Modalizing Mechanisms*

We have some epistemic access to the appearance of things. There is no problem in principle with the notion that the states of a bodily mechanism (our perceptual apparatus, in particular) might be attuned to (carry information about) the appearance of things.\(^1\) We also have some epistemic access to so-called modal facts -- facts whether something is possible or impossible, likely or unlikely, regardless of whether actual or not. In Peter Van Inwagen\(^2\)'s turn of phrase, we are able to modalize. This, on the other hand, is usually taken to be puzzling. In particular, it is often deemed unhelpful to model our epistemic access to the modal realm on the basis of perception, and postulate the existence of a bodily mechanism attuned to modal features of the world. Thus, Yablo's endorsement of the claim that "it is unclear how such a mechanism could work even in principle";\(^3\) and Peacocke's claim that the attuned-mechanism route amounts to the postulation of "dubiously intelligible faculties connecting the thinker with some modal realm".\(^4\)

The fact that many theorists look askance at modalizing mechanisms probably helps explain the prevalence\(^5\) of neo-rationalist approaches in contemporary epistemology of modality. In these approaches, mechanistic attunement is substituted, among other choices, by a suitably constrained faculty of conceiving,\(^6\) or by truth-tracking conditions on concept possession and understanding.\(^7\) Bueno and Shalkowski\(^8\) offer an empiricist alternative to this neo-rationalist mainstream but, in their account too, modal knowledge, and, in

---

* The author would like to thank Scott Page, and audiences in Girona, Munich and New York for helpful comments on earlier drafts. Financial support for this work was provided by the DGI, Spanish Government, research project FFI2011-26853, and Consolider-Ingenio project CSD2009-00056; and the Generalitat de Catalunya, under grant 2014-SGR-81.

\(^1\) At least once a number of comparatively exotic -- say, radically skeptic -- positions are set aside.


\(^3\) Stephen Yablo, “Is conceivability a guide to possibility?” *Philosophy and Phenomenological Research* 53, no. 1 (1993): 4. Yablo qualifies his acceptance of this claim: "[t]aken in a suitably flat-footed way, [it is] true enough" (*op. cit*). I will be defending that it is false in the suitably flat-footed sense.

\(^4\) Christopher Peacocke, *Being Known* (Oxford University Press, 1999), 163.


\(^8\) “Modalism and Theoretical Virtues.”
particular, knowledge of probabilities, depends on the presence of moderately sophisticated cognitive abilities, by means of which epistemic subjects can access the "underlying probability space" and thereby "draw appropriate inferences".  

It is unlikely that modal knowledge essentially depends on reasoning of this sort -- let alone on concept-possession conditions or faculties of conceiving of the sort appealed to by neo-rationalists -- if only because we know that mice and monkeys are able to estimate probabilities "in approximately optimal ways", and this suggests that the neural mechanisms associated with this ability are "phylogenetically ancient".

In this paper I show that there is no problem in principle with the idea of mechanisms becoming attuned to modal features of the world. I present and discuss a decision-theoretic model in which agents with severely limited cognitive abilities, at the end of an evolutionary process, have states which encode substantial information about the probabilities with which the outcomes of a certain Bernoulli process occur. Thus, in the model, a process driven by very simple, thoroughly naturalistic mechanisms eventuates in modal sensitivity. This result should help alleviate empiricist scruples about modality in general, and encourage philosophical exploration of the quasi-perceptual, mechanistic aspects of modal sensitivity.

Section I introduces the main idea, and describes the model in some detail. Sections II and III present and discuss some key results: in section II, the rationally optimal (i.e., payoff-maximizing) decision for agents in the model is discussed. It will turn out that optimal decisions depend systematically on the probabilities associated with a certain Bernoulli process: the rational agent

---

9Ibid., 682.
12I am assuming, here and throughout the paper, that (objective) probability is a modality in the relevant sense. This usage is reasonably standard (take, for example and from opposite ends of the philosophical spectrum, Peter Van Inwagen, “Why Is There Anything at All?” in Ontology, Identity, and Modality: Essays in Metaphysics (Cambridge University Press, 2001), 57–72; and Nora Berenstain and James Ladyman, “Ontic Structural Realism and Modality,” in ed. Elaine M. Landry, Structural Realism: Structure, Object, and Causality, (Springer, 2012), 149–65) but, should the reader take issue with it, hopefully the following will strike her as plausibly true: if something has nonzero probability, it is possible. Therefore, being sensitive (in the informational sense) to nonzero probabilities amounts to being sensitive to possibility.
15A Bernoulli process is, roughly, one with two possible outcomes, that occur with probability $p$ and $1-p$ respectively.
must, then, show modal sensitivity. Fully rational agents, on the other hand, are of limited interest for our current purposes. Section III shows that boundedly rational "agents" following blindly a prefixed strategy, after an evolutionary process, approximate the optimal decision discussed in section II.

As an illustration of how models of this sort can be brought to bear on contemporary discussions in the epistemology of modality, section IV briefly examines Williamson's\textsuperscript{16} naturalistic argument for counterfactualism. Section V offers some conclusions.

I. THE MODEL

I have suggested above that worries about the coherence of modalizing mechanisms might have been an important motivation in the development of neo-rationalist accounts in the epistemology of modality. A first step in the defense of such mechanisms is to note that sensitivity to the ways certain events might (but also might not) unfold can substantially impact the survival prospects of an agent, and hence be selected for. One important way in which it is good to keep an eye on alternative possible courses of events is when hedging one's bets with respect to these ways has a higher expected payoff than fully committing oneself to one of the alternatives.\textsuperscript{17}

Bet-hedging is extremely common in nature, and has been widely studied: plants,\textsuperscript{18} insects\textsuperscript{19} or birds\textsuperscript{20} hedge their bets in various ways. For a simple example of bet-hedging,\textsuperscript{21} consider trade-offs between egg size and number of eggs in a clutch. If it is certain that the weather is going to be bad in the following year, laying one or two big eggs is the right thing for birds to do: a bird hatched from a bigger egg has better possibilities of survival, and it is sensible to make sure that at least one or two manage to pull through. On the other hand, if it is certain that the weather will be benign, a larger clutch, with smaller eggs, should be risked: it can be hoped that chicks will do all right anyway, as food is abundant when the weather is good, and there will be more of them. If it

\textsuperscript{16}\textit{The Philosophy of Philosophy} (Wiley-Blackwell, 2007).
\textsuperscript{17}In fact, bet-hedging strategies with lower expected payoff can be selected for, if the payoff variance is also lower (Tom Philippi and Jon Seger, “Hedging One's Evolutionary Bets, Revisited,” \textit{Trends in Ecology \& Evolution} 4, no. 2 (1989): 41–44.) In the model to be presented in the sequel, bet-hedging has higher expected payoff than alternative courses of action.
\textsuperscript{21}Taken from ibid.
is uncertain whether weather will be good or bad, it often pays to follow the few-big-eggs strategy in some clutches and the many-small-eggs strategy in some others. Particular egg-laying profiles along these lines will be caused by, and carry information about, the fact that good- and bad-weather years occur with certain probabilities.

In the model to be presently described, the *Monster Hunt*, agents hedge their bets based on the probability associated with a certain Bernoulli process, and this increases their fitness. This is how modal sensitivity evolves.

**The Monster Hunt**

A certain kind of creature (we will call them *hunters*) makes a living out of hunting for monsters. These monsters come in two types: *air* and *sea* monsters, both of which develop, *Pokémon*-style, from the same prior, *inchoate* stage. That is: monsters start their lives in the inchoate stage and, after a short time, they change into their final air- or sea-monster form. The transition from inchoate to sea-or-air monster is indeterministic; and the probabilities of the two outcomes are $P_{\text{air}}$ and $P_{\text{sea}} = 1 - P_{\text{air}}$.

Hunters are able to see the type of the monster they are dealing with (sea, air, or inchoate); they then must do one of three acts: i) launching a *sea attack*, ii) launching an *air attack*, and iii) *preparing*, with a certain investment policy, for the eventuality of an ulterior attack. A *strategy* is a function that takes each of the three monster types to a probability distribution over the acts.

The only way to hunt an air (sea) monster is to launch an air (sea) attack, and the baseline benefit of a hunt is 10 payoff units. But the hunter can increase this payoff, if they have prepared for the right kind of attack. A preparation act takes the following form: the hunter splits 1 payoff unit into two parts -- e.g., 0.5/0.5, or 0.8/0.2. The first (second) fraction is invested in improving the prospects of an ulterior air (sea) attack, which will have a payoff that depends on the preparation investment through an inverted exponential, diminishing-returns function. In particular, the payoff of a "prepared" air attack$^{22}$ is

$$\frac{20 - 10}{e^{2 \text{inv}_{\text{air}}}} \quad \text{(Equation 1)}$$

where $\text{inv}_{\text{air}}$ is the fraction invested in the eventuality of an air attack. *Mutatis mutandis* for the payoff of a prepared sea-attack and $\text{inv}_{\text{sea}} = 1 - \text{inv}_{\text{air}}$. That is, if the hunter's investment policy is 0.5/0.5, then the payoff of both a prepared air attack and a prepared sea attack is $20 - 10e^{-1} = 16.32$. If their investment policy is 0/1 then the payoff of an air attack is 10, that of a sea attack is 18.6. Fig.

---

$^{22}$If it is launched against an air monster -- air attacks against sea monsters have payoff zero. I omit this *proviso* in what follows.
I will use "Prepare( \( inv_{air} \) )" to refer to the act of preparation in which the investment in the eventuality of an inchoate monster developing to air monster is \( inv_{air} \). The decision problem is played in one round -- the hunter sees a monster, acts, and the payoff is collected -- unless, in the first round, the hunter spots an inchoate monster and chooses to prepare. In that case, we wait until the monster has evolved, and play a second round: the hunter sees the monster type, acts, and the payoff is collected.

The full payoff matrix for the Monster Hunt is given in Tables 1 and 2: Table 1 gives the payoffs if the game is over in the first round. Table 2 gives the payoffs if there is a second round.
Table 1: First round payoffs

<table>
<thead>
<tr>
<th></th>
<th>Air Attack</th>
<th>Sea Attack</th>
<th>Prepare( ( inv_{air} ) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Monster</td>
<td>( 20 - \frac{10}{e^0} = 10 )</td>
<td>0</td>
<td>-1(^{23} )</td>
</tr>
<tr>
<td>Sea Monster</td>
<td>0</td>
<td>( 20 - \frac{10}{e^0} = 10 )</td>
<td>-1</td>
</tr>
<tr>
<td>Inchoate Monster</td>
<td>0</td>
<td>0</td>
<td>Play 2(^{nd} ) round</td>
</tr>
</tbody>
</table>

Table 2: Second round payoffs

<table>
<thead>
<tr>
<th></th>
<th>Air Attack</th>
<th>Sea Attack</th>
<th>Prepare( ( inv_{air} ) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Monster</td>
<td>( 20 - \frac{10}{e^{2inv_{air}}} - 1 )</td>
<td>-1</td>
<td>-2</td>
</tr>
<tr>
<td>Sea Monster</td>
<td>-1</td>
<td>( 20 - \frac{10}{e^{2(1-inv_{air})}} - 1 )</td>
<td>-2</td>
</tr>
</tbody>
</table>

This payoff structure is set up so as to make bet-hedging increase the fitness of hunters. If a hunter encounters an inchoate monster, and refrains from attacking straight away but prepares instead, then by the time the monster has evolved, a successful prepared attack will have a higher payoff than a successful unprepared attack would have otherwise had.

II. PAYOFF-MAXIMIZING STRATEGIES IN THE MONSTER HUNT

The Monster Hunt has 3 states \{Air Monster, Sea Monster, Inchoate Monster\}; and infinitely many acts (air and sea attacks, plus one act of preparation per each possible value of \( inv_{air} \), i.e., all reals between zero and one.) What’s the payoff-maximizing strategy for the hunter facing this decision problem? If she is seeing an air (sea) monster, she should launch an air (sea) attack -- this will get her 10 or more payoff units, and any other choice has zero payoff. If she is seeing an inchoate monster, launching any kind of attack has zero payoff. It is

\(^{23}\)The -1 and -2 sprinkled throughout the payoff tables keep track of the payoff units that have been invested in preparation.
easy to show that she should prepare, by letting \( \text{inv}_{air} \) depend on the probability of air-monster development, \( P_{air} \), in the following way:

\[
\text{inv}_{air} = \frac{1}{2} + \frac{1}{4} \ln \left( \frac{\text{prob}_{air}}{1 - \text{prob}_{air}} \right) \quad \text{(Equation 2)}
\]

In fact, the hunter cannot always invest in the way advised by Equation 2: \( \text{inv}_{air} \) cannot take values below 0 and above 1. We arrive at the payoff-maximizing \( \text{inv}_{air} \) by "clipping" \( \text{inv}_{air} \) :

\[
\text{inv}_{air} = \begin{cases} 
0 & \text{if } \text{inv}_{air} < 0 \\
\text{inv}_{air} & \text{if } 0 \leq \text{inv}_{air} \leq 1 \\
1 & \text{if } \text{inv}_{air} > 1 
\end{cases} \quad \text{(Equation 3)}
\]

A schema of the optimal decision is, thus,

\[
\begin{align*}
\text{Inchoate Monster} & \rightarrow \text{Prepare} \left( \text{inv}_{air} \right) \\
\text{Air Monster} & \rightarrow \text{Air Attack} \\
\text{Sea Monster} & \rightarrow \text{Sea Attack}
\end{align*}
\]

where \( \text{inv}_{air} \) is given by Equation 3. Fig. 2 shows the dependence of investment on probability in the optimal decision: it is obvious that \( \text{inv}_{air} \) carries a great deal of information about \( P_{air} \). In fact, for probabilities between 0.12 and 0.88, the optimal investment policy is perfectly informative about \( P_{air} \): we can recover without loss the latter from the former. The dependence is, moreover, nearly lineal, which would make the rough-and-ready policy of using the value of \( \text{inv}_{air} \) as an approximator of \( P_{air} \) a reasonably good one. For values of \( P_{air} \) between 0 and 0.12, or 0.88 and 1, \( \text{inv}_{air} \) provides no information at all about the probabilities with which inchoate monsters develop.
It is worth noting that the loss of informational connection between $inv_{air}$ and $P_{air}$ for very low and very high probabilities is not particular to the payoff function provided in Equation 1. Wherever payoff depends on investment in a way governed by exponentially diminishing returns (i.e., by a formula of the sort $payoff = a - b \cdot e^{-c \cdot inv}$), the payoff-maximizing investment will tend to minus infinity as the probability tends to zero, and to infinity as it tends to one. The necessary "clipping" to the payoff-maximizing investment will, of course, depend on the particular values for $a$, $b$, and $c$, and on the minimum and maximum possible investments in the situation at hand.

So, in creatures, if there are any, in which modal sensitivity has actually evolved through bet-hedging in the presence of diminishing returns, we should observe a systematic loss of precision in the evaluation of extreme probabilities. I do not claim that the Monster Hunt, or other similar models, are faithful descriptions of the way in which modal sensitivity actually comes about -- the model is designed to provide a reply to neo-rationalist in-principle rejections of

24The same thing happens if returns diminish potentially, rather than exponentially.
modalizing mechanisms. On the other hand, it should be said that, to the best of my knowledge, there is no empirical data against the contention that we are systematically worse at estimating very low or very high probabilities -- although there is empirical evidence\textsuperscript{25} that "clipping" should happen below 0.1 and above 0.9.

III. BOUNDEDLY RATIONAL AGENTS

The foregoing discussion of payoff maximization in the Monster Hunt shows how modal sensitivity could possibly come about: rational, payoff-maximizing agents are compelled to let their investment policy depend on the probabilities with which inchoate monsters evolve -- they will end up encoding probabilities as a side effect of their payoff-maximization goals. But this result falls short of a response to neo-rationalist worries about modalizing mechanisms, for at least two reasons:

First, throughout the previous section we have been dealing with a rational payoff maximizer. The neo-rationalist might plausibly rejoin that, sure enough, rational agents are modally sensitive, but we knew that already: the main rationalistic insight\textsuperscript{26} is, perhaps, that a non-negligible degree of rationality might be a necessary pre-condition of modal sensitivity. Indeed, calculating the optimal degree of investment involves equating the derivative of the expected payoff to zero, and solving the equation. The cognitive resources needed for this calculation are not trivial.

Second, a certain strategy might be payoff-maximizing -- and, therefore, under a natural mapping of payoff onto fitness, susceptible to being selected for in an evolutionary process -- while still be unreachable by evolution, or otherwise out of bounds for the interested agent. Compare: telekinesis, it is amply shown in the X-Men canon, would confer a huge advantage on human mutants; yet it is not meant to be.

I now report on the results of a batch of simulations in which a population of hunters with severely limited cognitive resources play the Monster Hunt. The hunters in this model are "agents" only in the least demanding sense: each of them follows a prefixed strategy (a function from states to a probability distribution over acts), blindly and without changes throughout their lives. They do not learn, and have no behavioral flexibility beyond what is afforded by their strategy. Still, by the end of a simulation run, a close approximation to the payoff-maximizing strategy typically emerges -- and, consequently, hunters in the final population snapshot have evolved to encode information, in their investment policy, about the probabilities with which inchoate monsters

\textsuperscript{25}Balci, Freestone and Gallistel, “Risk Assessment in Man and Mouse.”
\textsuperscript{26}Perhaps tacitly shared by empiricists such as Bueno and Shalkowski, “Modalism and Theoretical Virtues.”
develop. This shows, first, that modal sensitivity in no way depends on the presence of sophisticated cognitive abilities -- it is in fact a bit of a stretch to count hunters' dispositions in the model as cognitive at all -- and, second, that modal sensitivity, unlike telekinesis, is within the reach of a straightforward evolutionary process.

Simulations are set up as follows: initially, 50 hunters (each of which follows a different, random strategy) and 50 monsters are randomly placed on a grid. Hunters are paired with the closest monster, and a hunt takes place. When the hunt is over, hunters are randomly relocated. The structure of the hunter population evolves via a discrete analogue of the replicator dynamics with mutation: hunters which obtain higher payoffs reproduce more often, sometimes hatching an individual with a slightly mutated strategy.

Simulations are run for each value of $P_{\text{air}}$ from 0 to 1, at 0.005 increments, with 5 different simulations per each such value. In total 201 (values of $P_{\text{air}}$) times 5 (runs per value) simulations were run (1005 simulations). Modal sensitivity evolves systematically: Fig. 3 plots the mean fraction invested in preparation for an air attack, $\text{inv}_{\text{air}}$, in the hunter population, against the probability that an inchoate monster evolves to air monster, $P_{\text{air}}$, for each of these simulations. Pearson's $r$ between $P_{\text{air}}$ and $\text{inv}_{\text{air}}$ (which measures the linearity of this dependence) is 0.97 (with $p<0.001$). There is, then, a very strong, linear, positive correlation between these two variables -- otherwise obvious from Fig. 3.


28The NetLogo 5.1 model (Uri Wilensky, “Netlogo,” *Center for Connected Learning and Computer-Based Modeling*, 1999) used for this paper, with further implementation details and the full source code, can be downloaded from https://github.com/manolomartinez/modalizing-mechanisms.
This result vindicates the coherence of (non-rational) modalizing mechanisms: in the model, at the end of the simulated evolutionary process, there is a robust informational connection between the strategy followed by hunters and a genuinely modal feature of their world (the probability associated with inchoate-monster development). This connection builds up in an entirely non-mysterious way -- indeed, in a way that can be specified in perfectly explicit terms in 345 lines of code.

It is likely that neo-rationalist misgivings about modalizing mechanisms stem at least partially from the following combination of views: on the one hand, the view that mechanism-based knowledge is causally-based knowledge -- that, for example, the existence of perception-based knowledge depends on the fact that perceptually-knowledgeable states of affairs are able to causally affect our sensory organs. And, on the other hand, the view that merely possible states of affairs are causally inert -- what has not happened cannot cause anything. As the Monster Hunt model shows, one can accept both of these claims and still maintain that there can be mechanism-based modal knowledge: the

Figure

3: Probability of an inchoate monster evolving to air monster vs. the mean fraction that hunters invest in preparation for the eventuality of an air attack.
modalizing mechanism need not be causally affected by merely possible states of affairs; it can be causally affected, for example, by the presence of an indeterministic process, or by reliable evidence thereof. In particular, no hunter has ever perceived an inchoate monster evolve both to sea and to air monster, but this has not prevented the emergence of investment policies that carry information to the effect that both these outcomes are possible, and likely to different degrees.

IV. WILLIAMSON'S NATURALISTIC ARGUMENT FOR COUNTERFACTUALISM

Williamson has developed an influential treatment of the epistemology of modality according to which our knowledge that certain propositions are metaphysically possible depends on a more basic competence in counterfactual reasoning. At various places, Williamson argues in favor of the naturalistic credentials of this account, along the following lines:

Far from being sui generis, the capacity to handle metaphysical modality is an “accidental” byproduct of the cognitive mechanisms that provide our capacity to handle counterfactual conditionals. Since our capacity for modal thinking cannot be isolated from our capacity for ordinary thinking about the natural world, which involves counterfactual thinking, skeptics about metaphysical modality cannot excise it from our conceptual scheme without loss to ordinary thought about the natural world, for the former is implicit in the latter.

The idea is that one gets epistemic access to the more ambitious metaphysical possibilities (involving, perhaps, inverted spectrum scenarios, or Lewisian miracle worlds) for free, as a byproduct of an ecologically useful ability to evaluate everyday counterfactual situations. I will not take issue here with Williamson’s suggestion that knowledge of the metaphysical modalities depends on a prior ability to counterfactualize, but, in any event, Williamson’s naturalistic argument can be recast in more general terms: first, the evolutionary usefulness of our sensitivity to everyday modal facts explains that we are so sensitive; second, such sensitivity cannot be excised from a sensitivity to more remote modal features of the world; therefore, the evolutionary usefulness of our sensitivity to everyday modal facts explains that we are sensitive to more remote modal features of the world.

29 The Philosophy of Philosophy, chap. 5.
30 Ibid., 162.
31 But, as I pointed out in the introduction, we have good empirical evidence that knowledge of probabilities is phylogenetically very old. Thus, it deserves investigation as a likely precursor of our knowledge of possibilities -- at least a crude version of the latter could be derived from the former via a principle such as Something is possible if there is a time at which it has nonzero probability.
The models discussed in this paper vindicate the first premise in this argument: they describe a situation in which sensitivity to everyday modal facts is evolutionarily useful for a certain agent. They also provide a mechanism -- bet hedging in the presence of reliable evidence of an indeterministic process -- by which such a useful ability could be had.

On the other hand, they also show the second premise in Williamson’s argument to be false in general: sensitivity in the Monster Hunt is perfectly excised from any and all modal features that are not of immediate concern to hunters. It is, of course, possible that human agents are different from monster hunters, and that our modalizing ability has the kinds of inextricable ties to the modally remote that Williamson envisages. But, the model shows, this cannot be simply assumed. It should rather be argued for, precisely by first identifying, then probing, the mechanism by which everyday modalizing is done.

V. CONCLUSION

The Monster Hunt model provides, pace Yablo, a proof of possibility for mechanisms attuned to the modal aspects of reality, and does so in a way that is empirically well motivated. In a nutshell: agents in the model find it useful to encode modal information because this is the way to maximize expected payoff, in an bet-hedging scenario with diminishing returns on investment. Furthermore it is possible for them to encode modal information because they have epistemic access to evidence that an indeterministic process is about to occur -- in the model, evidence that an inchoate monster is about to develop to its final stage. The conditions under which modal sensitivity evolves (bet hedging, diminishing returns, etc.) are, to be sure, specific, but by no means contrived -- we have seen that there appear to be many instances of similar conditions in the natural world. On the other hand, I do not claim that this must be the way modal sensitivity has evolved; alternative routes are in all likelihood possible.

MANOLO MARTÍNEZ

Universitat Autònoma de Barcelona