in (2.1) when the outcomes and probabilities are *objectively known*.[3] Consider for example, whether you would rather have a fixed sum of money \$10 or be willing to play the lottery $L_1 = (4, 0.4; 16, 0.6)$ in an experiment where the experimenter implements the lottery L_1 as follows. He puts 4 green balls and 6 red balls, that are otherwise identical in all respects, in an opaque urn and shakes the urn to thoroughly mix the balls. A ball is then randomly drawn. If a green ball comes up, you get \$4 and if a red ball comes up, you get \$16.

3. *Subjective uncertainty*: These are choices among lotteries of the form given in (2.1) when the outcomes are either *objectively* or *subjectively* known. However, the probabilities in (2.1) are *subjective probabilities* that are non-negative and add up to one. In this case, we have a situation of subjective uncertainty, sometimes just referred to as uncertainty.[4] As an example, suppose that we alter the experiment under risk in 1 above as follows. The lottery L_1 takes the form of a Savage act, $f = (4, \text{green ball}; 6, \text{red ball})$ so that if a green ball comes up on a random draw of the ball from the urn, you get \$4 and if a red ball comes up, you get \$16. However, the catch is that the green balls are larger than the red balls. In this case, it is no longer the case that the individual has precise knowledge of objective probabilities of the events, green ball and red ball. However, when the individual ponders over the probabilities, he is confident that a green ball and a red ball are equiprobable. So he believes that he is facing a lottery of the form $L_2 = (4, 0.5; 6, 0.5)$. This is a situation of subjective uncertainty.

4. *Ambiguity*: These are choices among lotteries when neither objective nor subjective probabilities can explain choices. The leading example is the *Ellsberg paradox* that can be explained as follows.[5] Suppose you have two urns, a known urn U and an unknown urn K. In Urn K are 50 Red and 50 Green balls. In Urn U are 100 balls that are either red or green but the proportions are unknown. You are promised a prize of \$10 if a red ball comes up. You assign an objective probability of 0.5 of a red ball being drawn from Urn K. Also, using the principle of insufficient reason, you may assign a probability 0.5 of a red ball being drawn from Urn U. Yet, when you

[3] We use the term 'objectively known' in the sense that it is used in standard microeconomics books such as Mas-Colell et al., (2015, p.168). This implies that there is common agreement among people on the magnitude of these probabilities. For instance, a probability of 0.5 of a heads in an unbiased coin toss, or a probability of 1/6 of the number 3 coming up in a throw of a fair, six-sided, dice.

[4] A more formal and satisfactory definition in terms of Savage acts and probabilistic sophistication (Dhami, 2016; Section 1.3) can be given.

[5] The Ellsberg paradox can be explained in many ways (Dhami, 2016, Ch. 4), but here we choose the simplest example, the so called two-colors example.

are asked which urn you will bet on, you prefer Urn K to Urn U, despite assigning equal probabilities of drawing a red ball from each.

In this case one may form *source-dependent probabilities* that depend on the underlying source of uncertainty. Source dependent probabilities is an idea that goes back to Amos Tversky and recent empirical research suggests that they can explain the Ellsberg paradox (Dhami, 2016, Chapter 4). However, this situation has not played a direct role in the literature on heuristics.

There is fifth class of extremely important situations on which economic theory is almost completely silent and has no predictions to offer.

5. *True uncertainty*: A situation of true uncertainty (or Knightian uncertainty) arises when the outcomes and probabilities are unknown or unimaginable and objective/subjective estimates of these outcomes and probabilities are not available. In terms of the Ellsberg paradox explained above, true uncertainty would arise, if you had no information about how many colors are there in Urn U, and in addition, you might not even know how many balls are there in Urn U.[6] There might be no obvious way of inferring subjective probabilities of these events either, so that we may not be able to turn this into a situation of subjective uncertainty.

   There is some overlap between ambiguity and true uncertainty, but the latter is a much broader class. Ambiguity and true uncertainty coincide when probabilities are unknown but all outcomes (and the number of outcomes) are known. In the case of ambiguity, economic theory does have emerging predictions to offer. These predictions are based on correlating measures of ambiguity aversion arising from Ellsberg sort of experiments with a range of human choices that include the following: stock market participation; fraction of financial assets held in stocks; foreign stock ownership; ownership of stocks in own-company; selling behavior of stock in the financial crises; see, for instance, Dimmock, Kouwenberg, et al. (2016) for a study on US households and Dimmock, Kouwenberg, and Wakker (2016) for a study on Dutch households. However, since true uncertainty also involves unknown outcomes (as shown in Example 1 below), we prefer to keep the terms ambiguity and true uncertainty separate.

[6] The human eye can distinguish between 10 million colors and the number of balls can be very large. So, if you are a classical purist who believes that this situation can be accommodated within classical subjective uncertainty, then try to imagine the possible colors and the Cartesian product of colors and the number of balls over which you must form a subjective distribution. It is far fetched to imagine that humans could do this sort of thing, even if in principle we could allow for this possibility and the use of subjective uncertainty. As a practical matter, it is better to classify such a situation as one of true uncertainty.

In Example 1 below, we present a series of examples that are potentially consistent with situations of true uncertainty. These examples may also be formally consistent with subjective uncertainty if people can foresee the entire path of all possible events in the future and can assign subjective probabilities to them; which is unlikely.

Example 1 *: True uncertainty seems to be reflected in the judgements of even experienced market participants. Thomas Watson, President of IBM is said to have made the following statement in 1943, "I think there is a world market for maybe five computers." Indeed it would have been impossible to predict all possible eventualities (outcomes, technological breakthroughs, tastes) relevant to this question in 1943. Consider some other examples that can plausibly be seen to reflect true uncertainty.*[7]
"Fooling around with alternating current (AC) is just a waste of time. Nobody will use it, ever." (Thomas Edison, 1889)
"Television won't be able to hold on to any market it captures after the first six months. People will soon get tired of staring at a plywood box every night." (Darryl Zanuck, 1946, 20th Century Fox).
"Nuclear powered vacuum cleaners will probably be a reality within 10 years." (Alex Lewyt, 1955, President of the Lewyt Vacuum Cleaner Company).
"There is practically no chance communications space satellites will be used to provide better telephone, telegraph, television or radio service inside the United States." T.A.M. Craven, 1961, Federal Communications Commission (FCC) commissioner.
"Remote shopping, while entirely feasible, will flop." (Time Magazine, 1966).
"There's just not that many videos I want to watch." (Steve Chen, 2005 CTO and co-founder of YouTube expressing concerns about his company's long term viability).

In Example 1, Thomas Watson might have, in 1943, foreseen everything possible related to the computer industry in the future (e.g., all possible innovations, supplies, demands, tastes, and prices) for all future time periods and assigned subjective probabilities to all these events. He might have then have used expected utility to make a considered prediction. This would be consistent with BRA but we are very doubtful of this possibility. It is more likely that he faced true uncertainty and chose his best guess.

Arguably, some of the most important problems that decision makers face belong to the domain of true uncertainty. As possible candidates, consider investing in the stock market for most, except perhaps the most professional investors[8]; choosing a partner to

[7] We have drawn all these quotations from an article by Robert J. Szczerba in Forbes, Jan 5, 2015, titled "15 Worst Tech Predictions Of All Time". See also a nice list of such situations in Gigerenzer (2014, p. 41-42).

[8] Investment firms use complicated algorithms, but the assumptions behind them often also take the form of heuristics, educated guesses, homegrown thresholds, and margins of errors.

marry; choosing a University course at age 19; investing in a pension fund at age 25 that matures when one is 60. Economics does not have clear predictions to offer in these cases. Yet, people often do make decisions in such cases. People do invest in the stock market; people do choose marriage partners and divorce; students do choose universities to study; and young people do invest in pension funds (albeit sometimes with the aid of appropriate nudges). How do people make these decisions? It is even more surprising that this limitation of the scope of economic theory is almost never mentioned in economics courses.

3. The KT&O approach

When economists deal with situations 1-4 above, they typically use the *Bayesian rationality approach* (BRA); see our opening remarks. It typically delivers sharp, testable, predictions, often in conjunction with other auxiliary assumptions.

To see what the BRA entails, we first describe an illustrative but fairly 'general problem' (GP) in economics.

Problem GP: Consider the problem of maximizing a strictly quasi-concave function $f(\mathbf{x}, \boldsymbol{\theta})$, $\mathbf{x} \in \mathbf{C} \subset \mathbf{R}^n$, where $\mathbf{C}$ is a compact set, subject to a convex constraint $g(\mathbf{x}; \boldsymbol{\theta}) \leq 0$. $\boldsymbol{\theta} \in \mathbf{R}^p$ is an exogenous vector of variables, parameters, and updated beliefs. Beliefs are always updated by *using Bayes' Law.* This standard problem in economics is easily extended to several constraints and can deal with problems of risk, uncertainty, and ambiguity, as well as decisions over time. Under the assumptions of problem GP, an optimal solution, $\mathbf{x}^*$, exists, and it is unique. Let $f(\mathbf{x}^*; \boldsymbol{\theta})$ be the maximum value of the objective function.

Let $\mathbf{x}^H \in \mathbf{R}^n, \mathbf{x}^H \neq \mathbf{x}^*$ be any other potential candidate solution, perhaps based on another objective function, e.g., some lexicographic heuristic rule. Then, necessarily,

$$f(\mathbf{x}^*; \boldsymbol{\theta}) > f(\mathbf{x}^H; \boldsymbol{\theta}) \text{ (} \because \text{ the optimum is unique)} \tag{3.1}$$

Thus, no other solution, under the conditions of Problem GP, can do better than the optimization benchmark.

Remark 1 *: Problem GP is general enough to deal with situations of certainty, risk, and subjective uncertainty; the vector $\boldsymbol{\theta}$ can include, respectively, objective and subjective probabilities of outcomes. Problem GP assumes that individuals use Bayes' Law to update information; have full information of all the relevant data to solve the problem; exhibit full attention to all relevant aspects of the problem; engage in static or dynamic mathematical optimization effortlessly; and engage in emotionless deliberation.*

Is the behavior of people consistent with the BRA? Even if people do not literally follow the BRA, do they behave 'as if' their behavior is consistent with the BRA? The second of these two questions is motivated by Milton Friedman's evaluation of economic theories by the accuracy of their predictions, rather than the realism of their assumptions. These questions are often unaddressed in economics courses (unless it is a course in behavioral economics or the methodology of economics). Consistency of human behavior with the BRA and the appropriateness of mathematical optimization are taught and accepted as articles of faith in many leading economics programs in the world. Political science, management schools, and law programs have also tried to assimilate the BRA approach.

Tversky and Kahneman (1971, 1974) prominently tried to answer the questions posed above. Their program, the *heuristics and biases* research program, is one of the most significant achievements in all of social science. Other researchers have taken up the baton and have also made important contributions to (and often modifications of) this program; for this reason, we use the shorthand, KT&O program. Kahneman et al. (1982) is a classic introduction to the subject. There are several comprehensive treatments of the KT&O approach available that incorporate the more recent literature and significant modifications and advances, both empirical and theoretical (Kahneman, 2011; Dhami, 2016, Part 7).

The KT&O program has shown that people do not act 'as if' they use the BRA. On mathematical optimization in the BRA, in particular, the Nobel laureate Herbert Simon (1978) notes: "But there are no direct observations that individuals or firms do actually equate marginal costs and revenues." Another Nobel laureate, Reinhard Selten, has argued on many occasions that economic problems are *NP-Complete,* a term that is borrowed from the computer science literature.[9] The solution to such problems is hard/impossible to obtain; while no analytical solution may obtain, numerical solutions may be possible with high powered computers and sophisticated algorithms. Yet, in the BRA, the person on the street is assumed to solve such problems in his head, in an instant.

To say that this work poses problems for the BRA, and by implication for neoclassical economics, is an understatement. Yet many economics courses say little about those problems.[10]

In recent decades, the KT&O approach has been massively influential. But it has come under intense criticism from another school within bounded rationality. The G&O approach identifies Herbert Simon's bounded rationality approach as its intellectual fountainhead and there is a large number of contributors to it.[11] G&O have focussed on

[9] See, for instance, the entertaining panel discussion on behavioral economics at the 2011 Lindau Nobel Laureate Meeting in Economic Sciences. The participants were George A. Akerlof, Robert J. Aumann, Eric S. Maskin, Daniel L. McFadden, Edmund S. Phelps, and Reinhard Selten.

[10] Reinhard Selten's work on subgame perfection plays a central role in courses in game theory in economics. However, Selten was aware that its emprical performance was poor and for that reason, he was skeptical of it. Yet, the area that he gravitated towards, bounded rationality, is hardly taught in courses on economic theory.

[11] We avoid the term SG&O that prefixes Herbert Simon's influence on this approach because we believe that the KT&O program has also been influenced by it. Indeed, Herbert Simon was appreciative of the

procedural rationality and highlighted the ecological rationality of heuristics.

The debate between KT&O and G&O has come to be known as the *great rationality debate in psychology* (Stanovich and West, 2000; Dhami, 2016, Section 19.5). We believe that several aspects of this debate appear to rely on confusions and misunderstanding of the relevant subject matter. We try to clarify several aspects of this debate below.[12]

3.1. Some preliminaries

Gigerenzer and Brighton (2009), reprinted as Gigerenzer and Brighton (2011) by way of introduction to the edited volume by Gigerenzer et al. (2011), describe the KT&O program as follows. "By the end of the 20th century, the use of heuristics became associated with shoddy mental software, generating three widespread misconceptions.

S1 Heuristics are always second-best.

S2 We use heuristics only because of our cognitive limitations.

S3 More information, more computation, and more time would always be better."

Consider Problem GP that is rich enough to incorporate certainty, risk, and subjective uncertainty. In the world of Problem GP, S1, S2, S3 are not misconceptions, as shown next.

1. From (3.1), heuristics are always second best.
2. We know that $f(\mathbf{x}^*;\boldsymbol{\theta}) > f(\mathbf{x}^H;\boldsymbol{\theta})$, for any $\mathbf{x}^H \neq \mathbf{x}^*$. Hence, starting from $\mathbf{x}^H \neq \mathbf{x}^*$ one can always improve upon the value of objective function. If an individual chooses $\mathbf{x}^H \neq \mathbf{x}^*$, then cognitive limitations are a possible contributing factor.
3. Starting from $\mathbf{x}^H$ and the associated information on which it was based, more information might produce an even better heuristic $\overline{\mathbf{x}}^H$ such that $f(\mathbf{x}^H;\boldsymbol{\theta}) < f(\overline{\mathbf{x}}^H;\boldsymbol{\theta}) < f(\mathbf{x}^*;\boldsymbol{\theta})$.

Indeed, we appear here to be in the world of an *accuracy-effort trade-off* (Gigerenzer and Brighton, 2011). In this canonical case studied in economics, it is never the case that 'less (information, computation time, or attention), is more'. Gigerenzer and Brighton (2011, p.5) write: "Even when information and computation are entirely free, there is typically a point where less is more." In the task described above, this is never the case.

work of KT&O (Gigernzer, 2008, p. 86).

[12] We offer limited comments on the relevant mental or neural processes that are to be implicated in the use of heuristics, or the appropriate normative benchmark models to be employed.The interested reader can pick up these aspects of the debate in Stanovich and West (2000) and Stanovich (2012).

Under what conditions might S1, S2, and S3 fail? This might be true, for example, if the model GP is mis-specified, e.g., incorrect objective function, f, or incorrect constraint, g. The BRA approach assumes that, in the world of problem GP, decision makers know the correct functions f, g; indeed, this is important in models of rational expectations that underpin all of modern macroeconomics. In richer situations than those depicted in problem GP, decision makers may have several candidate functions in mind, but they typically converge towards the correct one through some underlying process of learning. However, empirical evidence and some of the relevant theory shows that there is no guarantee that such a convergence will take place; see Dhami (2016, Part 5) and for the impossibility of learning on theoretical grounds, see Section 15.9 in Dhami (2016).

S1, S2, S3 might also fail when (1) the primary task is not optimization but prediction, or (2) we have true uncertainty, which falls outside the remit of Problem GP. We consider some of these cases later in the paper.

3.2. Ecological rationality

Why might some heuristics that use less information do better than more complex strategies? The G&O view is that the heuristics are better adapted to the environment as captured in the view: "The rationality of heuristics is ecological, not rational." Herbert Simon famously characterized bounded rationality as the two blades of a scissor. One of the blades is the mind, and the second is the environment. This implies that cognitive strategies cannot be looked at independently of their environment.

How we do operationalize ecological rationality? After all, no precise definition is given. In behavioral economics this is typically addressed by stressing that human decisions (including heuristics) are context, culture, history, time and frame dependent (Dhami, 2016; Gintis, 2017). By way of analogy, social norms are often adapted to the social context in which they are situated. Insofar as they use limited information, they are similar to heuristics. Consider the following heuristic: *If you do not know what stand to take on a complicated issue of policy, follow the views of an official whom you trust, and with whom you largely agree.* On plausible assumptions, that heuristic has ecological rationality.

An alternative formalization of this view, *the less is more effect*, is considered in Section 9. Gigerenzer and Brighton (2009) also require heuristics to take account of the *less is more effect* or the *bias-variance dilemma* (see Section 9, below) in satisfying the requirement of ecological rationality. We do not believe this to be an essential requirement for ecological rationality for the reasons that we specify in Section 9 below. By ecological rationality, we mean to refer to the context and frame dependence of preferences, decisions, and beliefs. In this sense, we believe that the KT&O program satisfies ecological rationality; this is also exemplified in the examples that we use to illustrate the heuristics from the KT&O

program below.

Another sense in which issues of ecological rationality arise is that lab experiments might have low external validity (Levitt and List, 2007). This issue has already been thoroughly addressed in several recent publications, which suggest that, to the contrary, a great deal of lab evidence has a high degree of external validity (Camerer, 2015; and Section 3 of the introductory chapter in Dhami, 2016). Further support for the ecological rationality of the KT&O heuristics comes from the behavior of experts (see Section 6 below).

4. The KT&O program with definitions and labels

A major criticism of KT&O by G&O is that the relevant heuristics are not stated formally, so it is not clear what they mean, and anything goes (Gigerenzer, 1991, 1996; Gigerenzer and Gaissmaier, 2011). For instance, talking of representativeness and availability, Gigerenzer (1991, p. 102) writes that these are "largely undefined concepts and can be post hoc used to explain almost everything." Kahneman and Tversky's (1996, p. 585) respond that since representativeness can be elicited experimentally, there is no need to define it a-priori.

G&O have also criticized KT&O for not providing precise *labels* for the heuristics that they use (Gigerenzer and Brighton, 2009). The term labels appears to us to mean two things: (1) precise definitions of the heuristics and (2) specification of exact models in which these heuristics are situated. In this section, we consider the first of these meanings, while Section 5.6 deals with the second.

It is reasonable, of course, to ask for precise definitions of heuristics. One of the hallmarks of the use of heuristics within behavioral economics has been an increasingly formal approach that relies on clear definitions of the underlying phenomena. However, this appears not to be widely recognized outside behavioral economics. For this reason, in this section, we give formal definitions of the main heuristics used in KT&O in one place; some of these definitions have been formalized here for the first time, as far as we are aware. This section also has a secondary objective of providing a concise introduction to the KT&O program, which is needed to understand the subsequent criticisms.

4.1. The representative heuristic

The representativeness heuristic is one of the most versatile and useful heuristics (Kahneman and Tversky, 1984). Gigerenzer and Brighton (2011, p.18) write: "On the other hand, there is the label "representativeness," which was proposed in the early 1970s and means that judgements are made by similarity– but how similarity was defined is left open. A label can be a starting point, but four decades and many experiments later, representativeness has still not been instantiated as a model."

The representativeness heuristic has been increasingly defined and formalized (Dhami, 2016; Section 19.2; Rabin, 2002). Consider the weak law of large numbers:

Theorem 1 *(Weak Law of Large numbers; Hogg et al., 2005) Let* $\{X_n\}$ *be a sequence of independently and identically distributed random variables, with a common mean* μ *and finite variance* σ^2. *Denote the sample mean based on* n *observations by* $\bar{x}(n) = \sum_{i=1}^{n} x_i/n$, *where* x_i *is the sample realization of the random variable* X_i. *Then* $\bar{x}(n)$ *'converges in probability' to the population mean,* μ, *i.e.,* $\lim_{n\to\infty} P\left[|\bar{x}(n) - \mu| \geq \varepsilon\right] = 0 \ \forall \varepsilon > 0$.[13]

Definition 1 *(Representativeness): An individual uses the representativeness heuristic, or subscribes to the law of small numbers, if the individual holds the belief that for a finite sample size, the sample mean is identical to the population mean.*

The main import of the representativeness heuristic is that people who use it believe that the sample proportions mimic the population proportions for a finite sample size. This definition suffices in most cases. In specific examples, one can give definitions of representativeness that are consistent with Definition 1.

Example 2 *: Consider the following experimental instructions from Tversky and Kahneman (1974). A town is served by two hospitals. In the larger hospital, about 45 babies are born each day, and in the smaller hospital, about 15 babies are born each day. As you know, about 50 percent of all babies are boys. However, the exact percentage varies from day to day. For a period of 1 year, each hospital recorded the days on which more than 60 percent of the babies born were boys. Which hospital do you think recorded more such days? In the sample, 53 students said that both hospitals are equally likely to have recorded such days and an equal split of 21 students chose the larger and the smaller hospital, respectively. The correct answer is the smaller hospital because as the sample size grows the sample proportions become closer to the population proportions. Thus, the choices of 74% of the respondents violates the law of large numbers.*

Example 3 *: Consider experiments reported in Camerer (1987, 1990, 1995). Suppose that there are two urns,* $i = A, B$, *and one is chosen randomly by nature. Urn A has 1 red and 2 black balls. Urn B has 2 red and 1 black ball. It is common knowledge that nature chooses urn A with probability 0.6, so* $P(A) = 0.6$ *and* $P(B) = 0.4$. *A sequence of three balls is drawn with replacement from one of the urns. Suppose that the sample is 1*

[13] The weak law of large numbers can be strengthened to the *strong law of large numbers,* which guarantees almost sure convergence of the sample mean to the population mean. However, its scope of application is less than that of the weak law.

red and 2 black balls; denote this sample by RBB (the exact order of the coloured balls is not important). Experimental subjects do not know which urn the balls are drawn from. What is the posterior probability that the sample came from urn A? Using Bayes' law, this can be computed to be $P(A \mid RBB) = 0.75$.[14] *In contrast to the Bayesian estimate, an individual who uses the representativeness heuristic assigns $0.75 < P(A \mid RBB) \leq 1$. Clearly, various 'degrees of representativeness' that depend on individual-specific traits are possible. This is consistent with Definition 1. If we were to hypothetically draw infinite number of samples of 3 balls with replacement from each urn, then we can be almost sure that a sample of one red and two black balls is likely to have come from Urn A. However, individuals who use the representativeness heuristic ascribe such similarity to even small samples.*

The formal model of small numbers due to Rabin (2002) also uses Definition 1 to study representativeness in a particular context; see Dhami (2016, Section 19.2.3).

4.1.1. The gambler's fallacy

One implication of the law of small numbers is that many people are unable to generate a truly random sequence of events. For instance, when asked to write down a random sequence of coin tosses, subjects alternate too much between heads (H) and tails (T). So, randomly generated sequences by humans have too much autocorrelation; the sequence H, H, H, H is very likely to be followed by T, although the a-priori chance of a H or a T is identical (Bar-Hillel and Wagenaar, 1991; Rapoport and Budescu, 1992, 1997).

Definition 2 : *If subjects produce negative autocorrelation when asked to produce a hypothetical random process, they are said to commit the gambler's fallacy.*

When subjects were shown a subset of 3 coin tosses from an underlying set of 150 coin tosses, they assign the conditional probability P that the fourth toss will be H as follows. (1) $P(H \mid HHH) = 0.30$. (2) $P(H \mid HTH) = 0.412\%$. (3) $P(H \mid THT) = 0.588$. (4) $P(H \mid TTT) = 0.70$ (Rapoport and Budescu, 1997; Rabin and Vayanos, 2010). In each case, the ex-ante probability of H in the fourth toss is statistically 0.50, but the behavior of subjects is consistent with the gambler's fallacy in Definition 2. Rabin and Vayanos (2010) propose a formal and more general model that can account for the gambler's fallacy when the outcomes are neither binary nor independently and identically distributed.

There is reduced betting on a winning number in subsequent lottery draws, which is consistent with the gambler's fallacy (Clotfelter and Cook, 1993). Much evidence for the gambler's fallacy has come from the study of betting behavior, e.g., horse races (Metzger,

[14] The necessary calculations are given in Section 19.2.1 in Dhami (2016).

1984), dog races (Terrell and Farmer, 1996; Terrell, 1998), and gambling in casinos (Croson and Sundali, 2005). There is now a large literature on applications of the gambler's fallacy to finance. For instance, there is a tendency to sell winning stocks and hold on to loss making stocks for too long (*disposition effect*). The idea is that people think that a stock that has risen in the past is now due for a fall and vice-versa, which is consistent with the gambler's fallacy (Dhami, 2016, p. 1354).

4.1.2. The hot hands fallacy

The hot hands fallacy is the statistical opposite of the gambler's fallacy.

Definition 3 *: If subjects produce positive autocorrelation when asked to produce a hypothetical random process, they are said to engage in the hot hands fallacy.*

The *hot hands fallacy* has been documented in many contexts. In a basketball game, a player might be particularly successful in a sequence of shots; such success might however arise from pure luck. Data from betting behavior indicates that, in this case, observers assign a high probability that such a player will be successful in making the next shot too (Camerer, 1989). In contrast, the underlying data has been found not to reveal a hot hands effect. Several studies showed that the observed sequence of successful basketball shots was statistically a random sequence (Gilovich et al., 1985; Tversky and Gilovich, 1989a,b).

Recent research has raised the possibility of actual hot streaks in basketball and baseball (Green and Zwiebel, 2013; Bocskocsky et al., 2014; Miller and Sanjurjo 2015). However, a statistical demonstration of the hot hands phenomena, a non-trivial statistical task, does not imply that individuals are able to also see through and discover the hot hands effect.[15]

Some of the clearest and most persuasive evidence for hot hands comes from a novel field experiment by Guryan and Kearney (2008). They find that sales at lotto stores that have sold a winning ticket soar in the immediate weeks following the lotto win. The effect lasts for an impressive 40 weeks following the win, even after controlling for the greater salience of buying lottery tickets in the surrounding areas, following the win.

4.2. Anchoring

Anchoring can be formally explained in terms of Problem GP that we outlined above. Here, our use of Problem GP as a strictly optimizing problem is purely for illustrative purposes. In several anchoring experiments (see below), Problem GP is replaced by the

[15] When testing for a mixed strategy Nash equilibrium, it has been discovered that there is serial correlation in tennis serves and in penalty kicks in football. Yet, opponents do not discover such serial correlation in actions (Dhami, 2016, Section 12.3.4).

simpler problem of forming beliefs about some event (e.g., how many African countries are there in the United Nations?). This does not change our basic argument.

Recall that an optimal solution, $\mathbf{x}^*$, to Problem GP exists and is unique. Suppose that we have instead the following anchoring problem, Problem A.

Problem A: This problem has the following two stages.

Stage 1: The decision maker is 'informed' that the solution to Problem GP is $\widehat{\mathbf{x}}$ (call this an 'anchor'), where, say $\widehat{\mathbf{x}} \leq \mathbf{a}$ and $\mathbf{a} \in \mathbf{R}^n$.[16]

Stage 2: The decision maker is asked to solve Problem GP.

Definition 4 : *Anchoring is said to exist if in Problem A, (1) subjects give the solution to Problem GP as some $\mathbf{x}^A \neq \mathbf{x}^*$, and (2) when two different anchors $\widehat{\mathbf{x}}_1$ and $\widehat{\mathbf{x}}_2$ are given in Stage 1 to two different individuals/groups, then they come up with two different solutions: $\mathbf{x}_1^A \neq \mathbf{x}^*$ and $\mathbf{x}_2^A \neq \mathbf{x}^*$.*

Note that nothing in Problem A requires the anchor $\widehat{\mathbf{x}}$ to have any relevance for the optimal solution, $\mathbf{x}^*$. The whole idea of anchoring is that completely irrelevant and uninformative anchors may influence the choices made by people.

Consider the following examples of anchoring.

Example 4 : *Tversky and Kahneman (1974) rigged a wheel of fortune with the numbers 1-100 to come to a stop at either of the two numbers: 10 or 65. Subjects were asked to write down the number that the wheel stopped at (this corresponds to $\widehat{\mathbf{x}}$ in Problem A). They were then asked the following two questions. (1) Comparative judgment question: Is the percentage of African nations among UN members larger or smaller than the number you just wrote? (2) Absolute judgment question: What is your best guess of the percentage of African nations among UN members? The answers to the second question were found to be anchored too closely on the irrelevant number that came up on the wheel of fortune. Those who observed the number 10 (respectively, 65) in the comparative judgment question, answered 25% (respectively, 45%) in the absolute judgment question. This clearly, suggests anchoring.*

Example 5 : *The subjects in the experiments of Englich and Mussweiler (2001) were experienced trial judges with an average experience of 15 years. Subjects were asked to consider the case of a fictitious shoplifter accused of stealing from a supermarket for the 12th time. They were also supplied with detailed case material including opinions from a psycho-legal expert, and testimonies by the defendant and a witness. Opinions sought from other experienced and independent legal professionals indicated that the case material was*

[16] Depending on the experiment, several variants of this condition may be given. For instance, $\widehat{\mathbf{x}} < \mathbf{a}$, $\widehat{\mathbf{x}} \in [\mathbf{a}, \mathbf{b}]$, $\widehat{\mathbf{x}} \in (\mathbf{a}, \mathbf{b})$, $\widehat{\mathbf{x}} \in (\mathbf{a}, \mathbf{b}]$, or $\widehat{\mathbf{x}} \in [\mathbf{a}, \mathbf{b})$, where $\mathbf{a}, \mathbf{b}$ are vectors in $\mathbf{R}^n$.

complete and realistic; these professionals called for a mean sentence of 5.62 months with a standard deviation of 2.57.
Subjects in the experiment were asked to determine an appropriate sentence. In two different treatments, the case material showed that the prosecutor had asked for a sentence of, respectively, 9 months (high anchor) and 3 months (low anchor). Subjects in the high anchor treatment chose a sentence of 8 months and those in the low anchor treatment chose a sentence of 5 months. Experience did not mitigate the anchoring effect.

The anchoring phenomenon is remarkably robust across a wide range of domains, contexts and frames. These include estimates of price, the probability of a nuclear war, the evaluations of lotteries and gambles, issues of legal judgment, and first offers in price negotiation. Anchoring also has the potential to explain a range of other phenomena. These include the hindsight bias, preference reversals, and non-linear probability weighting.[17]

4.3. Hindsight-bias

Suppose that one wishes to make a *postdictive judgment*. This is just the flip side of making a *predictive judgement*, but it is backward looking (e.g., what prediction did you make 10 months ago about the stock prices for the current period?). Neoclassical economics assumes that postdictive judgements are perfect in the sense that the predictions made in the past are perfectly recalled (this follows from the assumption of *perfect recall*). However, the evidence suggests that people exhibit a hindsight-bias, relative to the perfect postdictive judgment. In common parlance, one appears wiser in hindsight (Fischhoff, 1975). Hindsight bias might appear to be an inexact phenomenon, that lacks a definition, i.e., simply a label. However, its use in behavioral economics is perfectly rigorous, as we see next.

An individual is asked at time t to predict the value of a random variable, X, at time $t+j$, $j>0$. Let the information set at time t be denoted by I_t. Then, the conditional time t prediction, $E\left[X \mid I_t\right]$, gives the *predictive judgment*. The individual observes the time $t+j$ realization of X, given by x, and is asked to state the remembered time t prediction; this is the *postdictive judgment*. Statistically, this is given by $E\left[E\left[X \mid I_t\right] \mid I_{t+j}\right]$, where I_{t+j} is the date $t+j$ information set, and $I_t \subseteq I_{t+j}$. In neoclassical economics, the assumption of perfect recall guarantees that

$$E\left[X \mid I_t\right] = E\left[E\left[X \mid I_t\right] \mid I_{t+j}\right]. \tag{4.1}$$

However, evidence to the contrary is typically found. Consider the following definition and example of hindsight bias (Dhami, 2016, Section 19.8) that is perfectly rigorous.

[17] The interested reader can pursue the details and the relevant references for all the claims in this paragraph in Dhami (2016, Section 19.6.3).

Definition 5 *: An individual is said to suffer from hindsight bias, or creeping determinism, if the following conditions hold:*
(a) The predictive and postdictive judgments differ, i.e.,

$$E\left[X \mid I_t\right] \neq E\left[E\left[X \mid I_t\right] \mid I_{t+j}\right], \text{ and}$$

(b) the postdictive judgment is biased in favour of the actual realization of the random variable X.

Definition 5 can then be operationalized in a variety of ways, as shown in the next example.

Example 6 *: Consider the following simple rule due to Camerer et al. (1989):*

$$E\left[E\left[X \mid I_t\right] \mid I_{t+j}\right] = \alpha x + (1-\alpha)E\left[X \mid I_t\right]; \alpha \in [0,1]. \tag{4.2}$$

The individual has no hindsight bias if $\alpha = 0$ (this corresponds to the neoclassical case in (4.1)). The most extreme form of hindsight-bias arises if $\alpha = 1$, in which case one's postdictive judgement has no relation to one's initial prediction. This is an extreme form of self-delusion in which the individual asserts: 'I fully knew it all along'. Hindsight bias must be distinguished from learning. In learning, one learns to make better predictions of future events. However, under hindsight bias, the relevant issue is remembered past predictions, i.e., postdictive judgments. There could be degrees of hindsight-bias that are individual specific, as captured by different values of α. Using individual-specific data, regression analysis, or other statistical techniques, can help uncover the individual's underlying preference parameter, α.

We are perfectly aware that hindsight-bias is not a heuristic. The relevant heuristic is given by (4.2) and the hindsight-bias arises whenever $\alpha > 0$.

4.4. The availability heuristic

Suppose that at time t, an individual needs to compute the expected value of some random variable X_{t+j} for some future time period $t+j$. Let the information set I_t capture all possible information at time t, that is required by the relevant economic theory for predicting the value of X_{t+j}. For instance, I_t might include past values of all possible relevant variables, or even the construction of hypothetical scenarios that might influence X_{t+j}. Then, economic theory assumes that individuals compute $E\left[X_{t+j} \mid I_t\right]$, where E is the expectation operator.

The computation of $E\left[X_{t+j} \mid I_t\right]$ could be extremely cognitively challenging (see Dhami, 2016, Parts 7,8). For instance, individuals might not have immediate access to all possible relevant past information in I_t. Their memory might be limited, selective, or strategic, and their recall might be subjective. Or, it might be the case their attention is limited and some information in I_t may be relatively more salient or vivid (e.g., more vivid memories are more easily recalled). For all these reasons, in predicting X_{t+j}, individuals might use only a subset of the information $I_t' \subseteq I_t$, and potentially ignore the rest. Individuals in neoclassical economics will never use a subset of the available information.

Definition 6 *: An individual is said to use the availability heuristic if he/she uses the information set $I_t' \subseteq I_t$ in determining the expected value of X_{t+j} (where I_t', I_t are defined above).*

Remark 2 *: We often do not observe the information sets used by people, but we may use data on observables to determine if the availability heuristic is being used. For instance, once we know the prediction of the relevant economic theory, $E[X_{t+j} \mid I_t]$, we can use the actual prediction $E[X_{t+j} \mid I_t']$, and form an indirect inference that the individual uses the availability heuristic if $E[X_{t+j} \mid I_t'] \neq E[X_{t+j} \mid I_t]$. Alternatively, one may prime subjects with different information sets, I_t^1,I_t^2,I_t^3,..., say by altering the vividness of events/memories. One can then study observed behavior $E[X_{t+j} \mid I_t^1]$, $E[X_{t+j} \mid I_t^2]$, ... If the responses are different, then the availability heuristic is in operation.*

Evidence for the availability heuristic comes from many sources. The typical experiments show that direct experience of a particular event increases the probability that one assigns to related events. Following on from the work of Lichtenstein et al. (1978), Pachur et al. (2012) showed that there is a significant positive correlation between one's estimate of the annual mortality rate from various forms of cancer and the availability of information on cases of cancer from one's social network. Kuran and Sunstein (1999) introduce the idea of *availability cascades*, where the initial availability of some vivid/salient news leads to self-fulfilling bouts of emotional reaction followed by increased vividness of the news and so on. Eventually this may lead to the adoption of policies that might not initially have been justifiable in cost-benefit terms.

4.5. The conjunction fallacy

The conjunction fallacy is defined as follows.

Definition 7 *(Conjunction fallacy): Given any two sets A, B, if $B \subseteq A$ then the conjunction fallacy arises if a decision maker assigns a higher probability to the set B, i.e., $P(B) > P(A)$.*

The original experiments were conducted by Tversky and Kahneman (1983) using the well known *Linda problem*. Hertwig and Gigerenzer (1999) argued that the Linda problem is exhibited by only 15% of their subjects when the problem is presented in a frequency format, as compared to a probability format. A prominent criticism of the KT&O approach has been that presenting information in a frequency format relative to a probability format makes decision making more compliant with the prescriptions of classical statistics (see Section 5.3 below for a critical analysis). For this reason, we first present the Linda problem in its lesser known frequency format, due to Kahneman and Tversky (1996). They used a between-subjects design, and gave the Linda problem to three groups of subjects (Groups 1, 2, 3).

Linda is in her early thirties. She is single, outspoken, and very bright. As a student she majored in philosophy and was deeply concerned with issues of discrimination and social justice. Suppose there are 1000 women who fit this description. How many of them are:

(a) High school teachers? [Groups 1, 2, 3]

(b) Bank tellers? [Groups 1, 2]

(c) Bank tellers and active feminists? [Groups 1, 3]

The median response for choice (c) was statistically larger than the median response for choice (b), confirming the conjunction fallacy. Subsequent evidence in Tentori et al. (2004) also supports the view that presenting information in natural frequencies does not eliminate the conjunction fallacy. Charness et al. (2010) find that incentives and an increase in the group size that makes the relevant decisions, reduces the conjunction fallacy. However, Bonini et al. (2004) and Stolarz-Fantino et al. (2003) find that monetary incentives do not reduce the conjunction fallacy; for the details of these studies, see Dhami (2016; Section 19.3).

4.6. Regression to the mean

Several fallacies arise from ignoring the statistical phenomenon of *regression to the mean*. This is precisely defined in classical statistics; see Dhami (2016, Section 19.10.1). Most people ignore regression to the mean, hence, make inferences that are statistically incorrect, i.e., they exhibit a bias. One of our favorite examples of this bias is from Kahneman (2011, Ch. 17). Israeli air force instructors found that praising pilots for executing complicated maneuvers reduces subsequent performance, but criticism for bad performance improves subsequent performance. They inferred that praise ought to be withheld but criticism was healthy. However, random luck played a huge part in the success of these complicated maneuvers and regression to the mean implied that pilots who had done brilliantly (respectively, not so well) this time are likely to do less well (respectively, much better) next time. Praise or criticism is irrelevant here, and the mistaken inference arose from ignoring regression to the mean.

4.7. Necessary and sufficient conditions

An understanding of necessary and sufficient conditions is central to economic models; individuals who obey the BRA do not violate the laws of logic. However, human beings can be confused about the difference between necessary and sufficient conditions; the classic experiment is due to Wason (1968). This confusion can be formally defined, so it is not a mere label either. For instance, if $P \Rightarrow Q$, then a confusion between necessary and sufficient conditions arises if one concludes $Q \Rightarrow P$, or violates the negation $\sim Q \Rightarrow \sim P$ (where $\sim$ denotes the negation of a statement).

Indeed, professional economists, who are steeped in such logic, may also make similar errors. From the observation that most investors, including mutual and pension fund managers, find it very hard to beat the stock market (Malkiel, 1990), it is sometimes concluded that financial markets must be efficient (Rubinstein, 2001). However, Barberis and Thaler (2003) point out that this is a confusion between necessary and sufficient conditions. Market efficiency in finance (sometimes termed as the efficient markets hypothesis) implies that

"Prices equal fundamental values $\Rightarrow$ no arbitrage opportunities."

However, the absence of arbitrage opportunities is only a necessary, but not a sufficient, condition for stock prices to equal fundamental values. Hence,

"No arbitrage opportunities $\not\Rightarrow$ prices equal fundamental values."

Thus, the difficulty in beating the stock market cannot be taken as evidence that prices equal fundamental values, or that financial markets are efficient.

4.8. Confirmation bias

A central feature in most economic models is a specification of how information is updated over time. The BRA requires individuals to use Bayes' Law for this purpose. However, empirical evidence shows that people make substantial departures from Bayes's Law, which are systematic (see Section 5.3). In particular, people are subject to a *confirmation bias*; their posterior beliefs, based on a biased interpretation of the existing evidence, are too close to their prior beliefs. One of the most succinct descriptions of the confirmation bias has been given by Lord et al. (1979, p. 2099), which deserves to be read in full:

"Thus, there is considerable evidence that people tend to interpret subsequent evidence so as to maintain their initial beliefs. The biased assimilation processes underlying this effect may include a propensity to remember the strengths of confirming evidence but the weaknesses of disconfirming evidence, to judge confirming evidence as relevant and reliable but disconfirming evidence as irrelevant and unreliable, and to accept confirming evidence at face value while scrutinizing disconfirming evidence hypercritically. With confirming evidence, we suspect that both lay and professional scientists rapidly reduce the complexity of the information and remember only a few well-chosen supportive impressions. With disconfirming evidence, they continue to reflect upon any information that suggests less damaging "alternative interpretations." Indeed, they may even come to regard the ambiguities and conceptual flaws in the data opposing their hypotheses as somehow suggestive of the fundamental correctness of those hypotheses. Thus, completely inconsistent or even random data—when "processed" in a suitably biased fashion—can maintain or even reinforce one's preconceptions."

Confirmation bias is supported by the evidence (Dhami, 2016, Section 19.9.1). It has been formalized and defined in the model of Rabin and Schrag (1999); see Section 19.9.2 in Dhami (2016). In this model, the decision maker receives some signals about the true state of the world. Signals are correctly perceived if they accord with one's prior beliefs. However, if they do not accord with the prior beliefs, then the individual misinterprets the signal with some probability; the degree of misrepresentation then captures the departure from standard Bayesian updating. This also clarifies the difference of this heuristic from simple hindsight-bias. One of the results of this model is that it leads to overconfidence in one's predictive abilities because ex-post the individual finds that the (incorrectly computed) posterior beliefs are closer to the prior beliefs, relative to Bayesian updating. Indeed, among many applications, such an inference process can lead to long-lasting stereotypes and to racial prejudices.

4.9. Heuristics that lack a precise mathematical definition

Several heuristics are hard to define in a precise mathematical sense.

False consensus effect or evidential reasoning

Consider two individuals A and B who are trying to solve some problem P. The action sets of the two individuals are identical, for simplicity, and given by S. Suppose that, for whatever reason, individual A has a predisposition to choose action $s_A \in S$ when faced with problem P. Now individual A is asked: What action do you think B will choose. There are many aspects of $B's$ choice that A might not know. For instance, A might not know the exact preferences of B. If A can form probability distributions over the possible preferences of B, then in neoclassical economics, this is a problem of asymmetric information, and it is routinely analyzed. However, the asymmetric solution might be too cognitively challenging, or its preconditions might not be satisfied (e.g., A might not know the relevant probability distribution of preferences, or might not be able to use

Bayes' Law). In this case, A might simply use his own predisposition to choose $s_A \in S$ as diagnostic evidence that B is also likely to choose s_A. This reasoning on $A's$ part is particularly relevant when A believes that B is like-minded in some relevant dimension (Robbins and Krueger, 2005; al-Nowaihi and Dhami, 2015). This is the false consensus effect.

The *false consensus effect* is a simple heuristic that allows people to form quick and ready inferences about the actions or beliefs of others by using their own predispositions as a guide; thus, it is a fast and frugal heuristic. The false consensus effect arises from the use of *projection-bias* that has been well documented in psychology (Robbins and Krueger, 2005). It is sometimes also known as *evidential reasoning* and it has been formalized in the case of strategic interaction (al-Nowaihi and Dhami, 2015).

Example 7 *In their classic experiment, Ross et al. (1977) asked Stanford University students whether they would be willing to walk around the campus for 30 minutes wearing a sandwich board inscribed with the word "Repent". Those who agreed, estimated, on average, that 63.50 percent of their fellow students would also agree, and those who refused, estimated that 76.67 percent of their fellow students would also refuse.*

There is now a great deal of evidence on the false consensus effect from many different domains including legal judgement, opinions on climate change, and assessment of risk preferences of others; for the evidence, and the references see Dhami (2016, Section 19.10.2).

In the context of strategic interaction, al-Nowaihi and Dhami (2015) show that evidential reasoning can be defined in a precise mathematical sense and it generates testable empirical implications. It can, for instance, explain the cooperative outcome in a prisoner's dilemma game that is problematic to explain for most neoclassical and many behavioral theories of strategic interaction.

The affect heuristic

Neoclassical economics ignores the role of emotions in decision making. Often it assumes that human beings make cold, calculated decisions. There is now a burgeoning literature in behavioral economics that tries to highlight the salient role of emotions in many decisions (Dhami, 2016, Part 6 and Section 13.5). The affect heuristic highlights that events have emotional tags (positive or negative) associated with them, which influence choices, in addition to the associated monetary aspects.

Zajonc (1980) offers the following example that illustrates the affect heuristic: we do not just see a car, we see a *sleek car*, a *cool car*, an *ugly car*, depending on the *affect* of the car on us. Slovic et al. (2002, p. 398) write: "We sometimes delude ourselves that we proceed in a rational manner and weight all the pros and cons of the various alternatives. But this is probably seldom the actual case. Quite often "I decided in favor of X" is no more than “I liked X”...We buy the cars we “like,” choose the jobs and houses we find “attractive,” and then justify these choices by various reasons.” Kahneman (2011) notes that like other heuristics, the affect heuristic substitutes the answer to a hard question (What do I think about it?) by the answer to an easier question (How do I feel about it?). This conserves cognitive resources, particularly in the case of complex and difficult problems that must be decided within a time constraint.

The affect heuristic is not straightforward to formalize in a precise mathematical sense. However, two features stand out. (1) There has been a great deal of work in psychology and behavioral economics on empirical measures of emotions (Dhami et al., 2018; Dhami, 2016). For instance, Dhami et al. (2018) show that the emotions of guilt, surprise-seeking, inferring intentions of other, and reciprocal feelings all play a significant role in determining contributions in public goods games. (2) Psychological game theory formalizes the role of emotions in a mathematically precise sense, but the underlying framework is optimization-based (Dhami, 2016, 13.5).

In contrast, there has been less progress in building precise procedural models of heuristics that incorporate emotions, and this is likely to remain a challenging task.

5. An evaluation of the criticisms of the KT&O program

In this section, we consider the criticisms of the KT&O program, mainly levelled by the G&O program. Over twenty years ago, Kahneman and Tversky (1996, p. 584) offered a blunt assessment of the criticisms of their position: "The position described by Gigerenzer is indeed easy to refute but it bears little resemblance to ours. It is useful to remember that the refutation of a caricature can be no more than a caricature of a refutation." Many of these criticisms are discussed in Kahneman and Tversky (1996) and Gigerenzer (1996). Although the original exchange remains worth reading, the literature has greatly progressed over the years. One of our aims is to offer an up to date evaluation of this debate, with the benefit of recent research and findings.

5.1. One event probabilities, context, and errors

Some of the arguments against KT&O can be dealt with relatively easily.

1. The frequentist approach requires a large number of repetitions. Hence, experimental evidence involving probabilities based on *singular events* may be suspect (Gigerenzer, 1991). For instance, in the famous Linda problem, whether Linda is a bank teller or a feminist is a singular event. In this case Gigerenzer (1991, 1993) argued that one can speak only about subjective probabilities in a Bayesian sense. Since any subjective probabilities are possible, so the argument goes, the experimental results

are uninformative about the violation of a rationality-based model. Kahneman and Tversky (1996) note that this view of probability is not universally held, and that this issue applies to only 2 out of the 12 biases that they highlight. Furthermore, the responses show systematic biases in one direction, which weakens the case for arbitrary subjective probabilities.

2. KT&O have been criticized for ignoring the context, content, and framing effects in their demonstration of biases (Gigerenzer et al., 1988; Gigerenzer, 1993). In other words, it is claimed that they ignore ecological rationality. This claim is mistaken. Kahneman and Tversky (1996, p. 583) write: "the assumption that heuristics are independent of content, task and representation is alien to our position, as is the idea that different representations of a problem will be approached in the same way." Several examples can be given. KT&O showed that the conjunction fallacy is reduced in a frequency format (see Section 4.5). They also drew upon the distinction between causal and non-causal base rates, which explains when the base rate is or is not ignored (Tversky and Kahneman, 1980).[18] The frame dependence of preferences has generally been a major finding in the approach of Kahneman and Tversky (Dhami, 2016, Parts 1,7).

3. The heuristics and biases program has been criticized on the grounds that subjects may simply be making errors, could be inattentive, or may be suffering from temporary lapses of judgment; see Stanovich and West (2000) for a survey. Some of the more extreme suggestions that have been made to cast doubt on the KT&O program include temporary insanity of the subjects, difficult childhood of the subjects, and entrapment by the experimenter (Kahneman, 1981, p. 340). The pattern of biases discovered by KT&O does not describe random mistakes, but systematic mistakes; hence, these objections cannot be taken seriously. Furthermore, there is typically a great deal of heterogeneity among subjects who exhibit biases, which contradicts conformity with a single rational response.

5.2. "We cannot be that dumb" critique

Gilovich et al. (2002) refer to the '*we cannot be that dumb critique*,' which argues that stellar human achievements, e.g., discovering the structure of DNA and space flights, are not consistent with the idea that people might be using simple judgment heuristics. This is

[18] In the Linda problem, based on the description it is plausible that she is a member of the feminist movement (causal base rates). However, if substituted by a non-causal description (e.g., Linda drinks black tea in the morning), then the conjunction fallacy is unlikely to arise (see Section 19.3.3 in Dhami, 2016).

a misunderstanding about the nature/process of scientific discoveries and the methodology of science.

Scientists are likely to use simple heuristics to build initial intuitions about their problems (indeed, so do social scientists). This is typical of the nature/process of scientific discoveries; James Watson reportedly gained insights on the double helix structure of DNA in a dream, and penicillin was discovered by Louis Pasteur while studying why silkworms were dying because of a bacterial infection (thus inaugurating modern medicine). However, the methodology in science and the sociology of science are entirely different. Scientific methodology relies on refutability of theories, stringent empirical testing of theories, replication of the evidence, and transparency of the data. The sociology of science ensures, to a much greater degree as compared to economics, that the gatekeepers of science journals, the editors and the referees, publish only peer reviewed research that conforms to the scientific method. This has been a remarkably successful combination in producing great progress in science. Hence, there is no essential contradiction between scientific progress in science and the use of judgment heuristics by scientists.

5.3. Frequency versus probability format and the KT&O program

Several critics of the KT&O approach have argued that the biases created by heuristics in the KT&O program are eliminated or substantially reduced when data is presented in frequency format rather than probability format (Gigerenzer et al., 1988, Gigerenzer, 1991, 1996, 2004, 2008). Our own reading of the evidence is as follows:

1. An entire class of heuristics in the KT&O program is unaffected by the distinction between frequency and probability formats. This includes the representativeness heuristic (including gambler's fallacy, hot hands fallacy), anchoring, availability, hindsight-bias, regression to the mean, confirmation-bias, and the affect heuristic.

2. The frequency format might reduce biases in the case of some heuristics, particularly in the case of the conjunction fallacy and in applications of Bayes's Law. However, significant bias remains in applications of Bayes' Law. After reviewing the relevant evidence, Kahneman and Tversky (1996, p. 585) cautiously noted: "Contrary to Gigerenzer's unqualified claim, the replacement of subjective probability judgments by estimates of relative frequency and the introduction of sequential random sampling do not provide a panacea against base-rate neglect."

Tversky and Kahneman (1983) themselves first highlighted the difference in results when information is presented in a frequency format relative to a probability format. They found that in a within-subjects design, the frequency format leads to very few conjunction errors. However, the conjunction bias is found to be sufficiently high in a between-subjects treatment (Kahneman and Tversky, 1996); see the discussion in Section 4.5. The anchoring

Study	Information format and judgement domain	
	Probability	Frequency
Cascells et al., (1978)	18 (60)	--
Cosmides & Tooby (1996; Exp. 2)	12 (25)	72 (25)
Eddy (1982)	5 (100)	--
Evans et al., (2000; Exp1)	24 (42)	35 (43)
Gigerenzer (1996b)	10 (48)	46 (48)
Gigerenzer and Hoffrage (1995)	16 (30)	46 (30)
Macchi (2000)	6 (30)	40 (30)
Sloman et al., (2003; Exp 1)	20 (25)	51 (45)
Sloman et al., (2003; Exp 1b)	---	31 (48)

Table 5.1: Percentage of responses consistent with Bayes' rule in different empirical studies. Sample sizes in parenthesis. Source: Barbey and Sloman, 2007

heuristic was also found to be used when the information in Example 4 was presented in a frequency format (Tversky and Kahneman, 1973, 1974).

The use of Bayes' Law is critical to the BRA, and it is indispensable in game theory. However, its central role and acceptance in economics is an article of faith; no empirical evidence of conformity with Bayes' law is taught in economics courses. When attempts have been made to test it in experiments, conformity with Bayes' Law is low, and the typical finding is one of neglect of base rates. Consider the well-known cab problem.

Example 8 *: In Tversky and Kahneman (1980), subjects received the following information: There are only two cab companies in the city, Green and Blue; 85% of the cabs are Green. There was an accident last night. A witness comes forward to testify that the cab involved in the accident was Blue. In similar conditions, the reliability of the witness is 80%, i.e., the probability that the witness gets it wrong is 20%. What is the probability that the actual cab involved in the accident was Blue? The median and the modal response was* 0.8*, which is also the probability with which the testimony of the witness is correct. The statistically correct answer, using Bayes' rule is 0.414. In particular, individuals appear to underweight base rates, i.e., that only 15% of the taxis are actually Blue. For factors that influence the degree to which base rates are taken into account, e.g., when data is given a causal or incidental representation, or when people are encouraged to think like statisticians, see Kahneman and Tversky (1996) and Kahneman (2011).*

Table 5.1, which uses information from Barbey and Sloman (2007), shows the percentage of responses that are consistent with Bayes' rule in a probability format (second column) and a frequency format (third column). There is greater conformity with Bayes' rule when the information is presented in a frequency format.[19] Yet, less than half of the subjects conform to Bayes' rule in this cross section of studies. While the frequency format reduces base rate neglect, on average about 60% of the subjects across the studies still exhibit base rate neglect. Evans et al. (2000) report that the frequency format has been found, depending on the experiments, to worsen, improve, or leave unchanged the quality of the judgments made.

3. It is possible that humans might have evolved to understand natural frequencies relatively better than percentages (Cosmides and Tooby, 1996; Pinker, 1997). However, most real world economic data is presented in percentage terms, e.g., interest rates on mortgage borrowing; the forecasts of inflation and unemployment by central banks; various insurable risks; performance indicators on open days in Universities; and the rates of return on portfolios. Thus, as a practical matter, one needs to understand human judgment and decision making when information is presented in a probability format.

4. When conformity with Bayes' rule under a frequency format is claimed to be truly spectacular, it appears confounded, at least in some cases, by framing effects that favour the outcomes in a frequency format. Consider, for instance, the problems reported in Gigerenzer (2008, p. 16-18) given to 160 experienced gynecologists.

A. Probability format

Assume that you screen women in a particular region for breast cancer with mammography. You know the following about women in this region:

The probability that a woman has breast cancer is 1% (prevalence)

If a woman has breast cancer, the probability is 90% that she will have a positive mammogram (sensitivity)

If a woman does not have breast cancer, the probability is 9% that she will have a positive mammogram (false positive rate)

A woman who tested positive asks if she really has breast cancer or what the probability is that she actually has breast cancer. What is the best answer?

(1) "It is not certain that you have breast cancer, yet the probability is about 81%." [14].

(2) "Out of 10 women who test positive as you did, about 9 have breast cancer." [47]

(3) "Out of 10 women who test positive as you did, only 1 has breast cancer." [20]

(4) The chance that you have breast cancer is about 1%." [19]

The numbers in the square brackets give the number of subjects out of 160 who chose that particular response. The statistically most correct answer is (4) but it is chosen by only 19 subjects (11.9 % of the sample).

[19] The Cosmides and Tooby (1996) result is an outlier that Sloman et al. (2003) and Evans et al. (2000) could not replicate.

B. Frequency format

The italicized text in the probability format is now changed as follows (the remaining information is identical):

10 out of every 1,000 women have breast cancer

Of these 10 women, we expect that 9 will have a positive mammogram.

Of the remaining 990 women with breast cancer, some 89 will still have a positive mammogram.

Imagine a sample of women who have positive mammograms. How many of these women have cancer? ____ *out of* ___

In this case, 87% choose the correct answer.

There are two changes in this text relative to the one in the probability format. First, information is presented in a frequency format. Second, the problem is framed with a subtle difference that appears to clarify it, possibly significantly. The relative contribution of these two factors is not clear. When framing is absent, we get the comparison described in Table 5.1. However, it would appear that with the altered framing, compliance with Bayes' rule may increase. This in itself is a valuable result and there is now a growing literature that the framing of information, particularly financial information that enhances financial literacy, ensures that people can make better choices (Bertrand and Morse, 2011).

We believe the following to be the appropriate frame-equivalent analogue of the frequency format to the above problem when the information is presented in a probability format:

The probability that a woman has breast cancer is 1% (prevalence)

Of these 1% women who have breast cancer, we expect that 90% will have a positive mammogram (sensitivity)

Of the remaining 99% women without breast cancer, *some 9% will have a positive mammogram (false positive rate)*

A woman who tested positive asks if she really has breast cancer or what the probability is that she actually has breast cancer. What is the best answer?

5.4. Empirical counterparts to the heuristics

G&O criticize the availability heuristic for not specifying the exact empirical proxy for availability. As Gigerenzer and Brighton (2011, p.18) put it: "The use of labels is still widespread..., and it therefore worth providing a second illustration of how misled by their seductive power. Consider the "availability heuristic,"...this label encompassed several meanings, such as the number of instances that come to mind...the ease with which the first instance comes to mind...the recency, salience, vividness, memorability, among others..."

We now address this criticism. The availability heuristic is well defined in a mathematical sense and is eminently testable (see Section 4.4, above). Hence, it meets the standards of scientific rigor in economics. There are several examples in economics, where the exact empirical counterpart of a variable in an economic model is not fully specified. This arises because economic theories are necessarily a parsimonious description of complex reality. We give three examples from Nobel Prize winning work in economics to illustrate this methodological point.

1. Friedman and Schwartz (1963) argued that an increase in 'money supply' increases inflation with long and variable lags. Similar predictions also hold in New Keynesian models with frictions. However, there are at least four different empirical definitions of money supply (commonly known as M1, M2, M3, M4), but theory does not specify which of these definitions is the most appropriate one (neither can "long" and "variable" lags be precisely defined). However, in this case, the data guides us on the appropriate definitions of money supply to use and it also helps us uncover the length of the lags.

2. Lucas (1988) showed that human capital is an important determinant of economic growth, but never specified the precise measure of human capital (primary/secondary school education? University education? vocational training?). Similar predictions, without an exact empirical specification of the corresponding variable are offered in several modern growth theory models. It was left to the empirical work to discover which measures of human capital were the most useful (Barro and Sala-i-Martin, 2003).

3. Kahneman and Tversky (1979) introduced the concept of *reference dependence* into economics. They did not specify the exact empirical proxy for reference points. However, they did opine that the status-quo is a strong candidate for a reference point, and that other plausible candidates include a fair outcome, expected outcome, or entitlements based on norms (see Dhami, 2016, Section 2.4.4). It was again left to empirical work to discover the most suitable and appropriate reference point depending on the application one uses.

For these reasons, we do not view the criticism of a lack of exact empirical proxy for the availability heuristic (and the other heuristics) to be an important one.

5.5. How can heuristics explain events A and not A?

One can explain the gambler's fallacy and the hot hands fallacy in terms of the representativeness heuristic. This might sound paradoxical; see, e.g., Definitions 2 and 3. How can both negative and positive autocorrelation be explained by the same theory? Gigerenzer

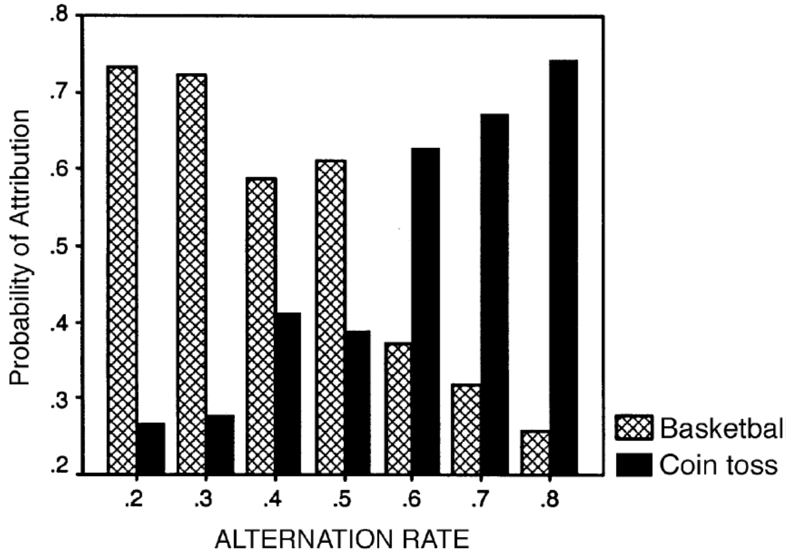

Figure 5.1: from basketball or a coin toss? Source: Ayton and Fischer (2004).

and Brighton (2011, p.18) criticize the representativeness heuristic thus: "No model of similarity can explain a phenomenon and its contrary; otherwise it would not exclude any behavior." They go on to write: " As it [*representativeness heuristic*] remains undefined, it can account even for *A* and non-A (Ayton and Fisher, 2004)."

We believe that this is likely to be a misunderstanding of how representativeness operates in these two cases. Consider the empirical evidence from the same paper that Gigerenzer and Brighton (2011, p.18) cite, Ayton and Fischer (2004). The key to providing a representativeness-heuristic based explanation for the gamblers and hot hand fallacies is to examine the domains in which these fallacies arise. Ayton and Fischer (2004) suggest that these domains are different. They propose that the hot hands fallacy is more likely to arise for *human performance*, as is the case with basketball and tennis players, rather than for *inanimate processes* such as throws of a dice, or coin tosses.

In Ayton and Fischer (2004), subjects were presented with various sequences of 21 binary outcomes (e.g., success and failure) that had varying degrees of positive and negative autocorrelation. The *alternation rate* is the number of times a sequence changes direction divided by 10. They were then asked to guess the source of these sequences: *human skilled performance* (e.g., arising from basketball shots or tennis serves) or an *inanimate process* (e.g., a coin toss, or a binary outcome roulette wheel).

The results for one of the comparisons is shown in Figure 5.1. Sequences with low rates of alternation were rated by subjects to have more likely come from basketball rather than coin tosses. Conversely, sequences with high rates of alternation were rated as more likely to have come from a coin toss. These results highlight when the gambler's fallacy is more likely to arise (high alternation rates arising from an inanimate process) and when the hot hands fallacy is more likely to arise (low alternation rates and the involvement of human skills).

Why does this difference in results arise? We believe that subjects are solving quite different inference problems in each case.

(a) In the case of inanimate processes, such as coin tosses (or a dice throw), subjects are sure of the underlying probabilities of H and T if the coin is fair (50% each). Thus, subjects may switch too often in small samples to preserve the population proportions, as suggested by the representativeness heuristic.

(b) In predicting the success rates of outcomes in professional sport, the underlying distribution from which the binary outcome (success, S, or failure, F) is drawn is unclear. The reason is that a streak of S or of F in a particular sports event, signals unobservable attributes about a player to observers, such as the level of confidence, fatigue, emotional stability, coaching, and the quality of training prior to the game. Here, the observer is trying to *infer the underlying model* that gives rise to the unobserved success rate in a particular game.

Suppose for instance, there are two competing models (Barberis et al., 1998). (i) A *trending model* in which the unobserved attributes are particularly conducive to good performance, so that the player has a high success rate. (ii) A *mean reverting model* in which success is intermingled with failure in such a way that the player achieves some mean level of success in the observed game. On observing a sequence of S, S, S, S, an observer who uses the representativeness heuristics is likely to infer that the observed sequence is representative of the trending model, and not the mean reverting model. Thus, the observer may assign a high probability that the fifth outcome will be S. In contrast the "true" underlying model might well be a random model (which is neither trending, nor mean-reverting) in which case, the correct statistical inference is a 50-50 chance of S. This data will immediately suggest a hot hands fallacy.

This reasoning potentially explains the evidence presented in Figure 5.1. The key to the explanation lies in the different inference procedures for the cases of inanimate and animate processes. In particular, to our mind, there does not appear to be any inconsistency in using the representativeness heuristic to simultaneously explain the gambler's fallacy and the hot hands fallacy.

5.6. The criticism of appropriate statistical norms

Gigerenzer (1996) questions the appropriate 'statistical norms' that apply to the problems in the experiments in KT. His argument is that the underlying model is not fully specified, which leaves unclear the appropriate statistical prediction. He writes (p. 592):

"A convenient statistical principle, such as the conjunction rule or Bayes' rule, is chosen as normative, and some real-world content is filled in afterward, on the assumption that only structure matters. The content of the problem is not analyzed in building a normative model, nor are the specific assumptions people make about the situation." This unease about rejecting heuristics without specifying the model in which they are situated is also implicit in the criticism of Gigerenzer and Brighton (2011, Section 3). This criticism potentially covers several points that we address below.

1. Normative standards of behavior: KT&O did not advocate, nor defend, a particular normative standard of behavior. They took the existing, and well established, normative standard of behavior in economics. Their objective was to test if people do actually conform to this normative standard.

   In economics, the normative standard is encapsulated in the BRA, which requires some of the following assumptions: People follow the law of large numbers; use expected utility theory; use all available information that the relevant economic model requires; follow Bayes' law to update their information sets; have perfect recall; are not influenced by irrelevant information; and their preferences are context/frame independent. In order to test the BRA, it is consistent with scientific methodology to test these individual assumptions directly, or alternatively, to test a model that uses these assumptions plus other auxiliary assumptions. The KT&O approach has often directly tested the assumptions behind the BRA framework, and they found that the normative standard was violated.

   This approach in KT&O also extends to the testing of theories such as expected utility theory. It can be shown that certain axioms of rationality (Definition 1.1 in Dhami, 2016, p.86) imply and are implied by an evaluation of risky outcomes in terms of expected utility theory (Proposition 1.1 in Dhami, 2016, p.86). One may then either test a model that uses expected utility theory (and other auxiliary assumptions), or directly test any of the axioms of rationality. In this case, it is simpler to directly test any of the axioms; indeed, testing of expected utility has largely employed tests of the independence axiom (Section 1.5.1 in Dhami, 2016). We see nothing objectionable in this approach, nor do tests of the independence axiom imply that the researcher accepts the normative standard implied by expected utility theory.

   A similar methodological approach was used for the testing of the main theory of decision making over time in economics, the exponential discounted utility model (Dhami, 2016, Ch. 9). In this case, the *stationarity axiom* has been directly tested to refute the exponential discounted utility model. An analogy from aerospace may help: The airframe of a proposed new airplane might be tested in a wind tunnel, without adding-on all the other components in the airplane. This is an acceptable procedure in science.

2. Ecological rationality: Implicit in the criticism of the KT&O program on the grounds of statistical norms is the view that the program does not take account of ecological rationality. Interpreting ecological rationality as context and frame dependence of preferences, we have argued above that this criticism does not hold; Kahneman and Tversky (1996) have themselves addressed this criticism directly.
   One may alter the environment within which a problem is embedded to analyze if the appropriate decision changes. For instance G&O have done so by examining differences in behavior between a probability format and a frequency format to study the ecological rationality of behavior (Section 5.3). This is interesting and useful, but of relevance to the KT&O approach is that the BRA assumes human behavior is identical in a frequency and a probability format. Hence, a demonstration of non-compliance with the predictions of BRA in either a probability or a frequency format is sufficient to disprove this normative benchmark. However, if it were to turn out that people conformed with the BRA in a frequency format for most of the cases that KT&O consider (which they do not), then a newer version of BRA may be proposed in a frequency format and re-tested. This is entirely consistent with how science progresses.

5.7. Systems 1 and 2

Both sides in the great rationality debate also differ about the appropriate models of the brain that may lead to the use of heuristics. This is not the focus of our paper and the interested reader may consult Stanovich and West (2000) and Stanovich (2012). But we offer some brief remarks on this issue because it has a bearing on some of the criticism of the KT&O program.

Kahneman (2011) devotes the first one-third of his book to developing a two systems model of the brain, Systems 1 and 2, that facilitates a deeper understanding of the biases. To be sure, there is no universal agreement among psychologists about the appropriate models of the brain in this context (Stanovich and West, 2000). Kahneman (2011) fully recognizes that System 1 and System 2 are useful concepts, not meant to correspond with specific brain areas (p.29), but aiding us in making greater sense of heuristics.[20]

On a widespread view, the quick, reactive, and automatic System 1 is responsible for many errors relative to the statistical benchmark. System 2 has been likened to a *lazy controller* in Kahneman (2011) and when it is called upon to intervene in an unusual

[20] Atoms and genes were first hypothesized as useful concepts long before their material counterparts were discovered.

situation, the agenda (e.g., affective emotions, recalled memory, associations) is chosen by System 1. System 1 tries to make sense of a situation even when the events may have been generated purely randomly. Kahneman (2011, p. 204-205) puts it as follows: "The sense-making machinery of System 1 makes us see the world as more tidy, simple, predictable, and coherent than it really is. The illusion that one has understood the past feeds the further illusion that one can predict and control the future. These illusions are comforting. They reduce the anxiety that we would experience if we allowed ourselves to fully acknowledge the uncertainties of existence." This gives rise, for instance, to the law of small numbers, the gambler's and hot hands fallacies, and overconfidence in the ability to explain the future from a hindsight-biased explanation of past events that were purely random. Indeed, this provides an altogether different view of the existence of heuristics that is not rooted in the bias-variance dilemma (see Section 9 below).

6. Experts and the KT&O heuristics

Do experts use heuristics that are identified in the KT&O program? Do they exhibit biases relative to the BRA? If the answer is in the affirmative, then, as noted above, it supports the ecological rationality of heuristics in the KT&O program.

There is now widespread evidence that many experts do exhibit biases (Dhami, 2016, Section 19.18). The evidence comes from multiple domains. When making predictions of political events, experts were only slightly more accurate than chance. However, they were better able to generate explanations for their predictions (Tetlock, 2002, 2006). Experts are typically more overconfident than lay people who lack similar experience and more experienced experts are more overconfident (Heath and Tversky, 1991; Glaser et al., 2007; Kirchler and Maciejovsky, 2002). The realized returns on stocks are within the 80% confidence intervals of the returns predicted by senior finance professionals in only 36% of the cases (Ben-David et al., 2013).[21]

Mathematical psychologists exhibit the law of small numbers (Tversky and Kahneman, 1973). The perceived riskiness of various hazardous substances by toxicology experts is consistent with the affect heuristic (Slovic et al., 1999). Clinical psychologists underweight base rates relative to a Bayesian calculation (Meehl and Rosen, 1955). A meta study establishes that decision makers who have the relevant expertise in the field suffer from the hindsight bias (Guilbault et al., 2004). Professional traders in a large investment bank were found to be hindsight-biased (Biais and Weber, 2009).

Evidence supports the important role of anchoring on a given list price by estate agents (Northcraft and Neale, 1987). Evidence of anchoring is also found in legal judgment

[21] See also Chapter 5 in Gigerenzer (2014) for more discussion on simple rules of thumb followed by experts in the financial markets that prevent good stock market predictions.

(Chapman and Bornstein, 1996; Englich and Mussweiler, 2001; Englich et al., 2006). Judges exhibit the false consensus effect (Solan et al. 2008). Finance professionals also exhibit a false consensus effect and they impute to others their own risk preferences (Roth and Voskort, 2014).

Experts, such as physicians and World Bank staff, exhibit framing effects, often of a similar magnitude to that observed with student populations (McNeil et al., 1982; Kahneman and Tversky, 1984; WDR, 2015). Wholesale car market dealers exhibit limited attention (Lacetera et al., 2012). WDR (2015) documents several kinds of biases among its professional staff. These include confirmation-bias, susceptibility to sunk costs, and the influence of framing. The WDR (2015, p.18) is candid in its assessment of expert-bias: "This finding suggests that development professionals may assume that poor individuals may be less autonomous, less responsible, less hopeful, and less knowledgeable than they in fact are." The WDR also suggests potential solutions to the problem of expert-bias. These include *dogfooding* (experts signing up and playing their own programs for real) and *red teaming* (having an adversarial outside team that tests the proposals).

Thus, the evidence suggests that many experts do use heuristics and exhibit biases relative to the BRA; this is not always eliminated by market experience.

7. The G&O program

Herbert Simon (1955) distinguished between *substantive rationality* (maximizing an objective function under constraints as in Problem GP), and *procedural rationality* (the process/quality of decision making). Simon's insight was that human beings may lack the information and cognitive ability to solve problems in BRA, such as Problem GP. But people do make decisions. Hence, he was interested in the cognitive processes that give rise to such decisions, or procedural rather than substantive rationality. Simon proposed the *satisficing heuristic* to operationalize procedural rationality, in which individuals set goals or targets, called aspiration levels. The individual then searches for alternatives that give rise to different payoffs. Once an alternative attains the aspiration level, it is deemed as satisfactory, and further search is terminated. The word "satisficing" is a neologism that alludes to the fact that such decision procedures are satisfactory and they suffice. Empirical evidence is supportive of the theory (Caplin et al., 2011). An important application of this approach is to explain cooperation in the prisoner's dilemma game, an outstanding problem for neoclassical game theory (Dhami, 2016, Section 19.14.1).

The G&O program focuses on procedural rationality. It makes an attempt to model the cognitive process by employing heuristics that are fast and frugal (economize on time and information). The KT&O program also focusses on fast and frugal heuristics and also uses procedural rationality, although in a different way from the G&O program.

For instance, in the KT&O program, procedural rationality is reflected in some of the following ways: Searching through the available information (availability heuristic); being influenced by prior information (anchoring); being influenced by the law of small numbers in making decisions (representativeness heuristic); being unduly influenced by one's own prior beliefs (hindsight-bias and confirmation-bias). In contrast, in the G&O program, procedural rationality is introduced by some of the following features: Using cues to make decisions, ranking cues according to their ecological validity, deciding on the basis of the first discriminatory cue, or tallying over cue values, or using a weighted average of the cue values.[22]

The class of *one-good-reason heuristics* is based on checking a series of cue-values that are arranged in decreasing order of *ecological validity* (these terms are described below). This class includes the *take-the-best* heuristic (Gigerenzer and Goldstein, 1996) and the *priority heuristic* (Brandstätter et al., 2006). Often these heuristics are used to answer quite simple questions, with binary answers. Let us use the following example: Which German city, in a sample of cities, has the highest population? Typically, the experimenter provides subjects with several cues and their ecological validity. For instance, some of the cues could be: Do you recognize the name of the city? Does the city have a football team? Does the city have an intercity train station? The cue can either differentiate between the cities or not.

Suppose that A and B are two cities in the sample and one poses the question: Is A more populous than B? A simple cue could be whether A and B are recognized (recognition heuristic). A cue value of 1 is given to the city that one recognizes, and 0 otherwise. Among two cities, the subject may (1) recognize one city but not the other, (2) recognize both cities, (3) not recognize any city. Only in the first case is the cue said to discriminate between the cities. The *ecological validity* of a cue is its overall ability to discriminate in pairwise comparisons between the cities in the sample of cities under consideration.

The take-the-best cue then proceeds in the following manner. (1) Search rule: The cues are examined in decreasing order of ecological validity. (2) Stopping rule: Terminate search when one encounters the first cue that can discriminate between the cities. (3) Decision rule: Conclude that the city which the cue has discriminated in favour of, is the larger of the cities. Gigerenzer and Goldstein (1996) used data on 83 German cities with 9 ecological cues and compared the take-the-best heuristic with other methods such as tallying (giving equal values to all cues and adding them up), weighted tallying (weighted values of cues are added), and logistic regression (what is the probability based on regression estimates that a city is more populous, conditional on the cue values?).

[22] In contrast to these methods, some of the most persuasive examples of simple rules of thumbs come in the form of checklists, such as hospital checklists to reduce infections, and other such simple algorithms that do appear to work (Gigerenzer, 2014, especially Ch.3).

Cues/ecological validity	Time 1	Time 3
Saturated fat (80%)	cake ? pie	cake > pie
Calories (70%)	cake > pie	cake > pie
Protein (60%)	cake > pie	cake > pie
Choice	cake	cake
Confidence	70%	80%

Table 7.1: Determining which of two foods has more cholestrol. Source: Hoffrage et al., 2000.

The authors show that in some cases, the take-the-best heuristic outperforms the other methods in terms of the number of cities successfully identified as more populous. In other contexts, Borges et al. (1999) applied the recognition heuristic to stock market choices, e.g., among two candidate stocks, pick the one that is recognized. The authors find that it outperformed several other methods of stock market investment such as mutual funds, market indices, and chance investment (or dartboard portfolios). Czerlinski et al. (1999) also find support for the take-the-best heuristic against other alternatives such as regression analysis, when the experimenter provides cues and their ecological validities. Subjects had to guess high school dropout rates in 57 Chicago public high schools. The cues, many with high ecological validities, included percentage of low income students, and average SAT scores.

Example 9 : *Consider the explanation of hindsight-bias in Hoffrage et al. (2000). A hypothetical consumer, Patricia, has to decide, at Time 1, which of two foods, a cake and a pie, has more cholesterol. She does not know the answer directly, so she uses cues to make an inference. She is provided by the experimenter with cues (saturated fat, calories, protein) and their ecological validities, stated as a percentage, just after the cue names; see Table 7.1. Ecological validity, the percentage of pairwise cases in which the cue can distinguish between the choices in the sample, can only be defined with respect to a reference class of foods. Here, the reference class selected by the experimenters is a random sample of foods from a Chicago supermarket. If a cue favours choice A over B, we write $A > B$, and if the cue cannot discriminate between the two choices, we write A ? B. For the moment, consider only Time 1. Saturated fat cannot determine if a cake has more cholesterol relative to a pie (cake ? pie), but calories can discriminate (cake $>$ pie).*

Using the steps described in the take-the-best heuristic above, at Time 1: The cue with the highest ecological validity cannot discriminate (cake ? pie), so Patricia is predicted to evaluate the next cue, calories, which successfully discriminates in favor of cake (cake $>$ pie). Hence, Patricia is predicted to choose cake. The confidence that Patricia has in her judgement is the ecological validity of the first cue that is able to discriminate, in this case *calories, so the confidence is 70%.*

Having made her choice, Patricia is now asked to make the same choice at a future time, Time 3. For whatever reason, she might have imperfect recall of her choice and the choice procedure at Time 1. It is now assumed by the authors that "...the cue values are not verdically remembered but show systematic shift towards feedback." We are not confident of the suitability of this inference and the reasoning preceding this inference; it appears to assume hindsight-bias, the very phenomenon that it wishes to explain. The authors use this assumption to construct the information in column 3 of Table 7.1. The inability of saturated fat to discriminate between cake and pie at Time 1, turns into discriminatory evidence in favour of cake at Time 3. Proceeding as above, cake is chosen with 80% confidence. This, the authors claim, explains hindsight-bias.

While this is a useful illustrative example of how take-the-best works, we are not persuaded by its ability to explain hindsight-bias.

Remark 3 *: Example 9 highlights some of the important features of the G&O approach. The decision is made in a fast and frugal manner (limited number of cues) and the procedure for choice is clearly spelt out (procedural rationality). However, the assumptions of this method, which are often unstated, need examination. Among different subjects, there is likely to be no agreement about (1) the reference class of foods that might be relevant for comparison, hence, no agreement about the precise figures for ecological validity, and (2) the relevant cues or their number.*

All this information is provided by the experimenter to the subjects, to ensure uniformity of available information. In the much more interesting case if subjects had to search for the relevant reference class and the cues, there is likely to be a separate prediction for each subject. This makes testing of the theory difficult. Furthermore, in an actual search problem, there is likely to be an infinite regress problem in search costs (Dhami, 2016, p. 1411). It might well be that some heuristics or some simple learning models are employed to search for cues; despite the stated objectives of the G&O program, limited progress has been made in this direction. Some human actions simply involve motor skills that we might have acquired over the course of human evolution, e.g., the use of the gaze heuristic to catch a ball (Dhami, 2016, Example 19.6). The more interesting set of questions for social scientists is the choice of heuristics to solve cognitive, social, economic and cultural problems. The G&O program is evaluated on these issues in Section 8.

Furthermore, we only observe Patricia's choices in Example 9, but not the mental process that she engages in. Any number of theories/mental processes might have led to the choice of cake as the relevant answer in this binary question. Thus, despite the claim to the contrary, this class of theories are 'as if' theories. As our final point, how do we accurately determine how much confidence Patricia has in her choices? Can we be sure that she is *not overconfident?*

A range of other heuristics are considered in the G&O program, such as LEX and LEXSEMI, which have been used to examine risky choices (Payne et al., 1993). Unlike the take-the-best heuristic, the LEX heuristic does not order cues by ecological validity, but by some other measure of importance. The LEXSEMI heuristic imposes a slightly more stringent requirement for cues to discriminate between choices by requiring that the difference between cue values exceeds a certain threshold (for instance, if one is comparing the outcomes in two lotteries, then the difference in outcomes must exceed a minimum threshold). The *elimination by aspects* (EBA) combines the LEX method with Herbert Simon's satisficing strategy.

Heuristics have also been used to make choices among objects with several attributes (Payne et al., 1993). Consider, for instance, the choice between different cars that differ in attributes such as reliability, fuel consumption, price, safety, and horsepower. It is likely that the desirability of various attributes may be in conflict, e.g., the car with the highest horsepower might also do worse on fuel consumption. Edwards and Tversky (1967) suggested the weighted adding strategy (WADD) in which the attributes of each car, subjectively weighted, are added and the car with the highest such score is chosen. But the subjective weights also make WADD difficult to test. One may simply add all attributes (equal weighing strategy, EQW). In both cases, the choice of attributes matters; adding or deselecting unobserved mental attributes of people changes the rankings. The LEX heuristic would advocate choosing the car with the highest attribute value on the most important attribute, say, pick the most 'reliable' car. The very fact that different people pick different cars, suggests that the ranking of these attributes in terms of importance is also subjective. But G&O do not have an underlying theory of how such a ranking might differ among individuals. Insofar as these attributes are subjective, and privately observed by the subjects, elicitation and predictions of such heuristics are vexed issues.

Simon's satisficing approach an be extended to multiple attributes as follows. Given a set of cues, one ranks the cues in some order of subjective importance and assigns each cue a satisficing level. Pick the first cue, and reject the alternatives whose cue value lies below the satisficing level. Then, move on to the next cue in order of importance and compare the surviving alternatives on their cue values relative to the aspiration level for this cue. Proceeding in this manner, if a unique alternative remains at some stage, it is chosen. If not, then some tie-breaking rule is used.

8. A critique of the G&O program

We have examined the criticisms of the KT&O approach in Sections 3 and 4. However, the G&O approach has faced much less criticism. Neoclassical economists have almost completely ignored the G&O approach.[23] Just as the KT&O approach has benefitted from sustained criticism, we believe that the G&O approach may also benefit from constructive criticism. In particular, some of the most useful criticisms that can sharpen and improve results/predictions in a program often come from people who work outside that program. This section in particular, and the rest of the paper engages in such an exercise.

8.1. What is the benchmark for comparison under risk and uncertainty

Economic theories typically do not assume that people use regression analysis to make choices. There is some use of this tool in theories of learning but it is quite limited in economics. In a leading book on economic theory (Mas-Colell et al., 1995) and in a recent analogue for behavioral economics (Dhami, 2016), the word "regression" does not even figure in the subject indices. Economists are typically interested in the general class of problems exemplified in Problem GP.

The comparison between regression analysis and heuristics is often used in G&O to demonstrate the superiority of heuristics over optimization for human decisions. Gigerenzer and Brighton (2011, p. 6) write: "Thus the scientific community would first have to rethink its routine use of multiple regression and similar techniques to facilitate accepting the idea that rational minds might not always weight but may simply tally cues".[24]

The general statistical point made in G&O is that regressions may overfit the data relative to tallying (unweighted addition of cue values) and other heuristics such as take-the-best, hence, their performance in out-of-sample predictions may be inferior. This is well recognized in econometrics (see the less is more effect below in Section 9) and is not problematic for economists. However, regression analysis is not the optimization benchmark in any area of economics that we are aware of.

When the domain under consideration is true uncertainty (unknown variables), then economics has no benchmark to offer, let alone regression analysis. Indeed, the use of regression analysis in such cases (which typically occurs in the G&O approach) is dubious because we cannot use variables in regression analysis that are unknown (see Section 11).

8.2. Empirical testing of heuristics in the G&O program

The underlying theory of heuristics in the G&O program does not predict the following elements in a heuristic choice: the choice of an appropriate heuristic from the adaptive toolbox (see Section 8.4, below), the actual selection of cues by individuals, and the reference class in which the ecological validity is to be determined. Furthermore, empirical tests have not been too successful in answering these questions.

[23] None of the core texts in any area of economics, with the exception of Dhami (2016, Part 7), give an adequate treatment of either the KT&O or the G&O approachs. This is clearly unsatisfactory.

[24] The word "weight" in this quote refers to the beta coefficients in OLS regression.

The take-the-best heuristic is possibly the most tested heuristic in the G&O program. There are two methods of introducing cues in tests of the take-the-best heuristic, both of which illustrate the enormous difficulties and pitfalls in testing the relevant theory.

1. *Inferences from givens*: Most empirical studies in the G&O program use the inferences from givens approach. In this approach, the experimenter directly gives cues to subjects and also reveals the ecological validity of the cues, thereby also choosing the reference class over which ecological validity is defined. As an experimental control, this is desirable because the experimenter can identify the information that is available to all subjects. However, it is unsatisfactory on two grounds.
   (i) It may lead subjects to the very decisions that the experimenter wishes to see by appropriate selection of cues and reference classes.
   (ii) It sidesteps the critical problem of how people search for cues. Indeed, if such subjective search for cues were employed, it might well be that there is substantial subjective variability in the cues employed by people, giving rise to different predictions for each individual. If the underlying theory is not guaranteed to have a unique prediction, then it is problematic for testing because any observed outcome can then be justified under the theory.

2. *Inferences from memory task*: This method corrects for the drawbacks of the first method, and is closer in spirit to the suggestion of Gigerenzer et al. (1999), namely, that cues should be drawn from memory. The only observable in this case is the actual decision made by a subject in the lab. As noted above, in this case there is likely to be no unique prediction of the relevant theory, and any prediction can be accommodated by some combination of cues and ecological validities.
   One potential solution, albeit an imperfect one, is to have two treatments, as in Bröder and Schiffer (2003). (1) In Treatment 1, subjects see the same information on a screen with all the relevant cues and their ecological validities. (2) In Treatment 2, subjects are narrated the relevant information, pictorially or verbally, and then they have to retrieve the relevant information from their memory. Two shortcomings of Treatment 2 remain. First, it solves the problem of cue-recall from identical information, but not the question of clue-search in the real world. Second, the experimenter only observes data on choices, so the problem of inferring the use of the take-the-best heuristic remain. We can never be sure what mental processes and solution methods were employed by the subjects. We stress these points because the G&O approach is predicated on procedural rationality.
   Consider some illustrative examples of the potential problems that arise from inferring the use of G&O heuristics based on observation of choice data. Suppose that the first cue in the take-the-best heuristic is the recognition cue (choose A over

B, if A is recognized and B is not). It could be that the recognition cue is highly correlated with other sensible methods of determining the choice between A and B. For instance, Scheibehenne and Bröder (2007) show that mere name recognition of players for the 2005 Wimbledon matches correctly predicted 70% of the matches. However, more highly seeded players, who are also more likely to win, are also more likely to be recognized. So choices between two players based on seedings of the players may be indistinguishable from choices based on the recognition heuristic. A similar problem exists in choosing among stocks based on the recognition heuristic (Borges et al., 1999). More recognized stocks are also more likely to belong to larger, more established firms, as compared to newer, transient, firms that are more likely to run into problems. So the choices made by someone who is simply more cautious are also likely to be indistinguishable from someone who uses the recognition heuristic.

In both methods described above (inferences from givens, and inferences from memory), and despite the shortcomings of the methods noted above, conformity with the take-the-best heuristic is relatively low. The temporal evolution of some of these studies is clearly described in Bröder (2011). We summarize some of these results next; for the details and references, see Bröder (2011).

1. The prediction that all subjects use the take-the-best heuristic all the time is rejected– only 5 out of 130 participants used it all the time.

2. When the percentage of subjects whose behavior best fitted with the take-the-best heuristic was used in a comparison with two other rules, Dawes rule and Franklin's rule,[25] 28% of the subjects conformed to take-the-best.

3. When the cost of acquiring cues (which were known and identical for all subjects) was added, as cues became more costly to acquire, the conformity with the take-the-best increased. Clearly if cues are expensive to acquire, subjects are more likely to make their decisions based on the first cue they encounter. However, in the real world, unless all the cues have already been searched for first, the individual cannot compare their costs.

4. When the performance of the take-the-best with costly cue values is compared with compensatory strategies (e.g., Franklin's rule), more people use the latter.

[25] Dawes' rule or tallying prescribes the following. In choosing between options A and B, consider the entire set of cues and tally the cue values in favour of each of the options. Then pick the option that has a higher number of cues in its favour. Franklin's rule prescribes a weighted combination of the cues to choose between the options.

5. When inferences from memory task are used, the predictions of 47% of the subjects are consistent with take-the-best in the verbal condition and 21% in the pictorial condition. However, as Bröder (2011), himself one of the authors of this study, notes (p. 375): "The results reported can also be accounted for by assuming people used a weighted additive strategy (e.g., Franklin's rule) that mimics take-the-best performance when the cue weights are noncompensatory..." We have already noted some of the other limitations of this study, above.

Clearly if these results are taken as direct evidence for take-the-best, it is a dubious victory.

Do subjects learn ecological validity of cues, if given a chance? Newell and Shanks (2003) gave cues to subjects and set the validity of these cues to 0.80, 0.75, 0.70, 0.69. They then gave subjects a chance to learn to order the cues by ecological validity. The authors set a learning trial period with 60 repetitions with feedback. They found that only 3 out of 16 subjects learned to order the cues in the correct order. Gigerenzer and Brighton (2011, p. 23) argue that the number of trials were too short and suggest that subjects would need at least 100 trials to learn to order all the cues correctly. This is an important limitation. In the real world, in many cases with significant consequences, opportunities to learn which cues to use might be limited. Consider, for instance, decisions such as marriage, divorce, number of children to have, choice of pension plan, choice of savings plans, choice of donating organs, and choice of a consumer durable. For firms, infrequent choices include merger decisions and capital restructuring decisions. It might well be that, under such circumstances, alternatives to the heuristics proposed by G&O may be used; examples include social norms, and public policy in the form of nudges.

Newell et al. (2002) report that a majority of their participants used some variant of frugal (in terms of the extent of information used) strategies. Using the criterion that subjects conform to all the features of the take-the-best (TTB) heuristic at least 90% of the time, they find that the behavior of only 33% of the subjects conforms to the criterion. In their commentary on these results on p. 381, the editors, Gigerenzer, Hertwig and Pachur (2011), object to the high figure of 90% used. The editors also criticize the methodology of testing single theories. They argue (p.381) in favour of running a horse race between models. However, when such competitive tests are run to show that TTB does better, the candidates in the horse race are not necessarily persuasive. For instance, Bergert and Nosofsky (2011) show that TTB does better than RAT (a weighted additive model). RAT works as follows: Suppose that a model has n features, or cues. Then assign weight w_i to the i^{th} cue such that $w_i = \log\left(\frac{v_i}{1-v_i}\right)$, where v_i is the ecological validity of the i^{th} cue. RAT requires choosing the model that maximizes $\sum_{i=1}^{n} w_i$. In order to get a unique prediction, this method requires common agreement about the validity of cues and a common set of cues across the subjects. We are not aware of mechanisms that would satisfy such stringent conditions.

Many of the problems employed in testing the take-the-best heuristic are relatively simple. For instance: Which city is more populous? Which of the players is likely to win a tennis match?[26] However, it is less clear which heuristics are used and how cues are generated to solve concrete economic problems such as the design of contracts, choice of savings and pension plans, choice of a mortgage, and the decision to donate organs. We do not have direct evidence of how individuals might search for cues in these cases, or how they might discover the ecological validity of these cues, or how they might even reach a common agreement about such things. In the absence of such an agreement, we are likely to have individual-specific predictions of take-the-best and other heuristics in the G&O program that pose immense problems for testing these theories.

8.3. Training people in using statistics

Gigerenzer (2008, p. 16-18) has advocated better training for people in statistical inference. Indeed, he has personally engaged in training medical doctors in the use of Bayes' rule in a frequency format. Building on this idea, some people have argued in favor of "boosts," understood as efforts to enable people to exercise their own agency. Greater statistical literacy is a boost. We are certainly in favor of statistical literacy, and we agree that it is important to improve people's ability to exercise agency. We will return to the question whether a prime minister, a president, or a national legislature could make significant progress on national priorities through training people to use statistics. For now, we focus on narrower questions.

Gigerenzer (2008) reports that doctors and children do better in statistical inference when trained and the learning effects of a frequency format are longer lasting than a probability format. While one would expect human beings to get better in statistical inference, the more they are trained, there are several concerns.

1. Training people in statistics is costly. Economists would ask for a computation of the opportunity cost before justifying the policy on cost-benefit grounds. For instance, could the money be better spent by (1) providing public information, and (2) by using regulation to require more transparent selling of financial products? For instance, patients using the National Health Service (NHS) in the UK are often provided direct information on the likely increase in life expectancy from taking various drugs, such as statins or blood thinning drugs. Governments require financial firms to directly provide information to consumers on the annual percentage rates (APR), an all-inclusive figure, to compare various financial products. By not comparing the costs

[26] See Section 11 for a more complex problem: Which stocks to invest in?

and benefits of a program of statistical training of people against other alternatives, it is difficult to judge its efficacy.

2. Programs for imparting statistical training to doctors teach them to compute conditional probabilities such as $P(A|B)$, where, for instance, A is the event that the patient has cancer and B is a positive result on a mammography test. In contrast, doctors are often interested in a range of factors that determine event A, hence, the appropriate medical diagnosis depends on more complex conditional probabilities of the form $P(A|B, C, D, E, \ldots)$. For instance: A : death from cardiac disease in the next 20 years; B : reading on the cholesterol test; C : did any of the grandparents die of a cardiac disease?; D : ethnic background; E : does the subject drink alcohol or smoke?
   The actual computation of $P(A|B, C, D, E, \ldots)$, is a cognitively and computationally challenging problem, even in the frequency format. As a result, an alternative, which appears to be in place in the NHS, enables doctors to compute $P(A|B, C, D, E, \ldots)$ directly on their computers, using a pre-designed software. Indeed, it is common for NHS patients tested for cholesterol (B) to be asked further questions $(C, D, E, \ldots)$ before the doctor presses a button on the computer to tell them the conditional probability that they might die in the next 10 years, with and without the use of statins.
   Arguably, of far greater importance is the issue of what doctors do with the number $P(A|B, C, D, E, \ldots)$ once they find it. If this turns out to be, say, 0.75, should they prescribe statins or not. Or if the patient is suspected of cancer, should they recommend chemotherapy or not. In the UK, this problem is solved by setting a threshold, so that if $P(A|B, C, D, E, \ldots)$ exceeds the threshold, the doctors undertake the treatment. The cutoff is apparently determined using the best available medical statistics and medical know-how. Insofar as medical statistics are not good enough, this affects the decision irrespective of whether one has used a frequency format or a probability format to compute the relevant conditional probability. In both cases, one would benefit from an improvement in medical statistics.

8.4. Does the G&O program tell us which heuristic to use?

In the G&O program, people draw upon an adaptive toolbox of heuristics and pick a heuristic that is appropriate to the environment (Gigerenzer et al., 1999; Gigerenzer and Selten, 2001). Which heuristic will be employed from the adaptive toolbox of heuristics, for a given problem? Gigerenzer (2008, p. 38, 39) and Gigerenzer and Brighton (2011) give identical answers. Both refer to a single paper by Rieskamp and Otto (2006). However, this paper provides a limited and unsatisfactory framework to answer one of the most fundamental questions in this literature. In our view, this remains an open question.

Rieskamp and Otto use a model of reinforcement learning in which people learn to use only one of two models. A take-the-best model and a WADD model (see Section 7 for description). The set of cues and their ecological validities is provided to the subjects; we have already noted the criticisms of experimenter provided cues and validities. Since the reinforcement learning model is an adaptive model, no other strategy than the ones initially considered can ever emerge from the model. The question that subjects have to answer is: Which of two companies is more creditworthy? The cues include financial flexibility, efficiency, capital structure, and own financial resources.

No underlying model is provided that explains how these factors translate into greater creditworthiness–certainly economics does not provide such a model. Thus, the particular sample of companies that is used has its own particular ecological validity of cues, and one cannot rule out the role of other cues/factors in determining creditworthiness.

Within this setup, take-the-best proves to be superior to WADD and gets better as more feedback is provided. It is now well known that the reinforcement learning model is dominated by a range of learning models and it provides misleading results on the speed of learning (Dhami, 2016, Part 5). For instance, Chmura et al. (2012) show that the reinforcement learning model gives the worst performance among the set of learning models that they consider.[27]

In a nutshell, we currently know very little about how heuristics are drawn from the adaptive toolbox.

9. Evaluation of the 'less is more' effect

Gigerenzer (2008) writes: "That simple heuristics can be more accurate than complex procedures is one of the major discoveries of the last decades." This is also often described as the *less is more effect*. The source of this finding is that in *prediction tasks*, there is a trade-off between unbiasedness and variance minimization, a well-known result in statistics (Gigerenzer, 2008, Gigerenzer and Brighton, 2011).

9.1. The less is more effect

To derive the less is more effect, suppose that the underlying relationship is give by

$$y = f(x) + \varepsilon, \tag{9.1}$$

where y is the dependent variable, x is the independent variable, and ε is a noise term. One may think of x, y, ε as vectors in the more general case, but this does not alter the

[27] Incidentally, Rienhard Selten, one of the authors of the idea that heuristics are drawn from the adaptive toolbox is a co-author on this study.

arguments. The data is given by pairs $(x,y) \in S$, where S is the sample. The true underlying function f is unobservable because of the noise term, ε. Suppose that we have some estimator $\widehat{f}$ of f. In the special case of linear regression analysis, we have $\widehat{f}(x) = a + bx$, where a, b are regression coefficients that are estimated from the sample S, by minimizing mean squared errors.

Suppose now that we are given some data $(x,y) \in S_1$ from another sample, S_1. We wish to find the expected mean square prediction error $E_{S_1}\left(y - \widehat{f}(x)\right)^2$, where E_{S_1} is the expected value with respect to the sample S_1. However, we wish our estimator to be robust to all possible samples, not just the sample S_1. To take account of this, suppose the set of all possible samples is given by some universal set of samples, U. Then, we wish to choose $\widehat{f}(x)$ so as to minimize $E\left(y - \widehat{f}(x)\right)^2$, where E is the expectation operator with respect to all possible samples U. The noise term ε is independently and identically distributed with $E[\varepsilon] = 0$ and $V[\varepsilon] = \sigma^2$, where V is the variance operator with respect to U.

Proposition 1 *: The expected prediction error is given by*

$$E\left(y - \widehat{f}(x)\right)^2 = Bias\left(\widehat{f}(x)\right)^2 + V\left(\widehat{f}(x)\right) + \sigma^2, \tag{9.2}$$

where $\sigma^2 = E\left[(y - f(x))^2\right]$, $Bias\left(\widehat{f}(x)\right) = f(x) - E\left[\widehat{f}(x)\right]$.

Proof of Proposition 1: In the steps below, we use (9.1), $V[z] = E[z^2] + E[z]^2$, and $E[y] = f(x)$, $V[y] = \sigma^2$.

$E\left(y - \widehat{f}(x)\right)^2 =$

$E\left[y^2 + \left(\widehat{f}(x)\right)^2 - 2y\widehat{f}(x)\right] = E[y^2] + E\left[\left(\widehat{f}(x)\right)^2\right] - 2E\left[y\widehat{f}(x)\right]$

$=V[y] + E[y]^2 + V\left[\widehat{f}(x)\right] + E\left[\widehat{f}(x)\right]^2 - 2E\left[(f(x) + \varepsilon)\widehat{f}(x)\right]$

$= V[y] + V\left[\widehat{f}(x)\right] + \left(f(x)^2 + E\left[\widehat{f}(x)\right]^2 - 2f(x)E\left[\widehat{f}(x)\right]\right)$

$= \left(f(x) - E\left[\widehat{f}(x)\right]\right)^2 + V\left[\widehat{f}(x)\right] + \sigma^2$. ■

If one is interested in prediction errors, as in machine learning, then the importance of Proposition 1 is that algorithms that are designed to minimize prediction errors must pay joint attention to the bias and the variance of the estimator function.

Consider the following example. Suppose that we have a set of data points, and we wish to fit a polynomial through these points. This polynomial is the relevant estimator of some true underlying function that describes the sample data in a noisy manner. The higher is the degree of the fitted polynomial, the better is the in-sample fit, i.e., more unbiased is the estimator within the sample. However, the sample dispersion is partly determined by random noise. Fitting the in-sample noise-inclusive data too closely by a higher order polynomial, reduces the ability to fit out-of sample data in a prediction task; such polynomials will exhibit higher variance in fitting data out-of-sample. In other words, while higher degree polynomials will have lower bias, they might have significantly higher variance; on net, the RHS of (9.2) may turn out to be higher for higher degree polynomials, beyond a certain degree. In this case, a lower degree polynomial that has higher bias in-sample but lower variance out-of-sample could make lower prediction errors; this is the essence of the less is more effect.

Proposition 1 lies at the heart of the less is more critique of G&O. The bias-variance tradeoff is statistically sound, but the relevant issue is whether it is useful to study human cognition. Gigerenzer and Brighton (2011, p. 12) write: "Our cognitive systems are confronted with the bias-variance dilemma whenever they attempt to make inferences about the world." However, we are not aware of any direct incontrovertible evidence in support of this assertion.[28] In the absence of such an assertion, it is a leap of faith to assume that the human mind is so hardwired as to develop and use algorithms that consider the bias-variance tradeoff. By implication, we also have no direct observations that the less is more effect is central to human cognition and decision making. We do not deny the immense importance of Proposition 1 in other areas such as machine learning and in public policy.

9.2. A critique of the bias-variance tradeoff

We now examine other objections to the use of the bias-variance dilemma.

1. Loss aversion and the objective function: The objective function in Proposition 1 gives equal weights to positive and negative forecast errors of the same magnitude. In contrast, loss aversion is now well-established as an empirical phenomena, not just in the domain of choice under certainty, risky, and uncertainty, but also in the domain of temporal choice where losses are discounted less than gains (Dhami, 2016, Parts 1 and 3). Thus, it is plausible to conjecture that when engaging in prediction tasks, people will give different weights to negative and positive prediction errors; one would expect negative forecast errors to be more salient. In this case, there is no presumption that Proposition 1 will hold.

2. Suitability of the error structure: The error structure underlying the derivation of Proposition 1 assumes that $E[\varepsilon] = 0$ and $V[\varepsilon] = \sigma^2$. In contrast, we now know from a great deal of evidence from behavioral economics, and in many applications in behavioral finance, that people make systematic prediction forecasting errors, not

[28] The two sources cited in the previous quote as supporting evidence, Griffiths and Tenebaum (2006) and Oaksford and Chater (1998) offer no direct evidence that the cognitive system considers the tradeoff between bias and variance.

random ones. For an extended period of time, such forecast errors can cumulate in the same direction which leads to systematic divergence of stock market prices from the fundamental values.

3. Mental models: In actual practice, there might be a range of other objectives than simple minimization of prediction errors that is reflected in Proposition 1. There is a growing literature on how people form *mental models* to understand complex reality in a manner that is unlikely to be the outcome of a bias-variance tradeoff. For instance, women in areas of India were found to give less water to children with diarrhoea, based on the mistaken analogy of a leaky bucket (WDR, 2015). This mental model, a heuristic in the fast and frugal category, has disastrous consequences; the problem is easily treated with a low cost solution of salt and sugar. For other examples of mental models and models of coarse categorization that illustrate the general point made above, see Dhami (2016, Section 19.12, 19.13).

4. Non-stationary environments and behavioral factors: Unlike the assumption made in Proposition 1 of a stationary (though stochastic) underlying environment, the real world environment typically lacks stationary. In many cases, people do not know the true underlying model, but they try to form inferences about the correctness of competing models by examining the data. For instance, does stock market data come from a financial model that is mean-reverting, or one that is a trending model? What if investors believe that the world switches between these two models depending on some underlying Markov transition process? Barberis et al. (1998) consider investor inference in such a scenario. This can explain stock market underreaction and overreaction relative to the rational benchmark. In other words, economic models require individuals to form beliefs of the relevant economic model and then, update their beliefs by using observed data. It is possible they never learn the true model and keep updating in the wrong direction (Dhami, 2016, Section 15.9). Furthermore, confirmation-bias, hindsight-bias, and overconfidence may furnish other powerful reasons why they cannot successfully learn the underlying models; see Dhami (2016, Section 19.8) for a formal model. However, all these critical considerations are missing from a purely statistical and mechanical exercise of minimizing prediction errors in (9.2).

5. Most of the tasks that economists are interested in are not *prediction tasks*; rather, these are *decision tasks* that require an appropriate decision theory, such as prospect theory, or hyperbolic discounting. For instance, how much should a tax evader evade, given the probabilities of detection and fines? How much should one invest in a risky asset relative to a safe asset, given that the distribution of returns of the risky asset

is known objectively/subjectively? As a taxi driver who chooses the hours of work, how many hours should the taxi driver work, conditional on a given wage? Should one buy a gym membership on a pay-as-you-go basis or on a fixed-fee basis? Should I introduce a new product before my competitor firm, given that I have access to reliable data on demand and prices? For concrete and empirically supported answers to these questions from a behavioral economics perspective that is independent of the bias-variance dilemma, see Dhami (2016).

9.3. Implications of the less is more effect for the KT&O program

An important criticism of KT&O by G&O is that they do not take account of the second term on the RHS of (9.2), i.e., variance, so a demonstration of biased inference in KT&O gives an incomplete/misleading picture. However, there are several weaknesses with this argument.

KT&O were interested in potential biases that decision makers may exhibit relative to the predictions of the BRA framework. It is important to realize that these biases are not necessarily statistical biases in the sense that they are used in (9.2). For instance, an important element of the BRA is the assumption that people have complete and transitive preferences. A demonstration of non-transitivity, as in Lichtenstein and Slovic (1971), is a bias relative to the BRA, but not in a statistical sense.

Most heuristics in the KT&O program do not use the bias-variance tradeoff as their justification, and as noted above, there is strong empirical support for these heuristics. Rather, they appear to reduce the cognitive dissonance that would arise from an imperfect understanding of the world and from being shown up as incompetent in some domain (see Section 5.7, above). Confirmation-bias and hindsight-bias are not about predictive judgements but postdictive judgements, so they are backward looking. Anchoring is about a previously given anchor; availability is about predicting on the basis of available information from memory; the representativeness heuristics is about inferring how likely a small population is to have come from some parent population (this includes the gambler's fallacy and the hot hands fallacy); not recognizing regression to the mean and the conjunction fallacy are also not about minimizing predictive error either. Thus, the existence of an entire body of empirically successful, fast and frugal, ecologically rational, heuristics that does not rely on the bias-variance dilemma, suggests that its importance is overemphasized in the G&O program.

10. On mathematical optimization and 'as if' theories.

G&O have criticized economic theories, particularly prospect theory, on the grounds that theories based on optimization are 'as if' theories (Gigerenzer, 2008). In other words, these theories do not explicitly model the mental processes that lie behind actual decision making and assume that humans behave ‘as if’ they followed these theories. The G&O position is summarized in the following quotation from Gigerenzer (2016, p. 38): "Although behavioral economists started out with the promise of greater psychological realism, most have surrendered to the as-if methodology. Cumulative prospect theory, inequity-aversion theory, hyperbolic discounting are all as-if theories. They retain the expected utility framework and merely add free parameters with psychological labels... which is like adding more Ptolemaic epicycles in astronomy. The resulting theories are more unrealistic than the expected utility theories they are intended to improve on. Behavioral economics has largely become a repair program from expected utility maximization."

Optimization is simply a tool used both by neoclassical economics and behavioral economics. But behavioral economics and its approach, methodology, results, and conclusions are very different from the canonical approach taken in neoclassical economics.[29] Optimization and as-if theories are not bad or undesirable per-se, unless they are rejected after stringent empirical testing. In physics, the sun and the earth are often modelled as points (e.g., in solving the two-body and three-bodies problems), which is as-if approach; however, some of these theories are remarkably accurate. As-if optimization based models are untenable if they conflict with the evidence. However, the success rates of behavioral theories such as cumulative prospect theory, inequity-aversion theory, and hyperbolic discounting, are impressive and overall, in their respective domains, arguably there do not exist alternatives that better explain the totality of evidence. If there are better alternatives, they should be embraced.

Of course it is appropriate to criticize as-if optimization based models if their predictions are inconsistent with the relevant evidence. Indeed, such a criticism is the starting point of behavioral economics. It has shown, conclusively, we believe, that several standard theories that are central to optimization based models in economics, such as expected utility, exponential discounted utility, self-regarding preferences, and several refinements of Nash equilibrium, are inconsistent with the evidence (Dhami, 2016).

Behavioral economics is certainly not a rescue project for expected utility. A great deal of evidence gathering effort has gone into behavioral economics to show that expected utility is untenable. Prospect theory is an ‘as if’ theory but overall, one that does best in situations of risk/uncertainty/ambiguity, as understood in economics (Dhami, 2016, Part 1). It is a leading example of a tenable ‘as if’ theory in behavioral economics. To be sure, there are other models, including models of heuristics, such as the priority heuristic, that

[29] Here is an extreme analogy: Just like the classical Greeks recognized the smallest indivisible unit of matter that we now call atoms, so does modern science. This should not be taken to mean that modern science and classical Greek approaches to understanding the physical world have any fundamental similarity or that one is a repair program for the other; after all the Greek approach was not an experimental science at all, whereas this is the hallmark of all modern science.

can explain some of the puzzles that prospect theory explains (Brandstätter et al., 2006). However, the priority heuristic has been criticized on the grounds that its predictions do not hold outside the dataset used in the paper and that its use of statistical techniques is questionable (Birnbaum, 2008). As of now, the range of phenomena explained by prospect theory is too vast to be accounted by any other theory.[30]

Recent theoretical work also suggests that there is a link between heuristics based approaches in time discounting and optimization models (Dhami, 2016, Section 10.4, page 627). The relevant heuristics in this case are non-compensatory and lexicographic.[31] Yet choice behavior in the presence of these heuristics can be shown to be equivalent to an optimization based delay-discounting model (al-Nowaihi and Dhami, 2018). This suggests the two approaches might be closer than one imagines them to be, although these issues have been insufficiently explored. Similarly, heuristics have been applied in dealing with strategic situations to provide an alternative to a Nash equilibrium in static games (al-Nowaihi and Dhami, 2015). This suggests that behavioral economics is not averse to the absence of optimization and does not necessarily ignore procedural rationality.

G&O have also expressed strong reservation to optimization based theories on the grounds that they have free parameters (see quote above). But there are statistical tests that enable us to determine whether these parameters add enough in the way of explanatory power that justifies their use or not (e.g., Akaike information criteria and Bayesian information criteria). G&O require that the parameters so estimated in these models should be held fixed in every context and frame in order to make predictions.[32] For instance, they argue that the parameter of loss aversion should be held fixed in all cases. This is unnecessarily stringent and not supported by the evidence. We do know that human behavior is context and frame dependent and examples abound in all areas of behavioral economics (Dhami, 2016). Behavior also depends on age, culture, gender, moods, and emotions. For instance, when disgust is induced, or when individuals have just experienced a prior loss, measured loss aversion increases. These are predictions that can be tested (and have been successfully tested), and allow one to take account of the richness of human behavior. We see nothing objectionable in this approach.

[30] For instance, it is not immediately clear how any current heuristic can account for some of the following behaviors that are easily explained under prospect theory (Dhami, 2016, Part 1): What explains the equity premium puzzle? Why do drivers quite too early on a rainy day in New York? Why are incentive contracts low powered? Why do firms exist? What accounts for the endowment effect? Why is human behavior so sensitive to goals? Why is skewness in returns to assets priced? What explains the Ellsberg paradox?

[31] Compensatory heuristics are those that give equal weight to the cues. Non-compensatory heuristics allow for unequal weights to the heuristics.

[32] I am grateful to Gerd Gigerenzer for highlighting this point in conversation.

11. On distinct domains of choices in KT&O and G&O.

As noted above, economics makes no predictions under true uncertainty, but individuals do make choices under true uncertainty.[33] How do they make such choices? This question lies at the frontiers of social science, and relatively little is firmly established. We speculate below on some of the possibilities, and in doing so, also argue that a great deal of the work of G&O and KT&O is best seen as operating in different domains.

1. Overconfidence: Even under true uncertainty, overconfident individuals may believe, incorrectly, that they can see through all the possibilities and assign subjective probabilities to all these possibilities. One can then use any of the standard decision making models such as prospect theory (where special cases include expected utility and rank dependent utility), or some heuristics based model based on the insights of, say, the priority heuristic.

2. Conservation of cognitive effort: Individuals may not believe that they can see through all the possible outcomes and probabilities, yet exerting cognitive and search costs to explore possible unknowns and their probabilities, is too costly. Individuals may then simply assign a vanishingly low probability to the unknowns and choose to focus on the known outcomes only and choose any of the standard decision making models as in 1. For an analogue of such sort of thinking under risk, where people simply ignore events of very low probabilities and optimize over the rest, see Dhami (2016, Section 2.11).

3. Norms: A tantalizing possibility is that individual behavior under true uncertainty might be determined by social or personal norms. Consider, for instance, marriage, arguably a problem in the domain of true uncertainty. One may decide to have a trial live-in relationship, and make the marriage decision based on an evaluation of the live-in period. This is a norm in some societies but not others. In traditional Indian arranged marriages, historically several social norms appeared to be at work in making this decision.[34]

4. Heuristics: Individuals may realize that they just do not have enough data to make any optimization based decisions, so they may resort to heuristics to make the choice.

There is no presumption that our four speculative categories are non-intersecting. For instance, heuristics may themselves be determined by social norms, which might in turn have been optimized to conserve cognitive effort in order, partly to prevent overconfident individuals from making choices that are suboptimal. When we speak of true uncertainty, below, we only consider the very last potential explanation of behavior, i.e., a heuristic-

[33] For readers steeped in the modern literature on ambiguity we note again that ambiguity is a special case of true uncertainty; the latter is much more general.

[34] These included checks of various proxies of desirable behavior. Proxies might include observations on whether cattle and other animals in the family were well looked after (possibly signaling caring and responsibility attitudes); whether the family was well regarded within their village; and whether the family was of a similar socio-economic background. In many societies, the social norm is to choose from within a particular social or ethnic group.

driven account of true uncertainty, as if it were a stand-alone explanation.

Our main claim in this section is that the domain of decisions in the KT&O and the G&O accounts of heuristics is, in many cases, non-intersecting. However, in some cases, the domains intersect, as in the use of the priority heuristic to deal with risk, and the efficacy of the frequency format relative to the probability format in reducing biases relative to the BRA. We have already dealt with these issues elsewhere in this essay. For this reason, here, we focus exclusively on the difference in domains between the two approaches.

The KT&O demonstration of heuristics and biases mainly considers the domains of *risk and uncertainty alone*. In several of their most important demonstration of biases, they explicitly provide data on outcomes and probabilities to subjects and in other cases data on probabilities an outcomes is not needed. This includes their demonstration of the law of small numbers (which includes the gambler's fallacy and hot hands fallacy); base rate neglect and violation of Bayes' Law; Conservatism or underweighting of the likelihood of a sample; hindsight-bias; confirmation-bias; regression to the mean; false-consensus effect; conjunction fallacy; and, confusion between necessary and sufficient conditions.

In other cases used in KT&O, the distinction between risk/uncertainty and true uncertainty is less clear; this includes the affect heuristic and the anchoring heuristic. To see this, consider the experiments reported in Kahneman (2011). Subjects were asked the following two questions:

1. Is the height of the tallest redwood more or less than x_0 feet?
2. What is your best guess about the height of the tallest redwood?

Two different anchor values of x_0 ($x_0^H = 1200ft$, $x_0^L = 180ft$) were used. Subjects in the high anchor condition (x_0^H) guessed the height to be substantially greater, thereby confirming the anchoring effect. One might wonder if in this case, subjects who did not know the answer faced true uncertainty, i.e., they might not have known the possible heights and the associated probabilities for redwood trees.

However, several other demonstrations of the affect heuristic and the anchoring heuristic clearly fall within the domain of risk/uncertainty, yet one still observes the relevant biases. Here are some examples. The affect heuristic was demonstrated with members of the British toxicology Association who were aware of the relevant risks (Slovic et al., 1999). The anchoring heuristic was demonstrated with property experts who were given all relevant information to evaluate the relevant properties (Northcraft and Neale, 1987); poetry reading sessions where self-evaluation was used also exhibited anchoring (Ariely et al., 2006); and experienced trial judges exhibited anchoring in legal judgment (Englich and Mussweiler, 2001).

One can, therefore, safely conclude that the relevant domain that KT&O used to demonstrate their biases was either *certainty, risk, or subjective uncertainty* but not *true uncertainty*. We argue below that the domain of problems that G&O deal with typically lies within the domain of *true uncertainty*, which they describe as *large worlds*. This observation on the G&O program is important not only because we need to understand human decision making in such environments, but also to note that many of the contributions of G&O and KT&O lie in non-intersecting domains. In light of this, the characterization of their positions as adversarial, as has been the case in the *great rationality debate*, can be misguided and misleading.

The research agenda in G&O speaks directly to problems of true uncertainty. They distinguish between *small worlds* and *large worlds*, drawing on terminology originally introduced by Savage (1954). Small worlds corresponds to our use of the terms 'risk' and 'subjective uncertainty' and large worlds corresponds to our use of the term 'true uncertainty.' In introducing their research agenda, and summarizing their approach Gigerenzer et al. (2011, p. xviii) write: "How should we make decisions in the large world–that is, when Bayesian theory or similar optimization methods are out of reach?...In sum, the accuracy-effort trade-off indeed exists in a small world. In a large world, however, where there is no general accuracy-effort tradeoff, heuristics can be more accurate than methods that use more information and computation, including optimization methods. Thus, one reason why people rely on heuristics is that these allow for decisions in both small and large worlds." In this quote the greater accuracy of heuristics relative to other methods that rely on more computation/information is predicated on the *less is more effect* that we have already analyzed above.

A good example of this approach for economists is in financial market investment. It is probably impossible to predict all possible outcomes and probabilities in stock markets, so it is, arguably, a problem in true uncertainty. The standard assumption in financial economics is to treat this as a problem in risk and uncertainty, yet even in this case there may be no unique prediction (see Example 11 below). Candidate optimization benchmarks that are typically used in the G&O program, to evaluate their heuristics against, such as logistic regression require information that is simply not available under true uncertainty. But people do invest in the stock market. How do they do it? We draw on two examples that are frequently cited in the G&O program as examples of successful heuristics (Gigerenzer et al., 2011).

Example 10 *: Borges et al. (1999) study the performance of the recognition heuristic in choosing among stocks (i.e., among two stocks, choose the one that can be recognized) against the performance of other mutual funds. In 6 out of 8 tests, the performance of the recognition heuristic was better than the performance of selected mutual funds. Economists would need more data and tests to be persuaded by the results; the data was over 1 year in the midst of a strong bull market that might have favoured stocks that are more easily recognized. More tests would need to be done to disentangle the predictions* *of the recognition heuristic in this case from other predictions of models in finance.*

Example 11 : *Consider the following simple investment strategy. In a portfolio of N assets, invest a share $\frac{1}{N}$ in each asset. This is sometimes known as the $\frac{1}{N}$ heuristic. Treating the problem of financial investment as one of risk and uncertainty (but not true uncertainty), DiMiguel et al. (2009) compare the performance of the $\frac{1}{N}$ heuristic with 14 other optimal investment theories often recommended in finance.*[35] *They use calibrated values and compare the performance of the alternative methods (based on Sharpe ratios) over a long simulated length of time. Optimization based investment theories require the estimation of statistical moments of the underlying statistical distributions, e.g., the mean and the variance-covariance matrices; but these are estimated with error. In contrast, there are no measurement errors involved with the use of the $\frac{1}{N}$ heuristic.*
Thus, the tradeoff is this: the $\frac{1}{N}$ heuristic is not sophisticated enough to take advantage of potentially greater profit opportunities that come from a more nuanced distribution of the portfolio, while the use of optimization based methods involve measurement errors that are reflected in lower ex-post profits. The authors then show that the $\frac{1}{N}$ heuristic outperforms all the optimal investment models; this also illustrates that more information and computation do not necessarily produce better results. However, the transmission mechanism based on measurements errors in the optimization based methods but not in the $\frac{1}{N}$ heuristic is different from the one invoked in the less is more effect in Section 9.1.
We offer three main comments on this work. (1) No evidence is given that people (as distinct from investment firms) do indeed use any of the optimal investment strategies described in the paper. (2) The $\frac{1}{N}$ heuristic does not belong to the class of take-the-best heuristics in the G&O approach (e.g., no cues are used at all). (3) The problem of investing in financial markets is arguably a problem in true uncertainty. For these reasons, DiMiguel et al. (2009) address an entirely different class of issues: Namely, that under true uncertainty, the $\frac{1}{N}$ heuristic performs better than the optimization benchmarks that were developed under risk and subjective uncertainty.

As our final point in this section, it is a tantalizing possibility that people may employ the heuristics in the KT&O program when dealing with true uncertainty, although they were not explicitly designed for this purpose. Suppose one faces a problem in true uncertainty. Then one may actually be influenced by an externally provided anchor or suggestion (anchoring heuristic); or by the availability of apparently similar information in the past (availability heuristic); or by the emotional tags association with the given information (affect heuristic). It might also be the case that Herbert Simon's aspiration adaptation theory may be used in such cases. We believe that it is worth exploring this

[35] This work is reprinted as Chapter 34 in Gigerenzer et al. (2011).

possibility in future research. This would provide a different account of decision making under true uncertainty, as compared to the G&O program.

12. Nudges and nudging

In neoclassical economics, observed choices reflect the underlying preferences of individuals. In this standard view, people's choices are assumed to promote their welfare. The reason is that such choices represent sufficiently careful and considered judgements about the welfare effects of actions; the preferences underlying these choices are sometimes termed as *normative preferences*. At the same time, individuals are often believed to be self-regarding, which gives rise to a role for public policy to internalize the externalities caused by self-regarding actions. Examples of relevant policy interventions include Pigouvian taxes to regulate the emission of pollution by firms, subsidies to consumers to buy electric cars, and tax-exempt charitable contributions. Regulation (such as fuel economy standards) might also be used to internalize externalities, though it is a crude response.

Do individuals really know what is in their best (including the long-term) interests all the time? Are they able to make choices that are in their long term interests? Consider the following empirically documented examples, F1–F6, of human behavior that by no means exhaust the scope of our argument.

F1. Individuals may overconsume some goods (sugar, saturated fat, alcohol, tobacco) and suffer ill health, a shorter lifespan, and a lower quality of life.

F2. Individuals may procrastinate too much, fail to complete their projects in time, fail to consolidate their finances in a timely manner that might be more conducive to their long term interests, not enroll in suitable pension plans, and take up annual gym memberships when a pay-as-you go membership could have saved them money.

F3. Individuals might under-save for their retirement (the data indicates a sharp drop in consumption at retirement (Bernheim et al., 1997)); or retire too soon.

F4. Individuals might make marriage and divorce decisions too hastily. They might also make purchase decisions on impulse, in a hot state, and regret it afterwards.

F5. Individuals might be misled by manipulative or deceptive, but not untruthful, advertising by firms that cleverly frames the options (Akerlof and Shiller, 2016).

F6. Individuals might not purchase fuel-efficient vehicles or energy-efficient appliances. Focussing on the short-term, they save money in the present but lose money over a period of years, and with any reasonable discount rate, they are likely to be making an unjustifiable choice.

Under normative preferences, we should not observe these phenomena. For example, individuals rationally choose to consume excessively unhealthy foods; never procrastinate; perfectly smooth consumption over their lifetime; choose to marry, divorce, and make their purchases rationally; and make frame-invariant choices, provided the frames are informationally equivalent.

We do not deny that in cases that appear to show mistakes (as in F1–F6), it is possible that some people, or most people, are doing exactly what they should do. We agree that third parties, including public officials, should hesitate before concluding that people have erred. But in many cases, people's decisions are not in their interests. In actual practice, a combination of public policy and self-imposed control mechanisms might be used to remedy the problems described in 1-6, above. Such remedies may take the form of (and in order): (1) Corrective taxes (public policy), and individuals who enlist private organizations (such as Alcoholic's Anonymous) or voluntarily check into rehab at a substantial financial cost (self-control mechanisms). (2) Internally or externally imposed deadlines to undertake the appropriate action such as to fill one's tax forms, and submit an assignment by a due date. (3) SMarT ('save more tomorrow') savings plan that commits individuals to save a fraction of future increases in their incomes (Benartzi and Thaler, 2004). (4) Cooling-off periods for divorce and a 21 day return policy for impulse purchases. (5) Regulation requiring financial firms to state the annual percentage interest rates (APR) on their financial products, and clear, color-coded declaration of the key nutritional facts of food products, such as calories, saturated fats, and salt content.(6) Clear labels, so that people can see, and appreciate, the potential economic and social benefits of fuel-efficient automobiles and energy-efficient appliances.

F1–F6 are easily explained within a behavioral economics framework. These and several other problems can arise from a range of judgement heuristics and biases (Thaler and Sunstein, 2009; Dhami, 2016). Many individuals report that they would prefer to save more, or prefer the median portfolio to their own, and others pay to kick a bad habit, indicating that they recognize their choices to be suboptimal (Sunstein and Thaler, 2003; p. 1169). Many individuals exhibit self-control problems as a result of *present-biased preferences*; this typically leads to temporally suboptimal outcomes such as inadequate savings for retirement, obesity, procrastination, and drug use (Dhami, 2016, Part 3). US data shows that many individuals do not utilize the full limits of their 401(k) pension plans. People also diversify their portfolios inadequately, and often use the $\frac{1}{N}$ heuristic in forming portfolio. This skews their pension portfolios (e.g., they might end of holding too much equity of their own company). Coupled with overconfidence in their own companies, they might hold too much company equity, with possibly adverse consequences (e.g., Enron employees). Despite the steep interest rates on payday loans, these are used excessively.

Individuals may suffer from limited attention (the BRA assumes unlimited attention); thus, they react more to salient taxes, and in the used car market there are discontinuous drops in the prices of used cars at 10K thresholds (Dhami, 2016, Section 19.17). Emotions play an important role in individual economic decisions, such as buying consumer durables.

Often these decisions are taken in a hot state and regretted in a cold state (Dhami, 2016, Part 6); admittedly, the normative issues are especially complicated in some of these cases. Individuals might also be heavily influenced by the framing of the same problem in informationally equivalent ways. Indeed, framing plays an important role in the advertising industry, which spends more than 500 billion per year. The frame and context dependence of preferences has been actively studied in behavioral economics (Kahneman and Tversky, 2000; Dhami, 2016).

In light of these departures from the assumptions of neoclassical economics, what welfare or normative significance should one attach to individual choices, and how must we modify classical welfare analysis?[36]

Issues of paternalism in behavioral economics are often captured under the umbrella term: *soft paternalism*. This includes *libertarian paternalism* (Thaler and Sunstein, 2003, 2009; Sunstein and Thaler, 2003), *asymmetric paternalism* (Camerer et al., 2003), and *light paternalism* (Loewenstein and Ubel, 2008). Such approaches tend to share two important features: (1) they allow people to go their own way, that is, to reject the direction suggested by the paternalist; and (2) they reflect a form of *means paternalism*, rather than *ends paternalism*, in the sense that they are respectful of the chooser's view of the ultimate destination. It is for this reason that a GPS device is a defining example of libertarian paternalism.

In the relevant cases, the policy is introduced by a *planner* or a *choice architect* (who may work in the private or public sector). The objective is to enlist policies that do not distort (or distort minimally) the choices of fully rational individuals, but at the same time nudge others in a direction that is in their *considered best interests (as judged by themselves)*. By considered best interests is meant the decisions that would be made by individuals if they had complete information, unlimited cognitive abilities, and no lack of self-control. In other words, individuals are enabled to make choices that make them better off, again as judged by themselves (Thaler and Sunstein, 2009, Ch. 5).

Admittedly, it can be challenging to know whether this criterion is met (Sunstein, 2017). If so, it may be best to insist on a particular form of choice architecture: *active choosing* (Sunstein, 2017). But in some contexts, it is difficult or perhaps impossible to avoid some kind of steering of individual choices. Any website has to have some kind of layout. Any form has to have some questions first and some questions last. Individuals are given choices and information by firms, other individuals, and the government. These are necessarily framed in a particular manner and already embed a certain choice architecture. Many examples can be given (Thaler and Sunstein, 2009) from doctors who advise patients

[36] Camerer et al. (2003) view the relaxation of normative preferences in welfare economics as a natural progression of the relaxation of other restrictive assumptions in economics, such as perfect competition, perfect information, certainty etc.

on choice of a treatment to architects who advise clients about house designs, to interior decorators who advise people on alternative interior designs.

Example 12 *(Sunstein and Thaler, 2003): Consider the arrangement of food items in a cafeteria in a certain order. On plausible assumptions, boundedly rational consumers,who pay special attention to items placed at eye level, could benefit from the placement of healthy options at eye level, relative to unhealthy snacks; this is an example of a* nudge. *Consumers with normative preferences are not influenced by the arrangement of items, so little cost is imposed on them. The intervention of the choice architect is a form of paternalism. But, ultimately it is the consumer who exercises choice; in this sense, the proposal respects* consumer sovereignty.

Example 13 *(Dhami, 2016, p. 1593): Cash machines dispense cash only when the relevant bank card has been taken out first; nozzles for dispensing different kinds of car fuel, e.g., diesel and petrol, are often of a different size or color to prevent incorrect use of fuel; many cars produce a beeping sound if the driver is not wearing a seat belt; automatic electric switches in cars and offices are often used to conserve energy, or prevent batteries from running out; default options in savings and pension plans enable people to invest better; government warnings on cigarette packs.*

Of great interest to behavioral economics are cases, traced above, that arise from self-control problems. These problems are often a product of *present-biased preferences* in conjunction with a model of *multiple selves*; this is shown rigorously in (Dhami, 2016, Part 3). Since this appears to be such a common and repeated source of misunderstanding of the rationale for nudges, we outline some of the basic machinery; for a textbook treatment see Dhami (2016, Parts 3, 8).

The idea of multiple selves, well known from psychology, may be explained by the following simple example that illustrates several features of the failure of self-control together with imperfect awareness of the possibility of such failure.

Example 14 *: Suppose an individual sets an alarm during the night (the night-self) to wake up in the morning (the morning-self). Individuals may often find it difficult to precommit to waking up on the alarm, because the night-self might not be able to predict accurately how the morning-self will react to the alarm. The morning-self may switch the alarm off, with possibly negative consequences for both selves. In this case, the night-self might benefit from a commitment technology that disciplines the morning-self. Manufacturers of alarm clocks seem aware of the problem, so they build-in a commitment device into the alarm clock through a snooze button that reactivates the alarm to set-off after a short interval, thereby acting as a disciplining device on the morning-self.*

An even stronger precommitment from the night-self would be to place the alarm clock away from the bed, increasing the cost of the morning-self to reverse the night-self's decision. Perhaps the night-self does not often engage in this strong precommitment because it underestimates the degree of the self-control problem in its future selves.

The simplest form of present-biased preferences is the quasi-hyperbolic form; for details see Dhami (2016, Ch. 11). Consider time $t \in \Gamma = \{0, 1, 2, \ldots, T\}$ and real valued consumption c_t at time t. Let there be multiple selves of the same individual– one for each period. Each self who makes a decision in any time period $t \in \Gamma$ has *quasi-hyperbolic preferences* given by.

$$U_t\left(c_t, c_{t+1}, \ldots, c_T\right) = u\left(c_t\right) + \beta \sum_{\tau=t+1}^{\tau=T} \delta^{\tau} u\left(c_{\tau}\right),\ 0 < \beta < 1. \tag{12.1}$$

The parameter β in (12.1) creates a bias towards immediate gratification for each self, and $\delta > 0$ is the discount factor in the *exponential discounted utility* (EDU) model. Setting $\beta = 1$ in (12.1) gives rise to the EDU model, which is the main model of decision making in neoclassical economics. In making economic decisions, the self at time $t \in \Gamma$ must form beliefs about the behavior of all the future selves. Such beliefs reflect the *degree of self-awareness* of the current self at any time $t \in \Gamma$ (see Example 14). Self t has beliefs that the preferences of all the future *selves* at time $k > t$ are given by

$$U_k\left(c_k, \ldots, c_T\right) = u\left(c_k\right) + \widehat{\beta} \sum_{\tau=k+1}^{\tau=T} \delta^{\tau} u\left(c_{\tau}\right),\ t < k \leq T-1, \tag{12.2}$$

where $\widehat{\beta}$ is an estimate of the actual β value of a future self. The relation between the actual impatience parameter of self t, β, and the estimate, $\widehat{\beta}$, for future selves is an important element in understanding many applications of libertarian paternalism.[37] O'Donoghue and Rabin (2001) propose the following classification.

1. *Time consistents*: People with standard time consistent preferences behave as if they are using the EDU model, so
$$\beta = \widehat{\beta} = 1. \tag{12.3}$$
These individuals imagine that none of their future selves ever has a present-bias (over and above that caused by the presence of δ). This is the standard case in economics. In this case, and given the other assumptions in the BRA, the empirical phenomena listed in F1–F4 above are unlikely to arise.

[37] In a more general model, one's beliefs about the present bias of all the future selves may be heterogenous. Thus, there could be a sequence of $\widehat{\beta}_k$ values $t < k \leq T-1$, but we simplify by setting $\widehat{\beta}_k = \widehat{\beta}$ for all k.

2. *Sophisticates*: Sophisticates know that they have a bias towards immediate gratification, and believe that their future selves will also have an identical bias, i.e.,

$$\beta = \widehat{\beta} < 1. \tag{12.4}$$

3. *Partial naifs*: Partial naifs know that they have a bias towards immediate gratification, but believe (mistakenly) that their future selves will have a smaller, but non zero, bias, i.e.,

$$\beta < \widehat{\beta} < 1. \tag{12.5}$$

4. *Naifs*: Naifs know that their current selves are present-biased, but believe that their future selves will have no bias towards immediate gratification, i.e., they believe that future selves are time consistent, hence

$$\beta < \widehat{\beta} = 1. \tag{12.6}$$

Evidence indicates that most people are partial naifs, i.e., they fall within the two extremes of sophisticates and naifs.

Remark 4 *(Dhami, 2016, Ch. 11): People who have present-biased preferences will exhibit the behavior outlined in F1–F4 even if they satisfy all other features of the BRA. The degree to which they exhibit such behavior, and whether the behavior is to procrastinate or preproperate, will depend on their degree of self-awareness, as captured by the three cases in (12.4), (12.5), (12.6).*

Critics have objected to several features of libertarian paternalism. Sugden (2008, 2013) has used the *contractarian approach* in which appeals are directly made to individuals, stressing the personal advantage to each individual from taking a particular action. The interested reader can get acquainted with some of the features of alternative proposals in Dhami (2016, Ch. 22)

12.1. Political economy of nudging

Many people have raised the possibility that nudging might be carried out by choice architects who are not benevolent entities. We agree that choice architects may not be benevolent. We would add that they may not have important or necessary information. This concern applies far more strongly to mandates and bans than to nudges, which preserve freedom of choice. Of course it applies to any government regulation in which regulatory capture, say, by lobbyists, is a possibility, whether it can be classified as a nudge or not. Economists have long been aware of the political economy of regulation and it is an established field in economics (Laffont and Tirole, 1993).

Alert to the risk of government error, Gigerenzer offers the following example. Letters sent out to women for mammography screening (a form of nudge) state that early detection reduces breast cancer mortality by 20%. However, in absolute terms it reduces mortality from 5 to 4 for every 1000 women after 10 years. Gigerenzer (p. 363) views the information in a percentage form as a misleading nudge to benefit the mammography industry (political economy considerations) and prefers education to nudging so that people can make a more informed decision. On p. 364, he writes: "...democratic governments should invest less in nudging and more in educating people to become risk savvy."

We are not sure that in thinking about appropriate tools, it is the best approach to contrast a misleading nudge with a perfect, and evidently hypothetical, educational approach (perfect by stipulation). No one should favor misleading nudging. Nor is there any contradiction between nudging people and educating them to be more risk savvy. Such efforts can go in parallel, which might be even more efficacious. In addition, the same safeguards that would need to be applied to prevent potentially 'misleading education' may also be used to prevent potentially 'misleading nudges.' Choice architecture is inevitable in both cases.

12.2. The rationale for nudges

Gigerenzer appears to believe that the rationale for nudges lies in arguments for a lack of rationality, which he calls "latent irrationality" (p. 361). In his view, those who believe in LP think that human beings are "hardly educable" (p. 361). They "prefer nudging to educating people." He adopts a definition of nudging, which he calls "its original meaning," though it has been used only by its critics: "a set of interventions aimed at overcoming people's stable cognitive biases by exploiting them . . ." (p. 363). Continuing this theme on p. 364 he writes: "As I will argue in some detail, the dismal picture of human nature painted by behavioral economists and libertarian paternalists is not justified by psychological research." In his view, "Nudging people without educating them means infantilizing the public " (p. 379). He adds, "The interest in nudging as opposed to education should be understood against the specific political background in which it emerged. In the US, the public education system is largely considered a failure, and the government tries hard to find ways to steer large sections of the public who can barely read and write. Yet this situation does not apply everywhere."

We have several comments.

(1) A GPS device is useful for human beings, even if we do not have a dismal picture of human nature. It helps people to get where they want to go. Many nudges have that characteristic; consider reminders, disclosures, and warnings (Sunstein, 2013).

(2) Though people can define terms however they wish, the stated definition of nudges

has not been embraced by any public official who has engaged in nudging, and it is inconsistent with its original meaning. Many nudges are explicitly educative (Sunstein, 2016), so it is difficult to understand the claim that those who embrace nudging "prefer nudging to educating people." (Is a double-sided setting for office printers a way of "infantalizing the public"?)

(3) Nudges, actually adopted by governments or under serious consideration, do not depend on controversial psychological research. Whatever their force, the objections to KT&O are irrelevant to most nudges; return to reminders, disclosure, and warnings. Default rules are a prominent kind of nudge, and they are usually (not always) powerful. That claim does not depend on a "dismal picture of human nature."

(4) One of us (Sunstein) worked in the U.S. government for an extended period of time, and he did not try hard, or at all, or ever, "to steer large sections of the public who can barely read and write." On the contrary, many of the nudges in which he was involved were explicitly educative; they involved disclosure of information, which requires a capacity to read (Sunstein, 2013). Some of those nudges involved default rules, as in the case of automatically enrolling poor children in free school meals programs (to which they were legally entitled) (ibid.). We are not aware of any U.S. official who has tried "to steer large sections of the public who can barely read and write" (for a catalogue of nudges, see Halpern, 2015).[38]

(5) Nations all over the world – including the United Kingdom, Canada, Ireland, Japan, Australia, the United States and the Netherlands, to name just a few – have created behavioral insights teams, or nudge units, to help solve policy problem (Halpern, 2015). The resulting initiatives have rarely, if ever, depended on controversial psychological research.

(6) People do use fast and frugal heuristics and they often work well. However, even when we account for ecological rationality (i.e., there is a clear context and frame in which the decision is situated), heuristics create biases relative to the BRA benchmark. We have already covered this in our discussion above. Correction of those biases, through educative nudging (or efforts to teach statistical literacy), may be a good idea.

(7) To this we add a brief note on overconfidence that also creates departures of behavior from the BRA benchmark (Kahneman and Tversky, 1996; Dhami, 2016). In his criticism of nudges, Gigerenzer (2015, p. 371) gives a prominent role to overconfidence, and discounts most of the existing evidence on overconfidence. Although there are difficulties in measuring overconfidence, behavioral economics has documented impressive evidence on overconfidence, using novel methods, not cited in Gigerenzer (2015); see for instance, the surveys in Malmendier and Taylor (2015), and Malmendier and Tate (2015). This evidence indicates that overconfidence may harm people's self-interest. In contrast,

[38]Some people might note that it is ususally not the most wonderful idea to insult nations or their people.

Gigerenzer (2015, p. 371) takes the view that when overconfidence exists, it may help people's self-interest, and writes: "If you earn your money by forecasting exchange rates or the stock market, you had better be overconfident." The evidence from behavioral economics indicates that more overconfident people may churn their portfolios more and lose money; and, more overconfident managers of firms may engage in greater merger activity, which is characterized by a high rate of eventual failure (Dhami, 2016, Ch. 21).

Gigerenzer (1993) has argued that if the relevant questions/events are sampled randomly, then overconfidence disappears. For instance, if subjects are asked an IQ question and the IQ question is randomly drawn from the set of all possible IQ questions, then the G&O claim is that overconfidence disappears (i.e., the subjects' assessment of the accuracy of their answers is close to the actual number of correct answers). Kahneman and Tversky (1996) already address this issue in their two-fold reply. First, not all individuals are overconfident in all possible situations. Second, even when the questions are randomly sampled, empirical evidence shows that overconfidence remains, provided the questions are not too easy. There is also the non-trivial problem of defining the "set of all possible IQ questions" from which random sampling is to be undertaken.

(8) Some of the problems that nudges have successfully tackled in actual practice comes from self-control problems due to present-biased preferences (Thaler, 2016). Such preferences, in the form of hyperbolic discounting are found not just in humans, but also in close primate relatives, suggesting that these were inherited from a common evolutionary ancestor (Dhami, 2016).

12.3. Other remedies for the same problems

Gigerenzer (2015, p. 367) writes: "There are multiple other reasons for harmful behavior, including the fast food and tobacco industry's multi-billion-dollar advertisement campaigns to seduce people into unhealthy lifestyles and the failure of educational systems worldwide to teach children statistical thinking. Libertarian paternalists, like the behavioral economists they draw upon, make a case against the human mind and thereby turn a blind eye to flaws in human institutions."

Nothing could be further from the truth. Some libertarian paternalists have spent most of their working lives on flaws in human institutions. The remit of libertarian paternalism is to provide an additional tool for public policy that imposes no restrictions on individual liberty. It does not deny the role of other regulations that might achieve a similar purpose. It adds to the menu of policy choices for policymakers. For instance, in order to reduce smoking, the government can use prohibitions and fines (ban smoking/penalize manufacturers) or use a nudge (appropriate warning on cigarette packages). More choices may sometimes be worse than fewer, but libertarian paternalism can hardly be blamed for adding to the menu of choices.

We have mentioned that some people argue for "boosts," such as an increase in statistical literacy, on the theory they improve people's capacity for agency (Hertwig, 2017). On one view, boosts are an alternative to nudges, and in some ways better. We cannot fully address that view in this space, and we agree that boosts can be a good idea, so we restrict ourselves to five points. First, some nudges are meant to improve people's capacity for agency, and belong in the same family as boosts; consider information (about caloric content, for example); warnings (about the consequences of paying late, for example); and reminders (that a doctor's appointment is coming up, for example). Second, some nudges make life simpler, and so allow people to put their focus and attention where they wish (consider a GPS device or a default rule). In that sense, they promote agency.

Third, nudges have an established track record (Halpern, 2015), with impressive cost-effectiveness (Benartzi et al., 2017); the empirical evidence for boosts is less robust, certainly for the most important public policy issues. Fourth, it is not easy to identify boosts that could provide significant help in addressing the principal problems faced by the world's governments (though we agree that statistical literacy would be a step forward). If one is advising the Prime Minister of the United Kingdom or the President of the United States, what boosts would successfully address (for example) highway safety, opioid addiction, obesity, greenhouse gas emissions, or immigration? Fifth, boosts and nudges can be complements, not alternatives.

13. Conclusions

The heuristics and biases approach, begun by Daniel Kahneman and Amos Tversky, is one of the most important research programs in social science and certainly one of its stellar achievements; we call it the KT&O program. It showed that the empirical evidence is not consistent with the Bayesian Rationality Approach (BRA) in economics. There has been increased formalization of the KT&O program. It has been intensely scrutinized and criticized, which we believe has benefitted it. We have presented a formalized version of the various heuristics in the KT&O program in one place, evaluated the criticisms of the approach, and described how these criticisms can be answered, and existing understanding improved, by a fuller consideration of the evidence and the emerging models.

The BRA approach deals with situations of certainty, risk, subjective uncertainty, and ambiguity. In contrast to these situations, many interesting and important real life problems belong to the domain of true uncertainty. Furthermore, any real world problems are NP complete, i.e., the time required to solve them grows too fast as the problem becomes more complicated. In fact, problems like this, such as the travelling salesman problem, cannot be solved in polynomial time. How do people solve such problems?

Decision making under true uncertainty is a vexed problem. Despite progress, we believe that the G&O program has not yet been able to provide a persuasive account of decision making under true uncertainty. It is not clear what benchmark to use to assess the performance of the heuristics in the G&O program; in contrast the benchmark is clear in the KT&O program. There remain severe problems in stringently testing the heuristics proposed in the G&O program because they rely on unobserved mental processes.

Finally, we consider an application to an approach in behavioral welfare economics, libertarian paternalism or nudging. This has recently been criticized in the G&O program. We show that libertarian paternalism does not depend on controversial psychological claims.

We offer one final note. The neglect of true uncertainty in economics is truly astonishing. The other social sciences do not provide a coherent framework in this regard either. More research is needed on how people make decisions in this case. We also need to have a basic agreement on what constitutes a benchmark in these situations against which proposed alternatives can be evaluated. Should this benchmark take an ex-ante perspective (prior to the resolution of true uncertainty) or an ex-post perspective (after the resolution of true uncertainty)? Heuristics in the KT&O program that were designed for evaluating the BRA may, in our view, be also potentially interesting candidates to think about decision making under true uncertainty. Or it could be that mental models and social norms might have evolved to deal with such cases. It is staggering that the black box of true uncertainty remains so impenetrably black.

References

[1] Akerlof, G. A. and Robert J. Shiller, R. J. (2015). Phishing for Phools: The Economics of Manipulation and Deception. Princeton University Press: Princeton.

[2] al-Nowaihi, A. and Dhami, S. (2015). Evidential equilibria: Heuristics and biases in static games of complete information. Games. 6: 637-676.

[3] al-Nowaihi, A. & Dhami, S. (2018). Foundations for Intertemporal Choice. CESifo Working Paper Series 6913, CESifo Group Munich.

[4] Ariely, D., Loewenstein, G. and Prelec, D. (2006). Tom Sawyer and the Construction of Value. Journal of Economic Behavior and Organization, 60: 1-10.

[5] Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? Memory and Cognition, 32, 1369 - 1378.

[6] Bar-Hillel, M. and Wagenaar, W. A. (1991). The perception of randomness. Advances in Applied Mathematics, 12, 428-454.

[7] Barberis, N., & Thaler, R. H. (2003). A survey of behavioral finance. In G. Constantinides, M. Harris and R. Stultz (Eds.) Handbook of the economics of finance (pp.1051-1119). Amsterdam: North-Holland.

[8] Barberis, N., A. Shleifer, and R. Vishny, 1998, A model of investor sentiment, Journal of Financial Economics 49, 307–343.

[9] Barbey, A. K. and Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. Behavioral and Brain Sciences 30 (3): 241–54.

[10] Barro, R.J. and Sala-i-Martin, X.I. (2003) Economic Growth. Second edition. MIT Press: Massachussetts.

[11] Ben-David, I., Graham, J., Harvey, C. (2013). Managerial Miscalibration, Quarterly Journal of Economics 128 (4) 1547-1584.

[12] Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker, W., Congdon, W. J. and Galing, S. (2017). Should Governments Invest More in Nudges? Psychological Science. 28: 1041-55

[13] Benartzi, S. and Thaler. R. H. (2004). Save More Tomorrow: Using Behavioral Economics to Increase Employee Savings. Journal of Political Economy. 112: S164-S187.

[14] Bergert, F.B., Nosofsky, R.M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. Journal of Experimental Psychology: Learning Memory and Cognition. 331:107–29.

[15] Bernheim, B., Skinner, D., Skinner, J., and Weinberg, S. (1997). What accounts for the variation in retirement wealth among U.S. households? NBER. Working Paper 6227.

[16] Bertrand, Marianne, and Adair Morse. (2011). Information Disclosure, Cognitive Biases, and Payday Borrowing. The Journal of Finance, 66 (6), pp. 1865-1893.

[17] Biais, B. and Weber, M. (2009) Hindsight Bias, Risk Perception and Investment Performance. Management Science, 55(6): 1018-1029.

[18] Birnbaum. M. H. (2008) Evaluation of the Priority Heuristic as a Descriptive Model of Risky Decision Making: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). Psychological Review. 115: 253–262.

[19] Bocskocsky, A., Ezekowitz, J. and Stein, C. (2014). The Hot Hand: A New Approach to an Old Fallacy, Paper presented at the MIT Sloan Sports Analytics Conference.

[20] Bonini, N., Tentori, K., & Osherson, D. (2004). A different conjunction fallacy. Mind & Language, 19: 199-210.

[21] Borges, B., Goldstein, D. G., Ortmann, A., & Gigerenzer, G. (1999). Can ignorance beat the stock market? In G. Gigerenzer, P. M.

[22] Todd & the ABC Research Group (Eds.) Simple Heuristics That Make Us Smart, pp. 59–72. New York: Oxford University Press.

[23] Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Choices without tradeoffs. Psychological Review, 113, 409–432.

[24] Bröder A, Schiffer S. (2003). Take the best versus simultaneous feature matching: probabilistic inferences from memory and effects of representation format. Journal of Experimental Psychology: General. 132: 277–93.

[25] Bröder, A. (2011) The quest for take-the-best: Insights and outlooks for experimental research. In Gigerenzer, G., Hertwig, R. and Pachur, T. (eds.) Heuristics: The foundations of adaptive behavior. Oxford University Press: Oxford.

[26] Camerer, C. F. (1987). Do Biases in Probability Judgment Matter in Markets? Experimental Evidence. American Economic Review, 77, 981-97.

[27] Camerer, C. F. (1989). Does the Basketball Market Believe in the 'Hot Hand'? American Economic Review, 79, 1257-61.

[28] Camerer, C. F. (1990) Do Markets Correct Biases in Probability Judgment? Evidence from Market Experiments, in J. Kagel and L. Green (Eds.) Advances in Behavioral Economics, 2, Northwood, NJ: Ablex Publishers, 125-172.

[29] Camerer, C. (1995) Individual decision making, in J. Kagel and A. E. Roth (eds) Handbook of Experimental Economics, Princeton, NJ.

[30] Camerer, C.F. (2015). The Promise and Success of Lab-Field Generalizability in Experimental Economics: A Critical Reply to

[31] Levitt and List. in Fréchette, G.R., and Schotter, A. (eds.) Handbook of Experimental Economic Methodology. Oxford University Press: Oxford. pp. 249-295.

[32] Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., Rabin, M. (2003). Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism. University of Pennsylvania Law Review 151: 1211-1254.

[33] Caplin, A., Dean, M. and Martin, D. (2011). Search and Satisficing. American Economic Review, 101(7): 2899-2922.

[34] Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. Applied Cognitive Psychology, 10, 519-540.

[35] Charness, G., Karni, E., and Levin, D (2010) On the Conjunction Fallacy in Probability Judgment: New Experimental. Games and. Economic Behavior, 68: 551-556.

[36] Chmura, T., Goerg, S.J., and Selten, R. (2012). Learning in experimental 2x2 games, Games and Economic Behavior. 76(1): 44–73.

[37] Clotfelter, C. T. and P. J. Cook. (1993). The gambler's fallacy in lottery play, Management Science. 39, 1521-25.

[38] Conlisk, J. (1996): Why bounded rationality? Journal of economic literature, 34(2), 669–700.

[39] Cosmides, L. and Tooby, J. (1996) Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition, 58:1-73.

[40] Croson, R. and Sundali, J. (2005). The gambler's fallacy and the hot hand: Empirical data from casinos. Journal of Risk and Uncertainty, 30:195-209.

[41] Czerlinski, J., Goldstein, D. G., & Gigerenzer, G. (1999). How good are simple heuristics? In Gigerenzer, G., Todd, P.M., & the ABC Research Group. Simple heuristics that make us smart. New York: Oxford University Press.

[42] Datta, S. and Mullainathan, S. (2014). Behavioral Design: A New Approach to Development Policy. Review of Income and Wealth 60 (1): 7-35.

[43] Dhami, S. (2016) Foundations of Behavioral Economic Analysis, Oxford University Press: Oxford.

[44] Dhami, S. and al-Nowaihi, A. (2018) Rationality in Economics: Theory and Evidence. CESifo Working Paper Series No. 6872. Forthcoming in Handbook of Rationality. MIT Press: Massachusetts.

[45] Dhami, S., Wei, M. and al-Nowaihi, A. (2018) Public goods games and psychological utility: Theory and evidence. Forthcoming in the special issue on Psychological Game Theory in Journal of Economic Behavior and Organization.

[46] DeMiguel, V., Garlappi, L. and Uppal, R. (2009) Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? Review of Financial Studies. 22: 1915-1953.

[47] Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., and Peijnenburg, K. (2016). Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence. Journal of Financial Economics 119: 559–577.

[48] Dimmock, S. G., Kouwenberg, R., and Wakker, P. P. (2016). Ambiguity Attitudes in a Large Representative Sample. Management Science 62, 1363–1380.

[49] Edwards, W. and Tversky, A., (1967) Decision Making: Selected Readings. Penguin.

[50] Englich, B., Mussweiler, T. and Strack, F. (2006). Playing Dice with Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making. Personality and Social Psychology Bulletin 32 (2): 188-200.

[51] Englich, B., and Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom. Journal of Applied Social Psychology, 31, 1535-1551.

[52] Evans, J. Handley, S.J., Perham, N. Over, D.E. and Thompson, V.A. (2000) Frequency versus probability formats in statistical word problems. Cognition 77: 197-213.

[53] Fischhoff, B. (1975). Hindsight âÂ■ foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1: 288-299.

[54] Friedman, M., and Schwartz, A. J. (1963) A Monetary History of the United States, 1867–1960. Princeton University Press: Princeton.

[55] Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases." In W. Stroebe & M. Hewstone (Eds.), European review of social psychology, (Vol. 2, pp. 83-115). Chichester, England: Wiley.

[56] Gigerenzer, G. (1993). The bounded rationality of probabilistic mental models. In K. I. Manktelow & D. E. Over (Eds.), Rationality: Psychological and philosophical perspectives (pp. 284-313). London: Routledge.

[57] Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. Psychological Review, 103, 592-596.

[58] Gigerenzer, G. (2008). Rationality for mortals: How people cope with uncertainty. New York: Oxford University Press.

[59] Gigerenzer G. (2014) Risk Savvy: How to make good decisions. Viking: New York.

[60] Gigerenzer G. (2015) On the Supposed Evidence for Libertarian Paternalism. Review of Philosophy and Psychology. 6: 361–383. |

[61] Gigerenzer, G. (2016). Towards a rational theory of heuristics. In R. Frantz & L. Marsh (Eds.), Minds, models, and milieux: Commemorating the centennial of the birth of Herbert Simon (pp. 34-59). New York: Palgrave Macmillan.

[62] Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. Topics in Cognitive Science, 1, 107–143.

[63] Gigerenzer, G., & Brighton, H. (2011). Homo heuristicus: Why biased minds make better inferences. Appetizer in Gigerenzer, G.,

[64] Hertwig, R. and Pachur, T. (eds.) Heuristics: The foundations of adaptive behavior. Oxford University Press: Oxford.

[65] Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. Annual Review of Psychology, 62, 451–482.

[66] Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. Psychological Review, 103, 650-669.

[67] Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. Journal of Experimental Psychology: Human Perception and Performance, 14. 513-525.

[68] Gigerenzer, G., Hertwig, R. and Pachur, T. (2011) Heuristics: The foundations of adaptive behavior. Oxford University Press: Oxford.

[69] Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). Heuristics: The foundations of adaptive behavior. New York: Oxford University Press.

[70] Gigerenzer, G., & Selten, R. (Eds.). (2001). Bounded rationality: The adaptive toolbox. Cambridge, MA: MIT Press.

[71] Gigerenzer, G., Todd, P. M., & the ABC Group (1999). Simple heuristics that make us smart. New York: Oxford University Press.

[72] Gilovich, T., Griffin, D. W., and Kahneman, D. (Eds). (2002). Heuristics and Biases: The Psychology of Intuitive Judgment. New York: Cambridge University Press.

[73] Gilovich, T., Vallone, R., and Tversky, A. (1985). The Hot Hand in Basketball: On the Misperception of Random Sequences. Cognitive Psychology, 17, 295-314.

[74] Gintis, H. (2009). The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences. Princeton: Princeton University Press.

[75] Gintis, H. (2017) Individuality and Entanglement: The Moral and Material Bases of Social Life. Princeton University Press: Princeton.

[76] Glaser, M., Langer, T., Weber, M. (2007). On the trend recognition and forecasting ability of professional traders. Decision Analysis 4: 176-193.

[77] Green, B. S. and Zwiebel, J. (2013) The Hot Hand Fallacy: Cognitive Mistakes or Equilibrium Adjustments? Evidence from Baseball. mimeo Stanford University.

[78] Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. Psychological Science,17 (9), 767–773

[79] Guilbault, R. L., Bryant, F. B., Brockway, J. H., & Posavac, E. J. (2004). A meta-analysis of research on hindsight bias. Basic and Applied Social Psychology, 26, 103-117.

[80] Guryan, J., and Kearney, M. S. (2008). Gambling at Lucky Stores: Empirical Evidence from State Lottery Sales. American Economic Review, 98(1): 458-73.

[81] Halpern, D. (2015). Inside the nudge unit: How small changes can make a big difference. W. H. Allen: London.

[82] Heath, C., Tversky, A. (1991). Preference and belief: ambiguity and competence in choice under uncertainty. Journal of Risk and Uncertainty 4, 5-28.

[83] Hertwig, R. (2017). When to consider boosting: some rules for policy-makers. 1: 143-61.

[84] Hertwig, R., and Gigerenzer, G. (1999). The conjunction fallacy revisited: How intelligent inferences look like reasoning errors. Journal of Behavioral Decision Making, 12, 275-305.

[85] Hoffrage, U., Hertwig, R. and Gigerenzer, G. (2000) Hindsight-bias: A by-product of knowledge updating? Journal of Experimental Psychology: Learning Memory and Cognition. 26: 566-581.

[86] Hogg, R. V., McKean, J. W. and Craig, A. T. (2005) Introduction to Mathematical Statistics, 6th ed (Prentice Hall).

[87] Kahneman, D. (1981) Who shall be the arbiter of our intuitions? Behavioral and Brain Sciences, 4:339-40.

[88] Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes (commentary on Stanovich and West). Behavioral and Brain Sciences, 23, 681-683.

[89] Kahneman, D. (2011). Thinking fast and slow. New York, NY: Farrar, Strauss, Giroux.

[90] Kahneman, D., Slovic, P., Tversky, A. (Eds.). (1982) Judgement Under Uncertainty: Heuristics and Biases, Cambridge University Press, Cambridge.

[91] Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. Econometrica. 47(2): 263–291.

[92] Kahneman, D. and Tversky, A. (1984) Choices, values, and frames. The American Psychologist, 39: 341-35

[93] Kahneman, D. & Tversky, A. (1996). On the reality of cognitive illusions: A reply to Gigerenzer's critique. Psychological Review, 103, 582-591.0.

[94] Kahneman, D. and Tversky, A. (2000) Choices, values, and frames. Cambridge University Press: Cambridge.

[95] Kahneman, D., Slovic, P. and Tversky, A. (1982). Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press: Cambridge.

[96] Kirchler, E., Maciejovsky, B. (2002). Simultaneous over- and underconfidence: evidence from experimental asset markets. Journal of Risk and Uncertainty, 25: 65-2585.

[97] Kuran, T., and Sunstein, C. (1999) Availability Cascades and Risk Regulation, Stanford Law Review, 51(4): 683-768.

[98] Lacetera, N., Pope, D. G. and Sydnor, H. (2012). Heuristic thinking and limited attention in the car market. American Economic Review, 102(5) 2206-2236.

[99] Laffont, J.-J., and Tirole, J. A. (1993) Theory of Incentives in Procurement and Regulation. Cambridge MA: MIT Press.

[100] Levitt, S.D., and J.A. List. (2007). What do laboratory experiments measuring social preferences reveal about the real world? Journal of Economic Perspectives 21(2), 153-174.

[101] Lichtenstein, S., & Slovic, P. (1971). Reversals of preferences between bids and choices in gambling decisions. Journal of Experimental Psychology, 89: 46-55.

[102] Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. Journal of Experimental Psychology: Human Learning and Memory, 4: 551-578.

[103] Loewenstein, G. and Haisley, E. (2008) The Economist as Therapist: Methodological Ramifications of 'Light' Paternalism, in A. Caplin and A. Schotter (Eds.), Perspectives on the Future of Economics: Positive and Normative Foundations, volume 1 in the Handbook of Economic Methodologies, Oxford: Oxford University Press.

[104] Lord, C. G., Ross, L. & Lepper, M. R. (1979) Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. Journal of Personality and Social Psychology 37(11):2098-109.

[105] Lucas, R. E. (1988) On the Mechanics of Economic Development. Journal of Monetary Economics. 22: 3–42.

[106] Malkiel, B. (1990). A random walk down wall street. New York: W.W. Norton.

[107] Malmendier, U. and Tate, G. (2015) Behavioral CEOs: On the Role of Managerial Overconfidence. Journal of Economic Perspectives, 29(4): 37-60.

[108] Malmendier, U. and Taylor, T. (2015) On the Verges of Overconfidence. Journal of Economic Perspectives, 29(4): 3-8.

[109] Mas-Colell, A., Whinston, M. and Green, J. R. (1995) Microeconomic Theory. New York: Oxford University Press.

[110] McNeil, B. J., Pauker, S. G., Sox, H. C., Jr., and Tversky, A. (1982). On the elicitation of preferences for alternative therapies. New England Journal of Medicine 306:1259-62.

[111] Meehl, P. E., & Rosen, A. (1955) Antecedent probability and the efficiency of psychometric signs, patterns or cutting scores. Psychological Bulletin, 52, 194-216.

[112] Metzger, M. (1984). Biases in Betting: An Application of Laboratory Findings, Psychological Reports 56, 883-888.

[113] Miller, J. B., and Sanjurjo, A. (2015). Surprised by the Gambler's and Hot Hand Fallacies? A Truth in the Law of Small Numbers. IGIER Working Paper No. 552.

[114] Newell B.R., Shanks, D.R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision making. Journal of Experimental Psychology: Learning Memory and Cognition. 29:53–65.

[115] Newell B.R., Weston N.J., Shanks D.R. (2002) Empirical tests of the a fast-and-frugal heuristic: Not everyone "takes-the-best". Organization Behavior and Human Decision Processes. 91:82-96.

[116] Northcraft, G. B., and Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. Organizational Behaviour and Human Decision Processes, 39, 84-97.

[117] Oaksford, M., & Chater, N. (1998). Rational models of cognition. Oxford, England: Oxford University Press.

[118] Pachur, T., Hertwig, R., & Steinmann, F. (2012). How do people judge risks: Availability heuristic, affect heuristic, or both? Journal of Experimental Psychology: Applied, 18, 314-330.

[119] Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). The adaptive decision maker. New York: Cambridge University Press.

[120] Pinker, S. (1997) How the mind works (New York: Norton).

[121] Rabin, M., (2002) Inference by Believers in the Law of Small Numbers, Quarterly Journal of. Economics, 117, 775-816.

[122] Rabin M, and Schrag J (1999) First impressions matter: a model of confirmatory bias. Quarterly Journal of Economics. 114(1):37-82

[123] Rabin, M. & Vayanos, D. (2010). The Gambler's and Hot-Hand Fallacies: Theory and Applications, Review of Economic Studies, 77: 730-778.

[124] Rapoport, A. & Budescu, D.V. (1992) Generation of random binary series in strictly competitive games. Journal of Experimental Psychology, 121, 352-364.

[125] Rapoport, A., and Budescu, D. V. (1997). Randomization in individual choice behavior. Psychological Review, 104, 603-617.

[126] Rieskamp J, Otto P. (2006). SSL: a theory of how people learn to select strategies. Journal of Experimental Psychology: General. 135:207–36

[127] Robbins, J.M., and Krueger, J.I. (2005). Social projection to ingroups and outgroups: a review and meta-analysis. Personality and Social Psychology Review. 9(1): 32.47.

[128] Ross L., Greene D. & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. Journal of Experimental Social Psychology 13, 279-301.

[129] Roth, B. and Voskort, A. (2014). Stereotypes and false consensus: How financial professionals predict risk preferences. Journal of Economic Behavior & Organization, 107: 553-565.

[130] Rubinstein, M. (2001) Rational markets: yes or no? The affirmative case, Financial Analysts Journal (May-June) pp. 15-29.

[131] Savage, L. J. (1954). The Foundations of Statistics. New York: John Wiley and Sons. Revised edition 1972, New York: Dover Publications.

[132] Scheibehenne B, Bröder A. (2007). Predicting Wimbledon 2005 tennis results by mere player name recognition. International Journal of Forecasting. 23:415–26.

[133] Simon, H. A. (1955). A behavioral model of rational choice. Quarterly Journal of Economics. 69: 99-118.

[134] Simon, H. A. (1978). Rational decision-making in business organizations. Nobel Memorial Lecture, 8 December, 1978.

[135] Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. Organizational Behavior and Human Decision Processes, 91, 296-309.

[136] Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.) Heuristics & biases: The psychology of intuitive judgment (pp. 397-420). New York: Cambridge University Press.

[137] Slovic, P., MacGregor, D. G., Malmfors, T., & Purchase, I. F. H. (1999). Influence of affective processes on toxicologists' judgments of risk (Report No. 99-2). Eugene, OR: Decision Research.

[138] Solan L., Rosenblatt T., Osherson D. (2008) False Consensus Bias in Contract Interpretation. Columbia Law Review, 108: 1272-1300.

[139] Stanovich, K. E. (2012). On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning. In K. Holyoak & R. Morrison (Eds.) (pp. 343-365) The Oxford handbook of thinking and reasoning. New York: Oxford University Press.

[140] Stanovich, K.E., & West, R.F. (2000). Individual differences in reasoning: Implications for the rationality debate [Target article and commentaries]. Behavioral and Brain Sciences, 23, 645-726.

[141] Stolarz-Fantino, S., Fantino, E., Zizzo, D. J., & Wen, J. (2003). The conjunction fallacy: New evidence for robustness. American Journal of Psychology, 116, 15-34.

[142] Sugden, R. (2008) Why incoherent preferences do not justify paternalism. Constitutional Political Economy, 19: 226-248.

[143] Sugden, R. (2013) The Behavioural Economist and the Social Planner: To Whom Should Behavioural Welfare Economics Be Addressed? Inquiry: An Interdisciplinary Journal of Philosophy. 56(5): 519-538.

[144] Sunstein, C. R. (2013). Simpler. New York: Simon and Schuster.

[145] Sunstein, C. R. (2016). The council of psychological advisers. Annual Review of Psychology, 67(1), 713-737.

[146] Sunstein, C. R. (2017). Default rules are better than active choosing (often). Trends in Cognitive Sciences, 21, 600-606.

[147] Sunstein, C. R., and R. H. Thaler. (2003) Libertarian Paternalism is Not an Oxymoron. University of Chicago Law Review 70 : 1159–202.

[148] Terrell, Dek. (1998). Biases in Assessments of Probabilities: New Evidence from Greyhound Races, Journal of Risk and Uncertainty 17, 151-166.

[149] Terrell, Dek and Amy Farmer. (1996). Optimal Betting and Efficiency in Parimutuel Betting Markets with Information Costs, The Economic Journal 106, 846-868.

[150] Tentori, K., Bonini, N., and Osherson, D. (2004). The conjunction fallacy: A misunderstanding about conjunction? Cognitive Science, 28, 467-477.

[151] Tetlock, P.E. (2002). Cognitive biases in path-dependent systems: Theory driven reasoning about plausible pasts and probable futures in world politics. In T. Gilovich, D. W. Griffin, & D. Kahneman. (Eds.). Inferences, heuristics and biases: New directions in judgment under uncertainty. New York: Cambridge University Press.

[152] Tetlock, P. E. (2006). Expert political judgment: How good is it? How can we know? Princeton, NJ: Princeton University Press.

[153] Thaler, R.H. (2015) Misbehaving: The making of behavioral economics. W. W. Norton and Company: New York.

[154] Thaler, R. H., and Sunstein, C. R. (2009) Nudge: Improving Decisions about Health, Wealth, and Happiness. New Haven: Penguin books.

[155] Thaler. R. H. and Sunstein, C. R. (2003) Libertarian Paternalism. American Economic Review: Papers and Proceedings 93 : 175–9.

[156] Tversky, A., and Kahneman, D. (1971). Belief in the law of small numbers. Psychological Bulletin. 76: 105-110.

[157] Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5 (2) 207-232.

[158] Tversky, A., and Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. Science. 185 (4157), 1124-1131.

[159] Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), Progress in social psychology (pp. 84-98). Hillsdale, NJ: Erlbaum.

[160] Tversky, A.,& Kahneman,D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90, 293-315.

[161] Tversky, A. and Gilovich T. (1989a) The Cold Facts about the 'Hot Hand' in Basketball, Chance, 2, 16-21.

[162] Tversky, A. and Gilovich T. (1989b) The Hot Hand: Statistical Reality or Cognitive Illusion? Chance, 2, 31-34.

[163] Wason, P.C. (1968) Reasoning about a rule, Quarterly Journal of Experimental Psychology, 20, 273-281.

[164] World Development Report (2015) Mind, society, and behavior. The World Bank: Washington DC.

[165] Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. American Psychologist, 35(2) 151-175.