Why Be Random?*

Thomas Icard

INTRODUCTION

Imagine you are approaching a stop sign at an empty intersection. It is clear that no one else is on the road. You would like to pass right through, to save time, gasoline, and wear on your breaks. But you also worry about the small chance that a patrol officer may be watching, and you do not want a fine. In such a circumstance, would it ever make sense to base your decision—whether to skip or stop—on the flip of a coin?[1]

On a familiar picture associated with Bayesian decision theory (e.g., Savage 1954), such randomized decisions are rationally permissible only in cases where the options look equally attractive. For instance, the costs associated with stopping should be on a par with the prospect of either avoiding those costs or paying a fine, weighted by the appropriate probabilities. A number of authors have argued that situations like these—the case of Buridan's donkey being the most famous—present us with positive reason to want access to a randomizing device in order to break the apparent symmetry (see Rescher 1959, and Ullman-Margalit and Morgenbesser 1977 and the large ensuing literature for critical discussion), though from a decision theoretic perspective any way of breaking the symmetry is acceptable, random or not. Some have declared such dilemmas to be extremely rare, if not inconceivable: there will always be some discernible difference between options, and further reflection will inevitably tip the balance.[2] A moment's thought will reveal either skipping or stopping as clearly the better option. Whatever one concludes about these issues, it is apparent that the rational role of randomization on this traditional picture is marginal at best. The sentiment was nicely summarized by economist Robert Aumann, who wrote: 'Practically speaking, the idea that serious people would base important decisions on the flip of a coin is difficult to accept' (Aumann, 1987, 15).

*Draft of May 4, 2019. This article has been accepted for publication in *Mind* (OUP).

[1]The example is from Godfrey-Smith (1996), where it is suggested that, intuitive though it may at first seem, flipping a coin is never to be recommended in such a scenario (p. 213).

[2]Leibniz famously held such a view. See the many references in Rescher (1959) and Ullman-Margalit and Morgenbesser (1977) for more examples.

An empirical observation about human beings, as well as other organisms, is that their behavior does often appear indeterminate, even random. This holds at multiple levels of organization—from the subcellular to the group level—and encompasses simple choice situations like the one described above (see Glimcher 2005 for a review). Meanwhile, in designing intelligent artifacts, engineers routinely add random noise to otherwise deterministic algorithms, and this often results in improvement. For example, the world champion automated Go program, AlphaZero (Silver et al., 2018), randomizes its decision about what actions to explore in determining its next move. Observations like these raise the question: is randomized behavior a mere heuristic, resulting in suboptimal organisms and artifacts, or is there some deeper normative justification that eludes the familiar Bayesian decision theoretic analysis?

The most celebrated arguments in favor of randomization have come from three arenas: game theory, experimental design, and reinforcement learning.[3] Some game theorists propose that randomization may be helpful in strategic situations. The 'gold standard' in experimental design is to use randomized, controlled trials. In reinforcement learning it is often argued that balance between exploitation of known rewards and exploration can be achieved by adding noise to action selection. All three claims have met Bayesian resistance.[4] These debates highlight a useful sharpening of our question: if the Bayesian argument against randomization seems convincing, but there are nonetheless situations where randomizing seems prudent, what assumptions in the argument might be violated in these situations? And what can we learn from this about when randomizing one's decision does promise to be helpful, even for a Bayesian?

The thesis of this article is that there are essentially two compelling rationales for randomization. The first is banal but unassailable, and as we shall argue, accounts for common intuitions in favor of randomizing: when unsure of what to do, if the best-looking known option involves randomization, choose this option. The second is equally unassailable, but also more interesting: access to a randomizing device is provably helpful for an agent burdened with a finite memory. This second rationale is robust, in that it applies equally to agents who are in every other way idealized. That is, even for an agent who can costlessly determine the best course of action, if that agent has bounded memory, then access to a random device still confers an advantage. Unifying previous results from computer science and statistics, we describe the general conditions under which this holds. In addition to making the positive case for these two rationales, part of the aim is to show how they together help alleviate the tension between the Bayesian prohibition on randomization and the observations noted above about randomness in people and programs.

---

[3]Notably, all three originated in their modern formulation within about a decade of one another: Borel (1923), Fisher (1935), and Thompson (1933).

[4]On games see, e.g., Kadane and Larkey (1982). For a recent argument against randomized experiments see Worrall (2007). Early representative work toward a Bayesian approach to exploration can be found in Bellman (1957).

After clarifying our question and rehearsing some familiar arguments in favor of randomized behavior, we then present the Bayesian counterargument. In that context we briefly discuss the first rationale and explain how it already illuminates many of the cases where randomization seems prudent. We turn next to matters of computational resources. A prominent approach to computational complexity so far reveals little motivation for randomizing. Nonetheless, it is possible to show that there is a deep connection between randomization and memory limitations. Having addressed the normative question of when randomization can be rational, we return to consider how the resulting story accords with some noteworthy empirical observations about random behavior in the biological world.

CLARIFYING THE QUESTION

Consider a decision situation in which an individual agent must choose from among a set $\mathcal{A}$ of possible actions. The agent's action, together with the state of the world, determines an outcome. We will assume that we can measure how good each outcome is by a real-valued utility $U : \mathcal{A} \times \mathcal{S} \to \mathbb{R}$, where $\mathcal{S}$ is the set of possible states. Let us suppose that the agent has costless access to a randomizing device. The agent can freely observe independent outcomes of a Bernoulli coin flip, using a coin with any given bias between 0 and 1. One might think of this as expanding the choice set from $\mathcal{A}$ to $\Delta(\mathcal{A})$, the space of all probabilities defined on $\mathcal{A}$. We can now pose the question: what ways of filling in more details of the scenario would make it rational for the agent to base a decision on outcomes of this randomizing device, to choose a nontrivial randomized action from $\Delta(\mathcal{A})$? More dramatically, when, if ever, should the agent arrange things so that the action causally depends on one or more outcomes of the device, thereby relinquishing at least some agential authority over the action that results (cf. Bratman 2001)?

This way of posing the question is deliberately presumptuous in what we admit as a randomizing device. We are assuming that outcomes are unpredictable in a strong and relatively objective sense, and in particular that the device is well modeled using the standard probability calculus.[5] The expected outcomes of such a device will satisfy a number of dissociable properties: unpredictability, independence with the states in $\mathcal{S}$, laws of large numbers, statistics at various orders, and so on. An argument for randomization could highlight a proper subset of these as especially important. At the same time, some of these properties may be easier to come by than others. Even producing *pseudorandom* bits is evidently far from costless (Vadhan, 2012). The reason for making this strong assumption is that the Bayesian argument purports to show why the agent would not strictly benefit from access to a device with all of these properties. At the end, after having argued the contrary, we will return to the question of where an agent might locate suitably random sources, should such be desired.

[5] A compatible account of what it means for a physical process to be random would be that of Eagle (2005). However, for present purposes it is not even necessary that the randomizing device be physically realizable.

A second assumption we will make is that the agent's choice is causally independent of the states—what state in $\mathcal{S}$ obtains does not causally depend on the choice selected from $\Delta(\mathcal{A})$—and there is no suspicion otherwise on the part of the agent. Violations of this assumption often constitute cases where randomization trivially becomes beneficial. For instance, imagine you are playing rock-paper-scissors and your adversary will come to know your strategy before they choose theirs. That is, the adversary will first learn which distribution from the set $\Delta(\{\mathsf{rock}, \mathsf{paper}, \mathsf{scissors}\})$ you have chosen, and then make their choice. It would obviously be desirable to opt for the fully random strategy that plays each of the three actions with equal probability. In fact, the more random (viz. unpredictable) your strategy is, the better you expect to fare. By assumption, the state (i.e., the adversary's action) is completely determined by your choice, which effectively trivializes the decision problem. From the agent's perspective, carrying out an action $a \in \mathcal{A}$ should give the same utility no matter whether a was chosen deterministically or as a result of the randomizing device; in particular, the relevant state that determines this utility should not depend on how a was selected. Adopting this causal assumption means that we are not considering scenarios in which randomization itself is associated with inherent cost or benefit.

Importantly, the causal assumption falls short of the stronger assumption that the agent believes its choice is *evidentially* independent of the state.[6] For reasons other than direct causal dependence, our agent might well assume that its own deliberation could reveal important information about the underlying state, e.g., by virtue of a common cause. These scenarios do not obviously trivialize. For instance, even if you do not think your adversary will be able to observe your strategy in rock-paper-scissors, they may nevertheless be especially adept at predicting what you will do. Merely finding yourself inclined toward an option may give you information, as your adversary may well have anticipated this very inclination. A number of authors have argued that cases like these can favor randomized actions.[7] Although the adversary may well have anticipated that you would play a randomized strategy, they cannot effectively capitalize on this fact. To the extent that such examples in fact occur, and to the extent that we find the recommendation compelling, we will need to show how they can be subsumed under our two rationales.

Finally, a related assumption that we will not be making is that the agent possesses a representation of uncertainty in the form of a probability function on $\mathcal{S}$. We would like to begin neutral on this matter, as again, one way of filling in more details of our scenario is that the agent lacks any such probability function, and this in turn may play into a putative justification of randomization. To foreshadow, the Bayesian argument will obviously hinge on whether we can

[6] An assumption of this sort is famously made by Savage (1954).

[7] Though decisions involving evidential dependence often drive a wedge between causal and other formulations of decision theory, both proponents and critics of causal decision theory have argued that randomization can be rational, e.g., in the famous Death in Damascus scenario (Harper, 1986; Ahmed, 2014). Related arguments arise in game theory; see discussion below.

assume the agent has an adequate prior probability function, and many arguments for randomization will focus attention on this very point. It is noteworthy that our second rationale—limited memory—is unrelated to this issue.

RATIONAL RANDOMNESS?

With so much motivation and background, it is worth briefly reviewing some of the familiar arguments for randomization. The details of these specific debates will not be as important as the broader intuitions that emerge from them.

Game Theory. The distinctive feature of game theoretic scenarios is that they involve multiple agents. The relevant state of the world in $\mathcal{S}$ is chosen by other intelligent agents whose interests, in the most extreme case, are perfectly antagonistic. One's win is another's loss. Even simpler than rock-paper-scissors, consider a version of the 'matching pennies' games in a penalty kick at a football match. The kicker wants to go left if the goalie is going right, and right if the goalie is going left. The goalie wants to go the same direction as the kicker.

Consider the goalie's deliberation. They should try to assess whether the kicker is more likely to go left or right, and then simply opt for that direction. Suppose the kicker went right the last three times, so the goalie predicts they will go right again. The kicker might suspect that the goalie will come to this conclusion, which would lead them to opt for left this time. Anticipating this, the goalie decides instead to go left, only to worry that this decision itself will have been anticipated by the kicker. One begins to suspect that there is no stable deliberative equilibrium that the goalie can ever hope to reach. In particular, the goalie's own deliberation seems to frustrate any transient prediction about what the kicker will do. A similar dilemma confronts the kicker.

Traditional game theory deals with scenarios like these under an assumption that each agent is 'ideally' rational: the agents both know what game they are playing and neither agent should simply be able to outwit the other. The operative normative concept in this context is that of a Nash equilibrium, a situation in which no agent has any motivation to deviate from their strategy, given the strategies of all the other players. Though such equilibria are guaranteed to exist, the strategies must in general be 'mixed', that is, randomized.[8] For example, in the football match scenario, the unique equilibrium is a situation in which each player randomly chooses left or right with equal probability. Given that each is playing such a strategy, neither has any motivation to deviate. Moreover, in this particular situation there is no other pair of strategies that enjoys such stability.

Aside from its role in idealized equilibrium theory, the recommendation to randomize in situations like these enjoys a certain air of plausibility. Particularly if there is a serious concern that the adversary may in fact be more shrewd or resourceful, randomization seems like a good *safety strategy.* It can provide a

[8] See, e.g., Leyton-Brown and Shoham (2008) for the proof. The formalism itself does not require interpreting mixed strategies as involving randomization at all, however. For other interpretations see, e.g., the discussion in (Osborne and Rubinstein, 1994, §3.2).

mechanism for protecting oneself against being outwitted. Rather than try to outguess the goalie, the kicker may be satisfied in knowing that they are at least as likely as not to pick the right direction. In their early textbook on game theory, Luce and Raiffa (1957) identify the virtue of mixed strategies in that they do not 'permit us to fall prey to our human frailty' (p. 75).

Experimental Design. As with game theory, the issues surrounding randomized controlled trials are complex. A helpful distinction to make up front is between *experiments to learn* and *experiments to prove* (Kadane and Seidenfeld, 1990). Consider an experiment to determine the causal effect of a proposed treatment for some medical condition. In an experiment to prove, the goal is to convince some third party, e.g., a government agency, of the causal effect. Typically, the third party will have access to the experimenter's strategy—including any randomization employed—and might only make their decision (accept or reject the claim, together with a cited justification) upon learning the strategy. In other words, such situations tend to violate our assumption of causal choice-state dependence. If this third party is adversarial, they may be inclined to find a flaw in any particular deterministic experimental design, or may simply have a blanket policy of rejecting any non-randomized design. Perhaps unsurprisingly, there are Bayesian arguments for randomization in experiments to prove (e.g., Berry and Kadane 1997; Banerjee et al. 2017).

Experiments to learn, by contrast, concern a sole individual who wants to make the best possible inference about the causal effect. Given a set of participants sampled from the population, the task might be to determine which participants will receive the treatment and which will receive a placebo. A worry is that there may be variables in the population—so called confounders—that exert independent causal influence on outcomes. Thus, if we observe that the treatment correlates highly with recovery, we want to make sure this is due to the treatment, and not, say, to the average age in the treatment test group being significantly lower. The existence of confounders frustrates the researcher's attempt at isolating the causal effect of the factor of interest.

Everyone will agree that any known confounder should be explicitly controlled, by appropriately balancing between the control and test groups. But what should be done once we account for these? A guiding intuition among proponents of randomization is that we simply cannot foresee all the possible confounders ahead of time, and it is quixotic—with serious potential consequences—to assume otherwise.[9] By letting the assignment depend on an external random source, it is guaranteed (in the limit, as the number of participants increases) that there will be no problematic confounders. In general, this is argued to make experimental design more objective, again less susceptible to human fallibility.

[9]Fisher (1935) argues that, 'it would be impossible to present an exhaustive list of such possible differences appropriate to any one kind of experiment, because the uncontrolled causes which may influence the result are always strictly innumerable' (p. 21).

Reinforcement Learning. A third, rather different intuition for randomization stems from the need to balance *exploitation* with *exploration* while learning about a dynamic environment by trial-and-error (e.g., Kaelbling et al. 1996). An illustrative example is ordering at a restaurant. Having ordered a certain dish before, you might know that it is likely to be good. You thus have some reason to exploit this known reward. However, for all you know, there could be much better items on the menu, which gives reason to explore for even higher rewards. The puzzle is how to balance the two. An influential idea is to select each option with probability proportional to how good you expect that option to be. This simple suggestion simultaneously guarantees that known high-reward actions will be exploited, but also that unknown actions will eventually be explored. It is also an easy strategy to carry out. More or less the same motivation—adding noise to escape 'local minima'—has been influential in other areas of learning as well.[10]

The reinforcement learning setting is inherently sequential. If it is your last visit to the restaurant, it may be most sensible to opt for what you think will be best. The intuition behind exploration is that it might be worth sacrificing this one meal for the sake of many future meals. Thus, one possibility, highlighted by tasks like reinforcement learning, is that randomization becomes especially useful when the action space actually consists of sequences of actions, perhaps interleaved with observed outcomes.

WHEN RANDOMIZATION PROVABLY CANNOT HELP

Given an action space $\mathcal{A}$—whether of moves in a game, assignment strategies in an experiment, or sequences of choices in a learning task—we would like some way of assessing how good a given action is, including any randomized action. Assume for the moment that $\mathcal{A}$ is finite, so that $\Delta(\mathcal{A})$ is the n-dimensional simplex on $\mathcal{A}$, and every decision $\delta \in \Delta(\mathcal{A})$ is a finite probabilistic mixture of non-probabilistic actions: $\delta = r_1 a_1 + \cdots + r_n a_n$, where each $r_i \in [0,1]$ and $\sum_{i \le n} r_i = 1$. This setup admits useful necessary conditions on a value function $V : \Delta(\mathcal{A}) \to \mathbb{R}$ for randomization to be strictly beneficial. V is said to be *convex* if for any two $\delta_1, \delta_2 \in \Delta(\mathcal{A})$, and every $r \in [0,1]$, we have

$$V\big(r\delta_1 + (1-r)\delta_2\big) \quad \le \quad rV(\delta_1) + (1-r)V(\delta_2). \tag{1}$$

If V is convex, then since any randomized strategy is a mixture of deterministic strategies, it will have value no greater than the greatest of these strategies. Hence, any argument to show that randomization is strictly useful will have to be based on a non-convex value function.

An important example of a convex value function is the expected utility expression from Bayesian decision theory. If we have a probability p on $\mathcal{S}$, then the function $E_p U(\delta) = \sum_s p(s)\big(\sum_a \delta(a) U(a,s)\big)$ is convex (even linear, meaning the

[10]See, e.g., Kirkpatrick et al. (1983). An early argument for randomization in machine learning appears in Turing (1950), who proposes 'it is probably wise to include a random element in a learning machine' (p. 459). In the same passage Turing anticipates the potential advantage of randomized strategies in memory efficiency, to be discussed further below.

inequality in (1) is an equality), ruling out the possibility that randomization will ever be strictly useful. A similar argument can be given when $\mathcal{A}$ and $\mathcal{S}$ are both infinite, indeed even when p is continuous, using Jensen's inequality (Lehmann and Casella, 1988, Corollary 7.9). This is the Bayesian argument against randomization. As soon as we have a prior on $\mathcal{S}$, assessing actions by their expected utilities renders randomization useless.

This observation buttresses all of the Bayesian counterarguments to the three proposals just canvassed. A Bayesian will take everything they know about the situation and devise a prior that codifies their uncertainty. The injunction is then simply to maximize the utility expected under that probability, which, as just observed, never requires randomization and often prohibits it.

In a strategic situation this means consolidating everything one knows about an adversary into a probability over the possible actions (which, recall, are the states in our setup). The fact that the situation involves another agent with opposed interests may help to sharpen this probability, but it does not pose any insurmountable difficulties distinct from cases where states are not chosen by another intelligent agent.[11] Playing a mixed equilibrium strategy is acceptable (again, not required) if your beliefs happen to align with the other player's equilibrium strategy. For instance, if you believe that your adversary is equally likely to play rock, paper, or scissors, then playing a uniformly random strategy is as good as (but also no better than) any other choice. Any deviation from this specific probability assignment, however, will generally result in the equilibrium strategy becoming strictly suboptimal. Indeed, if you think the adversary is even slightly more likely to play rock than paper or scissors, then deterministically choosing paper is strictly better than any randomized choice.

In experimental design the experimenter has a goal to learn as much as possible about the causal relation between treatment and outcome. Given their quantified uncertainty about this relation, and about all the possible confounders in the population, some actions (i.e., assignments) will be expected to provide more information than others. The Bayesian recommendation is therefore to choose an assignment that maximizes expected information gain.[12] As a very simple (if artificial) example, imagine an experiment with four participants. Suppose that our experimenter feels absolutely certain (believes with probability 1) that there is only one attribute that could possibly correlate with the outcome, and that exactly two of the four participants possess that attribute. A Bayesian experimenter would then insist on balancing this attribute between the test and control groups: one with and one without the attribute in each condition. Randomizing among the four ways of doing this is acceptable (by no means mandatory), while any further randomization in allocation threatens to provide strictly less information to the experimenter about the causal effect of interest.

Finally, in a learning context a Bayesian will have prior beliefs about the

[11]Kadane and Larkey (1982) present perhaps the most forceful argument of this sort, with a rejoinder in the same journal issue by game theorist John Harsanyi.

[12]This basic argument has been made many times over, early on by Savage (1954, §9.3).

reward distributions for the possible actions, as well as posterior beliefs given observations of those actions' consequences. This already means that there will be a Bayes-optimal strategy for balancing exploration and exploitation, guaranteeing the maximum expected long-run payoff (e.g., Bellman 1957). For example, if you can already foresee all possible sequences of orders at the restaurant and all ways those sequences of orders might play out (weighted by their probabilities), you should simply choose dishes in a way that achieves the highest overall expected reward. After all, balancing exploration and exploitation is only a means toward the end of performing the best possible sequence of actions given your uncertainty. Why should an agent do anything other than maximize overall expected reward?

REASONING WITH LIMITED RESOURCES

However one reacts to these debates, it must be admitted that the Bayesian response does not address a common feature in all three arguments for randomization, namely the focus on human limitations. Quantifying one's uncertainty in a sensible way is not always easy, and this is not to mention the challenges involved in calculating expected utilities for all the possible actions. Game theory, experimental design, and reinforcement learning all involve scenarios where attempts to formulate a good prior are somehow frustrated. Competitive situations with clever adversaries raise the possibility of evidential choice-state dependence. In experimental contexts, even when one does have prior convictions about the possible alternative causes of the effect, the experimenter will (or should) often lack higher-order confidence in these convictions.[13] Devising a sensible prior in reinforcement learning settings can be difficult, and at any rate it is computationally intractable to solve sequential decision problems exactly (specifically the problem is in the class PSPACE; see Papadimitriou and Tsitsiklis 1987).

These considerations do not by themselves show randomization is ever to be preferred, but they do undermine the Bayesian argument against it. Of course, this is in no interesting sense incompatible with the traditional Bayesian stance (at least in statistical decision theory). The broad desirability of codifying one's uncertainty in the form of a coherent probability measure is perfectly compatible with the obvious fact that decisions must sometimes be made without sufficient resources to formulate such a measure.[14] At the point when a decision must be made, the agent will inevitably have to settle on whichever option appears best at that point. Could such an option ever involve randomization? It could, and we have already identified several candidate scenarios where the argument sounds compelling. For instance, our football kicker will not have the luxury of devising a strategy that would reasonably guarantee better than one-half chance of success, while this level is guaranteed by the fully randomized strategy.[15]

[13] Robustness against misguided priors is one of the main arguments in favor of randomization in the influential paper by Rubin (1978).

[14] Essentially all of the founders of Bayesian decision theory were quite clear on this point. See, e.g., the candid discussion in Savage (1967).

[15] Interestingly, both goalies and kickers do appear to employ mixed strategies; see Chiappori et al. (2002). We will return to empirical evidence about randomizing behavior below.

The specific features of randomized strategies that make them good candidates as 'baseline' defaults for a given problem—unpredictability, simplicity, independence, etc.—are of course important and potentially complex. But the general principle is not especially deep. When should an agent randomize? The first answer is simply: when there is no further opportunity to think more and the currently best-looking option involves randomizing. The expectation that further thought and reflection, ideally leading to a sensible probability measure, would result in deterministic choice does nothing to call this principle into question.

If the most influential arguments for randomization are special cases of this general principle, it seems to raise the distinct possibility that randomization is nothing more than an expedient heuristic, appropriate only for those situations when further deliberation about what to do is impossible or otherwise frustrated. Jaynes (2003) gave voice to this very attitude: 'Whenever there is a randomized way of doing something, there is a nonrandomized way that yields better results from the same data, but requires more thinking' (p. 532). Is this true, or is there yet a deeper justification for randomization? We have so far focused on obstacles to formulating a prior and more generally to figuring out the best course of action. A natural suggestion is that randomization might help because carrying out a course of action can itself be costly. We now turn to this possibility.

COMPUTATIONAL COMPLEXITY

Having already observed that randomization can be reasonable when an agent lacks resources for ideal deliberation—and that this comports with prominent arguments in its favor—let us now imagine that deliberation is costless. Suppose, for instance, that we are designing an agent for some environment. We know the utility structure, and we might even have a prior over environments. From our perspective as agent designers, with all the time and resources we need, would we ever introduce randomization in the agent design? A natural hypothesis, which has animated an active subfield of theoretical computer science, is that randomization may render hard computational problems more feasible.

It will be helpful to refine our space of probabilistic actions $\Delta(\mathcal{A})$ to a space $\mathcal{M}$ of *probabilistic Turing Machines* (PTMs),[16] which we will think of as processing some input x (encoded as a binary string) and then (perhaps noisily) giving some output $a \in \mathcal{A}$, which will be taken as the chosen action. Thus, an element of $\mathcal{M}$ specifies both the action and the deliberation leading to it. Ultimately we are interested in the action a, which should be appropriate to the input x, but we also may care about the resources—space and time—consumed in determining a.

A PTM effects a mapping from inputs to probability distributions on actions. Call this a *behavior*. As long as we ignore processing costs, and the space of possible inputs is finite, the behavior of any PTM can be emulated by a probabilistic mixture of deterministic Turing machines. This result, known in game

[16] Putnam (1967), who forcefully introduced the very idea that the mind could be likened to a computing device, grounds his account on probabilistic Turing machines.

theory as Kuhn's Theorem,[17] shows that merely moving to the setting of Turing machines does not defeat the Bayesian argument against randomization: the performance of a probabilistic Turing machine can be no better than that of the best deterministic Turing machine in the behaviorally identical mixture.

However, the field of randomized complexity theory lifts both of these assumptions at once: processing costs now matter and inputs may be of any size. The second modification is already enough to undermine Kuhn's Theorem: even ignoring costs, there are infinitary behaviors of PTMs that cannot be perfectly imitated by a probabilistic mixture of deterministic machines.[18] The question is whether there is ever reason to design an agent that instantiates one of these quintessentially probabilistic machines.

A typical problem in computational complexity theory is to determine membership in some distinguished set X of inputs. A canonical notion of *feasible time approximation* (Gill, 1977) can be construed as defining a value function on machines, $V_X : \mathcal{M} \to \{0, 1\}$, for each X. Let us say that $V_X(M) = 1$ just in case:

1. For all $x \in X$: M returns 1 on x with probability at least .99.[19]
2. For all $x \notin X$: M returns 1 on x with probability at most .01.
3. There is a polynomial function $f : \mathbb{N} \to \mathbb{N}$, such that for all x: M always halts on x within $f(|x|)$ steps, where $|x|$ is the length of x.

When restricting to deterministic machines, this notion of feasibility collapses to the well-known class of polynomial time algorithms. It was long hoped that there would be problems X such that $V_X(M) = 1$ only for probabilistic machines—a conjecture known as P≠BPP—meaning that randomization truly helps in this setting. Yet, a number of developments in the field have convinced most researchers that P=BPP, that randomization does not help (Vadhan, 2012).

A similar question can be raised with respect to memory usage. If we replace 3 above by the requirement that there be a logarithmic function g such that M never uses more than $g(|x|)$ memory cells during the computation, this formalizes a canonical notion of feasible space approximation. The analogous conjecture for space complexity—known as L≠RL—is also widely believed to fail (see Vadhan 2012 for further discussion).

These considerations suggest we are unlikely to find a justification for randomization in this particular approach to complexity theory. It is sometimes concluded that this would mean, from a computational perspective, randomization has nothing but heuristic value.[20] This conclusion is premature.

[17]See Kuhn (1950). While the argument is not given in terms of Turing machines, it is straightforward to see that the proof relativizes to the computable setting.

[18]Consider a PTM that on input n returns each n-bit number with probability 2^{-n}. Any distribution over deterministic Turing machines will assign some M probability $\epsilon > 0$. Now choose n with $2^{-n} < \epsilon$. On input n the machine will return $M(n)$ with probability greater than 2^{-n}. This observation was already made by Meggido (1994).

[19]The threshold of .99 is arbitrary. Any threshold greater than 0.5 can be chosen.

[20]In the canonical textbook on artificial intelligence, for example, the authors write, 'In single-

REASONING WITH A BOUNDED MEMORY

A distinctive challenge facing virtually any intelligent agent is that the amount of data they will receive is vastly greater than the memory space they have available. From a decision theoretic perspective, this means that violations of the Principle of Total Evidence—according to which one's probability estimates and decisions should be based on all available information—will be inevitable. What should an otherwise perfect Bayesian agent do if they know they will not be able to process all of the data they will receive? As it turns out, this sole relaxation of the idealized picture is already sufficient to rationalize randomized behavior.

As a first illustration, consider a hypothesis testing scenario with K states $\mathcal{S} = \{1, \ldots, K\}$, each $i \leq K$ associated with a Bernoulli probability p_i of 'heads' (and $1 - p_i$ of 'tails'). After some observed coin flips, the aim is to guess the right hypothesis. Hypothesis testing is a central and ubiquitous task and has been argued to characterize fundamental cognitive problems ranging from concept learning (Fodor, 1975) to perceptual object recognition (Kersten et al., 2004).

When each hypothesis is equiprobable, the optimal solution is of course to guess the hypothesis that assigns highest likelihood to the observed sequence. Yet for an agent with a limited memory even this simple strategy is generally unavailable. Such an agent can only be in one of finitely many different states. It is therefore natural to model them as a (probabilistic) *finite state automaton.*

Whereas a Turing machine possesses an infinite tape that it can use as an unbounded memory buffer, a finite state automaton is simply (equivalent to) a Turing machine with a finite work tape and thus a fixed upper bound on memory space (see, e.g., Minsky 1967).[21] In the present context we think of the agent as starting in some initial state and then transitioning to other states depending on the data point most recently observed. Each state of the agent, we can assume, is associated with exactly one hypothesis $i \in \mathcal{S}$ ('The current guess is i'), though for each hypothesis $i \in \mathcal{S}$ there may be many possible agent states in which i is the current guess. For a fixed number n, we want to ask: what is the best automaton with exactly n states for a given hypothesis testing problem? It was shown in early work by Hellman and Cover (1970, 1971) that allowing stochastic transitions in the automaton is strictly helpful for hypothesis testing.[22]

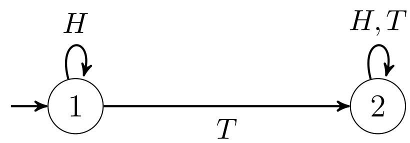
To see the intuition, consider a toy example of just two hypotheses, with $p_1 = 0.99$ and $p_2 = 0.9$. In neither case do we expect to see many tails. We nonetheless expect to see it ten times more often if hypothesis 2 is true. Suppose the agent can be in one of only two mental states at any given moment. What

agent environments, randomization is usually *not* rational. [...] In most cases we can do much better with more sophisticated deterministic agents' (Russell and Norvig, 2010, p. 50).

[21]Note that Putnam (1967) originally referred to probabilistic Turing machines as 'probabilistic automata', though he did not impose any restriction on memory.

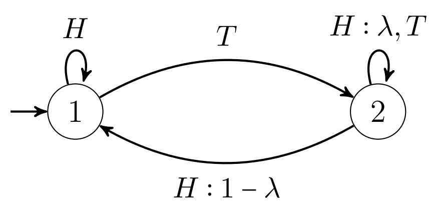
[22]The notion of optimality in these papers is asymptotic. An agent's performance is given by the expected limiting proportion of correct guesses, a value that always exists for finite state automata. This work has been extended along a number of lines, including more general decision problems, recently in Wilson (2014).

is the optimal strategy for solving this problem? The best a deterministic agent with only two states can do is to begin by guessing hypothesis 1, and then the first time tails is ever observed guess hypothesis 2 forever thereafter. That is, the following automaton is optimal for this problem:



The agent begins on the left hypothesizing 1, remaining there as long as heads is observed. Upon observing tails, the agent transitions to the right state, remaining there forever. On average, after 100 observations such an agent expects to make the correct guess about 68% of the time, certainly well above chance. After 200 observations this drops to 57%, and by 500 observations we expect any two-state deterministic automaton to be scarcely better than chance.

Suppose now that our agent can undergo probabilistic transitions between states. Is there any way to improve upon the performance of the best deterministic agent? A general result by Hellman and Cover shows that the following probabilistic automaton is optimal, where in this case $\lambda \approx 0.97$.[23]



The agent again begins on the left, guessing 1 as long as heads are observed. Upon observation of the first tails, the agent transitions to the right state, now guessing 2. Unlike in the deterministic automaton, from state 2 a heads observation will, with very small probability ($1 - \lambda \approx 0.03$), lead back to guessing 1.

The role of randomization here is similar to typical learning settings, in that it provides an inexpensive means of escaping hypotheses that are decreasingly likely given the data. Unlike all deterministic automata of its size, this probabilistic automaton in this example is expected to be correct at least 77% of the time, for any number of data points above a few dozen, again assuming p_1 and p_2 are equally likely. The same overall pattern arises for any finite numbers of states and hypotheses and other prior distributions on hypotheses.

The hypothesis testing problem is already reasonably general, but the underlying rationale for randomization in this setting is much more general and can be explained more abstractly. Fixing the space of actions $\mathcal{A}$ and a finite space of possible observations $\mathcal{O}$, recall a behavior is a function from $\mathcal{O}$ to $\Delta(\mathcal{A})$. Kuhn's

[23] In general $\lambda = 1 - \left(\frac{(1-p_1)(1-p_2)}{p_1 p_2}\right)^{\frac{1}{2}}$. See Hellman and Cover (1970, 1971).

Theorem tells us that any behavior encoded by a PTM can be achieved by a probabilistic mixture of deterministic machines. But now suppose we restrict attention to the space $\mathcal{F}_n$ of probabilistic finite automata of size at most n. Kuhn's Theorem now fails, in that there will be automata in $\mathcal{F}_n$ whose behavior cannot be emulated by any mixture of deterministic machines in $\mathcal{F}_n$. The example above illustrates this failure. If $\mathcal{O}$ includes all possible sequences of 500 coin flips, then the behavior of every deterministic two-state machine leads to chance performance. Hence, no mixture of such machines could reach the performance of 77% achieved by the best probabilistic machine.

If we think of assessment as now being defined over the space of finite state machines, there will often be value functions $V : \mathcal{F}_n \to \mathbb{R}$ whose maxima are concentrated around randomized machines whose behaviors are unachievable by deterministic machines in $\mathcal{F}_n$. Thus, the failure of Kuhn's Theorem defeats the Bayesian convexity argument against randomization.[24] Whether we are in such a case will of course depend on a number of details, and in particular on the nature of the underlying value function. It is worth giving several further examples of the general phenomenon to illustrate how widespread it is.

The same phenomenon can be witnessed in mundane, small-scale problems. Recall our opening example of the choice between skipping or stopping at a stop sign. Imagine the agent will meet Y stop signs, and the probability of a patrol officer at each is independent, estimated at 0.1. Each ticket would provide disutility -20. Suppose the disutility of stopping is not additive but multiplicative—wear and tear on the car is worse with each stop—so that stopping K times gives disutility $-(1.1^K)$. In a one-shot case, i.e., when $Y = 1$, it is easy to calculate that the agent should stop. For the case of $Y = 35$, however, the Bayes-optimal strategy is to stop 32 times and skip 3 (in any order)—the risk of tickets on those three occasions is outweighed by the compounded cost of stopping.

Carrying out any version of this (deterministic) strategy requires remembering previous actions. Suppose our agent has no memory for past actions whatsoever. Then a randomized strategy of skipping each time with probability 0.1 ($\approx \frac{3}{35}$) dominates both memoryless deterministic strategies (always skip or always stop). The reason is again simply that randomization affords the agent behaviors that it otherwise could not effect. The aim is to skip roughly 10% of the time, and the only way to come even close to this behavior (with high probability) without memory is by appeal to the randomizing device.

One final, somewhat different example of this same phenomenon, where limited memory renders randomization strictly beneficial, is in the area of estimating statistics from large data streams. Imagine an agent will observe many data points, coming from some class of possible observations $\mathcal{O}$, and the task is to

[24] Early in the development of modern game theory it was realized that imperfect recall could lead to failure of Kuhn's Theorem, and could even rationalize randomized behavioral strategies, e.g., Isbell (1957). See also Piccione and Rubinstein (1997). The broader significance of these observations to decision-making more generally, and their connection to conceptually quite different computational issues, seems not to have been noticed.

estimate the number of distinct members of $\mathcal{O}$ observed. Envision, for instance, a restaurant owner who expects to see many thousands of customers throughout the year, many of whom visit multiple times. By the end of the year, the owner wants to know how many distinct customers came into the restaurant.

Let us assume that a practical solution to this problem can only use an amount of memory sublinear in the size of $\mathcal{O}$ and the length of the stream. Thus, the restaurant owner cannot simply keep a tally with the number of distinct customers. What would be a good strategy to guarantee a reasonable estimate at the end of the year? It is known that there is no deterministic approach to this problem that even approximates the correct answer for every stream of a given fixed size, whereas a uniformly close approximation is possible if we allow randomization.[25] This highlights another known instance where a desirable probabilistic behavior is not simply a mixture of comparably memory-efficient deterministic behaviors. An agent facing a problem of this character would want to execute a good randomized strategy, if possible.

In sum, any concrete agent will have a finite memory, indeed a memory that will typically appear paltry in comparison with the vast amount of data they will have to confront. This relatively mild presumption is already enough to show how and why randomization may be strictly beneficial. As the examples make apparent, the theoretical possibility is plausibly common, from statistical inferences to sequential decision problems to basic data processing tasks.

WHENCE RANDOM BITS?

We have identified two reasons to randomize: (1) establishing the best course of action (including formulating a prior on states) can be costly, and (2) memory space is finite. Both evidently affect typical agents, whether biological or artificial, and we should thus expect randomization to be beneficial for a wide variety of agents and circumstances. Granting (1) and (2) as compelling rationales, we can now ask two further questions. First, how could an actual agent hope to carry out a randomized strategy? Second, do we already find evidence of intelligent organisms responding to (1) and (2) with appropriately random behavior?

The first question highlights the multitude of characteristics that a perfectly random source will embody, only some of which will be requisite for a given task. Fooling an adversary requires unpredictability; unbiased assignment in an experimental trial requires lack of correlation with traits in the population; and so on. For some tasks very simple ad hoc chance devices such as two-sided coins, or parity of the second-hand on a clock, will be sufficient. Other tasks are more demanding in what counts as suitably random. As an extreme example, cryptographic problems, where strategies must be unpredictable by virtually any conceivable computing agent, seem to demand methods whose very existence depends on ma-

[25] This and related landmark results are due to Alon et al. (1999). One good randomized strategy for this problem is quite simple. Begin by randomly choosing a 'hash' function $h : \mathcal{O} \to [0,1]$. Initialize a variable $Z \coloneqq 1$. On each new data point o from $\mathcal{O}$, set $Z \coloneqq \min\big(Z, h(o)\big)$. At the end return $1/Z$.

jor unproven conjectures in computational complexity theory. As Donald Knuth famously quipped about pseudorandom number generators, 'Random numbers should not be generated with a method chosen at random' (Knuth, 1969, 6).

It would be convenient, if quixotic, to imagine that human beings and other animals were capable of behaving more or less randomly without any external aid. To be sure, the kicker does not have recourse to coins or even hands of a clock during that fraction of a second when a choice must be made. Dating back as far as Reichenbach (1949) (§3, 'Normal Sequences'), many have been skeptical of the idea that people could act at random, and the skepticism was buttressed by empirical work claiming as much (Wagenaar, 1972). Reflecting this common wisdom Arntzenius (2008) writes, as a matter of course, 'It is not as if one has a chance device stored away in some convenient part of one's brain' (p. 292).

Yet there is evidence of something like this. Cortical neurons are known to encode magnitudes such as orientation by their average firing rates, and these firing rates are a fairly reliable indicator of the stimulus. When it comes to the precise firing pattern underlying this average, however, a number of studies have found the pattern to be best modeled as a genuinely random process.[26] Some studies have probed further in an effort to ascertain what kinds of unmeasured deterministic processes might ground this apparently stochasticity. One notable finding is that firing patterns seem to be a deterministic function of membrane potential, as established by inputs from neighboring synapses (Mainen and Sejnowski, 1995), which raises the possibility that the apparent stochasticity might be traced all the way to the molecular level. Such findings paint a rather different picture, on which random behavior is a kind of default that the nervous system must somehow mitigate in the direction of more determinate and deliberate activity. At the same time, the extent to which this randomness in neural firing manifests in observable behavior, not to mention whether it is in some sense available to an organism for strategic purposes, is admittedly very much an open question.[27]

In any event, despite the common wisdom that people are unable to act randomly, there are more direct behavioral studies on humans and non-human animals suggesting otherwise. In one study with human participants (Rapoport and Budescu, 1992), researchers found that people will generate sequences that pass relatively stringent statistical tests provided they are put in an adversarial situation involving monetary payoffs. In most previous experiments participants were explicitly asked to produce 'random' sequences, resulting in far too many easily detectable patterns, e.g., too few long runs of the same move. In these experiments, by contrast, the task is not explicitly to produce a sequence with a given property, but rather to gain as much reward as possible throughout the sequence of plays. Such strategic scenarios elicit behavior that largely avoids most of the easily detectable patterns.

[26]Specifically, the patterns are best modeled by a Poisson distribution fixed only by the average rate itself. That is, holding the average rate fixed, neural firing patterns seem to be completely random. The first work demonstrating this was Tolhurst et al. (1983).

[27]See Glimcher (2005) for an illuminating discussion of the issue, and many other references.

It is perhaps telling that a small number of participants in Rapoport and Budescu's experiments were able to achieve as good or better performance, compared to the random strategy, by relying on fairly simple deterministic strategies. When one's adversary is playing a fully randomized strategy, any strategy is as good as any other. At the same time, any detectable pattern that deviates from the fully random strategy could inform a reasonable probability about what they are likely to do next. If one is confident in the pattern, as discussed earlier, a deterministic strategy will be strictly better. While most participants seem to have attempted a defensive safety strategy, some appear to have believed (in some cases correctly) that they could improve on this baseline by capitalizing on predictable deviations from the safety strategy. (Of course, this also means that some non-random strategies performed worse than the random safety strategy.)

Even more remarkable behavior has been observed in non-human animals. Using frequency-dependent reward schedules, researchers have been able to test whether *variability* in behavior could be reinforced and encouraged in pigeons (Page and Neuringer, 1985). The study described here is from Machado (1993). Given two choices—press the left key (L) or the right key (R)—in succession, a bird will come to produce a sequence. The goal is for the sequence to be *normal* at level k, meaning that the proportion of all k-length subsequences should be uniform. One can think of this on the model of a search problem. Imagine there is some 'prize' k-length sequence, such that producing that sequence gives positive probability of a large prize. Assuming all k-length sequences are equally likely, the goal is to produce each subsequence the same number of times so as to maximize the overall chance of a prize. In these experiments, reward in the form of food is delivered throughout the trial, with probability of reward proportional to how balanced the produced sequence is among subsequences of length k.

There exists an optimal deterministic strategy for this task, provided by so called de Bruijn sequences. In the case of $k = 1$, the optimal strategy is simple alternation, LRLRLRLR..., guaranteeing the same number of L as R throughout. Pigeons routinely learn this strategy. For $k = 2$, the optimal strategy is slightly less obvious: alternating LLRRLLRR... balances all subsequences of length 2. With slightly more difficulty, pigeons are able to learn this strategy (Machado, 1993, Exp. 1). The answer for $k = 3$, however, is not at all obvious: repeating the sequence LLLRLRRR is the optimal strategy. No pigeons in this study demonstrated such behavior. More interesting, however, is the finding that almost all of the birds in this experiment behaved in a way that was eventually indistinguishable from a random Bernoulli process. The produced sequences passed a barrage of statistical tests, suggesting that they had learned to achieve variability by acting more or less randomly (Machado, 1993, Exp. 2).

A basic fact about the de Bruijn sequence LLLRLRRR is that repeating any strictly shorter sequence will necessarily omit at least one 3-length sequence altogether. This means that any deterministic strategy for this problem which uses strictly less memory than what is required for the optimal strategy will necessarily lead to significant imbalance. As shown by the cases of $k = 1$ and $k = 2$, pigeons are adept at finding optimal deterministic solutions when those are feasible. But remarkably, when the memory requirements become too demanding, the pigeons in this study latch on to the optimal alternative, which in this case is simply to act randomly, all but guaranteeing approximate balance for long enough sequences. Other studies have shown very similar patterns.[28]

Our two proposed rationales, (1) and (2), may certainly both be at play in a given scenario. In this case, for instance, not only is the memory demand high for carrying out the optimal strategy, but identifying the de Bruijn sequence itself can be difficult. If an agent's search is biased toward memory-efficient strategies in the first place, as certainly seems reasonable, then the resulting behavior may be naturally explained by appeal to both rationales.

To the extent that an agent demonstrates random behavior in a decision problem with causal choice-state independence, and to the extent that this random behavior seems rationally defensible, the claim is that we should expect such behavior to be explained by appeal to at least one of the two proposed rationales. The empirical phenomena discussed here provide some support for the claim. When feasible and clearly favorable, deterministic strategies are preferred. But when memory requirements become too demanding, or when concerns about being outwitted frustrate formulation of a compelling prior, organisms as diverse as people and pigeons are evidently able to utilize (approximate) randomization. Whether this behavior is ultimately traceable to stochasticity underlying neural firing patterns, or improvised association with some other suitably indeterminate and independent source, the end effect is close enough to what one would expect, rationally speaking, even given costless access to a perfectly random device.

CONCLUSION

It is sometimes assumed, following a familiar Bayesian decision theoretic argument, that randomization is never rationally required, aside perhaps from its marginal role in tie breaking. One of the founders of decision theory was nevertheless quick to point out, 'The need for randomization presumably lies in the imperfection of actual people' (Savage, 1962, 34). This admission leaves open which 'imperfections'—that is, which assumptions in the Bayesian argument—are most pivotal.[29] Difficulties in formulating a prior highlight an important class of cases, which fuel many of the familiar arguments in favor of randomization, from game theory, experimental design, and the theory of learning. While the many debates continue about exactly when and how one can sensibly codify one's

[28] For instance, Page and Neuringer (1985) had pigeons produce a series of separate sequences of length 8, and they were rewarded for the number of distinct 8-length sequences produced. While there are of course many optimal deterministic solutions to this problem, these all require significant memory. This task again induced apparently stochastic behavior.

[29] Savage himself drew attention to the multiagent setting, continuing that the need may lie 'in the fact that more than one person is ordinarily concerned with an investigation' (p. 34). As discussed earlier, situations involving causal choice-state dependence are widely acknowledged to legitimize randomization. Our interest has been in situations without such dependence.

uncertainty by means of a prior probability measure—and thus about when randomization should even be considered as a viable alternative—it is undeniable that there will be instances when a good 'satisficing' solution (Simon, 1956) to a problem involves randomizing.

Perhaps more striking is the fact that, even if we idealize away the difficulties involved in deliberation, the mere observation that agents are limited by a finite memory—which in turn limits the strategies they may be able to carry out—is enough to undermine the argument against randomization. This is not only due to cases of absent-mindedness or mundane forgetting. It arises as a fundamental facet of real-world agency: the amount of potentially relevant information confronting an agent is vastly greater than the amount of memory available. Thus, even from the perspective of an ideal (resource-unlimited) Bayesian agent choosing an automaton for a given task, there will be cases where one would strictly prefer a probabilistic automaton over any deterministic one. It should therefore be no surprise that even carefully constructed artificial agents, honed through decades of research and practical experience, would still involve randomization at multiple levels. The same can be said for biological agents. There is a precise sense in which certain randomizing agents could not be improved in any way, short of augmenting them with additional memory capacity.

The focus in this article has been on strategic aspects of randomization, viewed as a potential means toward an end. Lifting the assumption we have been making throughout, that the relevant states do not causally depend on the agent's choice, reveals further subtleties around the question of when random behavior is desirable. For many purposes randomization is itself seen as inherently either valuable or objectionable. One of the very features of randomization that makes it initially questionable from a decision theoretic perspective—that the agent is essentially relinquishing authorship, and perhaps also responsibility, for the resulting action—can be a virtue. Randomized decisions are often seen as more fair, and in some contexts less agonizing, precisely because any potential link between the resulting action and a possible motivating reason is severed (Gauthier, 1965). By the same token, echoing the sentiment expressed by Aumann and others, even in situations where two options are judged to be on a par, which we might expect to yield indifference, people are nevertheless reluctant to relegate their decision to a chance device if the possible outcomes are significant and morally charged (Keren and Teigen, 2010).[30] The issues here are complex, involving responsibility, reasons for action, agential authority, and other weighty topics. To assess the proper role of randomization in human agency broadly—not to mention its desired role in artificially engineered agents—these rich complexities must be confronted and ultimately reconciled with the strategic aspects examined in this article. Understanding the purely instrumental value of randomization is an

[30]How exactly participants in these experiments manage to break the apparent symmetry is another empirical question, also bearing on significant normative questions around the role of agency in parity cases, cf. Chang (2009).

important step toward that end.[31]

REFERENCES

Ahmed, A. (2014). "Dicing with death." *Analysis*, 99(4), 587–592.

Alon, N., Matias, Y., and Szegedy, M. (1999). "The space complexity of approximating the frequency moments." *Journal of Computer and System Sciences*, 58(1), 137–147.

Arntzenius, F. (2008). "No regrets, or: Edith Piaf revamps decision theory." *Erkenntnis*, 48, 277–297.

Aumann, R. (1987). "Correlated equilibrium as an expression of Bayesian rationality." *Econometrica*, 55(1), 1–18.

Banerjee, A., Chassang, S., Montero, S., and Snowberg, E. (2017). "A theory of experimenters. NBER Working Paper 23867.

Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.

Berry, S. M. and Kadane, J. B. (1997). "Optimal Bayesian randomization." *Journal of the Royal Statistical Society, B*, 59(4), 813–819.

Borel, E. (1923). "Sur les jeux où interviennent le hasard et l'habileté des joueurs." *Théorie des probabilités*, J. Hermann, ed., Paris: Librairie Scientifique.

Bratman, M. E. (2001). "Two problems about human agency." *Proceedings of the Aristotelian Society*, 101, 309–326.

Chang, R. (2009). "Voluntarist reasons and the sources of normativity." *Reasons for Action*, D. Sobel and S. Wall, eds., Cambridge University Press, 243–271.

Chiappori, P.-A., Levitt, S., and Groseclose, T. (2002). "Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer." *American Economic Review*, 92(4), 1138–1151.

Eagle, A. (2005). "Randomness is unpredictability." *The British Journal for the Philosophy of Science*, 56(4), 749–790.

Fisher, R. A. (1935). *The Design of Experiments*. Edinburgh: Oliver and Boyd.

Fodor, J. A. (1975). *The Language of Thought*. Harvard University Press.

[31] This article benefited greatly from conversations with numerous individuals, including R.A. Briggs, Rosa Cao, Frederick Eberhardt, Konstantin Genin, Joseph Halpern, Wesley Holliday, Krzysztof Mierzewski, Brian Skyrms, Jacob Steinhardt, and Kevin Zollman. Thanks also to audiences at Cambridge University, Columbia University, Fresno City College, and UC San Diego. I am especially indebted to Percy Liang and Mary Wootters for formative discussion.

Gauthier, D. P. (1965). “Rule-utilitarianism and randomization.” *Analysis*, 25(3), 68–69.

Gill, J. T. (1977). “Computational complexity of probabilistic Turing machines.” *SIAM Journal of Computing*, 6(4), 675–695.

Glimcher, P. W. (2005). “Indeterminacy in brain and behavior.” *Annual Review of Psychology*, 56, 25–56.

Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge University Press.

Harper, W. (1986). “Mixed strategies and ratifiability in causal decision theory.” *Erkenntnis*, 24(1), 25–36.

Hellman, M. and Cover, T. (1970). “Learning with a finite memory.” *The Annals of Mathematical Statistics*, 41(3), 765–782.

Hellman, M. and Cover, T. (1971). “On memory saved by randomization.” *The Annals of Mathematical Statistics*, 42(3), 1075–1078.

Isbell, J. R. (1957). “Finitary games.” *Contributions to the Theory of Games, Vol. III, Annals of Mathematics Study*, M. Drescher, A. Tucker, and P. Wolfe, eds., Vol. 39, Princeton University Press.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press.

Kadane, J. B. and Larkey, P. D. (1982). “Subjective probability and the theory of games.” *Management Science*, 27, 113–120.

Kadane, J. B. and Seidenfeld, T. (1990). “Randomization in a Bayesian perspective.” *Journal of Statistical Planning and Inference*, 25(3), 329–345.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). “Reinforcement learning: A survey.” *Journal of Artificial Intelligence Research*, 4(1), 237–285.

Keren, G. and Teigen, K. H. (2010). “Decisions by coin toss: Inappropriate by fair.” *Judgment and Decision Making*, 5(2), 83–101.

Kersten, D., Mamassian, P., and Yuille, A. (2004). “Object perception as Bayesian inference.” *Annual Review of Psychology*, 55, 271–304.

Kirkpatrick, S., Gelatt, C., and Vecchi, M. (1983). “Optimization by simulated annealing.” *Science*, 220(4598), 671–680.

Knuth, D. (1969). *The Art of Computer Programming*, Vol. 2. Addison-Wesley.

Kuhn, H. W. (1950). "Extensive games." *Proceedings of the National Academy of Sciences*, 36(10), 570–576.

Lehmann, E. and Casella, G. (1988). *Theory of Point Estimation.* New York: Springer Verlag.

Leyton-Brown, K. and Shoham, Y. (2008). *Essentials of Game Theory: A Concise Multidisciplinary Introduction.* Morgan & Claypool.

Luce, D. and Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey.* Wiley.

Machado, A. (1993). "Learning variable and stereotypical sequences of responses: Some data and a new model." *Behavioral Processes*, 30, 103–130.

Mainen, Z. F. and Sejnowski, T. J. (1995). "Reliability of spike timing in neocortical neurons." *Science*, 268, 1503–1506.

Meggido, N. (1994). "On probabilistic machines, bounded rationality, and computational complexity." *Essays in Game Theory: In Honor of Michael Maschler*, N. Meggido, ed., New York: Springer Verlag.

Minsky, M. (1967). *Computation: Finite and Infinite Machines.* Prentice Hall.

Osborne, M. J. and Rubinstein, A. (1994). *A Course in Game Theory.* MIT Press.

Page, S. and Neuringer, A. (1985). "Variability is an operant." *Journal of Experimental Psychology: Animal Behavior Processes*, 11(3), 429–452.

Papadimitriou, C. H. and Tsitsiklis, J. N. (1987). "The complexity of Markov decision processes." *Mathematics of Operations Research*, 12(3), 441–450.

Piccione, M. and Rubinstein, A. (1997). "On the interpretation of decision problems with imperfect recall." *Games and Economic Behavior*, 20(1), 3–24.

Putnam, H. (1967). "Psychological predicates." *Art, Mind, and Religion*, W. H. Capitan and D. D. Merrill, eds., Pittsburgh University Press.

Rapoport, A. and Budescu, D. V. (1992). "Generation of random series in two-person strictly competitive games." *Journal of Experimental Psychology: General*, 71(3), 352–363.

Reichenbach, H. (1949). *Theory of Probability.* University of California Press.

Rescher, N. (1959). "Choice without preference." *Kant Studien*, 51(1-4), 142–175.

Rubin, D. B. (1978). "Bayesian inference for causal effects: The role of randomization." *The Annals of Statistics*, 11(1), 34–58.

Russell, S. and Norvig, P. (2010). *Artificial Intelligence: A Modern Approach.* Prentice Hall.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley and Sons.

Savage, L. J. (1962). "Subjective probability and statistical practice." *The Foundations of Statistical Inference*, M. Bartlett, ed., London: Methuen, 33–34.

Savage, L. J. (1967). "Difficulties in the theory of personal probability." *Philosophy of Science*, 34(4), 305–310.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., and Hassabis, D. (2018). "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play." *Science*, 362(6419), 1140–1144.

Simon, H. A. (1956). "Rational choice and the structure of the environment." *Psychological Review*, 63(2), 129–138.

Thompson, W. R. (1933). "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples." *Biometrika*, 25(3-4), 285–294.

Tolhurst, D. J., Movshon, J. A., and Dean, A. F. (1983). "The statistical reliability of signals in single neurons in cat and monkey visual cortex." *Vision Research*, 23, 775–785.

Turing, A. M. (1950). "Computing machinery and intelligence." *Mind*, 59, 433–460.

Ullman-Margalit, E. and Morgenbesser, S. (1977). "Picking and choosing." *Social Research*, 44(4), 757–785.

Vadhan, S. P. (2012). "Pseudorandomness." *Foundations and Trends in Theoretical Computer Science*, 7(1-3), 1–336.

Wagenaar, W. A. (1972). "Generation of random sequences by human subjects: A critical survey of literature." *Psychological Bulletin*, 77(1), 65–72.

Wilson, A. (2014). "Bounded memory and biases in information processing." *Econometrica*, 82(6), 2257–2294.

Worrall, J. (2007). "Why there's no cause to randomize." *The British Journal for the Philosophy of Science*, 58, 451–488.