Credal dilemmas

Sarah Moss
ssmoss@umich.edu

Formal epistemologists have traditionally modeled rational belief states with probability measures, functions that map each proposition to a single real number. But many theorists have recently rejected this model as excessively precise. Sometimes your evidence for a certain claim is unspecific. And many have claimed that the rational response to unspecific evidence is for you to have *imprecise* credences, represented not by a single probability measure but by a set of measures.[2] Fans of imprecise credences face a formidable challenge, however. They must say how someone with imprecise credences should act when faced with various decision problems. In particular, imprecise credence fans have come under harsh criticism for failing to provide a decision theory that protects agents who follow it from foregoing sure money. Agents with imprecise credences seem doomed to act irrationally in diachronic cases, where they are called to make decisions at earlier and later times.[3]

In this paper, I respond to this challenge on behalf of imprecise credence fans. In fact, we have subtle and variegated intuitions about what actions are rational in various diachronic cases. Once we appreciate the complexity of these intuitions, we can see that contrary to the criticism described above, diachronic cases are in fact *evidence for* the essential claims motivating imprecise credence models. In §1, I review the case for accepting that agents have imprecise credences. In §2, I explain the case against imprecise credences in more detail. In §3, I investigate our intuitions about norms governing agents who act as though they are torn between different precise credal states. These agents face *credal dilemmas*, epistemic analogs of moral dilemmas.

---

2. This idea first surfaced among economists but has since been defended by a number of philosophers; classic defenses in the philosophical literature include LEVI 1974, JEFFREY 1983, VAN FRAASSEN 1990, WALLEY 1991, KAPLAN 1996, and JOYCE 2011.

3. For representative criticism, see WEATHERSON 2008, DORR 2010, ELGA 2010, and WHITE 2010.

I argue that recognizing the strong similarities between moral and credal dilemmas is a desideratum for any theory of the latter. In §4, I describe our complex intuitions about diachronic norms governing agents in credal and moral dilemmas. I argue that predicting these intuitions is a second desideratum of our decision theory. In §5, I use imprecise credences to model the doxastic states of agents facing credal dilemmas, and I give a decision theory for agents with imprecise credences. I argue that my decision theory satisfies both desiderata just described. In §6, I place my paper in the context of the debate over whether we should represent agents as having imprecise credences, explaining why my §5 arguments constitute support for imprecise credence fans.

This paper is a step forward in a more general movement towards "time-slice first" norms in epistemology. I describe this movement in §7. In a nutshell: rather than endorsing diachronic rules that inexorably bind the earlier and later opinions of a temporally extended agent, we should endorse variants of these rules that take time-slices of agents to be the fundamental subjects of epistemic evaluation. I argue that when we theorize about credal dilemmas, a time-slice first orientation prompts us to endorse just the right norms. In particular, my time-slice first theory does well in comparison with standard decision rules when it comes to acknowledging and explaining a simple and familiar way in which you can rationally change your mind.

1 The case for imprecise credences

Suppose you see a stranger start pulling objects out of a bag. He pulls out a tube of toothpaste, a live jellyfish, and another tube of toothpaste. And as you are watching, he is starting to pull out a fourth object. ELGA 2010 uses this example to motivate the imprecise credence view. Fans of imprecise credences are especially interested in your opinions about what the stranger will do next. According to the imprecise credence fan, you may not be able to say exactly how likely you think it is that the fourth object will be a tube of toothpaste. In fact, you may not even be able to say whether you think it is more likely than not that more toothpaste comes next. It may be correct to say that you simply have no settled opinion on the question. In addition, you may rest assured that your reluctance to have a settled opinion is appropriate. At best, having some exact real number assessment of the likelihood of more toothpaste would be a foolhardy response to your unspecific evidence. Some would go so far as to say that having that sort of response is simply epistemically impermissible. To sum up: intuition dictates that in some cases, alternatives to precise credal states are actual, appropriate, or even required.

Imprecise credence models enable us to respect all of these intuitions. In a case where you do not feel as if you have a specific real valued credence in a proposition, your opinions are best represented by a set of probability measures that assign different values to that proposition. Following VAN FRAASSEN 1990, let us call this set your *representor*. Having multiple probability measures in your representor may be the appropriate or required response to unspecific evidence.

It is important to recognize that certain intuitions are not sufficient to motivate imprecise credences. For instance, you may sometimes have latitudinarian beliefs about your opinions, such as the belief that there is no particular credence that you are rationally bound to assign to the claim that more toothpaste comes next. In addition, you may feel that this latitudinarian second-order belief is epistemically appropriate, or even required. But opponents of imprecise credences can respect the intuition that latitudinarian beliefs are sometimes actual, appropriate, or required. They may say that all agents are rationally bound to have precise credences, while denying that there are ever particular precise credences that agents are rationally bound to have. At the heart of the imprecise credence revolution is the intuition that some agents are deeply torn between multiple precise credal states, where simply ascribing latitudinarian second-order beliefs does not adequately capture this deep feeling of division. Standard decision theorists simply cannot do justice to the intuition that your opinions are symmetrically related to multiple probability measures. By contrast, imprecise credence fans can easily make equal use of multiple probability measures in representing your opinions, by saying that each of the relevant measures is contained in your representor.

2 The case against imprecise credences

It is fair to say that concerns about decision theory have been the biggest impediment to general acceptance of imprecise credences. DORR 2010 complains that "there is no adequate account of the way unsharp credences should be manifested" in decision making. WHITE 2010 criticizes some candidate decision theories. And the central theme of ELGA 2010 is that no decision theory for imprecise agents can accommodate basic intuitions about rational action. In particular, it is hard to say how imprecise agents should respond in diachronic decision cases, when they are offered bets at earlier and later times.

For example, say we start to bet on whether the fourth object pulled out by the stranger will be a tube of toothpaste. I offer you a bet at better than even odds: if the object is a tube of toothpaste, I give you ten dollars; if not, you give me nine

dollars. Just after that, I offer you a second bet: if the object is a tube of toothpaste, you give me nine dollars; if not, I give you ten dollars. Let us stipulate that you do not learn anything in between being offered these bets. And let us stipulate that you feel torn between a rather wide range of credences about whether toothpaste comes next. Assuming standard decision theory, some of those credences would recommend you take the first bet, and some would recommend you reject it. Some would recommend you take the second bet, and some would recommend you reject it. Imprecise credence fans must tell you what to do when you face these sorts of decisions, i.e. when different actions maximize expected utility according to different members of your representor.

ELGA 2010 separates decision rules for imprecise agents into three categories. *Strict rules* and *global rules* both constrain imprecise agents to act in ways that make them indiscriminable from agents with precise credal states.[4] A strict rule says that for any imprecise agent, there is a unique precise credal state such that the agent must evaluate all current and later bets as she rationally would if she were currently in that precise credal state. Elga complains that endorsing a strict decision rule undermines the motivation for accepting imprecise credences. Insofar as rational agents feel torn between credal states, rational agents feel torn between options. Insofar as unspecific evidence rationally precludes having precise credences, it precludes acting as if you have precise credences: "if your evidence doesn't nail down an exact probability…it would be very strange if it *did* nail down a completely precise pattern of rational betting odds" (6). By contrast, global rules allow an imprecise agent more flexibility in her decision making. There is no precise credal state that determines how she must evaluate current bets. Instead she may accept any bet sanctioned by some member of her representor without thereby having done anything irrational, as long as she has not already made decisions prohibited by that same representor member.[5] Unfortunately, global rules suffer from their own shortcomings. If current decisions are constrained by past decisions, different actions will be permissible for agents who share intrinsic properties when they face a certain decision. Elga argues that this is intuitively unacceptable. Whether you accepted some bet on a proposition should not

---

4. ELGA 2010 characterizes strict rules by analogy with the *midpoint rule*. He claims that "according to the midpoint rule, your evidence does nail down a completely precise pattern of rational betting odds" (6). I find it hard to make sense of this claim, since the midpoint rule prescribes incoherent betting behavior. Hence I am not adopting precise definitions of 'strict' and 'global' from ELGA 2010; rather, I am using these expressions to refine informal notions that Elga introduces. See the appendix for further details.

5. For those keeping careful score: I intend this definition of 'global' to subsume the *Caprice* rule from WEATHERSON 2008. By rejecting the second bet, you thereby reject both bets, which is irrational according to Caprice. Accordingly, rejecting the second bet is sanctioned only by representor members that prohibit your rejection of the first bet. The narrowing proposal from ELGA 2010 is not among my 'global' rules, but I shall concede the former as having been discredited by arguments that discredit the latter.

immediately constrain your opinions about that proposition, nor should it constrain the rational actions available to you as an agent with those opinions.

I am defending fans of imprecise credences. For the sake of argument, I will grant their opponents that strict and global decision rules are unacceptable. This is actually a rather insignificant concession, for two reasons. First, arguments in this paper will expose flaws of strict and global rules. Second, the most natural and widely advocated decision rules for imprecise agents are neither strict nor global rules. A majority of imprecise credence fans go in for a third option: *permissive* decision rules. According to a permissive rule, you may perform an action without thereby having done anything irrational just in case that action is permissible according to some member of your representor. In the toothpaste case, rationality does not legislate your response to the first bet. It is okay for you to accept it, but also okay for you to reject it. Similarly, it is okay for you to accept the second bet, but also okay for you to reject it. Neither action will commit you to having done anything irrational.

Permissivism seems like the most natural response when you are torn between accepting and declining a bet. Intuitively, if you are of two minds about a decision, rationality alone cannot break symmetries in your credal state. Each action must be at least as permissible as the other. Hence as long as some action is okay, either is okay. There is just one problem for permissive decision rules. They excuse apparently irrational behavior. For example: according to a permissive decision rule, you may rationally reject the first toothpaste bet. And after that, you may also reject the second toothpaste bet, without thereby having done anything irrational. But opponents of imprecise credences protest: if you had simply accepted both bets, you would have gained one dollar, no matter what the stranger pulled out of his bag. Hence by the time you reject the second bet, you will have foregone sure money for nothing. Foregoing sure money for nothing is clearly irrational.

Opponents of imprecise credences conclude: permissive decision rules are incorrect. There is no viable decision theory for imprecise agents. And this should inspire great pessimism about whether agents indeed have imprecise credences at all. After all, standard decision theory has no problem predicting that it is impermissible for any agent to reject both bets. Say you have a precise credal state and you rationally reject the first bet. Then rejecting the first bet maximizes expected utility according to your credal state. By stipulation, you do not learn anything between rejecting the first bet and considering the second. So when considering the second bet, you have the same precise credal state as before. Since rejecting the first bet maximizes expected utility according to your credal state, so does accepting the second bet. Hence it is

impermissible for you to reject the second bet after rejecting the first.[6]

In more generality, standard decision theory entails the following claim:

(DDB)[7]   If accepting each of a pair of bets would guarantee you sure money, and you rejected the first and have not since learned anything, then you cannot rationally reject the second bet.

Furthermore, (DDB) entails our intuition that you should not forego sure money for nothing in the toothpaste case. Hence standard decision theory entails our intuition about the toothpaste case. Permissive decision rules do not. To sum up: it appears that the score so far is 1–0, Sharpies.

3 Moral dilemmas and credal dilemmas

The score is far from settled. So far we have explored our intuitions about one example. But we should be more careful about what intuitions we take as data for our decision theory. It is true that in many cases, it is intuitively bad to reject each of a pair of bets that guarantee sure money. And when a case involving such bets is underdescribed, we naturally fill in the details for ourselves so that rejecting both bets seems bad. But it is not always intuitively bad to reject both bets, even when you do not get additional evidence between rejecting the first bet and considering the second. In fact, we can describe cases where it is intuitively okay to reject each of a pair of bets that guarantee sure money.

Just imagine a case where you feel deeply torn between two very different options. The best way to describe your mental state is to say that you wish you could split into two people and try both. After some amount of agony, you commit to one half of yourself over the other. In particular, you do something that the other half would not, by foregoing some benefit that the other half prefers. But afterwards, you have second thoughts. You do not learn anything. You simply change your mind. The other half prevails after all. Of course, now you regret what you did before. But like a rational agent, you are determined to ignore sunk costs. Having now identified yourself with your other half, you act accordingly. In other words, you do things that your first half would not, like foregoing benefits that your first half prefers.

---

6. This argument highlights assumptions made for ease of exposition: I set aside problems involving memory loss or *de se* updating, and I assume that standard decision theory is accompanied by the rule that you must update your credence distribution by conditionalizing it on the information you learn.

7. As in *don't decline both* or—more loosely speaking—*diachronic Dutch book*. This sort of argument against imprecise credences is developed at length in ELGA 2010. For further discussion, see WEATHERSON 2008, WHITE 2010, JOYCE 2011, and WILLIAMS 2012.

For a more concrete illustration of how you might rationally change your mind, consider the following example:

> Your elderly mother must move in either with you or with your sister across the country. In saying where she should live, the only thing you care about is that your mother is as happy as possible. But it is impossible for you to judge where she would be happier. There is nothing to decide the question. After many sleepless nights, you make up your mind to have her live with you. Just a few hours later, a friend offers you 5000 frequent flyer miles in exchange for five local bus tickets. But since your mother would enjoy using the bus to get around, it makes more sense for you to keep your tickets.

> Before calling your family and the moving van company, you decide to sleep on your decision. The next morning, you wake up with a pit in your stomach. It is not that you have gotten relevant evidence overnight. The decision is just as intractable as before. You are simply having second thoughts. Eventually you completely change your mind. Of course, now you wish you had taken your friend up on the frequent flyer miles. By coincidence, another friend offers you six local bus tickets in exchange for 5000 miles. But since you plan on visiting your mother and sister often, it makes more sense for you to keep your miles.

If you had first traded for the miles and then traded for the bus tickets, you would be one bus ticket ahead. But intuitively, your rejecting both trades does not make you irrational. In any case where you are torn between credal states, it is natural to change your mind before settling on a plan. In the case just described, losing out on a bus ticket is a consequence of that sort of behavior. The most intuitive assessment of the case is that your actions are a rational reflection of your badly divided opinions. To judge otherwise fails to recognize a simple and familiar way in which we can rationally change our minds.

In my initial abstract description of a case where you change your mind, I did not specify why you feel torn between options. In fact, there are two ways for a decision to be difficult. It could be that you feel torn between multiple credal states, as in the example just given. But it could also be that you feel torn between multiple value functions. In fact, the latter yields a more philosophically familiar variety of indecision. In our ordinary descriptions of moral dilemmas, we often mention conflicting values that recommend different actions. And moral dilemmas offer further intuitive examples of cases where agents can rationally change their minds. Consider the student from SARTRE 1946 who is torn between joining the Free French in England and helping his mother at home. Suppose he elects to stay with his mother. He may still have second thoughts. If he does eventually leave for England, he is not irrational in virtue of having missed the earliest possible convoy. In a similar vein: if you change careers, you are not irrational in virtue of having acquired some unnecessary experi-

ence at your first job. If you change your order at a restaurant, you are not irrational in virtue of having said something else at first. It is intuitively permissible for you to reconsider your decision about your elderly mother because of a multiplicity of values, rather than opinions. Perhaps such reconsideration is best described as a change of heart, rather than a change of mind. Either sort of change may be unfortunate, but intuitively neither is rationally forbidden.

A brief excursus about terminology: I use 'moral dilemma' without taking a stand on whether there are moral dilemmas in the technical sense developed by authors such as WILLIAMS 1965, McCONNELL 1978, CONEE 1982, and BRINK 1994. For instance, I do not presuppose that moral dilemmas involve conflicting obligations. Instead I follow authors such as FOOT 1983 and RAILTON 1996 in recognizing 'moral dilemma' as an ordinary language expression and using it with roughly its everyday meaning. BRINK 1994 explains that we may use 'moral dilemma' this way while remaining neutral on whether there are moral dilemmas in the technical sense, "provided we do not understand 'moral dilemmas' univocally in both contexts" (246). Similarly, I use 'incommensurable values' for values that agents feel torn between, such as those that they struggle with in ordinary moral dilemma cases. I do not presuppose that incommensurable values correspond to options that are "incomparable" in the sense of CHANG 1997b or "on a par" in the sense of CHANG 2002.[8] And to avert potential misunderstanding: my sense of 'moral dilemma' is subject-oriented rather than speaker-oriented, i.e. we should say agents face moral dilemmas in my sense when different actions are recommended by incommensurable moral values that they themselves have, as opposed to incommensurable moral values that we have. I presuppose throughout that our theory of moral dilemmas should follow from a general theory for agents deciding between actions recommended by incommensurable values of any sort, and I assume in §5 that it makes sense to use the sort of value functions I am talking about to compute the expected value of an action. All these assumptions fit best with literature on incommensurable values as opposed to moral dilemmas, but the latter is nevertheless a source of fruitful examples in the present context.

To sum up so far: in both moral dilemmas and credal dilemmas, sometimes it is intuitively impermissible to reject *good pairs of bets*, i.e. pairs of bets that guarantee sure money. ELGA 2010 gives one such example. But in both sorts of dilemmas, sometimes it is intuitively permissible to reject good pairs of bets. I have given several such examples. All these examples raise an important question: which intuitions should

---

8. CHANG 2005 claims that parity plays a "distinctive role" in rational choice because decisions between options on a par must be governed by a global decision rule (347). If Chang understands this feature as essential to the nature of parity, then I will not be discussing options that are on a par in her sense. However, if the essential feature of parity is that it answers a demonstrated need for a fourth value relation in CHANG 2002, my claims are compatible with my discussing options that are on a par.

we aim to vindicate in building our decision theory for agents in credal dilemmas? The upshot of reflecting on moral dilemmas is that our answer to this question should be guided by the following desideratum: *our theory of credal dilemmas should mirror our theory of moral dilemmas*.

It is not just that we have similar normative judgments about moral and credal dilemmas. In either case, it feels as though your decision necessarily falls outside the purview of standard decision theory. It is often said that moral dilemmas seem to demand some sort of arbitrariness from an agent, and the same has been said for credal dilemmas: "it must be really inexplicable why an agent chooses" according to one representor member instead of another (WEATHERSON 2008, 13). In light of these similarities, we would need to have some positive reason motivating us to theorize differently about sources of the same sort of practical indecision.

Furthermore, it is often unclear that you can even distinguish sources of practical indecision as reflecting multiple credences rather than multiple values. Deciding whether to care for your mother may call to mind incommensurable values, as noted by SARTRE 1946, RAZ 1986, and others. Changing your career may call to mind incommensurable values, but also different credences about what will make you and others most satisfied. It is not obvious that one can ultimately give solid grounds for attributing indecision to a multiplicity of values rather than a multiplicity of opinions about what is valuable. If different theories governed credal and moral dilemmas, one would not expect that we as theorists would have such a hard time discriminating credence-related and value-related sources of practical indecision.

WILLIAMS 1965 famously argues that we ought not assimilate conflicts of belief and conflicts of desire:

> (a) If I discover that two of my beliefs conflict, at least one of them, by that very fact, will tend to be weakened; but the discovery that two desires conflict has no tendency, in itself, to weaken either of them...(b) Suppose the conflict ends in a decision...The rejected belief cannot substantially survive this point, because to decide that a belief is untrue *is* to abandon, i.e., no longer to have, that belief, [while] a rejected desire, however, can, if not survive the point of decision, at least reappear on the other side of it on one or another guise.

This passage nicely brings out why *imprecise credences*, rather than *inconsistent beliefs*, should be the epistemic analog of incommensurable values. Being aware that you have multiple precise credal states in your representor does not have a tendency, in itself, to weaken any of them. In both moral and credal dilemmas, it is merely that some practical decision forces you to identify with one rather than the other. And in just the same way that rejected desires remain live, so do rejected credal states. "Should I have taken that bet about the toothpaste? Is it really more likely than not

that more toothpaste is next?" "Is my mother really better off with my sister? Should I have brought her to live here with me, after all?"

The arguments in this section yield a preliminary upshot for the debate over imprecise credences. In the moral dilemmas literature, a number of theories say that agents with incommensurable values may violate standard norms of rationality, say by rejecting good pairs of bets. This has prompted occasional skepticism about whether agents in fact have incommensurable values, but it has by no means settled the debate. In particular, it remains an open question whether we should aim to predict that rational agents simply cannot reject good pairs of bets. By comparison, it seems hasty to use diachronic cases to rule out analogous theories of credal dilemmas and thereby conclude that rational agents must have precise credal states.

4 How intuitions should inform our decision theory

If we can rationally reject good pairs of bets in moral dilemmas, we can rationally reject good pairs of bets in credal dilemmas. But should we use this as the premise of a *modus ponens* or *modus tollens* argument? Roughly speaking, CHANG 1997b opts for *modus tollens*. She describes a case where an agent loses money in virtue of having incommensurable values, and she concludes that "a pragmatic challenge to those who would oppose comparativism is to provide a well-motivated, non-ad-hoc account of how practical reason prohibits agents from becoming 'merit pumps'" (11). But we need not follow Chang here. In short, there are two strategies for explaining the intuitions canvassed so far. Strategy A: we give a decision theory that entails (DDB). Our theory is incompatible with our intuitions about cases where it seems agents may reject good pairs of bets without learning, like the case of the elderly mother in §3. But we explain away those intuitions by other means. Strategy B: we give a decision theory that does not entail (DDB). Our theory is compatible with our intuitions about cases where it seems agents may reject good pairs of bets. But we must find some alternative explanation for our intuitions about cases where it seems agents may not reject good pairs of bets, like the cases described by Elga and Chang. We cannot use (DDB) to explain those intuitions.

Standard decision theory commits us to Strategy A. But Strategy B seems preferable, *ceteris paribus*. Strategy A requires that we give an error theory of some ordinary judgments. Strategy B merely requires that we supplement our decision theory before predicting ordinary judgments. Strategy B allows that subtle differences between cases may settle whether agents may reject good pairs of bets without learning. By contrast, it seems like overkill to apply (DDB) to explain our intuitions across the board.

10

This preference for Strategy B is even more compelling once we appreciate the subtle contours of the normative intuitions we are aiming to explain. It is not just that we say that rejecting good pairs of bets is sometimes bad and sometimes okay. Our judgments are coordinated with specific features of dilemma cases. For instance, Elga and Chang both describe cases where agents vacillate immediately and act without any evidence of psychological effort. To give a relatively clean example: say I offer to bet you five cents that the stranger will pull out toothpaste next, and you say, "Of course." Just a second later, someone else says they will bet you five cents that the stranger will pull out toothpaste next, and you say, "No way." That seems irrational; do you want the bet, or not? However, notice that your actions seem more acceptable as we increase the time between bets, and as we fill in the case with evidence of your feeling torn over whether to accept each bet. If you hesitate and fret and ponder, we are less likely to criticize you for acting differently the second time around. And that is how we are apt to fill in the details of your decision about your elderly mother, the decision faced by Sartre's student, and many decisions about what career to pursue. In these decisions, it typically takes time for agents to reconsider decisions, and it typically takes effort for them to make decisions in the first place.

Our normative judgments are also coordinated with the frequency of alleged changes of mind. Advocates of (DDB) often focus on examples where agents frequently change from one frame of mind to the next. CHANG 1997b points out that it is intuitively irrational for an agent to repeatedly trade between ever more lukewarm cups of tea or coffee. It is intuitively irrational for Sartre's student to dither about on train platforms hopping between trains to join the Free French and trains back to his mother in Paris. Changes that appear strategic are still another cause for censure. Say the student perpetually has misgivings about the army just as his train is pulling into base camp, and misgivings about his mother only at the end of his long rides back to Paris, so that he never does anything for either side. For an epistemic analog, say that a conspiracy theorist is of several minds about the likelihood of his favorite theory. In response to a practical decision, he finally sides with some assessment of the likelihood. But as he later gets evidence that disconfirms his theory, he reconsiders his prior probabilities so that his assessment of the likelihood of his theory after conditionalizing on his evidence matches his earlier assessment of its likelihood. It looks for all the world like the itinerant student is failing to act on either of his professed values, and like the conspiracy theorist is failing to conditionalize. By contrast, a single genuine change of mind seems far more acceptable in either case.

This brings us to our second desideratum for a decision theory for agents in credal dilemmas: *our theory should predict the contours of our normative judgments about*

credal dilemma cases. The (DDB) rule states a sufficient condition for an action to be impermissible. If an action is impermissible according to the rules of decision theory, then it is impermissible simpliciter. But our ordinary normative judgments call for a theory that is more sensitive to the contexts in which agents reject good pairs of bets. In the next section, I outline such a theory. Having identified our second desideratum, we can reconceive the dialectic from ELGA 2010: ordinary judgments may count in favor of theories that let agents reject good pairs of bets, rather than against them.

This way of reclaiming a dialectical advantage is a distinctive sort of move that may prove useful in a wide range of debates. For example, compare the present discussion with the recent debate over reverse Sobel sequences in the philosophy of language. The dynamic semantic rules given by VON FINTEL 2001 and GILLIES 2007 state a sufficient condition for a sequence of counterfactuals to be infelicitous. If a sentence is infelicitous according to the rules of semantic theory, then it is infelicitous simpliciter. But in MOSS 2011, I argue that our ordinary language judgments call for a theory that is more sensitive to the contexts in which reverse Sobel sequences are uttered. In both debates, rigid theories are built to explain our reactions to generic or typical cases, and an appreciation for the contours of our ordinary judgments should prompt us to replace those rigid theories with more fitting alternatives.

5 A decision theory for agents in credal dilemmas

If we model agents as having imprecise credences, we can give a sensible decision theory for agents in credal dilemmas. In modeling your mental state, we shall use precise credal states and value functions as ingredients. It is simplest to conceive of both as functions on propositions, where a credal state measures the subjective likelihood of a proposition and a value function measures its value. But where possible, I remain neutral on details of implementation. For decision theoretic purposes, our model of your mental state consists of two elements: (1) a set of ordered pairs, each consisting of one precise credal state and one value function, and (2) one distinguished such ordered pair. The set of ordered pairs represents your imprecise credences and incommensurable values. The distinguished ordered pair represents the precise mental state that you identify with for purposes of action in a practical decision situation.[9]

In a decision situation, an agent must act to maximize expected value according to the precise mental state she identifies with. But an agent may rationally change which mental state she identifies with, so that she must then act to maximize expected value

---

9. It is natural to assume that (1) is the set of all ordered pairs constructed from a particular set of precise credal states and a particular set of value functions. It is also natural to assume that (1) contains (2) as a member. For ease of exposition, I adopt both simplifying assumptions.

according to some other part of her imprecise mental state. This is one way in which an agent might "change her mind," colloquially speaking. She may start to identify with another precise mental state even if she has already made decisions that did not maximize expected utility according to that state. That is why my theory does not entail (DDB). It is permissible for you to identify with a member of your imprecise credal state that rejects the first of a good pair of bets, and then with a member that rejects the second, whether or not you have learned anything in between.

This decision theory yields norms for agents with imprecise credal states and agents with incommensurable values. In what follows, I will focus on its consequences for agents in credal dilemmas. As a decision theory for such agents, our theory clearly satisfies the desideratum introduced in §3, namely that our theory of credal dilemmas should mirror our theory of moral dilemmas. For one thing, the theory itself gives a symmetrical treatment of imprecise credal states and incommensurable values. But more distinctively, our theory of credal dilemmas mirrors a number of extant theories of how agents should respond to moral dilemmas.

For instance, it is commonly suggested that agents can respond to moral dilemmas only by virtue of having some sort of bifurcated mental state with precise and imprecise components. BRATMAN 2006 says that agents in moral dilemmas adopt "an analogue of acceptance in a context, in the absence of belief" about what is valuable (712). RAZ 1997 says that an agent exercises his will in order to choose between incomparable options, so that "the ability and willingness to choose does not depend on valuing the chosen option more than the rejected one" (338). BLACKBURN 1996 says you may "plump" for an alternative when forced to act with incommensurable values. He is clear that plumping should not involve coming to feel that your decision was correct, as admiring such behavior would amount to "admiring a lack of understanding of our past selves" (132). All of these authors are clear that agents do not respond to moral dilemmas by adopting one of the precise credal states they are initially torn between.

It is also common for moral theorists to endorse permissive decision rules for agents with incommensurable values. NAGEL 1987 argues that agents in moral dilemmas can "proceed without further justification, but without irrationality either" (180). RAZ 1997 argues that when facing moral dilemmas, we choose from among many rationally eligible options, and "we are within our rights to change our minds" (119). HARE 2010 explicitly favors a permissive decision rule for agents deciding between options that are on a par. Finally, BROOME 2001 develops a permissive theory at length. According to Broome, you are required to respond to moral dilemmas in accordance with your intentions, but you are not required to have those intentions

rather than others. For example, given that Abraham intends to sacrifice Isaac, it is not the case that he should turn around and head for home after walking halfway up Mount Moriah. But nothing makes it the case that he should continue intending to sacrifice Isaac: "He has no reason not to repudiate [his intention], because he had no reason to form this particular intention in the first place. And by repudiating it he can release himself from the requirement it imposes on him. Provided he repudiates this intention, there is indeed nothing irrational about his turning back" (118). My theory of credal dilemmas resembles Broome's theory of moral dilemmas in several respects. For instance: what you ought to do in a particular decision situation depends on a psychological fact. This fact corresponds with some precise mental state such that you must act as that state recommends according to standard decision theory. But no rule of rationality demands that relevant psychological facts remain constant from one decision situation to the next. Hence it is rationally permissible for Abraham to change his mind about whether to sacrifice Isaac, even if he is already halfway up the mountain. In a similar spirit, Broome 2000 argues that you may rationally change your mind and "make the best of a bad job" by deciding to start a new career, as long as you repudiate your earlier career decision (34).

This brief digest of moral theories illuminates two very different sorts of intuitions about moral dilemmas. On the one hand, we must recognize that you cannot rationally make up your mind in more than one way at once. There is some sense in which your choice *for* the army is a choice *against* your elderly mother, and that is exactly why the former choice is so distressing. In just this sense, as Broome suggests, you repudiate an earlier action when you reject the second of a good pair of bets. On the other hand, we must recognize that when you have incommensurable values, you may bear equal allegiance to each value regardless of what actions you perform. There is some sense in which *nothing* you do should count as a choice against *either* of your values. In just this sense, as Blackburn suggests, you should remain neutral between accepting and rejecting bets that take sides on moral dilemmas, even after accepting or rejecting them. Just the same goes for agents in credal dilemmas. There is an intuitive sense in which imprecise agents make up their minds when they act, while there is an intuitive sense in which they do not. The theory I defend captures both senses. An agent identifies with a precise credal state when she acts, while she retains multiple credal states in her representor. An agent is judged according to what some distinguished precise credal state recommends, while every member of her representor is equally eligible for that position of distinction. To sum up so far: my theory recognizes deep similarities between credal dilemmas and moral dilemmas, reflecting results from a wide variety of literature about the latter.

The theory I am defending gives precise interpretations of notions like *making up your mind* and *changing your mind* and *having second thoughts*, and so it is natural to pair my theory with psychological theories of those notions as applied to agents facing credal or moral dilemmas. In particular, it is useful to reflect on what you might ordinarily take as evidence that someone did in fact change her mind about something. Here I do not mean conclusive evidence, or evidence sufficient for justified belief or knowledge, but rather any claim that raises your credence that some change of mind occurred. To take a simple example: it is intuitively more likely that an agent will change her mind over the course of a few days than that she will change her mind over the course of a few seconds. If an agent hesitates and frets and ponders over some decision, it is more likely that she has imprecise credences with respect to that decision, and so more likely that she will have second thoughts about the decision. On the other hand, it is unlikely that an agent will have second thoughts, followed by third thoughts, followed by fourth thoughts, and so on for a large number of iterations. For most agents, relentless reconsideration is theoretically possible but generally rare, perhaps in part because deliberation itself carries nontrivial costs.

In conjunction with these rather plain observations, my decision theory meets our second desideratum. That is, my theory predicts the contours of our normative judgments about credal dilemma cases. In a nutshell: my theory predicts that we should be more likely to say that an agent can rationally reject good pairs of bets insofar as we have evidence that she genuinely changes her mind between bets. Since an agent is unlikely to change her mind immediately or without any evidence of psychological effort, we are unlikely to say that an agent can rationally reject good pairs of bets under those circumstances. For the same reason, we are unlikely to judge that an agent is acting rationally if she is acting in ways that are permissible only if she is relentlessly reconsidering her decisions. Furthermore, my theory correctly predicts that our leniency increases in proportion with our evidence that agents are genuinely changing their minds.

In addition, my theory explains why we censure agents who vacillate strategically, like the itinerant student who has misgivings only at the end of long train rides, or the conspiracy theorist who has misgivings only when his favorite theory is disconfirmed. It would be a ridiculously unlikely coincidence if you genuinely changed which values or credences you identified with at exactly those moments where acting on such changes happened to best serve purely self-interested aims. Hence it is ridiculously unlikely that the soldier and conspiracy theorist are rational because they are constantly changing their minds and acting to maximize expected value according to each novel mental state they identify with. There are much better explanations

of their behavior available. It is most likely that the student does not actually have incommensurable values. He values self-indulgent concerns, and dithering on trains has high expected value according to that value function. It is most likely that the conspiracy theorist does not have imprecise credences. He is blindingly attracted to his favorite theory. His high credence in that theory is independent of his evidence, so he does not ever really update by conditionalizing on what he learns.

In the moral dilemmas literature, some have been extremely skeptical about vacillating agents, going so far as to say that genuinely identifying with values precludes frequently switching between them. For instance, Hsieh 2007 suggests that apparent frequent switching between values really amounts to adopting none of those values, perhaps because someone who switches between values fails to appreciate what makes each value choiceworthy (77). In the same spirit, one might claim that the conspiracy theorist could not possibly be identifying with precise credal states if the conditional probabilities of those states never actually dictate his responses to further evidence. However, it seems hasty to conclude that frequent switching between credal states is impossible. This sort of philosophical debate is familiar. In the abstract: there is some mental state that often coincides with some behavior. The behavior is good evidence that an agent is in that mental state. But one might argue for a stronger connection, saying that the behavior is a necessary condition for being in the mental state. These sorts of arguments typically produce strong conclusions: inferring according to *modus ponens* is necessary for understanding what 'if' means; being disposed to self-ascribe pain is necessary for being in pain; being disposed to correctly identify what sortal an object is in is necessary for being able to have thoughts about that object. But if special circumstances provide counterexamples for these strong conclusions, it seems best to endorse generic and defeasible claims in place of universal ones. In each case, a generic claim is still enough to support the practice of citing some behavior as evidence for some mental state, as in: "But he was never thinking about Emerson Hall; he was treating 'Emerson Hall' like the name of some moral philosopher!" or "But he was never a committed soldier; he just rode the train to wherever he wasn't needed!" It is possible to endorse these informal arguments while admitting that they are not logically valid. Promoting defeasible generalizations to conceptual truths risks mistaking exceptions for impossibilities. And this concern speaks in favor of allowing that someone might genuinely change his mind again and again, and even by coincidence thereby end up acting always in accordance with what best suits his self-interested concerns.[10]

---

10. Throughout this discussion, I have assumed that we can distinguish identifying with self-interested concerns from merely acting in accordance with them. The literature on rule-following demonstrates that it is notoriously difficult to spell out this sort of distinction precisely. I expect that puzzles about

So far I have argued that we blame an agent for rejecting good pairs of bets when we have evidence that she did not change her mind between bets. In more limited respects, we may sometimes blame an agent for rejecting good pairs of bets even if we have evidence that she did change her mind. And so we have a small arsenal of additional reasons why we may blame an agent for rejecting good pairs of bets, without appealing to (DDB). For instance, it could be irrational for you to move your elderly mother back and forth across the country every month, even if you did change your mind every month about whether she would most enjoy living with you or your sister. But that may just be because the cost of moving will eventually be high enough to outweigh the expected benefit of the move. For similar reasons, it could be irrational for a president to vacillate between deploying and withdrawing troops from a high-stakes conflict. It could be irrational for a scientist to vacillate between half-finished experiments, rather than seeing some project through to its conclusion. This need not have anything to do with whether she has changed her mind. The cost of setting up a new experiment may simply outweigh the relative benefit of running it. If you inherit a lab from someone, you may rationally prefer continuing some of his experiments even if you believe other experiments would have been slightly more fruitful.

The costs of changing your mind may not be so immediate or obvious. For instance, certain choices carry benefits that are realized only if you stick with them for a long while. The enduring nature of your longest relationships with friends, partners, or political causes may help those relationships have special features that you value. The same goes for enduring identification with precise credal states. A scientist may confirm some unlikely theory in virtue of identifying with some unusual rational response to her evidence. If her unspecific evidence does not decide between various credal states, identifying with one assessment of likelihoods may give her the resolve to pursue certain rewarding experiments. In a similar spirit, subjects have reported trying to prove or disprove some claim for long periods, only to quickly succeed after finding out whether the claim is true or false. In a nutshell: aligning yourself with some response to unspecific evidence can have practical advantages, and those advantages might ground some pragmatic preference for loyalty when it comes to identifying yourself with a particular precise credal state.

In addition, from a less pragmatic perspective, we may sometimes find fault with someone who rejects good pairs of bets, even when we think that rejecting both bets is rationally permissible for her. For instance, some virtues may be best expressed by

following rules are *a fortiori* puzzles about following decision rules. For present purposes, I rely on our intuitive grasp of the latter and set aside general worries about rule-following.

someone who sticks with some choices for a while. Just as certain moral virtues may be best expressed in the context of lasting relationships with people, certain epistemic virtues may be best expressed in the context of lasting identification with precise credal states. In short: ways of assessing evidence are like your friends. Just as you are not required to have any particular person as your friend, you are not required to identify with any particular precise credal state. But if you dedicate yourself in either sense, you will likely manifest honorable qualities. It requires special courage to maintain a lasting relationship with a person or a theory in the face of challenges. And loyalty is a virtue distinctively expressed in lasting relationships. Just as you are not obviously required to have lasting friendships, you are not obviously required to identify with the same precise credal state for a long time. But claims of permissibility and obligation do not seem to exhaust our intuitive judgments here. Even if an agent who often changes her mind is blameless, there may still be respects in which she fails to flourish.[11]

Could my theory also be used to explain why you should feel regret after facing credal dilemmas? For instance, one might claim that you should feel regret just in case you failed to maximize expected utility according to some member of your imprecise credal state and one of your incommensurable values. In response: I do not intend my theory as a contribution to the controversial literature on moral residue following from Williams 1965. I motivate my theory with intuitive judgments about what actions are permissible in various cases, and I thereby avoid relying on any controversial claims about whether we can trust our capacity to discriminate regret from remorse.[12]

In some respects, my discussion of diachronic cases resembles an argument developed in Christensen 1991. Christensen evaluates a short argument for the principle of Reflection from van Fraassen 1984. The argument goes: (1) If you are willing to accept a Dutch Strategy, you are irrational. (2) If you violate Reflection, you will be willing to accept a Dutch Strategy. Therefore, (3) if you violate Reflection, you are irrational.[13] Christensen objects that the argument begs the question against opponents of Reflection: "without an independent argument for the diachronic consistency of beliefs, the Dutch Strategy does not give us any reason at all to respect the princi-

---

11. At most, some transcendental argument might demonstrate that rationality requires some minimal stability in the credences and values you identify with, if having that much stability turned out to be a necessary condition of our being able to identify with any credences or values at all.

12. In this respect, my defense of imprecise credences is closest in spirit to defenses in Schoenfield 2012 and Williams 2012. Schoenfield argues that imprecise credences are needed to account for the rational permissibility of insensitivity to mild evidential sweetening, and Williams argues that imprecise credences best account for our judgments in forced-march Sorites cases, and in cases of indeterminate survival such as the Cabinet case from van Inwagen 1990.

13. For a careful definition of 'Dutch Strategy' and further background details, see Teller 1976, van Fraassen 1984, Christensen 1991, and Hájek 2008.

ple of Reflection" (242–3). In effect, some notable alleged counterexamples to (3) are equally plausibly counterexamples to (1). For instance, suppose I falsely believe that I will be given a drug that will alter my credence in some proposition. Then I will violate Reflection, and I will also be willing to accept a Dutch Strategy. Hence insofar as I am rational, I constitute a counterexample to (1) just as well as to (3).

The discussion of diachronic cases in this paper challenges a similar argument for the following claim from ELGA 2010:

(SHARP)   Perfect rationality requires one to have sharp degrees of belief.

The argument for (SHARP) goes: (1′) If you are willing to reject a good pair of bets, you are irrational. (2′) If you have imprecise credences, you will be willing to reject a good pair of bets. Therefore, (3′) if you have imprecise credences, you are irrational. I have argued that some counterexamples to (3′) are equally plausibly counterexamples to (1′), namely all those cases in which agents change their behavior after rationally changing their minds.

Christensen ends on a pessimistic note about diachronic norms of rationality: "since diachronic inconsistency is not in itself irrational or undesirable…it is hard to see any reason for thinking that a belief state which *guaranteed* diachronic inconsistency was something we should be concerned to avoid" (244). BRIGGS 2009 criticizes Christensen for claiming that we should reject diachronic norms. She points out that we need diachronic norms to distinguish rational changes of opinion from senseless or insupportable ones, and she claims that standard decision theory is just right for achieving that goal (67). I agree with Briggs that we should demand such explanations. The moral of my discussion of credal dilemmas is that standard decision theory does not obviously provide them. In particular, it is not obvious how standard decision theory can allow that agents facing credal dilemmas may rationally change their minds.

6 The upshot for imprecise credences

I promised support for imprecise credence fans. But what I have said so far does not constitute an argument for imprecise credences. I have argued that Strategy B decision theories are preferable to Strategy A theories. I pointed out that standard decision theory is a Strategy A theory. And I used imprecise credences to give a Strategy B theory. It is compatible with everything that I have said that there are other Strategy B alternatives to standard decision theory out there. It is even compatible what what I have said that you can give a Strategy B theory without supposing that agents have

imprecise credences, in which case arguments against Strategy A theories do not automatically constitute support for imprecise credence fans. In this section, I round out the dialectic by arguing for two claims. First: in a certain token respect, you can give a Strategy B theory while supposing that agents have precise credences. Second: in essential respects, that Strategy B theory already gives up the game to imprecise credence fans. The most plausible responses to my discussion of credal dilemmas still concede that they help rather than hurt the essential aims of those who posit imprecise credences in the first place.

It is helpful to start by further exploring the example from ELGA 2010 outlined in §2 above. Elga says that he will offer you a good pair of bets, each of which is sanctioned by some but not all members of your representor, according to the imprecise credence fan. The circumstances are as follows:

> First I'm going to offer you Bet *A* [on *H*]. Immediately after you decide whether to accept Bet *A*, I'm going to offer you Bet *B*. Since the bets will be offered in quick succession, you're sure that as regards the truth of *H*, your state of mind will remain exactly the same throughout the whole bet-offering process. That means: You won't get any new evidence relevant to *H*. You won't lose any evidence regarding *H*... You won't have any religious revelations, conversion experiences, or other kind of change in your opinion regarding whether *H* is true. (4)

Elga claims that intuitively, you cannot rationally reject both bets:

(AAO) Any perfectly rational agent who is sequentially offered bets *A* and *B* in the above circumstances (full disclosure in advance about the whole setup, no change of belief in *H* during the whole process, utilities linear in dollars) will accept at least one of the bets. (4)

Elga argues that (AAO) is incompatible with viable decision theories for agents with imprecise credences. Hence, Elga concludes, rational agents must have precise credences.

Elga provides strong circumstantial evidence that you identify with the same precise credal state when you consider Bet *A* and Bet *B*. He stipulates that the second bet is offered "immediately," and that the bets are offered "in quick succession." He does not recognize any sense in which you hesitate or fret or ponder your decision about either bet. From these facts, it is natural to infer that in the case described above, you do not change your mind between bets. Even more saliently, Elga stipulates that you don't have any "change in your opinion regarding whether *H* is true" and that there is "no change of belief in *H* during the whole process." Later Elga adds your "state of opinion on the bet proposition will remain absolutely unchanged" (6). According to

my theory, the proper response to Elga's example depends on a terminological question: does changing which precise credal state you identify with count as a "change in your opinion" for Elga? Is it possible, given Elga's description of his case, that our circumstantial evidence is misleading and you actually do change your mind between his bets, or is that ruled out by facts that Elga describes?

Suppose changing your mind counts as changing your opinion, or that for some more complicated reason, the circumstances that Elga describes are simply incompatible with your changing your mind between bets. Either way, my decision theory entails that it is impermissible for you to reject both bets. Elga argues that this is the intuitively correct verdict, so there is no problem for my decision theory here. On the other hand, suppose that changing your mind does not count as changing your opinion, and the circumstances Elga describes are compatible with your changing your mind between bets. In other words: suppose that for all Elga says, it could turn out that the circumstantial evidence is misleading, and you did actually start identifying with a different representor member between rejecting Bet A and considering Bet B. Then we should reject (AAO), the intuition on which Elga bases his argument. In some cases, you are intuitively perfectly rational, and you identify with a different precise credal state between rejecting Bet A and considering Bet B. And in some of those cases, you reject both bets. Hence (AAO) is intuitively too strong, and Elga cannot fault my decision theory for yielding intuitive verdicts incompatible with (AAO).

Either way you look at it, the pivotal example in Elga 2010 does not immediately threaten my decision theory. But at this point, opponents of imprecise credences may have a more nuanced objection in mind. They may complain: "Look, when you identify with a particular member of your representor for purposes of action, you do not count as having imprecise credences. Instead you have precise credences, namely the credences you identify with. In diachronic cases where agents rationally change their minds without learning, agents are simply changing which precise credences they have." To spell out a moral analog: suppose Sartre's student elects to stay with his mother, but then has a change of heart and joins the next convoy bound for England. Opponents of incommensurable values may say that it is not that Sartre's student has incommensurable values and merely changes which values he acts on. He never had incommensurable values at all. First he valued his mother more than the army. Then he valued the army more than his mother. End of story.

In other words, opponents of imprecise credences may respond to diachronic cases by defending their own alternative to standard decision theory. They can give a Strategy B theory without using imprecise credences, as long as they say that agents in credal dilemmas can rationally change their credences without learning anything.

As long as agents change their precise credences in cases where they rationally reject both bets, we can endorse non-revisionary theories about how they should act on those credences. To sum up: opponents of imprecise credences should admit that the score is not 0–1 in favor of precise credences. But they may complain that the score is not 2–1 in favor of imprecise credences, either. For it appears that either precise or imprecise credences could be used to predict our intuitions about cases where agents must not reject both bets, and also our intuitions about cases where they may. And so the opponent of imprecise credences may claim the final score is 2–2.

However, at this point, opponents of imprecise credences have essentially given up the game. According to my theory, your action is governed by the precise credal state that you "identify with." According to the alleged alternative, your action is governed by the precise credal state that you "have." This difference in vocabulary need not signal a difference in substance. The substance of each theory is located in how it departs from standard decision theory. According to my theory, you may rationally change which representor member you identify with, even if you do not learn anything. According to the alleged alternative, you may rationally change which precise credal state you have, even if you do not learn anything. This latter claim constitutes an equally radical revision of standard Bayesian doctrine. *And in essence, it is exactly the revision that imprecise credence fans wanted.* Recall from §1 that the idea at the heart of the imprecise credence revolution was the intuition that some agents are deeply torn between multiple precise credal states, such that second-order beliefs do not adequately capture this division. An essential goal of the revolution was to do justice to the intuition that your opinions may be symmetrically related to multiple probability measures. If we say that rational agents can shift between precise credal states at will, we have achieved that essential goal. We have established that there is a set of precise credal states such that each is a deeply eligible candidate for governing your rational action. We have established a deep sense in which even in the absence of learning, you may rationally change your mind. By comparison, whether we describe all those eligible credal states as being "contained in your representor" seems like a dispute without much substance. Hence my tentative conclusion: diachronic cases are in fact evidence for the essential claims motivating imprecise credence models.

However, we should consider one more response of last resort. The opponent of imprecise credences may reject the intuitions that have guided my discussion. I have aimed to do justice to our intuition that when rational agents face credal dilemmas, they may simply change their minds without getting more evidence. The most efficient way to reject such intuitions is to say that in any such case where you change

your mind, you must have actually gotten more evidence after all.[14] If you choose to sleep on a hard decision and you wake up having second thoughts, your changing your mind can be attributed to your getting more evidence. There are no intractable credal dilemmas. First you think it is more likely that your elderly mother would most enjoy living with you. Then you get some evidence that makes you think it is more likely that she would most enjoy living with your sister. End of story.

Diehard fans of imprecise credences will likely be unmoved by this response. I am somewhat sympathetic with it. For one thing, it may prove useful to model imperfect agents as getting evidence through self-reflection. And on a personal note, sometimes I fret over credal dilemmas in the hopes that I might thereby learn that I have been incoherent or unreflective about some relevant information. On the other hand, it is exactly these hopes that tend to drive my friends up the wall, especially if they are fans of imprecise credences. More generally, we can conclude: epistemologists are confronted with incompatible families of intuitions about whether agents in credal dilemmas can change their credences only insofar as they are merely conditionalizing on further evidence they have gotten about their situation. Hence my tempered conclusion about the relative advantage of imprecise credence models: alternatives to standard models best respect our straightforward and immediate intuitions about whether agents faced with credal dilemmas can change their minds without getting more evidence.

The central contribution of this paper to the debate about imprecise credences is defensive rather than offensive. Recall from §1 that we have strong independent reasons to model agents as having imprecise credences. Failing to spell out an adequate theory of diachronic decision cases has been taken to be the Achilles' Heel of imprecise credence proposals. I have countered that imprecise credence fans fare at least as well as their opponents with such cases. This should open the field for others to move forward with imprecise credence models of our mental states.

7 A brief coda on time-slice first epistemology

For the moment, suppose we move forward with the precise credence variant of the theory defended in §5. Agents have precise credal states, and they may rationally shift between precise credal states in ways other than by conditionalizing on incoming evidence. At a first glance, this counts as a challenge to the principle of Condi-

---

14. To be exact: your precise credal state must have undergone a rationally permissible change, i.e. a change that reflects some change in your evidential situation. As we have been setting aside cases involving memory loss and *de se* updating, we may assume that such rational credal changes will involve your coming to have more evidence, as opposed to strictly less evidence.

tionalization. And so understood, it reveals the same insights as other challenges to Conditionalization that we have been setting aside for the purposes of this paper. For example, suppose you suddenly lose a lot of your memories. Afterwards you are told that several years ago, some climatologists gave a large body of evidence to you and someone else who was exactly as smart and rational as you. In response, one of you judged that it was .78 likely that the global sea level would rise enough to submerge the Maldives by 2800, and the other judged that it was .83 likely. But you have no idea which judge you were. In this case, intuitively it is not required that you have the same credences about the Maldives as before. There is no steadfast rule that binds you to your previous opinions just because they were yours. If you still have all the first-order evidence at your disposal, then intuitively you need not even aim to have the same credences as before. In fact, it would seem at best rather strange for you to go about forming opinions about climate change by trying to figure out which judge you were before. The fact that your earlier counterpart believed some proposition is defeasible evidence for that proposition, but that fact does not also have *intrinsic* import as far as your current opinions are concerned.[15]

However, neither memory loss nor credal dilemmas must necessarily be understood as challenges to Conditionalization. The most charitable interpretation of Conditionalization is compatible with both. It is naïve to understand Conditionalization as a diachronic rule that says what credences you should have at a later time, given what credences you had at an earlier time, literally speaking. Instead we should understand it as a synchronic rule that says what credences you should have in light of some evidence, given what credences you have without that evidence. It may be a fact about humans that some interval of time passes whenever we acquire evidence, but that should be an irrelevant artifact of our psychology as far as decision theory is concerned. Understood as a rule about responding to evidence, the principle of Conditionalization has no immediate consequences for cases involving memory loss or credal dilemmas. Of course, one might claim that Conditionalization was originally intended as a literally diachronic rule, and that 'Conditionalization' should therefore be reserved for a rule that binds together the credences of different temporal slices of agents—but I am inclined to interpret the Founding Fathers charitably. Hence, to sum up the first lesson of our present discussion: we should not understand the principle of Conditionalization as a literally diachronic rule.

A second lesson is that we needn't understand Conditionalization as a literally diachronic rule in order to use it to help distinguish rational changes of opinion from

---

15. To make the case more vivid, we could suppose that the other judge was your intrinsic duplicate until you each formed your credences about the Maldives. I am grateful to David Manley for helping me appreciate the force of these sorts of permissive intuitions about diachronic rules of rationality.

senseless or insupportable ones. In Moss 2012, I argue for a particular theory of how your past credences should influence your present credences. In short, your current credences are not directly constrained by your past credences, but by your current credences about what credences you used to have. In this respect, memory resembles testimony, as your credences are not directly constrained by others' credences, but by your credences about what credences others have. In conjunction with my theory of updating as communication, the synchronic variant of Conditionalization outlined above yields significant rational constraints on credences for agents in normal diachronic decision cases. In cases where you suspect you have memory loss, your credences are constrained as they would be in conversations where you suspect your interlocutor is unreliable. And finally, in cases of credal dilemmas, my theory is appropriately permissive about how your past actions rationally constrain your present ones.

Both of these lessons are part of a general movement in epistemology. The first goal of the movement is uncovering reasons to reject literally diachronic rules of rationality. The second goal is showing how to replace these rules with better ones. At a very general level of description, these better rules generally do not constrain current opinions with past opinions, but with second-order opinions about past opinions. Here is a simpler way to conceive of the movement: we replace epistemic rules governing temporally extended agents with rules that take time-slices of agents to be the fundamental subjects of epistemic evaluation. As long as the intrinsic mental states of time-slices satisfy certain conditions, no rules of rationality will have been violated. Hence if I had to summarize the movement in two words, they would be: time-slices first.

8 Appendix

According to the *midpoint rule* introduced in ELGA 2010, an agent should evaluate bets on a proposition according to the arithmetic mean of the most extreme values assigned to that proposition by members of her representor. For example: "an agent whose probability in H is the interval [10%, 80%] should evaluate bets on H in the same way as an agent whose probability in H is exactly 45%" (6). Elga claims that the midpoint rule "require[s] agents to bet just as if they had precise probabilities" (6), and this feature of the midpoint rule is supposed to undermine our motivation for saying that it constrains the actions of agents with imprecise credences.

However, the midpoint rule does not require agents to bet as if they had precise probabilities. For one thing, the midpoint rule prescribes synchronically incoherent betting behavior. For example, say that the members of your representor R assign credences to propositions p and q as follows:

$$\{\langle C(p), C(q) \rangle : C \in R\} = \{\langle .2, x \rangle : .2 \leq x \leq .6\} \bigcup \{\langle y, .2 \rangle : .2 \leq y \leq .6\}$$

Furthermore, say that for each member C of your representor, we have $C(p \wedge q) = 0$. In more colloquial terms: your mental committee unanimously holds that p and q are mutually exclusive. From these two facts, it follows that your representor members assign the range of values $[.4, .8]$ to the disjunction $p \vee q$. According to the midpoint rule, then, you ought to place bets as if you had the following precise credences:

$$C(p) = .4$$
$$C(q) = .4$$
$$C(p \wedge q) = 0$$
$$C(p \vee q) = .6$$

But these credences are incoherent, so they are not the values of any precise probability measure. Any agent who acts just as if they had these credences will be subject to a synchronic Dutch book.

The midpoint rule also prescribes diachronically incoherent betting behavior. For example, say that the members of your representor R assign credences as follows:

$$\{\langle C(p \wedge q), C(p \wedge \neg q), C(\neg p \wedge q), C(\neg p \wedge \neg q) \rangle : C \in R\}$$
$$= \{\langle .4 - x, .2 - x, .1 + x, .3 + x \rangle : 0 \leq x \leq .1\}$$

Say you update on p as traditionally recommended by imprecise credence fans, namely

by conditionalizing each representor member on this proposition. Then the members of your resulting representor $R|_p$ will assign credence to q as follows:

$$\{C|_p(q) : C|_p \in R|_p\} = \left\{ \frac{.4 - x}{.6 - 2x} : 0 \leq x \leq .1 \right\} = [.\overline{6}, .75]$$

By the midpoint rule, you should therefore act as if you had credence $C|_p(q) = .708\overline{3}$.

Meanwhile, by the midpoint rule, you should currently act as if you had the following credences:

$$C(p \wedge q) = (.4 + .3)/2 = .35$$
$$C(p) = (.6 + .4)/2 = .5$$

But note that this entails that you should currently act as if $C(p \wedge q)/C(p) = .7$ rather than as if $C(p \wedge q)/C(p) = .708\overline{3}$. Hence if you act as recommended by the midpoint rule, there are no precise credences such that you act like you have those credences and update by conditionalizing on your evidence. Assuming your updating policy governs what conditional bets you accept, TELLER 1976 demonstrates that if you act just as if you had these credences, you will be subject to a diachronic Dutch book.

To summarize: the midpoint rule does not require agents to bet just as if they had precise probabilities. Similar arguments count against a variety of decision rules that might intuitively count as "strict" rules. Hence it is difficult to define 'strict rule' by reference to paradigm examples like those given in ELGA 2010.

References

Blackburn, Simon. 1996. "Dilemmas: Dithering, Plumping, and Grief." In Mason (1996), 127–139.

Bratman, Michael. 2006. "*Thinking How to Live* and the Restriction Problem." *Philosophy and Phenomenological Research*, vol. 72 (3): 707–13.

Briggs, Rachael. 2009. "Distorted Reflection." *Philosophical Review*, vol. 118 (1): 59–85.

Brink, David. 1994. "Moral Conflict and Its Structure." *Philosophical Review*, vol. 103 (2): 215–247.

Broome, John. 2000. "Incommensurable Values." In *Well-Being and Morality: Essays in Honour of James Griffin*, Roger Crisp, editor, 21–38. Clarendon, Oxford.

—. 2001. "Are Intentions Reasons? And How Should We Cope With Incommensurable Values?" In *Practical Rationality and Preference: Essays for David Gauthier*, Christopher Morris & Arthur Ripstein, editors, 98–120. Cambridge University Press, Cambridge.

Chang, Ruth. 1997a. *Incommensurability, Incomparability, and Practical Reason*. Harvard University Press, Cambridge.

—. 1997b. "Introduction." In Chang (1997a), 1–34.

—. 2002. "The Possibility of Parity." *Ethics*, vol. 112 (4): 659–688.

—. 2005. "Parity, Interval Value, and Choice." *Ethics*, vol. 115 (2): 331–50.

Christensen, David. 1991. "Clever Bookies and Coherent Beliefs." *Philosophical Review*, vol. 100 (2): 229–47.

Conee, Earl. 1982. "Against Moral Dilemmas." *Philosophical Review*, vol. 91 (1): 87–97.

Dorr, Cian. 2010. "The Eternal Coin: A Puzzle About Self-Locating Conditional Credence." *Philosophical Perspectives*, vol. 24 (1): 189–205.

Elga, Adam. 2010. "Subjective Probabilities Should Be Sharp." *Philosophers' Imprint*, vol. 10 (5).

von Fintel, Kai. 2001. "Counterfactuals in a Dynamic Context." In *Ken Hale: A Life in Language*, Michael Kenstowicz, editor. MIT Press, Cambridge.

Foot, Philippa. 1983. "Moral Realism and Moral Dilemma." *Journal of Philosophy*, vol. 80: 379–398.

van Fraassen, Bas C. 1984. "Belief and the Will." *Journal of Philosophy*, vol. 81 (5): 235–56.

—. 1990. "Figures in a Probability Landscape." In *Truth or Consequences: Essays in Honor of Nuel Belnap*, J.M. Dunn & A. Gupta, editors, 345–356. Kluwer, Dordrecht.

Gillies, Thony. 2007. "Counterfactual Scorekeeping." *Linguistics and Philosophy*, vol. 30: 329–360.

Gowans, Christopher, editor. 1987. *Moral Dilemmas*. Oxford University Press, New York.

Hájek, Alan. 2008. "Dutch Book Arguments." In *The Oxford Handbook of Rational and Social Choice*, Paul Anand, Prasanta Pattanaik & Clemens Puppe, editors. Oxford University Press, Oxford.

Hare, Caspar. 2010. "Take the Sugar." *Analysis*, vol. 70 (1): 237–47.

Hsieh, N. 2007. "Is Incomparability a Problem for Anyone?" *Economics and Philosophy*, vol. 23: 65–80.

Jeffrey, Richard C. 1983. "Bayesianism With a Human Face." In *Testing Scientific Theories*, John Earman, editor. University of Minnesota.

Joyce, James M. 2011. "A Defense of Imprecise Credences in Inference and Decision Making." *Philosophical Perspectives*, vol. 24 (1): 281–323.

Kaplan, Mark. 1996. *Decision Theory as Philosophy*. Cambridge University Press, Cambridge.

Levi, Isaac. 1974. "On Indeterminate Probabilities." *Journal of Philosophy*, vol. 71: 391–418.

Mason, H. E., editor. 1996. *Moral Dilemmas and Moral Theory*. Oxford University Press, New York.

McConnell, Terrance. 1978. "Moral Dilemmas and Consistency in Ethics." In Gowans (1987), 154–173. Originally published in *Canadian Journal of Philosophy 8*, 269–287.

Moss, Sarah. 2011. "On the Pragmatics of Counterfactuals." Ms., Department of Philosophy, University of Michigan. Forthcoming in *Noûs*.

—. 2012. "Updating as Communication." Ms., Department of Philosophy, University of Michigan. Forthcoming in *Philosophy and Phenomenological Research*.

NAGEL, THOMAS. 1987. "The Fragmentation of Value." In GOWANS (1987), 174–87.

RAILTON, PETER. 1996. "The Diversity of Moral Dilemma." In MASON (1996), 140–166.

RAZ, JOSEPH. 1986. *The Morality of Freedom*. Oxford University Press, New York.

—. 1997. "Incommensurability and Agency." In CHANG (1997a), 110–28.

SARTRE, JEAN-PAUL. 1946. "The Humanism of Existentialism." In *Jean-Paul Sartre: Essays in Existentialism*, WADE BASKIN, editor, 31–62. Carol Publishing Group, New York.

SCHOENFIELD, MIRIAM. 2012. "Chilling Out on Epistemic Rationality: A Defense of Imprecise Credences." Ms., Dept. of Linguistics and Philosophy, MIT. Forthcoming in *Philosophical Studies*.

TELLER, PAUL. 1976. "Conditionalization, Observation, and Change of Preference." In *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, W. HARPER & C.A. HOOKER, editors, 205–253. D. Reidel, Dordrecht.

VAN INWAGEN, PETER. 1990. *Material Beings*. Cornell University Press, Ithaca.

WALLEY, PETER. 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.

WEATHERSON, BRIAN. 2008. "Decision Making with Imprecise Probabilities." Available at `http://brian.weatherson.org/vdt.pdf`.

WHITE, ROGER. 2010. "Evidential Symmetry and Mushy Credence." In *Oxford Studies in Epistemology*, TAMAR SZABÓ GENDLER & JOHN HAWTHORNE, editors, vol. 3, 161–86. Oxford University Press, Oxford.

WILLIAMS, BERNARD. 1965. "Ethical Consistency." In GOWANS (1987), 115–37. Originally published in *Proceedings of the Aristotelian Society* 39, 103–124.

WILLIAMS, J. ROBERT G. 2012. "Decision Making Under Indeterminacy." Ms., Department of Philosophy, University of Leeds.