Usefulness Drives Representations to Truth

A Family of Counterexamples to Hoffman's Interface Theory of Perception

Manolo Martínez

An important objection to signaling approaches to representation is that, if signaling behavior is driven by the maximization of usefulness (as is arguably the case for cognitive systems evolved under regimes of natural selection), then signals will typically carry much more information about agent-dependent usefulness than about objective features of the world. This sort of considerations are sometimes taken to provide support for an anti-realist stance on representation itself. Here I examine the game-theoretic version of this skeptical line of argument developed by Donald Hoffman and his colleagues. I show that their argument only works under an extremely impoverished picture of the informational connections that hold between agent and world. In particular, it only works for cue-driven agents, in Kim Sterelny's sense. In cases in which the agents's understanding of what is useful results from combining pieces of information that reach them in different ways, and that complement one another (i.e., that are synergistic), maximizing usefulness involves construing first a picture of agent-independent, objective matters of fact.

1 Introduction

Research into the so-called "representational status" question (Neander 2017) seeks to explain what makes something a *representation*; i.e., what endows such entities as words and sentences in a public language; alarm calls and other animal-communication devices; or perceptual states thoughts and other mental states, with semantic properties—what makes them *mean*, or *be about* something. An idea defended by many participants in this research—indeed, perhaps the closest one can find to a consensus—is that it is illuminating to think of representations as *signals* sent by a sender, able to observe a certain type of events, to a receiver, who can act in a way that has consequences for both partners. This is proposed as a way of understanding basic cases of representation, for which sender and receiver are to be identified with extremely simple mechanisms (perhaps located within one and the same organism); and it would also be true for sophisticated cases, involving signals exchanged by intentional, perhaps full-blown human, agents.

In one popular version of this *signaling framework*, senders are able to detect world states, and produce signals as a result. These signals then guide receivers in producing certain effects

(overt behavior, or inputs to other downstream systems, as the case may be). If all goes well, receiver responses together with the state the world is in conspire to make the communicative arrangement useful to sender and receiver—see Figure 1.[1] It is this usefulness that grounds the distinction between mere signals and representations, though different theories will fill in the details differently.
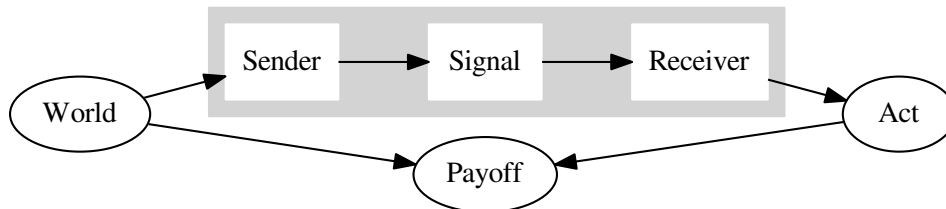

Figure 1: A sender-receiver pipeline

To name but a few of the theorists that have defended versions of this idea, in the *teleosemantic* approach to the representational-status question (Artiga 2019; Martínez 2013; Millikan 1984; Neander 2017; Papineau 1987; Shea 2018) representations mediate between a *producer* and a *consumer* (those are the names senders and receivers go by in teleosemantic quarters), and their semantic properties depend on the way they help fulfill their *teleofunctions*—i.e., whatever it is that producer and consumer are *supposed* to achieve. Such teleofunctions are, in turn, tied up with usefulness: the most common, *etiological* approach to the naturalization of teleofunction (Wright 1994; Millikan 2002) makes them depend on regimes of natural selection, learning, or design. In the partially independent work, spearheaded by Brian Skyrms and colleagues following early insights by Lewis (1969), on so-called *signaling games* in evolutionary game theory (Skyrms 2010; Zollman 2011; Huttegger et al. 2010; Wagner 2014; Martínez 2015; Martínez & Godfrey-Smith 2016), it is urged that representational content can be superseded by a more fine-grained informational content in the signals sent from a sender to a receiver (Skyrms 2010, chap. 3). Although informational content is ubiquitous, it will only play an explanatory role when it helps sender and receiver populations reach a dynamically stable configuration. Again here, the dynamics in question are driven by usefulness—by payoffs, in more common game-theoretic terms.

This dependence of representational status on usefulness is equally clear in Mitchell Green's investigation of what he calls "organic meaning" (Green 2007, 2017, 2018). Green aims at developing a minimal notion of representation which can help bridge the gap between Grice's (1957) *natural* and *non-natural* meanings. In particular, Green is concerned with circumventing two problems the Gricean program is typically thought to be saddled with: the *problem of analytic*

---

[1]Here and throughout the paper I am considering *cheap talk* models, where signals are costless.

priority and the *problem of cognitive load* (Green 2018, p. 3). Both difficulties exploit the fact that Griceans, in their attempt to derive non-natural meanings from natural ones, typically appeal to sophisticated intentions on the part of agents. Green aims at substituting the complicated intentional scaffolding that non-natural meanings typically have in the Gricean program with contents emerging from (non-intentional) sender-receiver interactions. The bearer of this simplest kind of content would be, thus, a *signal*, which Green defines as "a behavioral, physiological, or morphological characteristic *fashioned or maintained by natural selection because it serves as a cue to other organisms*" (Green 2018, p. 7, my emphasis).

The signaling framework, as the foregoing summary will have made apparent, has a very wide appeal, both in Dretskean quarters and, perhaps more surprisingly, Gricean ones—at least in Green's very interesting take on the Gricean program. There is, however, a significant problem with it: it is unclear in just which way the fitness of agents (or whatever other way *usefulness* is cashed out) is related to their ability to track truths about their environment. We expect the signaling approach to leave us with content attributions that roughly coincide with those made in the relevant sciences (say, behavioral ecology, developmental psychology or neuroscience) yet many theorists have pointed out, in different ways, that the usefulness of a signaling transaction doesn't always, perhaps ever, track the kinds of worldly properties that simple representations are routinely taken to track. The alternative, the argument goes, is that signals simply narrowly track whatever is useful for the signaling individuals, with very little regard for more objective demarcations between worldly states of affairs.[2]

Different factors complicate the relation between fitness and truth, in different ways: sometimes false beliefs turn out to be adaptive (McKay & Dennett 2009); sometimes avoiding misses (false positives) is more important than securing hits (Godfrey-Smith 1991). Another complicating factor, often discussed in relation to teleosemantics, is what Price (1998) calls the *Output Problem* (see also Martínez 2013): take the fly-catching behavior of a *Rana pipiens* frog. Here we can think of the frog's perceptual states as signals mediating between a producer, in the early visual areas, and the consumer, somewhere in the motor cortex. In the description of this system made by Lettvin et al. (1959), on which philosophers have typically relied, it is claimed that these perceptual states detect "accessible bugs" (Lettvin et al. 1959, p. 1951), yet Millikan's teleosemantics, given that "content falls out of normal conditions for *usefulness* of representations" (Millikan 1991, p. 163, original emphasis), will describe the same systems as a detector of "frog food" (*ibid.*)[3] There is a common pathology at play in all of these examples: evolution favors mechanisms that are attuned to usefulness. Truth-tracking only comes along for the ride, and will be dropped whenever necessary (Stich 1990).

Donald Hoffman and his colleagues (Hoffman 2009, 2016; Hoffman & Singh 2012; Hoffman & Prakash 2014; Hoffman, Singh & Prakash 2015; Mark 2013) have been pushing an influential game-theoretic argument that recapitulates some of the considerations offered in the previous

---

[2]A terminological point: in line with the literature I am engaging with in this paper, "truth" always denotes truth about subject-independent, worldly states of affairs. I will be taking issue with skepticism about evolved cognitive systems tracking truths in this sense. This skepticism is compatible with these systems being able to track truths about fitnesses, usefulnesses or other subject-dependent states of affairs.

[3]Furthermore, as I have argued in Martínez (2013), Millikan's teleosemantics is probably forced to describe the relevant detector in even more usefulness-related terms—perhaps as detectors of *something good for frogs*.

paragraph on behalf of the skeptic. This argument aims at showing that "veridical perceptions—strategies tuned to the true structure of the world—are routinely dominated by nonveridical strategies tuned to fitness." (Hoffman, Singh & Prakash 2015, p. 1480.) Hoffman's ideas have been around for some time, but they seem to have started gaining momentum as of late; both in academic circles (for example, the paper just cited is the target piece in a special issue that *Psychonomic Bulletin and Review* published in 2015) as well as in the popular press, where they have been taken up by outlets such as *Scientific American* ("Did Humans Evolve to See Things as They Really Are?", Michael Shermer, November 2015), or *The Atlantic* ("The Case Against Reality", Amanda Gefter, April 2016.)

Leaving aside the rather extravagant idealist views that Hoffman's argument tries to undergird, it offers an elegant and economical way of exploring the skeptical line discussed above. While, I will argue, the argument is unsuccessful, and Hoffman's "case against reality" not particularly strong, showing precisely what it gets wrong will help us get a sharper view on the kinds of circumstances in which usefulness drives representations to truth. My strategy will be to set up the same kind of models on which Hoffman and colleagues base their skeptical conclusions and show that, once we start moving away from the very extreme of cognitive and environmental simplicity, truth-tracking representation quickly becomes fitness maximizing. >>>>>>> grazer

In section 2 I reconstruct Hoffman and colleagues's main model, and the skeptical argument that depends on it, as presented in Hoffman (2009) and Hoffman, Singh & Prakash (2015). In the model, a sender observes a certain cue (say, the color of a fruit, or the amount of water in a puddle) which can take a number (or a continuous range) of values, and, following a deterministic strategy, produces one of a discrete set of signals. Receivers then use this signal to choose one among a number of alternative resources—hopefully thereby maximizing the utility[4] of the chosen resource. Hoffman and colleagues show that the sender strategy that maximizes the expected utility of the resources thus obtained is one that gives rise to signals that are highly informative about the utility of resources, but do not preserve the structure of the space of possible cues—signals that offer, that is, little to no access to the value of those agent-independent cues.

Hoffman's argument relies on the following crucial idealization: in the decision process, the agent choosing one resource among others relies on a single cue—they are wholly *cue-driven*, in Sterelny's (2003) sense. Somewhat more precisely: the agent relies on a body of sensory information that can be arbitrarily complex (and, for this reason, perhaps does not qualify as a cue as this notion is typically used in ethology) but such that there is a direct path (a Markov chain) taking from sensory information to signal to action to usefulness, without the need or the possibility to incorporate information coming from, say, other sensory modalities, or memory. In the sequel I will keep using "cue" to refer to the first link in these linear signaling pipelines—i.e., from sensory information to signal.

In any event, the situation faced by perceivers in ecologically realistic contexts is seldom this favorable: much more often, sender strategies for different sensory modalities need to evolve independently using as input cues that are only partially informative about utility. Utility-maximizing receiver strategies have to be calculated, subsequently, by putting together the pieces

---

[4] I use "utility", a more common term in discussions in economics and game theory, interchangeably with "usefulness".

of information provided by the different sensory modalities—action becoming, thus, *decoupled* from any single perceptual state.

In section 3 I describe a model (inspired by the decision problem faced by certain species of bark beetle) fully analogous to Hoffman's model, but such that agents have *two* relevant responses to signals, each one best in a different context. In this model, the sender strategy that maximizes utility is very informative about agent-independent cues. This is so, intuitively, because the sender does not have all the information needed to calculate which action will maximize utility in the current context, and is thus forced to "tell it like it is", and leave matters to the discretion of a receiver downstream. In summary, Hoffman's argument only works for extremely simple cognitive systems in informationally transparent ecological contexts. Typically, though, ecologically realistic contexts are informationally translucent. As a result, perception is typically decoupled from action, and utility-maximizing perceptual strategies typically track truth. In section 4 I offer some concluding remarks.

2 Hoffman's Model

In many sender-receiver models in the Lewis-Skyrms tradition the number of possible world states is comparable to the number of signals and of acts. This is adequate in cases in which the goal is to investigate the emergence of coordination between sender and receiver, and the flow of information derived from this. Hoffman and colleagues' aim is different: they are interested in how cues are categorized by a sender-receiver system, when the only goal of the system is to maximize utility. As a result, the space of "world states" that the sender (which in the intended interpretation is something like the visual cortex) is sensitive to is much bigger—indeed, in the model to be presently introduced, the cue can take any real value in the [0, 1] interval.[5] The signals that the sender produces, in their intended interpretation, count as the output of categorization of the cue—something like a percept. The receiver (which in the intended interpretation is something like the motor cortex in the same agent that hosts the sender) produces an action that is fully determined by this signal. Each pair of action and cue value will have an associated utility. Sender and receiver aim at maximizing the utility of the action they choose.[6]

Senders and receivers live in a world populated with *resources* they depend on, and try to make decisions about which such resource to choose, by relying on a cue that bears a (possibly non-linear) functional relation to utility. Think of bananas: the cue that senders detect is, say, their color, which we might encode as a real number from 0 (awfully green) to 1 (nice and yellow); and their utility be directly linked to their nutritional value in a way codifiable also as a real from 0 (not very nutritious at all) to 1 (very nutritious). For these idealized bananas, the function taking cue value to utility will be perhaps monotonically increasing, with higher detectable values (i.e., yellower bananas) corresponding to higher utilities (more nutritious ones).

---

[5]Of course, the results to be reported here depend on a computational implementation which calculates over double-precision floats—not at all infinite precision, but good enough.

[6]Although Hoffman's model aims at representing intra-organismic (roughly, perceptual) signaling, it applies without modification to inter-organismic signaling if we assume commonality of interests between sender and receiver.

In general, we can model any one such resource as an ordered pair, $\langle d, u \rangle$. Here, d is the value of a certain cue; in what follows, I will talk of *detectable values* to refer to values of the cue, and use D to refer to the set of possible detectable values—in this paper D will always be the set of reals between 0 and 1. The second member of the resource pair, u, is a utility; the set U of possible utilities will also be the reals between 0 and 1. There is a *resource function*, $r : D \rightarrow U$, that takes detectable values to utilities.

Senders in the model can send one from a discrete set of signals. I will conventionally refer to them as *Red*, *Green*, *Yellow* and *Blue*. Also, in general, I will use s to refer to individual signals, and S to refer to the set of all signals. A *sender strategy* is a function, $p : D \rightarrow S$, that takes detectable values to signals.[7] Hoffman and colleagues identify *realist*, truth-preserving, sender strategies with those that preserve world structure. This just means that there is a ranking among signals (say, *Blue* > *Green* > *Yellow* > *Red*) such that if a detectable value d_1 is mapped onto a signal s_1, and another detectable value d_2 is mapped onto another signal s_2, then $s_1 > s_2$ only if $d_1 > d_2$. This is a mouthful, but what it amounts to can be easily seen with an example. Figure 2 shows a realist strategy, with signals preserving the structure of detectable values in the sense explained above: e.g., any detectable value falling under *Blue* is guaranteed to be higher than any detectable value falling under *Green*. On the other hand, Figure 3 is not a realist strategy: there is no way to define a lineal ordering of signals that preserves structure, as both very low and very high detectable values are mapped onto *Red*. I will follow Hoffman and colleagues in calling such non-structure-preserving strategies *interface strategies*.[8]

In a world so characterized, we present senders with a decision problem. Suppose one of them is faced with three resources (each of them, remember, modeled as a pair of a detectable value and a utility). As above, there are four available signals, conventionally coded as colors: $S = \{Red, Yellow, Green, Blue\}$. The sender observes the three resources, and sends a signal following a fixed sender strategy. The receiver uses the signal to choose one of the resources, relying on the following *Simple* receiver strategy:

Simple: Choose the resource corresponding to the highest signal according to the perceptual ranking *Blue* > *Green* > *Yellow* > *Red*. In case of draw, choose randomly.

So, for example, if the three resources are mapped by the sender's strategy onto <*Blue*, *Red*, *Blue*>, the Simple receiver strategy mandates that they choose randomly between the first and third resource.

Finally, we need a resource function, r, connecting detectable values to utilities. Suppose, to begin with, that the resource function is the identity function: $r(d) = d$. That is, a resource with a detectable value of d yields a utility of d. It turns out that, in this case, the optimal signaling strategy is given by Figure 4.[9] This optimal signaling strategy is truth-preserving:

---

[7] I am assuming, with Hoffman, that senders are pure strategists. That is, they never follow strategies such as, e.g., "if the cue value is d then toss a fair coin and send the *Red* signal if heads and the *Yellow* signal if tails", or any other random combination of signals. Sender strategies are deterministic.

[8] Hoffman and colleagues offer a general taxonomy of perceptual strategies in (Hoffman, Singh & Prakash 2015). The foregoing explication reconstructs the subset of that taxonomy that is sufficient for our current purposes.

[9] In this figure, and all other analogous ones in the paper, the resource function is superimposed in translucent white, and the expected utility of the perceptual strategy is given in a box in the left half of the chart, at the height that
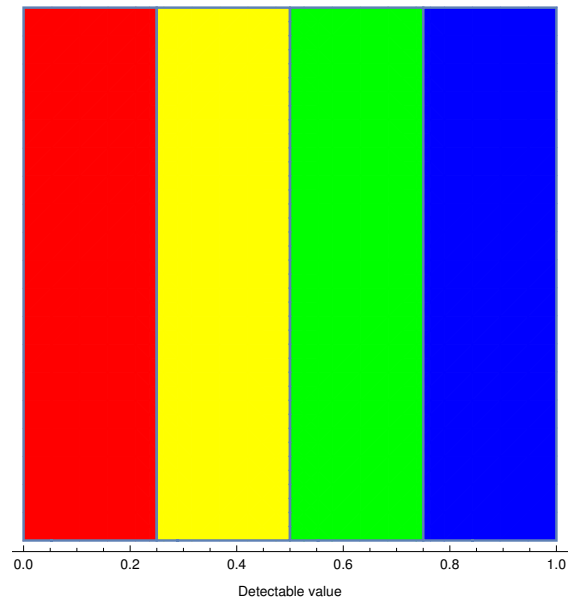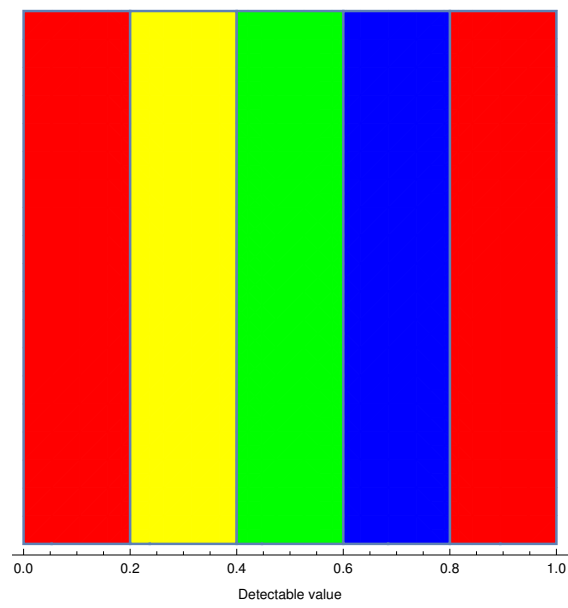
Figure 2: A realist strategy

Figure 3: An interface strategy

detectable values corresponding to a signal p_1 are guaranteed to be higher than any detectable value corresponding to a signal p_2 if $p_1 > p_2$. So far so good for the realist.
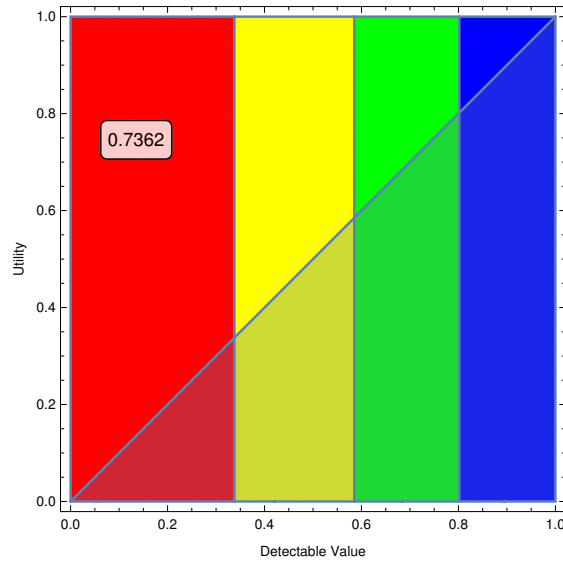

Figure 4: Optimal signaling strategy when utility is given by the identity function.

Now, the catch: the realist signaling strategy in Figure 4 is optimal only because the resource function r is monotonically increasing, but there is no reason to think that, in general, detectable values and utilities should bear this kind of relation. Consider a slightly more realistic banana: now cue values go from green to yellow to black (once it's gone bad), and the resource function is something of a Gaussian, maximally nutritious at some intermediate, yellow detectable value, and not very nutritious at the green and black extremes. As Hoffman, Singh & Prakash (2015) point out, many resources have a resource function of this sort.

Figure 5 shows the best *realist* (truth-preserving) signaling strategy when the resource function is a Gaussian probability density distribution, with mean $\mu = 0.5$ and standard deviation $\sigma = 0.15$, normalized so that all utilities lie between 0 and 1. This strategy is plainly not very good: signals are coerced to preserve the structure of detectable values (that is what truth-preserving implies), so that the most highly-ranked signal (i.e., the *Blue* range) must be used to map the maximum of the resource function, *plus everything to its right*. This means that a sender-receiver system following this strategy will foolishly choose a resource with a detectable value of 1 (and zero utility) over one with a detectable value of, say, 0.3 (and utility of 0.41.) That is, they will choose

---

corresponds to its value.

The way to read the sender strategy is as follows: the width of, e.g., the red region gives the detectable values for which *Red* will be sent. In this strategy, *Red* goes from a detectable value of 0 to one of 0.338. *Mutatis mutandis* for the rest of colors.

It should be noted that optimal strategies, here and throughout the paper, are calculated by numerical maximization of the relevant expected utility equations, and it is, thus, always possible for the numerical solver to get stuck into a local maximum. The Mathematica notebook I have used to calculate expected utilities and generate the figures in this paper can be downloaded from https://osf.io/m6h49/.
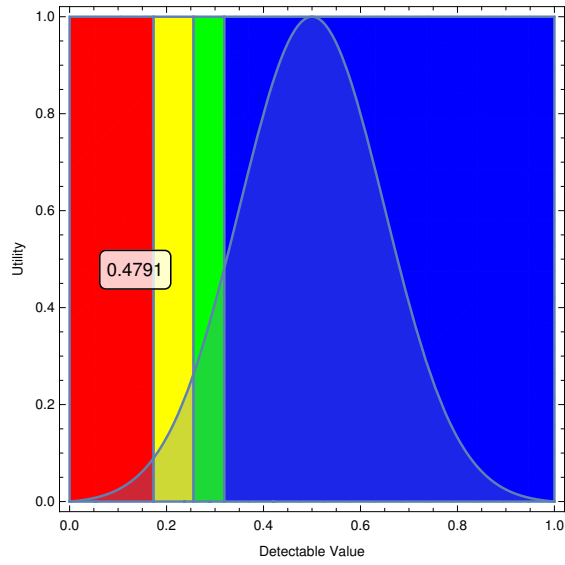
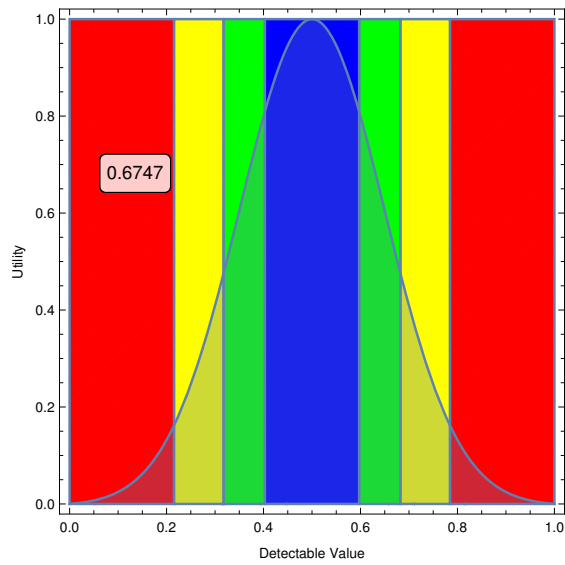Figure 5: Best realist signaling strategy for Gaussian utilities

Figure 6: Optimal signaling strategy for Gaussian utilities

9

black bananas over yellow ones simply because the former lie to the right of the latter in the designated color ordering.

Figure 6 gives the optimal sender strategy. It maps a narrow range of detectable values around the maximum of the utility function to *Blue*, and then pairs of regions symmetrically placed around the maximum for the rest of signals. This is an interface strategy in Hoffman and colleagues' sense: seeing, e. g. a *Blue* resource and a *Green* resource tells us *nothing* about which of the two underlying detectable values is higher. On the other hand, those signals tell us a great deal about the associated utilities, and, in particular, that we should pick *Blue*. We may now note that, in general, non-monotonic resource functions will often correspond to optimal perceptual strategies that are of the interface kind: it will often pay to follow the contours of local maxima and minima of utility when mapping detectable values to signals. This is, in a nutshell, the result that Hoffman and colleagues allude to when they claim that "[v]eridical perception escapes extinction only if fitness varies monotonically with truth" (Hoffman, Singh & Prakash 2015, p. 1480).[10][11]

More than its extravagantly anti-realist ramifications, the most useful theoretical product of Hoffman's work on this topic is this clear model where the truth vs fitness controversy can play out and be assessed formally. In the following section I will use a very similar model to show that, *pace* Hoffman, this formal assessment actually favors the truth camp. The skeptical results I have just reviewed depend crucially on there being a resource function univocally yielding a utility for every detectable value. This allows receivers to rely on what I have called the "Simple receiver strategy", which univocally yields an action for every signal. Agents in Hoffman's models are, thus, wholly *cue-driven*: detectable values univocally translate to a signal (through a sender strategy), and signals univocally translate into action (through the Simple receiver strategy.) On the other hand, if the relation between detectable values and fitness is one-to-many, utility-maximizing perceivers need to exercise a modicum of flexibility of response. In such cases, as we are about to see, realist sender strategies maximize expected utility.

3 Narrow-Band and Broad-Band Responses

One of the examples Hoffman uses to motivate his views on the relation of usefulness to truth is the male *Julodimorpha bakewelli*, a jewel beetle, who will happily choose a certain kind of Australian beer bottle (a "stubby") over a female to (attempt to) mate with (Gwynne & Rentz 1983).[12] This is so, presumably, because "[t]he shiny brown colour of the glass is similar to

---

[10]While in this paper I focus on the expected utility achieved by the different sender strategies, Hoffman and colleagues make their argument in terms of an evolutionary game played between interface and realist perceivers. In fact, while in general strategies with the best expected utility need not be evolutionarily stable, in the games discussed by Hoffman and colleagues the dynamics of the resulting game are only responsive to expected utilities (dominating strategies always evolving to fixation; see Mark 2013, p. 512), and the evolutionary game-theoretic spin is, as far as I can see, theoretically idle.

[11]According to O'Connor (2014, p. 849) the assumption that "there is little relation between real-world structure and payoff structure ... is unwarranted." While I think there is merit to this, here I accept, for the sake of the argument, Hoffman's claim that resource functions will typically be non-monotonic.

[12]Gwynne and Rentz were awarded the 2011 Ig Nobel prize in biology "for discovering that a certain kind of beetle mates with a certain kind of Australian beer bottle."

the shiny yellow-brown elytra of *J. bakewelli*" (Gwynne & Rentz 1983, p. 80) which in turn, according to Hoffman, suggests that "[t]he beetles' perceptions relied not on veridical information but rather on heuristics that worked in the niche where they evolved" (Hoffman, Singh & Prakash 2015, p. 1481).[13]

One thing that these jewel beetles have in common with the senders and receivers in the model of the foregoing section is that the relevant signal (for the beetle: a perception as of a shiny brown surface; for receivers in the model: one of *Blue*, *Green*, *Yellow* and *Red*) is rigidly followed by one concrete action (for the beetle: attempting to mate; for receivers in the model: the action determined by the Simple receiver strategy.) Kim Sterelny (2003, p. 34) calls this kind of response to perception *narrow-banded*, and distinguishes it from *broad-banded* response, in which agents have a "large menu of potential responses" to perceived features (*ibid.*).

Leaving human perceivers (who clearly have broad-banded-response capabilities in Sterelny's sense) aside, response strategies that are more flexible than the very narrow-banded receiver strategy in Hoffman's model are widespread in nature, and in all likelihood phylogenetically ancient. For example, for many bark beetles, the decision to use a certain tree as host for breeding and feeding depends not just on the tree giving out the right semiochemical cue (the right odor), but also on the color of its bark (Campbell & Borden 2006, 2009), suggesting that these beetles follow a receiver strategy of the following sort: "if the relevant semiochemical is present, then use tree as host if dark-barked, otherwise pass". Hawkmoths seem to follow an analogous foraging strategy: they will only pollinate a flower that gives out the right odor if it also presents the right visual features (Goyret, Markwell & Raguso 2007; Raguso & Willis 2005).[14]

Bark beetles and hawkmoths have developed such minimally broad-banded responses, presumably, because their ecological niche is not as accommodating as the model in section 2: semiochemical cues provides incomplete information about the suitability of a tree as host; bark color provides incomplete, but complementary, information about the same question, and the bark beetle has had to evolve a decision rule that incorporates these two inputs. The "band" in these responses is not much broader than the one that gave us beer-loving beetles, but, as we are about to see, it is already enough to make realist perceptual strategies utility-maximizing, in a model otherwise fully analogous to the one described by Hoffman.

One cue bark beetles are sensitive to is the concentration of a certain semiochemical, acquired through olfaction. Besides, another relevant piece of information (i.e., whether the relevant tree is of dark or light bark color) is made available to them via visual perception. Each cue maps in a one-to-many way to the fittingness of the tree as a host—to its utility as a resource. Knowing that the tree is dark-barked is not enough to conclude its fittingness. Same with knowing that it gives out the right odor. This is an example of what Sterelny (2003) calls *informationally translucent*

---

[13]Cohen (2015) points out that Hoffman's description of the jewel beetle case is tendentious, and that the correct theory of the content of the beetle's signals might conclude that the beetle is, after all, correctly representing the presence of a shiny brown surface, as opposed to incorrectly representing the presence of a female. While Cohen might well be right about this, the objection I will presently develop accepts, for the sake of the argument, Hoffman's preferred gloss on truthfulness as structure preservation.

[14]See Christensen (2010) for a compelling discussion of many more cases of broad-band responses in phylogenetically ancient, simple organisms.

environments: those in which "ecologically relevant features of [the perceivers's] environment … map in complex, one to many ways onto the cues [these perceivers] can detect" (*ibid*, p. 21). In particular, in this very simple case, detectable values (odors) map in a one-to-two way to utilities.

I will now consider a bark-beetle variant of the model presented in section 2, with a similar informational structure. The intuitive picture is one in which agents, just like hawkmoths or bark beetles, perceive (say, smell) the resources before them; and independently know (via a different sensory modality) whether the tree before them is light-barked or dark-barked. In consequence, the receiver within the agent will need to integrate the perceptual information coming from the sender with information (acquired independently) about the light-bark or dark-bark context they happen to be in.[15]

Formally, this means that there will be two resource functions taking detectable values to utilities, one for each bark color.[16] Both functions will be normalized Gaussians like the one in Hoffman's model. The first one, r_{dark}, will peak at $\mu = 0.2$ and have a standard deviation of $\sigma = 0.15$. This function will be operative in the dark-bark context. The second one, r_{light}, will have the same standard deviation, but peak at $\mu = 0.8$, in the light-bark context—see Figure 7. I will also assume that light- and dark-bark contexts are equiprobable, and that the receiver knows which context it is in at any given time (say, via a different sensory modality).
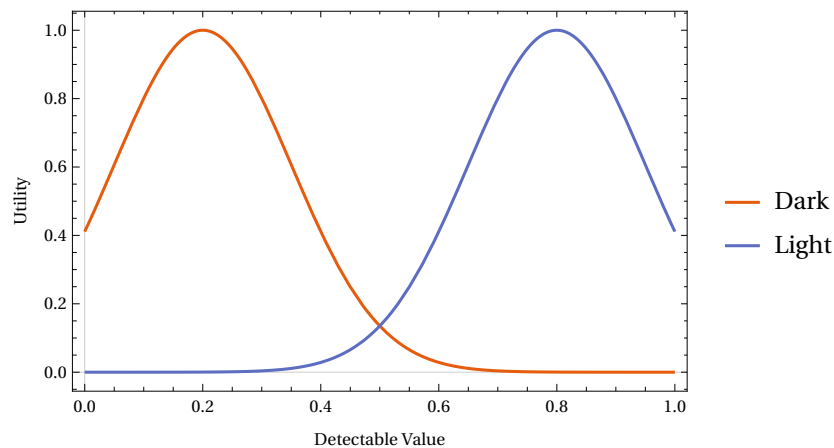

Figure 7: Two context-dependent resource functions

Strategies that try to follow the contours of the utility functions do not maximize expected utility in this model—they are too smart for their own good. First, trying to track the contours of *one* utility function at the expense of the other is obviously suboptimal: Figure 8 shows a strategy, one of two symmetric ones, in which perception has chosen to focus on the light-bark context, and neglect the dark-bark one. As a result, this perceptual system will show terrible performance half of the time. The expected utility of this strategy is 0.39.

Second, a strategy that tries to optimize both contexts at the same time, as in Figure 9, is somewhat,

---

[15]I would like to thank an anonymous reviewer for their very helpful comments on this paragraph.

[16]Or, equivalently, one resource function taking ordered pairs of a detectable value and a bark color to utilities.
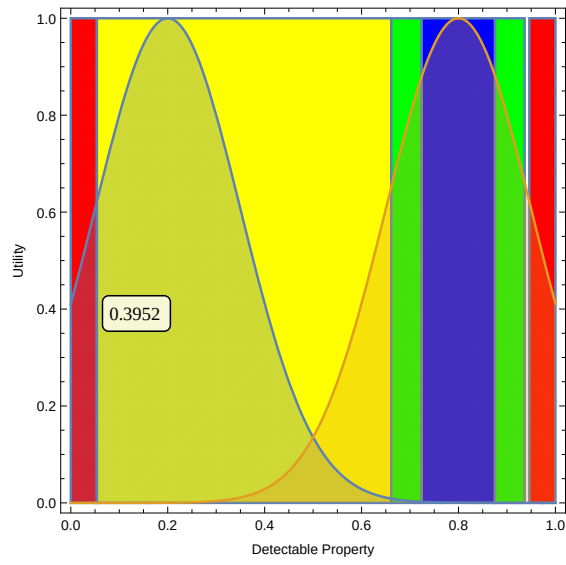
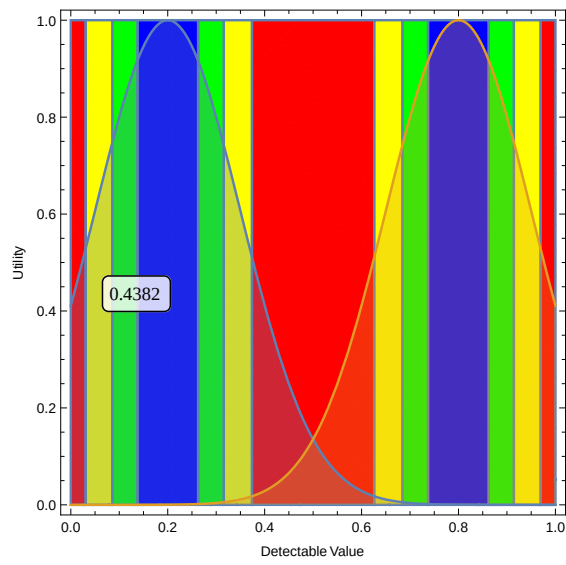Figure 8: Interface strategy following one resource function

Figure 9: Interface strategy following both resource functions at the same time

13

but not much, better. The two blue regions mean that, e.g., if the light-bark context is operative and there is a resource whose value falls into the dark-bark blue range, that resource will be chosen, even if its utility is almost zero. The expected utility for this strategy is 0.44.
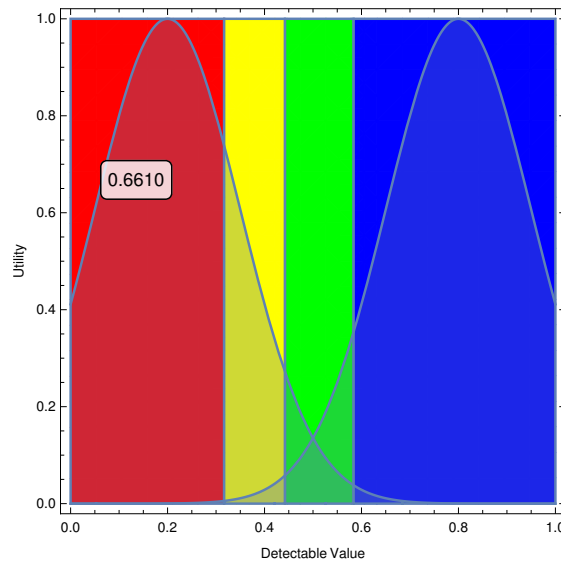

Figure 10: The optimal perceptual strategy for the translucent world

Now, the optimal perceptual strategy, in Figure 10, takes into account the fact that the receiver knows the color of the bark it is facing (light or dark). The receiver, that is, will not make a decision to act solely on the basis of the signal it gets from the sender, but rather combine it with what they know about the context they are in. Concretely, they will follow this minimally broad-band strategy, reminiscent of the decision strategy employed by bark beetles:

Broad-band: When the bark is dark, choose the resource corresponding to the *lowest* signal according to the ranking *Blue > Green > Yellow > Red*. In case of draw, choose randomly.

When the bark is light, choose the resource corresponding to the *highest* signal according to the ranking *Blue > Green > Yellow > Red*. In case of draw, choose randomly.

In a nutshell, the sender is not trying to second-guess the use to which information about detectable values will be put—it is not trying to track a resource value it only carries partial information about. It is rather "telling it like it is", so that information about detectable values can be combined with information about bark color into the Broad-band receiver strategy. This is a truth-preserving strategy, and it yields an expected utility of 0.66, 50% higher than the best interface, non-realist, strategy.

This result is not an artifact of the actual resource functions in Figure 7. Figure 11 shows what happens to expected utilities if we move the Gaussian curves (that give utility functions for the two contexts) further apart or closer together. The upshot is that as soon as the means of the two curves are separated more than one standard deviation (0.15), give or take, the truth-preserving

strategy wins. Only when the curves mostly overlap (that is, when the world is not in fact translucent) does a different, interface strategy become optimal.
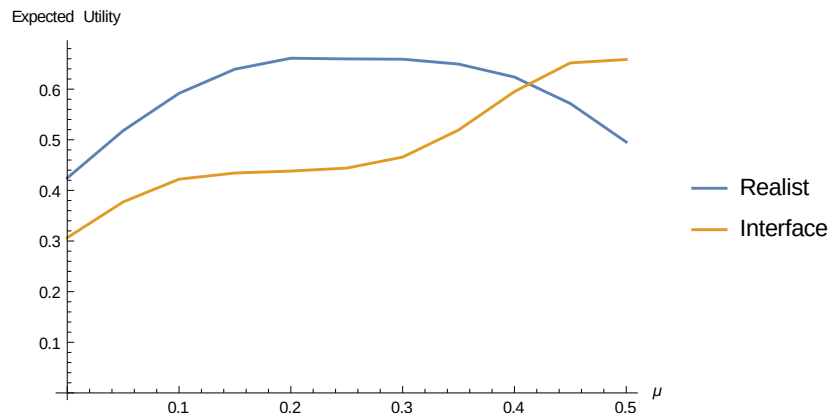

Figure 11: Expected utilities for realists and non-realists, for different arrangements of resource functions. μ marks the position of the mean for r_{dark}. The mean for r_{light} is placed symmetrically, at $1 - \mu$.

4 Conclusions

Hoffman's skeptical argument to the effect that usefulness trumps truth only works if the signals produced by the sender are the sole available cue in the maximization of usefulness. If, on the other hand, the sender-receiver system needs to build its sense of what is useful by combining information coming from different sources, things are not so clear. In particular, in the model based on the behavior of bark beetles presented in section 3, the way to maximize usefulness is by giving faithful information about detectable cues—not usefulness. This is because the relation between the two pieces of information that needed combining (the one coming from the sender; and what I called the "context" the sender-receiver system is in) are *synergic*: each on its own does not tell much about usefulness, but both combined do. *Whenever information about usefulness is presented in a synergistic fashion, truth-tracking signaling strategies will be necessary.*

The foregoing discussion holds, therefore, an important lesson for the signaling framework. Most of the work I reviewed in the introduction[17] discusses models with one source of information—one sender. I have offered evidence that suggests that the push to categorize the world on more or less objective grounds might depend, as Sterelny intimated some years ago, on the necessity to combine different pieces of complementary information. Incidentally, Sterelny appears to focus on *redundancy* as the informational property of stimuli that makes robust tracking (i.e., multi-cued detection) fitness conducive—see his (2003, sec. 2.4.) The above suggests that what information theoretists call *synergy*, (see, e.g., Williams & Beer 2010; Martínez 2017) somewhat

---

[17]Though not all, see, e.g., Skyrms (2009).

underexplored in Sterelny's writings, are probably more central than redundancy to the emergence of truth tracking from usefulness maximization.

5 Acknowledgements

References

Artiga, M 2019, 'Beyond Black Spots and Nutritious Things: A Solution to the Indeterminacy Problem'.

Campbell, SA & Borden, JH 2009, 'Additive and synergistic integration of multimodal cues of both hosts and non-hosts during host selection by woodboring insects', *Oikos*, vol. 118, no. 4, pp. 553–563.

Campbell, SA & Borden, JH 2006, 'Integration of visual and olfactory cues of hosts and non-hosts by three bark beetles (Coleoptera: Scolytidae)', *Ecological Entomology*, vol. 31, no. 5, pp. 437–449.

Christensen, W 2010, 'The decoupled representation theory of the evolution of cognitiona critical assessment', *The British Journal for the Philosophy of Science*, vol. 61, no. 2, pp. 361–405.

Cohen, J 2015, 'Perceptual representation, veridicality, and the interface theory of perception', *Psychonomic bulletin & review*, vol. 22, no. 6, pp. 1512–1518.

Godfrey-Smith, P 1991, 'Signal, Detection, Action', *Journal of Philosophy*, vol. 88, no. 12, pp. 709–722.

Goyret, J, Markwell, PM & Raguso, RA 2007, 'The effect of decoupling olfactory and visual stimuli on the foraging behavior of Manduca sexta', *Journal of Experimental Biology*, vol. 210, no. 8, pp. 1398–1405.

Green, M 2017, 'How much mentality is needed for meaning?', in K Andrews & J Beck (eds), *The Routledge handbook of philosophy of animal minds*.

Green, M 2018, 'Organic Meaning', in A Capone (ed.), *Further Advances in Pragmatics and Philosophy*, Springer.

Green, MS 2007, *Self-expression*, Oxford University Press.

Grice, P 1957, 'Meaning', *Philosophical Review*, vol. 66, pp. 377–388.

Gwynne, DT & Rentz, DCF 1983, 'Beetles on the bottle: Male buprestids mistake stubbies for females (Coleoptera)', *Austral Entomology*, vol. 22, no. 1, pp. 79–80.

Hoffman, DD 2016, 'The interface theory of perception', *Current Directions in Psychological Science*, vol. 25, no. 3, pp. 157–161.

Hoffman, DD 2009, 'The user-interface theory of perception: Natural selection drives true perception to swift extinction', *In*, Citeseer.

Hoffman, DD & Prakash, C 2014, 'Objects of consciousness', *Frontiers in Psychology*, vol. 5, p. 577.

Hoffman, DD & Singh, M 2012, 'Computational evolutionary perception', *Perception*, vol. 41, no. 9, pp. 1073–1091.

Hoffman, DD, Singh, M & Prakash, C 2015, 'The interface theory of perception', *Psychonomic bulletin & review*, vol. 22, no. 6, pp. 1480–1506.

Huttegger, SM, Skyrms, B, Smead, R & Zollman, K 2010, 'Evolutionary Dynamics of Lewis Signaling Games: Signaling Systems vs. Partial Pooling', *Synthese*, vol. 172, pp. 177–191.

Lettvin, JY, Maturana, HR, McCulloch, WS & Pitts, WH 1959, 'What the frog's eye tells the frog's brain', *Proceedings of the IRE*, vol. 47, no. 11, pp. 1940–1951.

Lewis, D 1969, *Convention: A philosophical study*, John Wiley & Sons.

Mark, JT 2013, *Evolutionary Pressures on Perception: When Does Natural Selection Favor Truth?*, University of California, Irvine.

Martínez, M 2015, 'Deception in SenderReceiver Games', *Erkenntnis*, vol. 80, no. 1, pp. 215–227.

Martínez, M 2017, 'Synergic kinds', *Synthese*, pp. 1–16.

Martínez, M 2013, 'Teleosemantics and Indeterminacy', *Dialectica*, vol. 67, no. 4, pp. 427–453.

Martínez, M & Godfrey-Smith, P 2016, 'Common Interest and Signaling Games: A Dynamic Analysis', *Philosophy of Science*, vol. 83, no. 3, pp. 371–392.

McKay, RT & Dennett, DC 2009, 'Our evolving beliefs about evolved misbelief', *Behavioral and Brain Sciences*, vol. 32, no. 6, pp. 541–561.

Millikan, R 2002, 'Biofunctions: Two Paradigms', in A Ariew, R Cummins & M Perlman (eds), *Functions: New Essays in the Philosophy of Psychology and Biology*, Oxford University Press, pp. 113–143.

Millikan, RG 1984, *Language, Thought and Other Biological Categories*, The MIT Press.

Millikan, RG 1991, 'Speaking up for Darwin'.

Neander, K 2017, *A Mark of the Mental: In Defense of Informational Teleosemantics*, MIT Press.

O'Connor, C 2014, 'Evolving perceptual categories', *Philosophy of Science*, vol. 81, no. 5, pp. 840–851.

Papineau, D 1987, *Reality and Representation*, Basil Blackwell.

Price, C 1998, 'Determinate functions', *Noûs*, vol. 32, no. 1, pp. 54–75.

Raguso, RA & Willis, MA 2005, 'Synergy between visual and olfactory cues in nectar feeding by wild hawkmoths, Manduca sexta', *Animal Behaviour*, vol. 69, no. 2, pp. 407–418.

Shea, N 2018, *Representation in Cognitive Science*, Oxford University Press, OXFORD.

Skyrms, B 2009, 'Evolution of signalling systems with multiple senders and receivers', *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 364, no. 1518, pp. 771–779.

Skyrms, B 2010, *Signals: Evolution, Learning & Information*, New York: Oxford University Press.

Sterelny, K 2003, *Thought In A Hostile World: The Evolution of Human Cognition*, John Wiley & Sons, Malden, MA.

Stich, S 1990, *The fragmentation of reason*, Cambridge, MA: The MIT Press.

Wagner, E 2014, 'Conventional Semantic Meaning in Signalling Games with Conflicting Interests', *British Journal for the Philosophy of Science*.

Williams, PL & Beer, RD 2010, 'Nonnegative decomposition of multivariate information', *arXiv preprint arXiv:1004.2515*.

Wright, L 1994, 'Functions', in E Sober (ed.), *Conceptual Issues in Evolutionary Biology*, The MIT Press. Bradford Books, pp. 27–48.

Zollman, K 2011, 'Separating Directives and Assertions', *The Journal of Philosophy*, vol. 11, pp. 158–169.