On the normative status of mixed strategies

Kevin J.S. Zollman*

August 12, 2020

Abstract

Flipping a coin to decide what to do is a common feature of everyday life. Mixed strategies, as these are called, have a thorny status in normative decision theories. This paper explores various ways to justify choosing one's actions at random. I conclude that it is hard to make sense of this behavior without dealing with some difficult consequences.

*This paper was prepared for a volume in honor of Teddy Seidenfeld. In addition to Teddy, the author would like to thank Lara Buchak, Annie Duke, Simon Huttegger, Thomas Icard, Jim Joyce, Howard Lederer, Brian Skyrms, and Greg Wheeler for helpful discussions relating to this paper.

1 Introduction

> Well, I got through the night and no one looked in the box. Not even me. The person who gave up her whole evening to watch it. A whole evening of TV gone. What a mockery of justice that I can't take even a little tiny peek . . . OK, heads I look, tails I don't. Oh, yeah. Oh, yeah! Heads! I mean, alright then. No! I have a duty not to look. Well, then again, I promised the coin I would. . .
>
> *Turanga Lela, Futurama episode "The Farnsworth Parabox"*

It is in a philosopher's job description to make the most mundane behavior seem odd and paradoxical. Nowhere is this more true than in decision theory. Almost everyone has found themselves flipping a coin to decide what dish to order in a restaurant, what book to read, or even what job to take. The idea of flipping coins captivates us. Fictional characters as diverse as Harvey Dent,[1] Anton Chigurh,[2] and Donald Duck[3] determined to make important decisions using coins.

For a decision theorist, this behavior is difficult to justify. Traditional decision theories allow this kind of randomization only in exceptional circumstances, and they never endorse coin flips as the uniquely correct action. But the pull of coins, dice, and the like continues unabated. How shall we make sense of this?

Decision theorist call opting to randomize one's decision a "mixed" strategy. This is in contrast to choosing one of the available options, which are

[1] Harvey Dent, better known as Two-Face, from the Batman comics and movies opts to use a coin to make many of his decisions. In the movie *The Dark Knight*, he remarks, "The world is cruel, and the only morality in a cruel world is chance. Unbiased, unprejudiced, fair."

[2] Chigurh is the killer from the book and movie *No Country for Old Men* by Cormac McCarthy who would determine whether to kill someone by flipping a coin. He explains that this is part of the arc of the world, telling one of his victims who lost the coin toss, "You can say that things could have turned out differently. That they could have been some other way. But what does that mean? They are not some other way. They are this way."

[3] In the comic strip "Flip Decision" Donald is convinced to make all difficult decisions by flipping a coin. After following the advice of his coin, and driving down the road the wrong way, Donald is fined $50, not because he drove the wrong way down the road, but for "letting a dime do your thinking for you."

called "pure" strategies. The terminology comes about because randomization is a type of blending whereby one combines two different actions into a single action. But rather than combining aspects of each action to make a compromise—another reasonable use of the word "mixed"—a mixed strategy blends by making each probable.

Standard utility theory only permits randomization between two (or more) strategies when the agent is indifferent between them. One is rarely indifferent, and this leads many decision theorists to consider mixed strategies an oddity. But, mixed strategies show up in important contexts. First is the everyday case discussed above. People in state of uncertainty about the best course of action will often choose to flip a coin rather than picking an option. Second, mixed strategies are theoretically important in game theory. If we restrict ourselves to considering only pure strategies, many games lack a Nash equilibrium: a point where every player is happy with her choice. If we expand to allow players to choose mixed strategies, however, we are guaranteed a Nash equilibrium under weak assumptions. This was John Nash's seminal contribution.

The third appearance of mixed strategies is in statistical inference. Randomized controlled trails are often called the "gold standard" of research. For the Bayesian, however, this might seem mysterious. Randomization in these trials is a species of mixed strategy. Take as a simple example a medical trial, where patients are randomized between the current treatment and an experimental one. For any given patient, instead of flipping a coin, I could deterministically assign them to one or the other treatment (pure strategies). Under traditional decision theory, you should only randomize when you are indifferent between the informativeness of the two strategies: assigning your patient to a treatment. This hardly underwrites the ubiquity of this statistical method (see Kadane and Seidenfeld, 1990, and citations therein).

While mixed strategies are useful in decision and game theory, I will restrict myself to a single question: under what conditions would we endorse an agent who randomizes? More precisely, when should a rational agent regard randomizing her behavior as strictly better than all other alternatives? I conclude that it is difficult to justify mixing: those situations where mixing is genuinely preferred are beset with difficult problems. In the end, decision theory has managed to fulfill the philosopher's obligation of making mysterious a common behavior while leaving a little room for randomization in special circumstances.

2 Classical decision theory

The classic problem in traditional decision theory is the one-person decision under uncertainty. Consider a traveler contemplating two flight itineraries. One is faster, but involves a minuscule layover in a large airport. The other is slower, but affords more time between flights. Should the first leg of the trip be on time, our decider would prefer the faster trip. But if the first leg is even a few minutes late, the quick trip will become much longer as the traveler waits for the next available seat. The longer trip is more robust to delay, but would involve more unpleasant time spent in airports.

Not only would our traveler like to avoid time in airports, but she also knows her proclivity to second guess her decisions. If she takes the longer flight, she will look to see how the other option fared. If the shorter flight was on time, she'll regret having booked the longer flight—engaging in constant self-criticism. On the other hand, if the short flight is late, she'll congratulate herself for a decision well-made.

Traditional decision theory carves this decision into three parts. First is the set of outcomes, O which represent all possible objects of value. Idealizing slightly from the real case, we might imagine four outcomes:

o_1: Short trip, no missed flights

o_2: Missed flight

o_3: Long trip, short would have worked out

o_4: Long trip, short would have gone wrong

In addition to outcomes, there is a set of states, S. In our idealized travel story, the two relevant outcomes would be "No delay in first leg" and "Slight delay in first leg" Finally, we assume that the agent has available to her a set of actions, A. Each element a of the set A is a function from S to O. The action function represents what outcome the agent expects to receive in each state. In the case of the airport there are two actions: book the short trip (a_1) or book the long trip (a_2). The long trip results in outcomes o_3 or o_4, while the short trip results in o_1 or o_2.

All of this can be summarized in the familiar decision matrix picture in figure 1. The traveler is choosing a row. The state is determined by a column. Jointly the two pick out a unique outcome.

	No Delay	Delay
Short flight	o_1	o_2
Long flight	o_3	o_4

Figure 1: A traveler's decision

Actions in this model are the "pure" strategies available to the agent. She can choose an action with probability one by booking an itinerary. To allow mixed strategies, we give our traveler a drawer full of coins of various (known) bias. Our agent can choose one of those coins and flip it, buying one itinerary if heads and another if tails.

Of course, it doesn't matter that these are coins and not dice. But it is critical that the randomization be exogenous, uncorrelated with the states of the world. Why? Because the underlying model assumes the agent cannot condition her action on the state of the world. If she could, the agent would most prefer the "mixture" that gives her the short trip if and only if she won't miss the connection. This would be a type of mixed action, but one that is effectively magical for the decision at hand.

The assumption of exogenous mixing devices like coins or dice is a useful heuristic device. The metaphysical assumptions of chance objects is not important. We only need a mechanism whereby the individual agent can generate uncertainty about her own act. This can occur via a randomizing device or by employing the services of a confederate who makes the choice on behalf of the agent. Whatever the source, when an agent chooses a mixed strategy, she finds herself uncertain about what she will do. We will take this to be the essential defining characteristic of a mixed strategy. An agent prefers a mixed strategy if she prefers to be less certain about her own action than she could otherwise be.

Not all ways of being uncertain about one's own actions count as mixed strategies. Skyrms (1990b) describes deliberation as a dynamic process of changing ones uncertainty about one's own action. Julie knows at some point tomorrow she will order lunch. But since Julie has not yet begun the process of decision, she doesn't yet know what her order will be. Her current uncertainty over her act tomorrow is an uncertainty about the outcome of her own process of deliberation. Thinking of Julie assigning probabilities in this case is controversial (Levi, 2000; Joyce, 2002), but I need not venture in. As of today, Julie doesn't count as adopting a mixed strategy because Julie is not choosing to be less certain than she could otherwise be. However, if the *conclusion* of her deliberation she remains uncertainty about her action, then I would say that Julie has chosen a mixed strategy.

Having defined mixed strategies, we might now ask: under what conditions might an agent opt to choose such a thing? When will our traveler reach into the drawer of coins?

To answer this question, we must model the agent's beliefs and values. In classic decision theory, we assume the agent behaves as-if she represents her uncertainty over states with a single probability function, $p : S \rightarrow [0, 1]$. Furthermore, she behaves as-if she values outcomes with a single utility function $u : O \rightarrow \mathbb{R}$.[4] The agent chooses an action which maximized her expected utility.

Suppose our traveler looks at the published delays for past flights and estimates the probability of bad weather. After careful reflection she concludes that the probability the first flight will be delayed is 0.2. Without losing any generality, we can also fix two points in the agent's utility function. Assuming our traveler is a human being familiar with air travel, we can conclude that her favorite outcome is a short flight without a missed connection (o_1) and her worst outcome is missing the flight (o_2). We can therefore assign $u(o_1) = 1$ and $u(o_2) = 0$. This leaves only the o_3 and o_4 outcomes which we will assign $u(o_3) = x_3$ and $u(o_4) = x_4$.

Our agent will prefer the short flight over the long one if $x_3 + 4x_4 > 1$. She will choose the long flight over the short one if $x_3 + 4x_4 < 1$. In either of these cases, the agent would regard any flip of a coin as inferior to her preferred choice. Why "mix" a worse option with a better one? The only time where our agent would entertain a mixed option would be when she is otherwise indifferent between the two, when $x_3 + 4x_4 = 1$.

This is not idiosyncratic to this example. An agent will only opt for a mixed strategy over some set of pure options when she is indifferent between all the options assigned positive probability in the mixture. The reason is straightforward: if one option is better, one should do it more often; it's better after all.

There is substantial disagreement about how common such situations are. Montaigne said, "It might rather be ... that nothing presents itself to us

[4]Henceforth I will dispense with using the "as-if" language and simply refer to the agent's probability and utility function. It is worth noting, however, that for many decision theorists it is not critical that a rational agent consciously construct a probability and utility function to come to her decisions. Nor is it critical that the agent have any mental objects or representations resembling probability or utility functions.

wherein there is not some difference, how little soever; and that, either by the sight or touch, there is always some choice that, though it be imperceptibly, tempts and attracts us" (de Montaigne, 1877, Book 2, Chapter XIV). Others, like Ullmann-Morgalit and Mogenbesser (1977), disagree.

But whether they are common or rare, they remain special in certain respects. If our traveler opts to flip a coin, she must be indifferent between the two trips. In such a case she is indifferent between choosing one pure option, choosing the other pure option, and all possible probability distributions over the actions, all the coins in her drawer. So while she may choose to flip a coin, rationality never requires her to do so. This stands in contrast to ordinary experience where people exert extra effort in order to flip a coin.

Nor can flipping a coin be seen as a mechanism for breaking the apparent symmetry between two otherwise equivalent options. Rather than reducing the number of options, the introduction of mixed strategies multiplies them (Ullmann-Morgalit and Mogenbesser, 1977). Instead of the two trips, our traveler now has an infinity of equivalent mixed strategies. This also stands in contrast to ordinary life. Not only do people experience the pull of a coin, but they often desire a fair coin—which assigns equal probability to the two options—over a mixture that weights one option more heavily than another.

This same issue can be put in slightly different language that shows the tight connection between mixed strategies and the "value of information." When an agent is opting for a mixed strategy, she is opting for a probability distribution over her own actions rather than knowing what she is choosing for sure. In a certain sense, she is opting for ignorance over certainty. She could be certain about her own action by choosing a pure strategy, but instead she is opting to be unsure. In classical decision theory, an agent should never pay to avoid free information,[5] and similarly an agent should never be willing to go out of her way to know less about her own actions than she could otherwise (cf. Skyrms, 1990b, chapter 4).

While all of this makes mixed actions mysterious, it also contains a kind of consistency. Mixed strategies, like other strategies in decision theory, have a kind of dual rationality to them. When an agent flips a coin or rolls a die, she evaluates the mixed action as at least as good as the pure actions *ex ante*—that is, before the coin has flipped or the die rolled. Ex ante, our decider is not yet aware which pure action she will take. But after the coin is flipped or the die rolled, she will come to know. We might ask, after learning

[5] Under appropriate conditions (see Kadane et al., 2008).

the outcome of the roll or flip, is our agent still happy with her choice *ex post*.

Since, in traditional decision theory, an agent is only willing to choose a mixed strategy ex ante if she indifferent between all the pure actions she might take, then she cannot regret the decision ex post. After learning the outcome of the flip or roll, she has moved from one alternative to another which she considers equally good. As we step away from classical decision theory to its nearby neighbors, we will—at points—lose this nice symmetry.

3 Non-classical decision theories

We are now in a position to see why classical expected utility theory has trouble making sense of people's desire to adopt mixed strategies. If one want to allow for this possibility, in one way or another, one must adopt a decision theory which is more permissive and allows for individuals to prefer one or more mixed strategy to at least one of the options that comprise it.

There are a number of candidates that meet this requirement. Before we investigate each in detail, we can sketch the form of a worry. Consider a mixed strategy M which assigns a non-zero probability to ultimately taking some pure action A and an agent who strictly prefers M to A. Now our agent faces an odd situation. Imagine an agent who chooses M. She then begins the process of implementing M by flipping her coin or rolling her die. Suppose she observes the outcome of her randomizing device which results in her taking action A. Our agent will now regard herself as worse off than she was before the coin was flipped or the dice rolled. From her subjective perspective, she would (a) prefer not to know the outcome of the randomizing device and (b) prefer to pick the coin back up and flip it again. Depending on how A relates to the other pure strategies in the mixture, it might also be the case that the agent now wishes she had chosen another pure action, perhaps B, instead of the mixture M. This later situation is a violation of the ex ante–ex post consistency discussed above.

This odd consequence of the preference for mixed strategies will feature in the objection to several of the particular non-classical decision theories that follow.

3.1 Imprecise probability

To see one such illustration, let us return to the story of our traveling academic. Suppose, our agent (being a careful researcher) looks into all the factors influencing the punctuality of her flight. She discovers this route is delayed 20% of the time. But this particular flight will not be a normal one, bad weather is forecast for the day of the trip. (Our absent minded academic has waiting too long to book her flight.) Looking at the average for flights during similar weather she finds that delays can be as as high as 70%. This airline is regarded as more reliable than others, but she doesn't know how to quantify this.

Traditional decision theory requires our agent compile all this information into a single number that represents her all-things-considered judgment about the probability of a delay. Some scholars feel this is too demanding. Our agent knows the probability of delay is greater than 20% and less than 70%, but insisting on more precision would force her to make arbitrary choices. Instead, these scholars often suggest that agents' uncertainty be represented with sets of numbers instead of one.[6]

Our traveler's beliefs might be represented by all numbers between 0.2 and 0.7. In doing so, we haven't yet solved our traveler's problem. x_3 and x_4 might not completely determine the best trip. Different trips might be optimal for different admissible probabilities. For illustration consider figure 2.

Merely expanding the representation of beliefs has not yet given us a full decision theory. To do so, we must specify how an agent will use these numbers in making their choices. Options abound. For illustration, let us consider one popular rule, often called Γ-maximin, where the agent prefers the act with the higher minimum expected utility (Gilboa and Schmeidler, 1989).

This decision rule is distinct from its traditional cousin, regular maximin. Traditional maximin requires the agents to ignore probabilities entirely and only focus on which action would yield the worst possible outcome. For regular maximin, our agent would always choose the long flight because the short flight contains the worst possible outcome, o_2. The probabilities of the states are ignored. Γ-maximin, on the other hand, recognizes that one should not be so cautious; one should take some risk. So the agent calculates the worst *expected* utility given what she knows.

[6]This idea sometimes called "imprecise," "indeterminate," or "interval" probabilities has a long history going back at least to Keynes (1921).

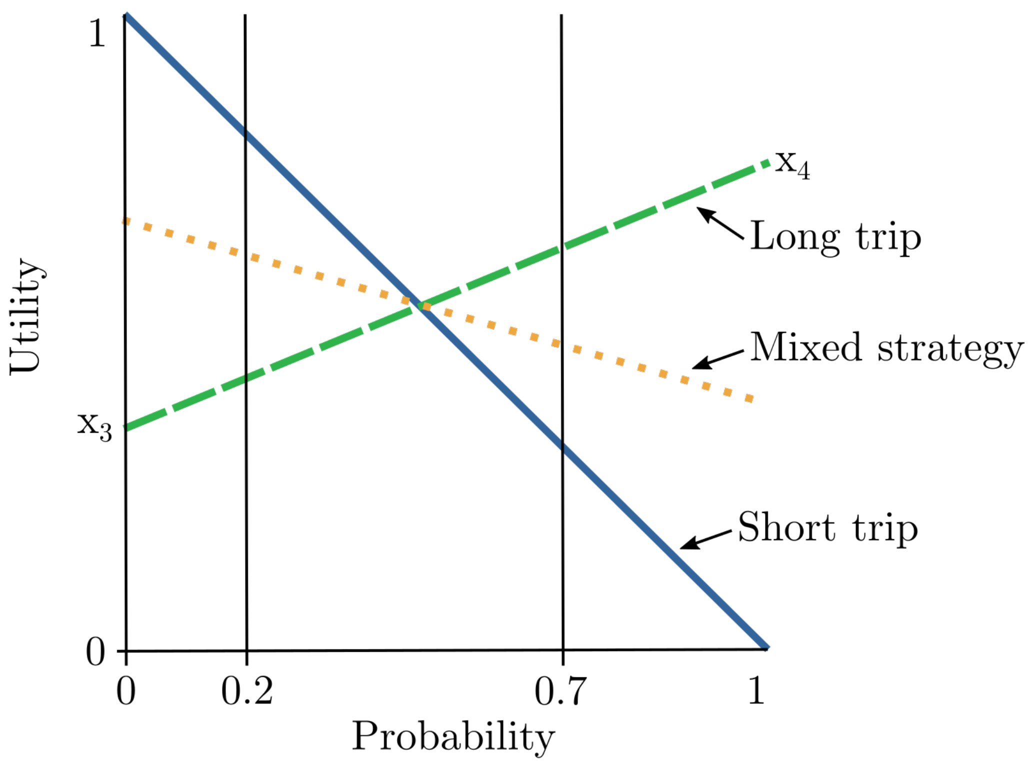

Figure 2: A representation of the various utilities for an agent with imprecise probabilities. The x-axis represents all the potential probabilities of a delay (from 0 to 1). The y-axis represents the expected utility of each potential action give the probability specified on the x-axis. Each line corresponds to a different action.

Our traveler should figure out what the worst expected utility would be for each action. In the case of the short layover, this would occur at the highest probability of delay: 70%. Here, her expected utility is 0.3. For the longer delay, her worst possible expected utility occurs at the lowest possibility of delay: 20%. Here her expected utility is $0.2x_4 + 0.8x_3$.

How about flipping a fair coin to decide which flight to book? Now all four outcomes are possible. The expected utility from that action is represented in figure 2. The worst outcome in this illustration will occur with a high probability of delay. For some values of x_3 and x_4 the expected utility for the mixed action will be higher than the comparable worst case scenario for booking the longer trip. Γ-maximin would choose the mixed strategy as the uniquely correct action!

Paradoxically, the mixed strategy introduces a little more precision into our imprecise world. In so doing, it improves the worst case at the expense of the best case. Since Γ-maximin pays attention to one, but not the other, the improvement is felt. The mixed strategy is chosen, even when the agent isn't indifferent between the two pure actions. What's more, with imprecise probabilities, one can now find situations where the mixed action is strictly preferred to either of the pure actions.

But this also provides an opportunity to illustrate the abstract worry presented above. Suppose our traveler asks her friend to flip a coin and book the short layover if heads and the long one if tails. Her choice maximizes her minimum expected utility, it is therefore best in this context. Her friend flips the coin and exclaims, "oh, it's heads!" Now, our traveler's evaluation of her situation has gotten worse! She knows she's booked the short layover which she regards as worse than the mixed option she chose.

This again shows the relationship between the value of information and preferring mixed strategies. If I strictly prefer a mixed option, then I would prefer *not* to learn the outcome of the mixing device; I want to be ignorant of my own action. This can occur in the case of Γ-maximin with imprecise probabilities. Our agent feels as though her situation has gotten worse when she comes to learn the outcome of her own mixing.

In addition, the imprecise case illustrates how mixed actions may lose their ex ante–ex post symmetry. After our traveler's friend informs her that the flip is heads, our traveler now knows she will book the short layover. Because the short layover has a lower minimum expected utility than the long layover trip, our agent now wishes she hadn't chosen the coin and instead chosen the pure strategy of booking the long layover. She regrets her choice even though she was happy with it prior to learning the outcome of the flip.

Γ-maximin is not the only possible decision theory for imprecise probabilities. In fact, several proponents of imprecise probabilities have rejected it in favor of other decision rules for broadly the reasons discussed here. Other decision rules allow for a wider variety of options, and this issue is more complicated in those settings (see Seidenfeld, 2004). Space will prevent a complete exploration of this issue here.

3.2 Incommensurable values

In the last section, our agent was entertaining several probabilities as potential beliefs. Symmetrically, she might also be unsure about the appropriate value for an option—the uncertainty now finds home in the utility function rather than the probability function.

For simplicity of presentation, let us consider a different academic traveler. He is also torn between a long a short layover, but his reasons for conflict are different. Our new traveler knows the probability that the short itinerary would result in delay, say this is 30%. If he gets delayed he will miss his talk and harm his career. The longer itinerary involves leaving home much earlier thus guaranteeing that he will make his talk. By leaving early, however, he will miss time in the morning with his family, something he also finds valuable.

Our new traveler must now balance the considerations of a career and family that he both values. One consideration (career) speaks toward leaving early and the other (family) encourages him to leave late. Like with probabilities, traditional decision theory requires that our agent balance these two considerations and come to a single number: an all-things-considered evaluation of the various benefits and costs. He must construct a method of comparison between family and career and form his preferences for flights on that basis.

Levi (see, e.g. 1990) has criticized the traditional approach. Instead, Levi argues, we should use an analog to imprecise probabilities, where our agent considers several ways of trading off the two considerations. Above an agent entertained different probability functions, here an agent utilizes different utility functions.

If our agent cared only for family, for instance, he would assign the short trip a utility of 1 and the long trip a utility of 0. If he only valued his career, he would reverse those numbers. We can represent ways of balancing these two considerations by considering different weights in averaging those two numbers. One might assign a higher weight to family or not. This is pictured in figure 3 (which bears a noticeable resemblance to figure 2).

Again one must specify a way to deal with a decision where an agent is unsure about the appropriate weighting between the two values. One might import Γ-maximin again and ask the agent to evaluate each action according to the worst value among all weightings the agent considers reasonable. If our agent is maximally uncertain about how the two values should be compared—if he considers all weightings as reasonable—each action has a minimum all-things-considered utility of 0. If he pays attention to family, his career might be the "right" thing to pay attention to and vice versa.

Now what if he flips a coin? Here, the value of the mixed option is constant across all weights. Whether only family, only career, or anything in between is the right way to aggregate the two dimensions, flipping a coin always results in an expected utility of 0.5.

For essentially the same reasons as occurred with imprecise probabilities, our agent might opt for a mixed action (assuming he adopts a decision rule like Γ-maximin in these cases as well). But all the same provisos remain. He would prefer a friend flip the coin, so he doesn't learn the outcome. And after learning the outcome he might wish he had chosen differently.[7]

3.3 Inherent value in the process of randomizing

In our description of the travelers, we assumed they neither like nor fear uncertainty. Perhaps real travelers are not like this. Most often, people are assumed to be afraid of risk—they prefer certain outcomes over uncertain prospects that have slightly higher value on average. But they might have other attitudes toward risk.

Traditional decision theory incorporates attitudes toward risk by manipulating the utility functions. As an example, let us return the first traveler with known probability and utility functions. Suppose she enjoys taking risk. A risk seeking traveler might choose the shorter trip, since that experiences a larger swing in potential outcome. Maybe she wants more. Perhaps she

[7]In our very rudimentary example, we cannot construct an analog to the regret faced in the imprecise probability setting. This example doesn't allow for a case where the agent, after finding out the coin flip, had wished that he had booked a trip different from the one given to him by the coin. But, with a little more sophistication we could achieve the same failure of ex post justification.

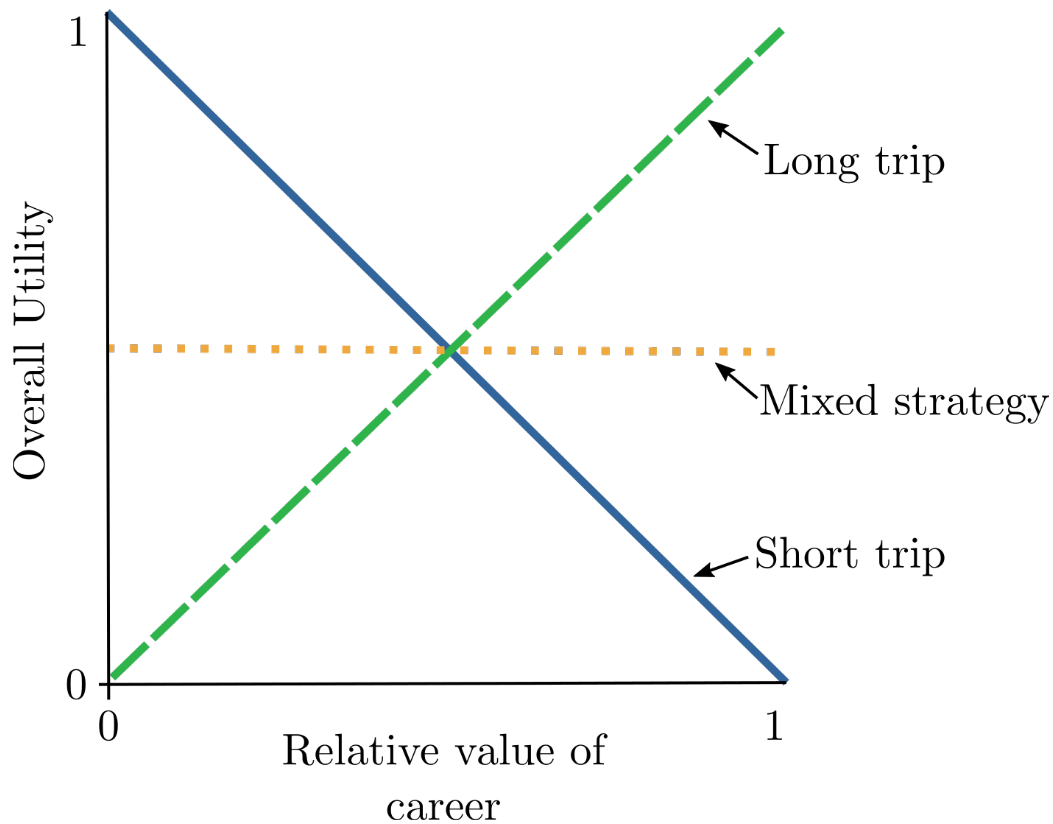

Figure 3: A representation of the various utilities for an agent balancing two different values. The x-axis represents all the potential weights that one might assign to one's career from no weight (0) to full weight (1). The y-axis represents the aggregate utility of each potential weighting given on the x-axis. Each line corresponds to a different action.

would enjoy the uncertainty over which trip she will book. In such a case, she might prefer a mixed strategy because it increases her uncertainty over outcomes: randomization makes all four outcomes possible.

How shall we account for this risk loving traveler? We might simply regard her as irrational. Or, one might account for her behavior by changing the way we define outcomes. The original specification of "outcomes" did not include everything of value to the traveler, one might argue. If she values uncertainty, we should build that into the outcomes. Under this modification of traditional decision theory, we would build "the joy of flipping the coin" into the outcomes. We would then have eight primitive outcomes: the original four achieved either by coin or by choosing a pure strategy.

This has the undesirable effect of destroying coin flipping as a mechanism to "mix" the outcomes. So, while our agent might prefer a coin over a pure option, we might no longer regard this as a mixed strategy: by choosing the coin she can access fundamentally different outcomes than those available without a coin. This is an issue we will return to in later sections.

Another option would be to change how our traveler computes the value of mixed options. Instead of calculating the linear average—as in computing expected utility—she might combine the value of outcomes differently.

Buchak (2013) has developed a normative decision theory where the value of mixed options is not a linear combination of the value of the pure options. In Buchak's theory an agent evaluates an action by calculating something she calls the "risk weighted expected utility." In this theory agents calculate expected utility but in a way that allows them to transform the probabilities by treating them as larger or smaller than in traditional expected utility.

There is more than one equivalent way to calculate expected utility. One way, necessary to develop Buchak's theory, is to sum by starting with the worst outcome. Then add the probability of getting an outcome as good as the second-worst outcome multiplied by improvement in value from the worst to the second-worst. Next, add the probability of the third worst multiplied by that improvement. Etc. For a simple three outcome decision this looks like this:

$$EU = u(w) + \big(p(m) + p(b)\big)\big(u(m) - u(w)\big) + p(b)\big(u(b) - u(m)\big) \qquad (1)$$

Where w is the worst outcome, m is the middle outcome, and b is the best outcome.

So far this calculates the same quantity as the traditional expected utility calculation. However, Buchak allows agents to weigh these probabilities

	$p_1 = 0.1$	$p_2 = 0.9$
A_1	-10	2
A_2	10	-2

Figure 4: A risk weighted utility gamble

differently. In her theory an agent chooses a function $r(\cdot)$ which transforms the probabilities into another number. The risk weighted Expected Utility of the same decision would be:

$$EU = u(w) + r\big(p(m) + p(b)\big)\big(u(m) - u(w)\big) + r\big(p(b)\big)\big(u(b) - u(m)\big) \quad (2)$$

If $r(p) < p$ this captures a notion of risk aversion and if $r(p) > p$ this captures a type of risk seeking.

Consider an agent who faces the decision problem pictured in figure 4 and who has $r(p) = \sqrt{p}$. For this agent A_2 is preferred to A_1. However, most of all the agent would like to adopt a mixture between the two options. For Buchak's agent, playing A_1 with probability of approximately 0.517 is optimal.

Buchak's decision theory allows—but does not require—an agent to prefer a mixed option over the pure options comprising it. Everything depends on the r-function and the underlying gamble the agent is confronting.

Buchak's theory provides a cleaner way to preserve the notion of "mixing" actions than modifying outcomes. But she is confronted by the same problems as Γ-maximin. Buchak's agent would prefer to remain ignorant of the coin's outcome. Since the mixture is preferred to either A_1 or A_2, the agent would regard her situation as worse where she to know her action. And her agent can maximally regret her choice. If she learns that the coin has given her A_1, she will regret her decision and wish she had chosen A_2 instead.

It is also worth noting another interesting contrast with Γ-maximin. In the previous section, an agent who is worried about risk, who looks at the worst case, might prefer a mixed strategy while an agent who looks at the best case could not. In Buchak's theory, an agent who enjoyed uncertainty or risk might opt for a mixed strategy because it is more risky, but a risk averse agent would not. This illustrates that there is not a robust relationship between our intuitive notions of risk seeking and risk aversion and mixed strategies.

3.4 Fairness

Imagine a single agent must allocate an indivisible good among two or more individuals. A parent, for example, may only be able to attend his daughter's softball game or his son's concert but not both. In such a setting a parent, who desires to allocate his time fairly, must find a mechanism that reflects his valuing his children equally.

These are not only small-scale problems. For example, one might be concerned that an autonomous vehicle not act so as to prefer the lives of some people rather than others. (Perhaps a BMW might be programmed to specifically decrease the risk to other BMWs.) In such a case, one might think the only fair outcome in trolley-problem like settings is one where the risk of death is distributed equally.[8] Or, perhaps a national healthcare system finds itself with more patients in need of a medical treatment than it can provide. Or a homeless shelter may find more seeking shelter than it has beds. Etc.

The allocaters (the parent, the AI system, the government, or the shelter) may try to find a principled reason to prefer one over the other. Perhaps the father went to the son's concert last time, or the government can assess who might benefit most, or the shelter can prefer the person who arrived first. But occasionally in these settings one cannot find the appropriate distinguishing features to chose one beneficiary over another. In such a setting, fairness considerations might drive one to use a randomizing device.

The randomizing device gives access to a payoff distribution that is impossible through other means. Our father would like to treat his children equally, but equality is not possible with pure strategies: either his daughter or his son will feel mistreated. But, the coin flip achieves—at least temporarily—an equalization of payoffs.

So far, this looks very much like a case where the agent has an intrinsic preference for randomization. The source is different, here it is a preference for equal treatment rather than a desire for uncertainty. But the same problem remains. Once the coin is flipped the father is left with an unequal treatment of his children. He (and one of his children) might now object that he should flip the coin again to regain equal expected payoffs. This issue is tightly connected with a debate in the social choice literature about the role of ex ante and ex post considerations of fairness where many of the same difficulties remain (cf. Fleurbaey and Voorhoeve, 2013).

[8] This example was suggested to me by Greg Wheeler.

3.5 Imperfect recall

So far we have retained some of the core assumptions of classic decision theory. We have assumed that agents face no bounds when calculating their decisions. Most notably, they can remember the entire past and make any possible plan of action for the future. This assumption has not really been mentioned because all the decisions we've discussed are single shot decisions that take place in an instant (they are in *normal form*). Decision theory is capable, however, of representing a broader class of decisions: those where one's choices stretch out over time. I do not decide, all at once, how I will teach each and every class this semester. Instead, I make an initial plan I modify—within constraints—over time. These are called extensive form decisions.

Traditional decision theory puts constraints on how extensive form decisions are modeled. We will focus on perfect recall. In extensive form decisions agents cannot forget, most critically they cannot forget their own actions. This is obviously a false idealization. While I chose what I ate for lunch thirty days ago, I could not possibly remember what it was. (Even yesterday is a stretch.) In many cases this idealization is harmless, but some decisions trade on it critically. Here is a famous example due to Piccione and Rubinstein (1997).

Suppose a philosophy professor with a simple commute. She must drive from her office through the countryside. To find her house, she must pass two otherwise identical barns and turn on the unlabeled road immediately after the second one. She would like to get home as quickly as possible, we will assign this outcome a utility of 3. If, she mistakenly turns after the first barn, she will end up hopelessly lost, an outcome we will represent with utility 0. If she fails to turn toward her house after passing the second barn she will also get lost, but will figure out her mistake far more quickly. In this case, she will get a utility of 1.

Assuming perfect recall, this is a trivial decision. The agent has three pure strategies: turn after the first barn, turn after the second, or never turn. Obviously the second is superior yielding a utility of 3, while the others are worse.

But, we are talking about a philosophy professor, and she is prone to get lost in thought. After coming across one of the two barns she cannot recall whether she passed one already: she's unsure if this is the right place to turn. Now our agent has two pure strategies: turn after seeing a barn or stay straight. Both pure strategies are unattractive: neither gets her home without getting lost.

However, with a mixed strategy she fares better. If she turns after seeing a barn with probability p, then she gets home most efficiently with probability $(1-p)p$. Choosing among all possible mixed strategies, her utility is maximized when she turns with probability 1/4. Adopting this strategy results in her receiving an expected utility of 9/8.

In this setting a mixed strategy is uncontroversially better. It makes possible one outcome (getting home via the quickest route) that was inaccessible otherwise. This is somewhat different from other cases of mixing where the outcomes made possible through the mixed strategy are the same as those available through the pure strategies. This allows for a victory for randomization but not really for "mixing" because the outcomes of the random strategy are not a mixture of the outcomes available to the pure actions. One prefers randomization only because it gives access to an outcome that is inaccessible otherwise.

Issues of consistency over time appear here as well. Piccione and Rubinstein (1997) illustrate that our agent might not be happy with her strategy when it comes time to follow through. Consider our philosophy professor who computes the optimal strategy for getting home: turn after seeing a barn with probability 1/4. Pleased with her decision theoretic prowess, she departs the office. She notices a barn. Remembering her prior plan (she managed not to forget that!), she computes the probability that she is at the correct turn for her house as 3/7. If she turns now – with probability 1 – she will get an expected utility of 9/7 which is higher than her expected utility of keeping to her plan.

Even though her plan, if carried through, has an ex ante higher utility than the pure strategy: turn after a barn, our agent will not carry it out. Or, more precisely, she cannot carry it out without at some point choosing an action she regards as irrational at the moment of choice.[9] Although the issue of the time consistency of her preferences is slightly different than in the preceding three examples: the general problem remains.

Even with all these concerns, the absent minded driver provides perhaps the strongest argument thus far for randomization. But, one might be con-

[9]This issue has connections to the notion of ratifiability in Jeffrey's decision theory (Jeffrey, 1990; Skyrms, 1990a). There is a substantial discussion about how to handle situations like this in related contexts. Space will prevent me from exploring it further here.

cerned that this example is somehow special. Perhaps only absent minded philosophers need to worry about such things. But the problem is far more general, Icard (2019) shows how randomization is useful in many settings that agents regularly confront and with less ridiculous bounds on their memories. While randomization might never be optimal for the gods, who can remember all, for us mere mortals it represents the best way around other constraints which we are unable to alter.

3.6 Act–state dependence

Sometimes an agent's choice might influence the probability of a state. This might occur because the action has a causal influence on the state. I may want to know if the soufflé has fallen. To learn the status of my dinner, I could open the oven. But, if the soufflé has not finished cooking, the act of opening the oven may itself cause the soufflé to fall.

To evaluate mixed strategies in these contexts, we must understand how the choice of the mixed strategy interacts with the underlying causal structure. In the simplest example, like the soufflé, mixed strategies function as they do in the classic decision theory context. If I flip a coin to decide whether to open the oven, the value of that decision is a simple linear combination of the value of choosing to open the oven or not. In these contexts, mixtures function in the same way they do in classical decision theory.

The situation need not be so clean, however. If the act of mixing one's actions has a different causal effect than the pure acts, things might change in important ways. Suppose a judgmental acquaintance offers me a tray containing two cookies, one much larger than the other. If he observes me taking the larger cookie, he will presume that I'm greedy. If I take the smaller, he will accuse me of unreasonable modesty. I cannot win. But, he might alter his view if he sees me first flip a coin to decide. Suppose the coin would improve his assessment of my character. If so, I would strictly prefer to flip to decide which cookie to take in order to secure a better judgment from this unpleasant dessert companion.[10]

In a case like this, I opt to mix because doing so has different causal effects than any of the pure acts under consideration. Given his peculiar

[10] This is not an implausible situation, Bicchieri and Chavez (2013) found that others are more forgiving of apparently unfair behavior if that behavior was the result of a flipping a coin *even when* the person benefited by the unfairness could have chosen to be fair without the coin.

view of fairness, the judgmental cookie provider views flipping a coin as the appropriate choice mechanism. The value of the mixed option is higher than either of the pure options as a result. This case has similar features as the case of the absent minded driver; by flipping a coin I make accessible and outcome (being judged a good person) that is unavailable through any pure action available to me

Beyond direct causal influence, there are other—more controversial—ways for acts and states to become correlated. Fantastical examples involving people who can perfectly predict the agent's behavior or desires are introduced in order to correlate an agent's action with the probability of the state.

Consider the famous Death in Damascus problem, introduced by Gibbard and Varian (1978). An agent must decide whether to go to Aleppo or Damascus. But she faces an evil oracle who will attempt to predict where she will go and send death to meet her there. The agent would prefer to avoid death but believes that the predictor is perfect. What should she do?

Here there appears to be no direct cause from the agent's action to the location of death, but they are correlated nonetheless. Much can, and should, be said about the appropriateness of these examples in decision theory, but for the moment I will set these aside (see, e.g. Skyrms, 1982; Joyce, 1999).[11]

How should an agent evaluate mixed strategies in this context? If the predictor can only predict that the agent will flip the coin but not how the coin will come out, then the agent has maximized her chance of survival by flipping. While each pure action leads to death for sure, flipping the coin has improved her chance of survival to 0.5 (see Ahmed, 2014).

In these contexts, is the agent satisfied with her choice after flipping the coin? In the case of the judgmental cookie provider, he will not care that I learn the outcome of the flip. The baker will not view me as being impolite or overly modest if I follow the directions of my coin. I will be happy to learn my cookie choice and endorse it after the fact.

We must fill in more of the fantastical story of the oracle to determine our agent's feelings after learning the flip of the coin. If I observe the flip and then follow the directions of the coin, can the oracle now predict my action? Can the oracle's mind reading abilities extend to the many individual decisions I must make in implementing the choice of the coin? If so, I will be immediately unhappy with my choice once I learn the outcome of the coin; I would reach for yet another coin. If the predictor could only see into my brain before the

[11] We will revisit some justifications for decisions of this type in section 4.3.

coin flip—but not after—the coin flip remains acceptable ex post just as it does ex ante.

This illustrates a way in which the superiority of mixed strategies in Death in Damascus comes about through a conflation of two different considerations. The mixed strategies are not superior because they induce uncertainty in *me* about my action. I am choosing to flip a coin not because of its randomization qualities, but because of its inherent unpredictability to someone else. If I could go to Aleppo with probability one but in a way that the oracle was unable predict, I would just as soon do that. The mixed strategy is only chosen because the situation dictated that I could not make someone else uncertain without becoming uncertain myself.[12]

The same criticism applies to the more realistic case of causal influence. If I could fool the judgmental cookie provider, I might opt to flip a coin and choose the larger cookie regardless of outcome. I could convince the provider that I was being fair and secure the larger cookie!

Hidden behind both examples is an assumption about how my behavior influences the decisions of another chooser (either my judgmental cookie provider or the oracle). Decisions of this sort are usually analyzed with the tools of game theory modeling all decision makers together (Skyrms, 1990b). We will find similar complex issues arise in that setting, and to it we now turn.

4 Game theory

4.1 Classical game theory

Two children, Ann and Bob, agree to play a game with two coins. Each will simultaneously place the coin heads or tails up on the table. After placing each will reveal their choice to the other. Should the coins match, Ann will win both; Bob wins otherwise.

A normal form picture of this game is presented in figure 5. Ann chooses a row, and Bob replaced nature in choosing a column. The first payoff in the individual cells represents Ann's payoff and the second gives us Bob's.

We can ask in this game is there any set of strategies that forms a mutual best response? That is, can we specify a strategy that satisfies these two conditions:

[12]My thanks to Jim Joyce for a discussion on this topic.

		Bob	
		Heads	Tails
Ann	Heads	$1, -1$	$-1, 1$
	Tails	$-1, 1$	$1, -1$

Figure 5: Matching pennies

(1) Ann is doing the best she can given Bob's strategy

(2) Bob is doing the best he can given Ann's strategy

No pure strategy can simultaneously satisfy both conditions. If the two coins match, then (2) is violated. If they don't (1) is. Conditions (1) and (2) define what it is to be a Nash equilibrium in this game, and therefore there is no pure strategy Nash equilibria in this game.

As in the previous section, we can expand the set of options available to each player by adding in exogenous randomizing devices that allow them to choose any probability distribution over their pure actions. In the case of decision theory, we required that the randomization device be independent of the choice of state. In a similar vein, we require that the randomization device available to Ann is independent of the one available to Bob. (But, more on this later.)

We have now expanded our game by adding an infinity of new strategies. Ann chooses a real number $s_a \in [0, 1]$, the probability she chooses Heads. Bob chooses $s_b \in [0, 1]$. Now we can ask, is there any pair of real numbers that satisfies conditions (1) and (2) above.

There is. If Ann and Bob both flip two different fair coins ($s_a = s_b = 0.5$) both conditions will be met. The calculation is easy in this case. If Bob has a 50% chance of choosing Heads, Ann is indifferent between Heads and Tails. The same is true for Bob. Thus, anything Ann does is best, and ditto for Bob.

Nash proved this trick is always available (Nash, 1950): when you expand games (satisfying very weak constraints) with mixed strategies, one will always find an equilibrium point satisfying conditions (1) and (2). As a matter of mathematical fact, there is little to add. Points of this form don't always exist in games without mixed strategies, but they do when mixed strategies are introduced.

As in decision theory, mixed strategies in game theory are controversial. First, as a normative matter, there is little to recommend them. If I know that the other players are playing their part of a mixed Nash equilibrium, I have no affirmative reason to play my part. I could, but I could just as well do any of an infinite number of other things.

This problem is identical to the issue discussed above: in classical decision theory there is no reason to prefer to be uncertain about one's own actions. And this problem reproduces itself in game theory as well. However, game theory reveals other methods for justifying mixed strategies. In fact, mixed strategies in game theory are rarely about being uncertain about one's own action and are instead about inducing uncertainty in others. They share this feature with the examples of act–state dependence discussed above.

It is worth noting: nothing from the previous section is lost when we move to game theory. All of the modifications to classical decision theory described above can be incorporated into games. But, nothing new is added. Instead in the sections that follow, I will discuss how one might justify mixed strategies in ways unique to game theory and not available in classical one-person decisions.

4.2 Putting beliefs into equilibrium

As defined in the previous section, beliefs play no role in equilibrium. A state is a Nash equilibrium if both players are best responding to one another, full stop. This way of characterizing equilibrium is parsimonious and allows the equilibrium concept to be applied to a wide domain of behaviors including animal and robotic interactions. But it provides little by way of normative justification. To do this, we can't simply focus on how the actions of the players correspond to one another. We have to consider the beliefs of the players to determine whether or not there is a good normative justification for their actions.

There is more than one way to do this, but for the purposes of this discussion I will define equilibrium in this way:

1. Each player has a conjecture about the (possibly mixed) strategy of all players, including herself.

2. Each player is choosing a strategy which is a best response relative to her conjecture about the other players.[13]

[13]This condition builds in assumptions about a player's knowledge of the game and about her rationality. We could break this condition into two conditions: (2a) the player knows

3. Each player's conjecture about all players is true.

These three conditions define a generalization of Nash equilibrium called a "correlated equilibrium" (Aumann, 1987). If we add the condition that all conjectures entail probabilistic independence between the actions of the players then we have a definition of Nash equilibrium (cf. Aumann and Brandenburger, 1995).

This definition makes clear a connection between game theory and the case of act–state dependence discussed above. Imagine we insist that these conditions obtain in every possible circumstance. If we then ask a player to consider altering her strategy, she is imagining that as she changes her strategy others' beliefs will track that change. Recall, condition 3 requires that all conjectures be true, so if one player changes her strategy, the other players must also change their beliefs.

Consider the case of Ann and Bob above. Suppose they begin in the Nash equilibrium where each player is playing Heads with probability 0.5. Ann considers changing her strategy, by say playing Heads with probability 0.75. In a world where she changes, Bob's belief must be correct (condition 3), so Bob must believe that Ann is playing with probability 0.75. Since Bob must choose a best response to his belief (condition 2), Bob must play Tails with probability 1. And this situation is worse for Ann than her current situation, she'll lose 75% of the time instead of 50%. So she does not want to change.

Described this way, condition 3 is as mysterious as the oracle in the Death and Damascus problem. Why should Ann suppose that when she changes her strategy Bob will somehow know it? Is Bob psychic?

There are a few justifications for this relatively strong assumption, but if one is to be found, it will occur in the most competitive of situations: zero-sum games.

4.3 Zero-sum games

4.3.1 Being out-thought

Conflict of interest is central to the game we described between Ann and Bob: when one gains the other loses. This category of games has a central place in game theory and is called a zero-sum game.

her own payoffs in the game and (2b) the player is rational (Aumann and Brandenburger, 1995).

In zero-sum games, Nash equilibria have a special safety property. If I play my part of a Nash equilibrium in a zero-sum game, I maximize my minimum expected gain. I establish the highest floor for my payoff, regardless of what my opponent does. Since my opponent's gain is also my loss, we can also describe this as minimizing my maximum loss, or minimax. This provides an interested additional justification for playing a mixed strategy in zero-sum games: one can guarantee that even if the opponent magically predicts what you will do you have done your best.[14]

In matching pennies, if Ann flips her coin to decide her strategy, she guarantees herself an expected payoff of 0 regardless of Bob's choice. Had she adopted a different strategy, say choosing heads for sure, then she risks Bob choosing his best response and leaving her with an expected utility of -1.

Because we are starting with traditional decision theory as a basis, safety considerations do not have normative pull. In traditional decision theory, the maxim to maximize expected utility theory is all there is to one's decision. However, with some supplement, the safety property of mixed strategies in zero-sum' games might serve to provide groundwork for a normative argument.

True zero-sum games are rare, but there is a prominent example: poker.[15] In poker, an individual's cards are not completely revealed but actions are taken in sequence. One player chooses whether to bet (and in some cases how much to bet), another observes this and then chooses how to respond, and so on. In order to play the game well one must occasionally bluff, i.e. bet without a strong hand.[16] Of course, one should not bluff every time one has a bad hand. One must choose when to bluff.

Poker players often have the uncanny feeling that no matter what they do, a better player has always managed to out think them. If they choose to bluff, the strong player knows it. If they try bluff in different circumstances, the better player is there too. Without having an explicit causal theory for how this occurs, early poker theorists appealed to the minimax theorem and the

[14]This fact does not generalize to other games, however. Maximizing your minimum gain can harm you and your opponent in non-zero-sum games.

[15]Lest the erudite philosopher believe that poker is not a serious concern: Fiedler and Wilcke (2012) estimated the world market size of poker in 2010 was above $3 billion.

[16]von Neumann and Morgenstern (1953); Gillies et al. (1953) prove this for idealized forms of poker.

benefits of randomization (Sklansky, 1978).[17] Equilibrium-based reasoning has remained a part of poker theory ever since.

As in the Death in Damascus problem above, this appeal is only valid if the opponent can infer that you are randomizing but not able to infer the outcome of your randomization device. So, if the strong player can pick up a "tell" (a physical mannerism that reveals whether one is bluffing) a randomization device does no good. But if the strong player is only capable of predicting one's thought processes, then one ought to randomize. There remains a gap in explaining how another player might come to model my reasoning, but this cannot be any more mysterious than problems like Death in Damascus.

A similar story is offered in a very different context by the statistician R.A. Fisher. Fisher is confronted by a woman who claims to be able to tell if milk was added to a cup before or after tea was poured into it. Finding this claim incredible, Fisher designs an experiment in order to test the claim. Fisher will present the woman with eight cups containing both milk and tea: four where the milk was added first and four where it was added last. The woman will be tasked to guess which cup is which. He explicitly recommends that the order of the cups be randomized to prevent the woman from cuing in on other factors missed by the experimenter (Fisher, 1960). (Despite the randomization, the woman was able to determine which cups where which accurately.)

4.3.2 Learning

There is a second justification also found in the poker strategy literature. While we might model a single hand of poker as a "game" in game theory, rarely do people quit after a single hand. Instead people spend hours, sometimes days, playing poker together. The same players may confront one another over periods extending for decades. As a result, they run the risk that another player might come to learn their strategy.

When we model a single hand of poker using game theory, we are idealizing away this repeated game structure. The fact that another might have played against me before and might play against me again is ignored. If there were no computational limitations, we would prefer to model the entire

[17]Noted poker player and author Doyle Brunson tentatively endorsed the possibility of ESP (Brunson, 2002), but as far as I'm aware he is among a very small minority of poker authors who endorse this peculiar causal theory.

"meta-game" (as it's called in poker) of millions of potential hands. But, alas, we cannot.

When the "meta-game" is projected down into the one-shot case, learning resembles being out-thought. Suppose I am forced to adopt a single strategy at the start of my poker playing career. I stick with that strategy, giving other players a chance to learn it and respond appropriately. In this, admittedly extreme, situation the other players will look like the oracle in Death and Damascus: they will always be one step ahead of me. This occurs not because they have ESP, but because they learn and adapt while I do not.

Of course, I am not forced to adopt one strategy for my entire card playing career. I might adapt as well. But the rates of adaptation might not always be equal. Professional poker player Annie Duke describes this exact justification for playing an opponent who knew her well. Whenever she tried to adapt her play, her opponent, Erik Seidel, was able to adapt faster. In the end, she intentionally minimize her maximum loss by adopting a strategy that reduced Seidel's strategic advantage. In so doing, Duke defeated Seidel (Duke et al., 2013). Justifications along this line can be found in several poker strategy guides (e.g., Sklansky 1978; Harrington and Robertie 2004, 52-3; Brokos 2019).

In suppressing the actual repeated game structure, we have provided a different justification for the questionable assumption that my opponent can out think me. These reasons are similar to those discussed by Icard (2019), we might suppress the repeated game structure because I am incapable of remembering all my past plays. For reasons similar to those behind the absent-mind driver game, I might opt to randomize in order to ensure that I closely approximate what would be the ideal solution if I was sophisticated enough to implement it.

4.4 Can randomization be eliminated

Zero-sum games appear to offer an opportunity to justify mixed strategies similar to Death in Damascus but without the somewhat implausible assumptions of an oracle. Someone being more intelligent or a faster learner in a repeated game might generate phenomena similar to the oracle and therefore justify mixing. However, several arguments have been advanced to suggest that randomization in these settings might not be needed. The arguments are general, and apply to non-zero-sum games as well, but for our purposes we will keep this focus for illustration.

4.4.1 Endogenizing randomization in poker

In a very early academic paper analyzing mixed strategy in poker, Bellman and Blackwell (1949) show one might achieve the same result through a deterministic strategy. Consider the following simplification of Bellman and Blackwell's poker game. One player, the responder, is given a card with a real number $y \in [0,1]$ written on it. This card is face up, the number is known by both players. Another player, the bettor, is given a card with another real number written on it, x. This card is visible to the bettor but hidden to the responder. x is drawn from a uniform distribution over $[0,1]$ and this distribution is common knowledge between the players.

Each player antes \$1, and then the bettor can either bet \$2 or check (bet \$0). The responder observes whether the first bets, and if he does, the responder can call (equal the bet) or fold. If the responder folds, the bettor wins the pot regardless of the value of his card. If the responder doesn't fold, the numbers are compared and the higher number wins the pot. (The pot is evenly split in the case of a tie.)

If the bettor has a higher card than the responder (if $x \geq y$), then there is no reason not to bet.[18] But what should a bettor do if the card is lower? First suppose that the bettor is considering mixing in this case, betting with probability a and checking with probability $(1-a)$. The optimal mixed strategy is

$$a = \begin{cases} \frac{1-y}{2y} & \text{if } y \geq \frac{1}{3} \\ 1 & \text{otherwise} \end{cases}$$

With that value, the responder is indifferent between calling or folding when $y \geq \frac{1}{3}$ and will always fold otherwise.

Here we seem to have a classic case for the benefit of mixing. The bettor should "bluff" (i.e. bet with a card he knows to be a loser) with frequency a. Bellman and Blackwell point out that even here, mixing is not necessary. One can instead, adopt a deterministic strategy that depends on the card. So, for example, one can divide all the low-card values into two contiguous regions one of size ay and the other of size $(1-a)y$. If the bettors card is in one region, he bets. If it's in another region, he checks. Or the bettor could divide the interval up into several different regions. So long as the total area of the "bet" region remains ay any way of dividing is acceptable. All these strategies are illustrated in figure 6.

[18] Speaking more exactly, betting weakly dominates checking.

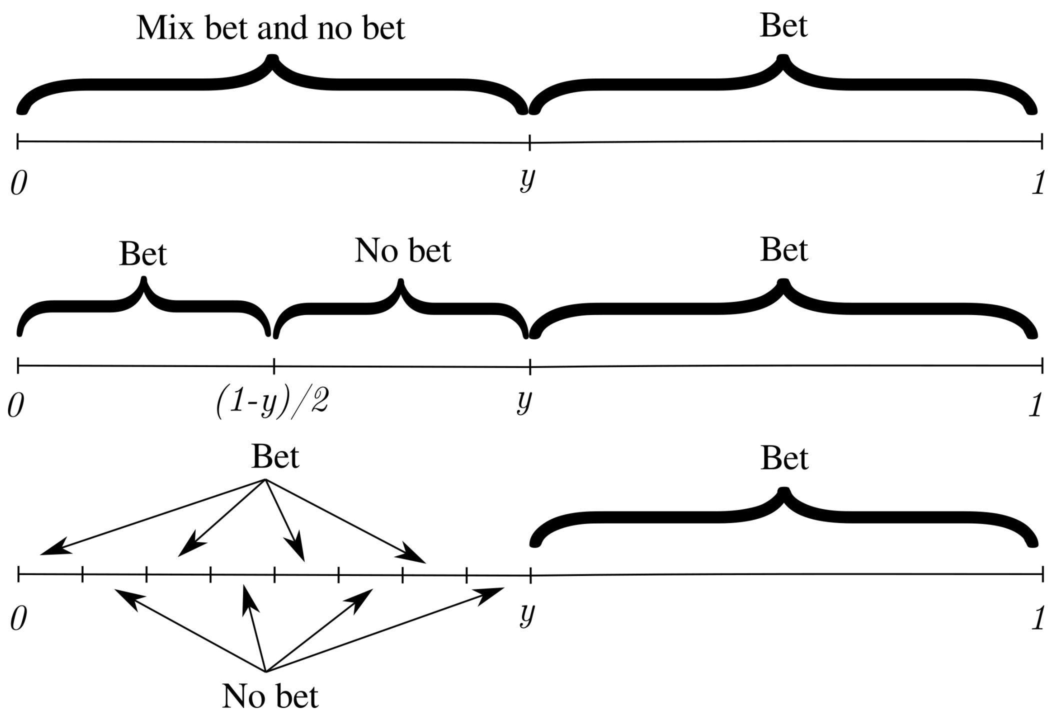

Figure 6: A representation of three equilibrium strategies in the simplified Bellman and Blackwell poker game. The top strategy involves explicit mixing. The bottom strategies involve utilizing the dealt card as a randomizing device.

Bellman and Blackwell consider a more general game, where the value of y is privately known by the responder, and show the same general conclusion holds (but with different strategies). Exogenous mixing in this case is unnecessary,

> This result, that only pure strategies are required in the continuous case, seems quite unexpected, and at first sight to run counter to the known (heuristic) theory of bluffing. However, the paradox is easily resolved. In a continuous game, one allows the game, which furnishes a random card, to do the bluffing. It turns out that if the cards are dealt at random, any further randomization furnished by mixed strategies on the part of the players is superfluous (Bellman and Blackwell, 1949, 601).

Independently, the poker strategy community had a similar realization. David Sklansky, a famous poker strategy author, wrote that a player might use his cards to furnish appropriate randomization in order to prevent his opponent from learning his strategy (Sklansky, 1978).

4.4.2 Other ways of being unlearnable

Although I am not aware of any formal analysis of this, the poker strategy community has offered a different justification. Several authors have pointed out that one need not even use the cards to furnish apparent randomization. In the context of suppressed repeated games, one must only be unlearnable. First and foremost, this is conditional on the learning strategy of one's opponent. One may not need to randomize against an opponent who is not actively attempting to learn (Miller et al., 2005, 252).

But even against a sophisticated opponent, randomization may be unnecessary. A player might condition her behavior on many game-relevant factors like how much money her opponent has, the strategies observed by opponents in the game, how long the player has been at the poker table, etc. If the opponent is unaware of exactly what a player is conditioning on, it will be difficult, if not entirely impossible, for them to learn what deterministic strategy the player is pursuing. Like the card in the example from Bellman and Blackwell, all these internal features of the game will present the opponent with an unlearnable—but pure—strategy (Duke and Vorhaus, 2011; Flynn et al., 2007).[19]

[19] My thanks to Annie Duke for a helpful discussion of this issue.

		Bob	
		Heads	Tails
Ann	Heads	$1+\epsilon\delta_A, -1+\epsilon\delta_B$	$-1+\epsilon\delta_A, 1$
	Tails	$-1, 1+\epsilon\delta_B$	$1, -1$

Figure 7: Perturned matching pennies

4.4.3 Harsanyi's purification argument

This general strategy of taking exogenous randomization and endogenizing it continued in game theory research for several decades. Harsanyi's "purification" theorem represents a kind of end-point for this research. Harsanyi imagined a base game subject to small, idiosyncratic payoff perturbations for each player. He assumed that these perturbations are private knowledge: each player only knows what are the actual perturbations of her payoffs not those of the other players.

Consider the perturbed matching pennies game in figure 7. There are two relevant sets of parameters. First are the noise parameters, δ_A and δ_B. These represents the perturbations of the parameters for each player. We assume that they are independent draws from (potentially different) distributions, and that each player knows the outcome of her draw, but not the others. Ann knows the realization of δ_A. Ann knows the distribution of δ_B, but does not know the realization of it. Similar for Bob.

The second parameter is ϵ. This represents the *amount* of noise, and is a fixed parameter known by both players. We will consider a sequence of games that approaches the traditional matching pennies game by considering what happens as $\epsilon \to 0$; what happens as the noise becomes a vanishingly small part of the game.

Since each player is aware of their own δ parameter, they can condition their behavior on that. But, since they don't know the other's δ parameter the other will appear to be adopting a mixed strategy. That is, even if Ann knows what Bob's strategy is—that is, what Bob's function from δ_B to his choice of H or T—she can't know whether he will play H or T on this particular game, since that will depend on δ_B.

For illustration, let's consider δ_A and δ_B are drawn from a uniform distribution on $[-1, 2]$. There is nothing special about this distribution, it just serves as an illustration. We could have chosen any other (subject to some minimal constraints).

Let's suppose that from Ann's perspective, Bob is going to play H with probability p. Then she will play H if:

$$\delta_A \geq \frac{2-4p}{\epsilon}$$

If we let q be the probability that Ann will play heads, this is equal to the probability that δ_A satisfies that constraint. Given our particular assumption about the distribution of δ_A, this means,

$$q = \frac{2-4p}{3\epsilon} + \frac{1}{3}$$

Now, suppose that Bob knows this to be true of Ann. From his perspective, Ann is implementing a mixed strategy, even though he knows that Ann is never explicitly mixing. Bob should play H if,

$$\delta_B \geq \frac{4q-2}{\epsilon}$$

Using this equation, we can calculate what p will be,

$$p = \frac{4q-2}{3\epsilon} + \frac{1}{3}$$

.

Solving for both p and q, we find that each player will play H with probability,

$$\frac{\epsilon^2 - 3\epsilon + 54}{3\left(\epsilon^2 + 36\right)}$$

Which approaches $1/2$ as $\epsilon \to 0$.

From each players perspective, the other player is playing the mixed strategy Nash equilibrium of the game (as the noise goes to zero). But from their own perspective each they are playing a pure strategy. Given the obtained value of δ_A, Ann will either play H or T with probability 1 and she is aware of this fact.

Harsanyi's theorem is very general, applying to most games of interest and many plausible distributions of the noise parameter. And it has important implications for game theory. From the modeling perspective, it justifies the use of mixed strategies in game theory. While neither Ann nor Bob are playing a mixed strategy, they will appear to by a modeler who is unaware of the values of the noise parameters. If we think of the noise as being generated by something internal to the players, a strange mood today or something, then it will look like randomization even though it's not.

But, what is good from the modeling perspective is bad from the normative perspective. What Harsanyi shows, like Bellman and Blackwell before, is that one need not actually mix to achieve the benefits usually attributed to mixed strategies. One can always play one's pure best response, but nonetheless be sufficiently unpredictable to prevent one's opponent from getting the upper hand.

In the discussion of act–state dependence in section 3.6, we saw that mixing qua mixing was not particularly important. Mixing appeared beneficial because it had the effect of being unpredictable to someone else, and unpredictability was the central value.

When we turn to game theory, the same thing holds true. What is critical in zero-sum games (and in some other games as well) is one's opponents' beliefs but not one's own beliefs. If the only way I can get my opponent to be unsure about my actions is to be unsure myself, then I would like to mix. But, if I can make my opponent unsure in other ways—by utilizing a feature of the game or by being unlearnable—then I might as well do that.

This way of viewing mixed actions has developed into a kind of consensus view, at least among those who work in the "epistemic program" in game theory. A central paper in this program describes this way of viewing mixed strategies:

> In recent years, a different view of mixing has emerged. According to this view, players do not randomize; each player chooses some definite action. But other players need not know which one, and the mixture represents their uncertainty, their conjecture about his choice. (Aumann and Brandenburger, 1995, 1162).

For our purposes in this paper, this view of mixed strategies abandons what motivated us to inquire about mixed strategy. For Aumann and Brandenburger no player is uncertain about her *own* action, she is uncertain about others. And this is all it means to adopt a mixed strategy for many contemporary game theorists. While this interpretation is a fruitful and interesting one, it steps away from our central focus because it no longer considers mixing as central.

5 Convincing others

In statistics, randomization is often required for various forms of statistical tests. For example, in randomized controlled trials, scientists must use a randomizing device to assign patients to one of two (or more) different treatments. This is a species of mixed strategy: rather than choosing for herself which treatment to assign to a patient, a scientist is opting to randomize.

There are many justifications for this behavior. Some are based on analyses from classical (i.e., non-Bayesian) statistics and will be left to the side here. Instead, we will focus specifically on Bayesian justifications where—as we have seen—the terrain is difficult.

The most common justification is that by randomizing the scientist removes potential confounders. If our scientist were to choose, for instance, to assign all men to treatment A and all women to treatment B, if the effectiveness of the treatment varied by sex then her choice would make the experiment less informative. For a Bayesian, this justification cannot go far because randomization admits the possibility (however small) that one treatment *will* contain all men and the other all women. If the scientists believes this would be a bad outcome, she ought to avoid it by choosing *deterministically* a different allocation of subjects to the two different groups. Kadane and Seidenfeld (1990) generalize this worry to other uses of randomization in statistics.

We have already discussed one potential defense when we talked about Fisher and the milked tea. If the scientist is suspicious that either a subject (or some other nefarious force) might observe the scientist's choice and act to harm the experiment, then randomization would be justified. (This is essentially the basis for the minimax-based statistical theory of Wald 1950.)

A second justification discussed in (Kadane and Seidenfeld, 1990) has a more distinctively social character. If the scientist is interested in convincing others—as well as herself—there may be a reason for randomization. Ann knows that Bob distrusts her. Bob believes that Ann might be willing to selectively place subjects into different treatments to increase the chance that one treatment appears effective. Bob may be suspicious of any data where Ann exhibited control over the assignment of subjects to treatments. In such a case, Ann has an incentive to choose a mechanism for assignment which takes the choice out of her hands. There are many such mechanisms, and not all of them involved randomness (Kadane and Seidenfeld, 1990).

That multiplicity can be a problem, however. Imagine Ann has at her disposal a large set of potential assignment mechanisms from which she could choose: perhaps she could use a certain type of pairing, randomization, or alternating assignment (first subject goes to A, second to B, etc.). Then, Ann might look to see which of those apparently "objective" assignment mechanisms would yield the sample most likely to produce her favored conclusion. (Or so a suspicious Bob might think.)

To avoid this problem, Ann and Bob would benefit, in the long run, to settle on a single mechanism for assignment. This would function like a convention (in the sense of Lewis, 1969). Ann would abide because she knew that if she didn't, Bob would ignore her study. Bob would distrust Ann if she didn't randomize because everyone else randomizes. In this setting Ann prefers a mixed strategy, not because of any effect it has on her, but because it is the only thing Bob will accept.

Kadane and Seidenfeld (1990) offer a slightly different justification. In the preceding paragraph, Ann randomized because it would convince Bob. We offered a Bayesian defense for how Bob might come to require such a thing. But why should we insist that Bob be a Bayesian? Bayesians must convince other Bayesians as well as non-Bayesians. If Bob is a follower of classical statistics, he may think that randomization is epistemically superior and in such a case Ann may want to do so in order to convince Bob.

While randomization functions as a convention in modern science, it might be a socially inferior one. Kadane and Seidenfeld (1990) provide ethical reasons to switch to a different convention which allows for the assignment of patients to treatments.

In this paper we are less interested in alternative social arrangements, but rather we are attempting to find justifications for randomization. In this setting we have found one, but it functions much like the cookie example in section 3.6. Ann randomizes not because she thinks its a good idea but because she thinks it will have a desirable effect on Bob. If she could choose a superior (in her mind) assignment of patients to different treatments but convince Bob that it was random, she would prefer to do that. Ethical and other constraints may prevent this, but the illustration shows the limit to value of randomization in this context.

6 Establishing rules

One last related justification for randomization comes from Smart's defense of utilitarianism (Smart and Williams, 1973). Smart imagines a setting where, in large community, the optimal action would be for a small fraction of the population to engage in some behavior while the rest refrains. Smart's actual example is somewhat misleading, so I will use a slightly more modern one.

The El Farol Bar problem, named after a bar in Santa Fe, New Mexico, imagines a bar that is nice when it's not too crowded. Many people are trying to decide whether, on a given Thursday night, they should go to the bar to see live music or hang out at home. If everyone goes to the bar, the bar will be too crowded and everyone will wish they stayed home. On the other hand, if no one goes, the bar will be empty and each person will have wished they had gone (Arthur, 1994).

There are many Nash equilibria for this game where some people go (with certainty) and others stay home (with certainty). However, the only *symmetric* Nash equilibrium—one where everyone adopts the same strategy—is a mixed equilibrium, where each person flips a biased coin to decide whether or not to go to the bar or stay home. As we've already discussed, when in equilibrium, no individual has a reason for preferring their mixed strategies over their pure choices.

The point that Smart raises is that the mixed strategy is the only policy or social norm that could be easily adopted by everyone (Smart and Williams, 1973, 57-62). Other norms, where some people go while others stay home will be (a) perceived as unfair and (b) would be difficult to implement, since it would require communicating a more complicated rule. ("If you are such-and-such, go to the bar, but if you are such-and-such stay home.")

This presents one last, somewhat limited, defense for mixed strategies. While they are not uniquely compelling from the perspective of those who are asked to implement them, they may have some appeal for those who wish to establish norms or policies. Smart considers this as a potential use for mixed strategies in rule-utilitarianism, where the rule that's endorsed involves a mixed strategy. There are thorny issues about the coherence of rule-utilitarianism that we needn't discuss here, but it is worth noting that this gives us one additional refuge where we might find some normative punch behind mixed strategies.

7 Conclusion

I have been studiously avoiding one common-sense use for coin flipping, expressed well by Danish poet Piet Hain,

> Whenever you're called on to make up your mind,
> and you're hampered by not having any,
> the best way to solve the dilemma, you'll find
> is simply by spinning a penny.
> No—not so that chance shall decide the affair
> while you're passively standing there moping;
> but the moment the penny is up in the air,
> you suddenly know what you're hoping.

This justification has little room in traditional decision theory. Decision theory normally assumes either (a) you know your preferences or (b) your preferences are arbitrarily determined by your choice behavior. In neither case can we come to learn our own preferences. But, non-ideal humans experience this frequently. We want something in a way that is inaccessible to us, and by flipping a coin we come to find out our preferences. Even Hein's justification only goes so far, however, it doesn't enjoin us to listen to what the coin tells us to do, only to convince ourselves that we will—momentarily—until we learn our true preference. After we've done that we can ignore the coin and do what we've wanted to all along.

There remain other interpretations of mixed actions. In evolutionary game theory, for instance, mixed strategies represent states of a population rather than an individual behavior. But even here, there is little normative endorsement of individual mixtures, merely statements about the state of a population.

Overall, there are severe constraints on when mixed strategies can be normatively endorsed in decision and game theory. In essence there is a single problem that cannot be avoided: if I prefer a mixture over the pure strategies it contains, then I will be unhappy—to a certain extent—when I find out the result of my mixing device. Any theory which leads me to prefer a mixed strategy must grapple with issues of coherence over time.

Even if one is happy to deal with this consequence, there are still only a few relatively limited situations where randomization is endorsed because of its inherent properties (as opposed to a conceptually distinct consequence of mixing, like influencing others' beliefs). There are a few modifications to decision theory that will endorse mixing, but only in particular circumstances.

We are back where we started: a common practice of using coins and the like to assist in decision making appears difficult to justify. And perhaps this is how it should be: the goal of decision theory is not just to tell us how great we are, but to point out that which does not stand up to scrutiny.

Dedication

I am overjoyed to be writing this paper in honor of my friend and colleague Teddy Seidenfeld. Since I came to Carnegie Mellon over a decade ago, Teddy has been a constant intellectual companion. Almost always over coffee or tea, Teddy has lent his considerable intellect and knowledge to me. Our conversations span a wide variety of topics, from politics to baseball to decision theory. But, somehow, no matter the topic he is always insightful. I have learned so much from him, and I cannot adequately express my gratitude for his kindness.

There are several different threads that run throughout Teddy's work. In this paper, I thought to grab one and pull hard. Readers familiar with Teddy's thought will see him throughout the paper, especially the first major section on decision theory. My talks with Teddy have so infected my brain, that I sometimes struggle to remember which ideas were mine and which were his. Nowhere is this more true than in this paper. What little I have to contribute can be found, principally, in the second section especially the connections with poker.

References

Ahmed, A. (2014). Dicing With Death. *Analysis 74*(4).

Arthur, W. B. (1994, may). Inductive Reasoning and Bounded Rationality. *The American Economic Review 84*(2), 406–411.

Aumann, R. J. (1987, jan). Correlated Equilibrium as an Expression of Bayesian Rationality. *Econometrica 55*(1), 1–18.

Aumann, R. J. and A. Brandenburger (1995, sep). Epistemic Conditions for Nash Equilibrium. *Econometrica 63*(5), 1161–1180.

Bellman, R. and D. Blackwell (1949). Some Two-Person Games Involving Bluffing. *Proceedings of the National Academicy of Science 35*, 600–605.

Bicchieri, C. and A. K. Chavez (2013). Norm manipulation, norm evasion: Experimental evidence. *Economics and Philosophy 29*(2), 175–198.

Brokos, A. (2019). *Play Optimal Poker: Practical Game Theory for Every Poker Player.* Independently Published.

Brunson, D. (2002). *Super System* (3rd ed.). Cardoza Publishing.

Buchak, L. (2013). *Risk and Rationality.* Oxford: Oxford University Press.

de Montaigne, M. (1877). *Essays of Michel de Montaigne.*

Duke, A., V. Russo, and J. Doody (2013). *Heads-up Tournament Poker: Hand-by-hand.* Cardoza Publishing.

Duke, A. and J. Vorhaus (2011). *Decide to Play Great Poker: A Strategy Guide to No-Limit Texas Hold 'Em.* Huntington Press.

Fiedler, I. and A.-C. Wilcke (2012). Online Poker in the European Union. *Gaming Law Review and Economics 16*(1-2), 21–26.

Fisher, R. A. (1960). *The Design of Experiments.* New York: Hafner Publishing Company.

Fleurbaey, M. and A. Voorhoeve (2013). Decide as you would with full information. In N. Eyal, S. A. Hurst, O. F. Norheim, and D. Wikler (Eds.), *Inequalities in Health: Concepts, Measures, and Ethics*, pp. 113–128. New York: Oxford University Press, USA.

Flynn, M., S. Mehta, and E. Miller (2007). *Professional No-Limit Hold 'em: Volume I.* Las Vegas: 2+2 Publishing.

Gibbard, A. and H. R. Varian (1978). Economic Models. *The Journal of Philosophy 75*(11), 664–677.

Gilboa, I. and D. Schmeidler (1989). Maxmin Expected Utility with Non-Uniqure Prior. *Journal of Mathematical Economics 18*, 141–153.

Gillies, D. B., J. P. Mayberry, and J. von Neumann (1953). Two variants of poker. In H. W. Kuhn and A. W. Tucker (Eds.), *Contributions to the Theory of Games (volume 2)*, pp. 13–50. Princeton: Princeton University Press.

Harrington, D. and B. Robertie (2004). *Harrington on Hold'em: Expert Strategy for No Limit Hold'em Tournaments. Volume I.* Las Vegas: 2+2 Publishing.

Icard, T. (2019). Why Be Random? *Mind (forthcoming)*, 1–23.

Jeffrey, R. C. (1990). *The Logic of Decision.* Chicago: University of Chicago Press.

Joyce, J. M. (1999). *The Foundations of Causal Decision Theory.* Cambridge: Cambridge University Press.

Joyce, J. M. (2002). Levi on Causal Decision Theory and the Possibility of Predicting one's on Actions. *Philosophical Studies 110*(May), 69–102.

Kadane, J. B., M. J. Schervish, and T. Seidenfeld (2008). Is Ignorance Bliss. *Journal of Philosophy 105*(1), 5–36.

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

Keynes, J. M. (1921). *Treatise on Probability.* London: Macmillan and Co.

Levi, I. (1990). *Hard Choices: Decisionmaking Under Unresolved Conflict.* New York: Cambridge University Press.

Levi, I. (2000). Review: The Foundations of Causal Decision Theory by James M . Joyce. *Journal of Philosophy 97*(7), 387–402.

Lewis, D. (1969). *Convention: A Philosophical Study.* Cambridge: Harvard University Press.

Miller, E., D. Sklansky, and M. Malmuth (2005). *Small Stakes Hold'em: Winning Big with Expert Play.* Las Vegas: 2+2 Publishing.

Nash, J. F. J. (1950). Equilibrium Points in N-Person Games. *Proceedings of the National Academicy of Science 36*, 48–49.

Piccione, M. and A. Rubinstein (1997). On the Interpretation of Decision Problems with Imperfect Recall. *Games and Economic Behavior 20*(1), 3–24.

Seidenfeld, T. (2004). A contrast between two decision rules for use with (convex) sets of probabilities: Γ-maximin versus E-admissibility. *Synthese 140*, 69–88.

Sklansky, D. (1978). *Sklansky on Poker Theory.* Las Vegas: GBC Press.

Skyrms, B. (1982). Causal Decision Theory. *The Journal of Philosophy 79*(11), 695–711.

Skyrms, B. (1990a). Ratifiability and the Logic of Decision. *Midwest Studies In Philosophy 15*(1), 44–56.

Skyrms, B. (1990b). *The Dynamics of Rational Deliberation.* Cambridge: Harvard University Press.

Smart, J. and B. Williams (1973). *Utilitarianism, For and Against.* Cambridge University Press.

Ullmann-Morgalit, E. and S. Mogenbesser (1977). Picking and Choosing. *Social Research 44*(4), 757–785.

von Neumann, J. and O. Morgenstern (1953). *Theory of Games and Economic Behavior.* Princeton: Princeton University Press.

Wald, A. (1950). *Statistical Decision Functions.* Wiley.