



Natural selection and veridical perceptions

Justin T. Mark*, Brian B. Marion, Donald D. Hoffman

Department of Cognitive Science, University of California, Irvine, CA 92697, USA

ARTICLE INFO

Article history:

Received 2 April 2010

Received in revised form

23 June 2010

Accepted 21 July 2010

Available online 24 July 2010

Keywords:

Evolution

Replicator dynamics

Interface games

Sensation

Game theory

ABSTRACT

Does natural selection favor veridical perceptions, those that more accurately depict the objective environment? Students of perception often claim that it does. But this claim, though influential, has not been adequately tested. Here we formalize the claim and a few alternatives. To test them, we introduce “interface games,” a class of evolutionary games in which perceptual strategies compete. We explore, in closed-form solutions and Monte Carlo simulations, some simpler games that assume frequency-dependent selection and complete mixing in infinite populations. We find that veridical perceptions can be driven to extinction by non-veridical strategies that are tuned to utility rather than objective reality. This suggests that natural selection need not favor veridical perceptions, and that the effects of selection on sensory perception deserve further study.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Students of perception often claim that perception, in general, estimates the truth. They argue that creatures whose perceptions are more true are also, thereby, more fit. Therefore, due to natural selection, the accuracy of perception grows over generations, so that today our perceptions, in most cases, approximate the truth.

They admit that there are, on occasion, illusory percepts; but these appear, most often, in psychology labs with contrived displays. And they acknowledge limits to perception; visible light, for instance, occupies a small portion of the electromagnetic spectrum, and visible objects inhabit a modest range of spatial scales. But they maintain that, for middle-sized objects to which vision is adapted, the colors, shapes and motions that we see are, most often, good estimates of the true values in the real world.

Their statements on this are clear. [Lehar \(2003\)](#), for instance, argues that “The primary function of perception [is] that of generating a fully spatial virtual-reality replica of the external world in an internal representation.” [Geisler and Diehl \(2003\)](#) say, “In general, (perceptual) estimates that are nearer the truth have greater utility than those that are wide of the mark.” [Palmer \(1999\)](#), in his textbook *Vision Science*, asserts that, “evolutionarily speaking, visual perception is useful only if it is reasonably accurate.” [Yuille and Bülthoff \(1996\)](#) concur, saying, “visual perception ...involves the evolution of an organism’s visual

system to match the structure of the world and the coding scheme provided by nature.”

If perception estimates the truth then, they argue, a natural model of perception is Bayesian estimation. In such a model for visual perception, the eye receives some images, I , and the brain then estimates the true values of scene properties, S , such as shape and color (see, e.g., [Knill and Richards, 1996](#)). To do so, it computes the conditional probability measure $P(S|I)$, and then selects a value for which this conditional probability is, say, maximized.

By Bayes’ rule, $P(S|I) = P(I|S)P(S)/P(I)$. The term $P(S)$ denotes the “prior” distribution of true values of physical properties, such as shapes and colors. The “likelihood” term $P(I|S)$ describes how the physical world maps to images at the eye. $P(I)$ normalizes all values to probabilities. These terms, as used in the computations of the brain, are shaped by selection, so that its estimates are accurate. As a result, the priors and likelihoods used by the brain accurately reflect the true priors and likelihoods in the world.

This account of perception and its evolution is, no doubt, appealing. But it depends crucially on the claim that truer perception is fitter perception. This raises two questions. Does evolutionary theory support this claim? And what, precisely, is meant by true perception?

We answer the second question, in the next section, by formalizing possible relations between perception and reality. Then, to answer the first question, we use evolutionary games to explore the relative fitness of these possible relations. We find that truth can fare poorly if information is not free; costs for time and energy required to gather information can impair the fitness of truth. What often fares better is a relation between perception and reality akin to the relation between a graphical user interface

* Corresponding author. Tel.: +1 720 234 7985.

E-mail addresses: jtmark@uci.edu (J.T. Mark), bmarion@uci.edu (B.B. Marion), ddhoff@uci.edu (D.D. Hoffman).

and the hardware of a computer (Hoffman, 1998). The icons on a desktop guide effective interaction with the computer, but the colors, shapes and locations of the icons do not, in general, resemble any properties of the software or hardware they represent. An interface promotes efficient interaction with the computer by *hiding* its structural and causal complexity, i.e., by hiding the truth. As a strategy for perception, an interface can dramatically trim the requirements for information and its concomitant costs in time and energy, thus leading to greater fitness. But the key advantage of an interface strategy is that it is not required to model aspects of objective reality; as a result it has more flexibility to model utility, and utility is all that matters in evolution.

2. Models of perception

Evolutionary games explore the fitness of competing strategies in frequency dependent selection, i.e., when fitness of a strategy can vary with the types and abundances of other strategies (Maynard Smith and Price, 1973; Maynard Smith, 1982; Weibull, 1995; Samuelson, 1997; Hofbauer and Sigmund, 1998; Fudenberg and Tirole, 1998; Nowak and Sigmund, 2004; Nowak, 2006; Szabó and Fáth, 2007; Sandholm, 2007; Antal et al., 2009). We are interested here in using evolutionary games to explore the fitness of perceptual strategies. A perceptual strategy is a relationship between perception and objective reality. What this means, precisely, is the topic of this section.

What this means, intuitively, is a helpful place to start. Since before Plato, philosophers have proposed many theories of the relationship between perception and reality. This topic, and its terminology, are controversial. But here are a few key theories (see, e.g., Heyer and Mausfeld, 2002; Radnitzky and Bartley, 1987; Schwartz, 2004).

One theory is *naive realism*. According to one version of naive realism, perception faithfully and exhaustively resembles reality. We see the truth, the whole truth and, most often, nothing but the truth. *Critical realism*, also known as *scientific realism*, weakens this claim: perception faithfully resembles a part of reality, but not all of reality. We see the truth, but not the whole truth, and sometimes something other than the truth. For instance, we see visible light but not ultraviolet or X-rays, and we can have misperceptions, such as optical illusions. Most students of perception today are critical realists.

The *interface theory* (or *desktop theory*) weakens the claim even further: perception need not, and in general does not, resemble any aspect of reality (Hoffman, 1998, 2009). Whereas naive and critical realism assert that perception is useful because, exhaustively or in part, it is true, the desktop theory asserts that perception can be useful because it is *not* true. Just as a Windows desktop hides the complexity of computer hardware, perception hides the complexity of the reality it represents, and instead provides a species-specific guide to action.

Metaphysical solipsism claims that there is no reality beyond my perceptions. Perceptions do not resemble reality, nor are they an interface to reality, because there is no reality other than my perceptions.

We now make such theories precise, beginning with a formal structure for perceptions. We assume that a collection of perceptions, say color experiences, can be represented as a set, X . The relationships among these perceptions can be represented using a subset, \mathcal{X} , of the power set, 2^X , of X together with a map $\phi: X \times \dots \times X \times \mathcal{X} \times \dots \times \mathcal{X} \rightarrow \Gamma$, where Γ is some set. For instance, if we wish to represent probabilistic relationships among color experiences, then \mathcal{X} can be a σ -algebra and $\phi: \mathcal{X} \rightarrow [0,1]$ a probability measure; or if we wish to represent similarities

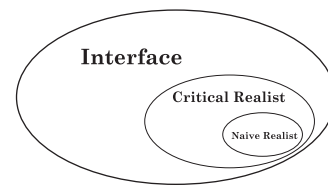


Fig. 1. Classes of perceptual strategies. Naive realist strategies are a strict subset of the critical realist strategies. Both are strict subsets of the interface strategies. Does natural selection necessarily favor these subsets? Or can it sometimes find solutions in the broader class of interface strategies?

among color experiences, then $\phi: X \times X \rightarrow [0,\infty)$ can be a distance metric. In short, we will write (X, \mathcal{X}, ϕ) for a collection of perceptions together with their relationships.

In similar fashion, we assume that the external world can be represented by a set, W , and that its relationships can be represented using a subset, \mathcal{W} , of the power set, 2^W , of W and a map $\psi: W \times \dots \times W \times \mathcal{W} \times \dots \times \mathcal{W} \rightarrow \Delta$, where Δ is some set. In short, we will write (W, \mathcal{W}, ψ) for the world and its relationships.

A *perceptual strategy* is a map $g: W \rightarrow X$. Different perceptual strategies differ in the properties of g .

For the simplest version of a naive realist strategy, $X=W$, $\mathcal{X}=\mathcal{W}$, $\phi=\psi$, and g is a bijection. Perception faithfully mirrors all of reality. Few perceptual researchers today are naive realists; nevertheless, we analyze this strategy for theoretical completeness.

We distinguish two types of critical realist strategies: *strong* and *weak*. Strategies of the strong type are a proper subset of strategies of the weak type. For the strong type, $X \subset W$, $\mathcal{X} \subset \mathcal{W}$, $\phi = \psi|_{\mathcal{X}, \mathcal{X}}$, and $g|_{\mathcal{X}}$ is a bijection. ($g|_{\mathcal{X}}$ denotes the restriction of g to \mathcal{X} .) Perception faithfully mirrors a subset of reality. For the weak type, $X \subsetneq W$ in general, and g is a homomorphism. Perception need not faithfully mirror any subset of reality, but relationships among perceptions reflect relationships among aspects of reality. Thus, weak critical realists can bias their perceptions based on utility, so long as this homomorphism is maintained. Because this is the view favored by most perceptual researchers, we will study only the weak type.

For the interface (or desktop) strategy, in general $X \subsetneq W$ and g need not be a homomorphism. Perception need not faithfully mirror any subset of reality, and the relationships among perceptions need not reflect relationships among aspects of reality. Our simulations in Section 8 indicate that the homomorphism constraint on the weak critical realist, requiring it to reflect aspects of the structure of reality, are typically less fit than interface strategies, which are not so constrained.

For metaphysical solipsism, $W = \mathcal{W} = \emptyset$ and g is undefined. We will not be concerned here with metaphysical solipsism.

Given these definitions, the naive realist strategies are a subset of the critical realist strategies, which in turn are a subset of the interface strategies, as illustrated in Fig. 1. We wish to study the fitness of these three classes of perceptual strategies. For this, we turn to evolutionary games. We first consider games in which naive realist and critical realist strategies compete, and find that critical realist strategies can drive naive realist strategies to extinction. The whole truth is not always more fit than partial truth. We then consider games in which interface strategies are added to the competition, and find that interface strategies can drive critical realist and naive realist strategies to extinction. The truth, in whole or in part, is not always more fit than perceptions that see no truth at all.

3. An elementary interface game

Consider an infinite population of agents, paired repeatedly at random to compete in a two-player game. In this game, each

agent must choose one of three territories. Each territory contains two resources (e.g., food and water) which may, or may not, be correlated. The quantity of each resource ranges from 1 to 100, and their sum is the utility of that territory. (Later we consider utility that is a nonmonotonic function of quantity.)

When an agent chooses a territory, the utility of that territory becomes the agent's fitness. The agent that chooses second must select between the remaining two territories. Each agent uses one of two perceptual strategies in its attempt to find the territory with greatest utility.

The *simple* strategy, which is a critical realist strategy, observes only one resource per territory, say food. If the quantity of that resource is above a threshold, it sees “green,” otherwise it sees “red.” If there is only one green territory, the agent chooses that territory. If there is more than one green territory, the agent chooses among them at random. If there are only red territories, it again chooses at random.

The *truth* strategy, which is a naive realist strategy, sees the exact quantity of each resource in each territory, and chooses the best available territory.

Seeing more data takes more time. So, in the simplest version of this game, *simple* chooses first when competing against *truth*. For more complex versions, priority of choice is settled probabilistically.

Similarly, seeing more data takes more energy, so *truth* requires more energy than *simple*. We subtract the cost of energy from the utility that each agent gets from its territory.

There are, of course, infinite variations on this simple game. The number of resources and their correlations can be varied. The number of territories can be increased. The *simple* agent might see three or more colors rather than just two. We explore some of these variations below, using Monte Carlo simulations. But first, to fix ideas, we start with a particularly simple variation that we can solve in closed form. We find that the costs in time and energy charged to *truth* can exceed the benefits it receives from perfect knowledge, so that *truth* ends up less fit than *simple*. We derive the payoffs to *truth* and *simple* with mathematical detail, in order to make the structure of interface games quite clear.

4. A single resource game

In this game, the external world has three territories, T_1, T_2, T_3 , and one resource, say food, that takes discrete values in the set $V = \{1, 2, \dots, m\}$. The possible values of food in the three territories form our external world space $W = V_1 \times V_2 \times V_3$, where each V_i is a copy of V . We assume that W has a measurable structure $\mathcal{W} = 2^W$, where 2^W is the power set of W ; each set $A \in \mathcal{W}$ is an event. Food is uniformly, and independently, distributed on each measurable space (V_i, \mathcal{V}_i) , $i = 1, 2, 3$, where $\mathcal{V}_i = 2^{V_i}$.

We consider two perceptual strategies: *truth* and *simple*. The perceptual space of *truth* is the measurable space $(Y, \mathcal{Y}) = (W, \mathcal{W})$. The perceptual space of *simple* is the measurable space (X, \mathcal{X}) , where $X = C_1 \times C_2 \times C_3$, for $C_i = \{0, 1\}$ and $\mathcal{X} = 2^X$. We think of 0 as the percept “red” and 1 as “green;” for mnemonic convenience, we sometimes write *R* instead of 0 and *G* instead of 1. When *simple* views a territory, it does not see the actual food value of that territory; instead it sees either red or green. It sees green if the actual food value exceeds some boundary $\beta \in V$, and red otherwise. This is, of course, a particularly simplistic form of perceptual categorization, adopted here for its mathematical simplicity; this evolutionary game can be modified to use more complex forms of perceptual categorization.

The sensory map from the world to the perceptual space of *truth* is the identity map, $g: W \rightarrow Y$. In the case of *simple*, for any boundary $\beta \in V$, we define a map $f_\beta: V^3 \rightarrow X$, induced by the map

$f_\beta^V: V \rightarrow C$ given by

$$f_\beta^V(v) = \begin{cases} 0 & \text{if } 1 \leq v \leq \beta, \\ 1 & \text{if } \beta < v \leq m, \end{cases} \tag{1}$$

for any $v \in V$. The map f_β^V assigns a color to a single territory V , and f_β uses f_β^V to assign a color to each of the three territories. This is illustrated in Fig. 2.

Observe that the perceptual space of the *simple* strategy, as defined above, can be written as

$$X = \{000, 001, \dots, 111\}, \tag{2}$$

or, in our *R* and *G* notation

$$X = \{RRR, RRG, \dots, GGG\}, \tag{3}$$

where, for instance, *RRG* means the first two territories are red and the other is green.

4.1. Simple

To use evolutionary game theory to analyze the competition between the *simple* and *truth* strategies, we must first compute the expected payoffs to each strategy when competing with itself and with the other strategy. These payoffs contribute to the evolutionary fitness of each strategy.

We now compute expected food payoffs for *simple*. If *simple* is first to choose a territory, then there are two events of interest: $G_0 = \{0 \text{ green territories}\} = \{RRR\}$, and its complement $G_{123} = \{1, 2, \text{ or } 3 \text{ green territories}\} = X - G_0$. If G_0 occurs, then *simple* must pick a red territory; otherwise *simple* picks a green territory, since it always picks green if possible. The probabilities for these events are

$$P(G_0) = \left(\frac{\beta}{m}\right)^3, \tag{4}$$

$$P(G_{123}) = 1 - P(G_0),$$

where β/m is the probability of a single red territory, and we use the fact that the territories are independently distributed.

The expected payoff of a green territory is

$$E(v|G) = \frac{m + \beta + 1}{2}, \tag{5}$$

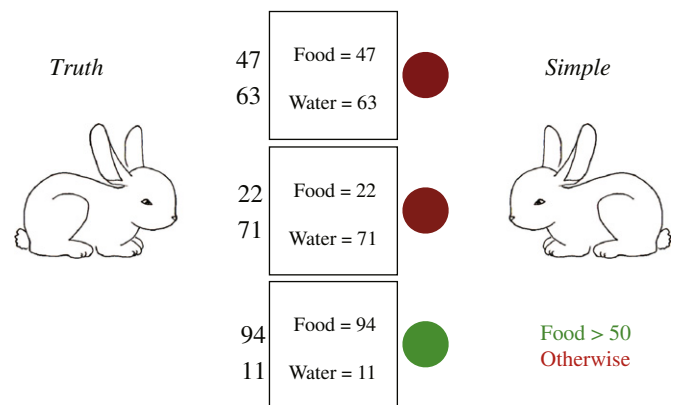


Fig. 2. Illustration of the *truth* and *simple* perceptual strategies. In the middle are three territories, each having a food value and water value between 1 and 100. The perceptions of *truth* are shown to the left of each territory. Since *truth* sees the exact values, its perceptions are identical to the food and water values of each territory. The perceptions of *simple* are shown to the right of each territory. Here *simple* sees the color green if the food value of a territory is > 50; otherwise it sees red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

because food is distributed uniformly on $[\beta + 1, m]$ and so has an expectation midway between $\beta + 1$ and m . The expected payoff of a red territory is

$$E(v|R) = \frac{\beta + 1}{2}, \tag{6}$$

because food is distributed uniformly on $[1, \beta]$ in red territories. Multiplying these payoffs by their probabilities, we find the expected payoff to *simple* when choosing first:

$$E_S^{(1)}(v|\beta, m) = P(G_0) \frac{\beta + 1}{2} + P(G_{123}) \frac{m + \beta + 1}{2}. \tag{7}$$

where the subscript S denotes the *simple* strategy and the superscript (1) indicates choosing first. Taking the derivative of (7) with respect to the boundary, β , we find that *simple* maximizes its expected payout if it has the boundary $\beta = m/\sqrt{3}$.

In this game, *simple* chooses first when competing against *truth*, so its expected payoff against *truth* is

$$E_{S|T} = E_S^{(1)}(v|\beta, m), \tag{8}$$

where $S|T$ denotes *simple* competing against *truth*.

We now compute the expected payoff for *simple* vs *simple*. One *simple* agent is randomly picked to choose first. We have just computed the expected payoffs for *simple* choosing first, so we now consider the payoffs for *simple* choosing second. In this case, the events of interest are $G_{01} = \{0 \text{ or } 1 \text{ green territories}\}$ and $G_{23} = \{2 \text{ or } 3 \text{ green territories}\}$. For the event G_{01} , *simple* must choose red (since there is at most one green territory, and the first agent took it) and for G_{23} *simple* chooses green. The probability associated with each event is then

$$P(G_{01}) = \left(\frac{\beta}{m}\right)^3 + 3\left(\frac{\beta}{m}\right)^2 \left(\frac{m-\beta}{m}\right), \tag{9a}$$

$$P(G_{23}) = 1 - P(G_{01}), \tag{9b}$$

where in (9a), the first term is the probability that all three territories are red, and the second term is the probability of exactly 1 green territory.

Using the expected payoffs for red and green territories from (5) and (6), we compute the expected payoff for *simple* choosing second to be

$$E_S^{(2)}(v|\beta, m) = P(G_{01}) \frac{\beta + 1}{2} + P(G_{23}) \frac{m + \beta + 1}{2}. \tag{10}$$

Using (7) and (10) we can now define the expected payoff to *simple* when competing against another *simple* agent to be

$$E_{S|S} = \frac{1}{2}[E_S^{(1)}(v|\beta, m) + E_S^{(2)}(v|\beta, m)], \tag{11}$$

which is simply the average of Eqs. (7) and (10).

4.2. Truth

In this simplified game, when *truth* and *simple* compete, *simple* chooses first but might not choose the best territory. The probability that it chooses the best depends on the number of green territories. For instance, if there is just one green territory, then *simple* will choose it and get the best. *Truth* then chooses the best of the remaining territories. Thus, the key events for calculating the expected payoffs for *truth* when competing against *simple* are $G_0 = \{RRR\}$, $G_1 = \{RRG, RGR, GRR\}$, $G_2 = \{GGR, GRG, RGG\}$, and $G_3 = \{GGG\}$ with probabilities

$$P(G_0) = \left(\frac{\beta}{m}\right)^3,$$

$$P(G_1) = 3\left(\frac{\beta}{m}\right)^2 \left(\frac{m-\beta}{m}\right),$$

$$P(G_2) = 3\left(\frac{\beta}{m}\right) \left(\frac{m-\beta}{m}\right)^2,$$

$$P(G_3) = \left(\frac{m-\beta}{m}\right)^3. \tag{12}$$

The payoff for each event can be found using (1) “order statistics” and (2) the probability that *simple* chooses the best territory. The r th order statistic of a statistical sample is its r th-smallest value. For instance, if one samples, with replacement, three numbers from a uniform distribution on the integers between 1 and 100, the first order statistic is the smallest of these three numbers.

In particular, for our purposes here, order statistics allow us to calculate the maximum value of a sample (see, e.g., Calik and Gungor, 2004). This value depends only on the number of red and green territories, because red and green territories have different uniform distributions. Red territories draw from the set $[1, \beta]$, whereas green territories draw from $[\beta + 1, m]$. Thus, let j be the number of red territories ($0 \leq j \leq 3$) and k the number of green territories ($0 \leq k \leq 3$), so that $j+k$ is the total number of territories (three in the example at hand). Then the cumulative distribution function (cdf) of the r th order statistic for V is either

$$F_{r;j}(v|R) = \begin{cases} \sum_{i=r}^j \binom{j}{i} \left(\frac{v}{\beta}\right)^i \left(1-\frac{v}{\beta}\right)^{j-i} & \text{if } 1 \leq v \leq \beta, \\ 1 & \text{if } \beta < v \leq m, \end{cases} \tag{13a}$$

$$F_{r;k}(v|G) = \begin{cases} 0 & \text{if } 1 \leq v \leq \beta, \\ \sum_{i=r}^k \binom{k}{i} \left(\frac{v-\beta}{m-\beta}\right)^i \left(1-\frac{v-\beta}{m-\beta}\right)^{k-i} & \text{if } \beta < v \leq m. \end{cases} \tag{13b}$$

Eq. (13a) describes the cdf of the r th order statistic for red territories. In particular, $F_{j;j}$ is the cdf of the maximum value drawn from j red territories. Eq. (13b) does the same for green territories. Eqs. (13a) and (13b) can be used to calculate the expected value of the r th order statistic for red and green territories (Calik and Gungor, 2004):

$$E_{r;j}(v|R) = \sum_{v=0}^m [1 - F_{r;n}(v|R)], \tag{14a}$$

$$E_{r;k}(v|G) = \sum_{v=0}^m [1 - F_{r;n}(v|G)]. \tag{14b}$$

$E_{j;j}(v|R)$ is the expected value of the best of j red territories; similarly $E_{k;k}(v|G)$ is the expected value of the best of k green territories. Using (14a) and (14b) we can compute the expected value of the best territory chosen by *truth* for any combination of j red and k green territories. For this computation, the most relevant events are again $G_0, G_1, G_2,$ and G_3 because they allow us to compute the probability that *simple* chooses the best territory in each event. The expected payoffs for *truth* associated with each event are

$$E(v|G_0) = \frac{2}{3} E_{3;3}(v|R) + \frac{1}{3} E_{2;3}(v|R), \tag{15a}$$

$$E(v|G_1) = E_{2;2}(v|R), \tag{15b}$$

$$E(v|G_2) = \frac{1}{2} E_{2;2}(v|G) + \frac{1}{2} E_{1;2}(v|G), \tag{15c}$$

$$E(v|G_3) = \frac{2}{3} E_{3;3}(v|G) + \frac{1}{3} E_{2;3}(v|G). \tag{15d}$$

Appendix A gives formulas for each $E_{i;j}$ in these equations. Eq. (15a) describes the event G_0 consisting of 0 green (i.e., three red territories). In this case *simple* chooses randomly from among the three territories. There is a $2/3$ chance that *simple* will not

choose the best territory, in which case *truth* will choose the best territory. This gives us the first term in (15a). There is a 1/3 chance that *simple* will choose the best territory, in which case *truth* will choose the second best territory. This gives us the second term in (15a). Eq. (15b) describes the event G_1 consisting of one green and two red territories. In this case *simple* chooses the green, i.e., the best, territory, forcing *truth* to choose the best of two red territories. Eq. (15c) describes the event G_2 consisting of two green and one red territories. *Simple* chooses randomly between the two green territories. Half of the time *simple* will choose the best territory forcing *truth* to take the second best territory. Otherwise, *simple* will choose the second best territory allowing *truth* to choose the best and giving the second term in (15c). For (15d), describing the event G_3 , there are three green territories and *simple*, once again, chooses randomly among them. Thus *truth* will take the best territory 2/3 of the time, and the second best 1/3 of the time.

Using the probabilities for the events G_0 through G_3 from (12), and their expected payoffs from (15), the expected payoff for *truth* when competing against *simple* is

$$E_{T|S}(v|\beta, m) = \sum_{i=0}^3 E(v|G_i)P(G_i). \tag{16}$$

When *truth* competes against *truth*, each chooses the best territory available. Thus, to find the expected payoff we need only compute the expected values of the best and second best territories, viz., $E_{3:3}(v)$ and $E_{2:3}(v)$. The expected payoff to *truth* when competing against *truth* is then

$$E_{T|T}(v) = \frac{1}{2}[E_{3:3}(v) + E_{2:3}(v)], \tag{17}$$

where

$$E_{3:3}(v) = \frac{3m^2 + 2m - 1}{4m}, \tag{18a}$$

$$E_{2:3}(v) = \frac{m + 1}{2} \tag{18b}$$

(see Calik and Gungor, 2004, Section 6).

5. Evolutionary dynamics

Given the expected payoffs from the previous section, evolutionary game theory can be used to study the long term interactions between *simple* and *truth*. For instance, when does one strategy drive the other to extinction, and when do the two stably coexist? To this end, we create a payoff matrix describing the competition between the strategies, as presented in Table 1.

The payoff matrix is defined as follows. *Simple* gets payoff a when competing against *simple* and payoff b when competing against *truth*. *Truth* gets payoff c when competing against *simple* and payoff d when competing against *truth*. The payoff to a strategy is taken to be the fitness of that strategy, i.e., its reproductive success. The fitness of a strategy in one generation determines the proportion of players using that strategy in the next generation. Asymptotically, i.e., after many generations, there are five possible outcomes for the two strategies, presented in Table 2, that are determined by the payoffs in Table 1.

Table 1
Payoff matrix.

	<i>Simple</i>	<i>Truth</i>
<i>Simple</i>	a	b
<i>Truth</i>	c	d

Table 2
Selection dynamics.

<i>Simple</i> wins	$a > c$ and $b > d$
<i>Truth</i> wins	$a < c$ and $b < d$
Bistable	$a > c$ and $b < d$
Stably coexist	$a < c$ and $b > d$
Neutral	$a = c$ and $b = d$

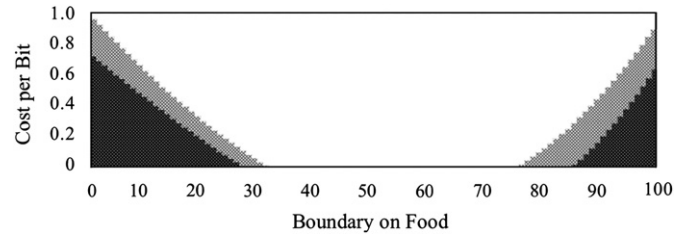


Fig. 3. Stable evolutionary outcomes of the single resource game.

A strategy “wins” if, asymptotically, it drives the other strategy to extinction regardless of the initial proportions of the strategies; a winning strategy is the best response to itself and the other strategy. Two strategies stably coexist if, independent of their initial proportions, those proportions approach asymptotically stable values; each strategy is the best reply to the other strategy, but not to itself. Two strategies are bistable if their initial proportions determine which strategy drives the other to extinction; each strategy is the best response to itself, but not to the other strategy. Two strategies are neutral if their initial proportions, whatever they happen to be, are preserved asymptotically.

The payoffs in Table 1 for *simple* and *truth* are the expected values calculated previously (Eqs. (8), (11), (16), and (17)) minus the cost of information, c_e , by the number of bits used. We think of c_e as the energy expenditure that each strategy uses to acquire and process information, and to choose a territory. This is central to the analysis of perceptual strategies, since processing more information takes, on average, more time and energy. Accounting for information costs, the entries for the payoff matrix are

$$\begin{aligned} a &= E_{S|S} - 3c_e, \\ b &= E_{S|T} - 3c_e, \\ c &= E_{T|S} - 3c_e \log_2(m), \\ d &= E_{T|T} - 3c_e \log_2(m). \end{aligned} \tag{19}$$

Simple sees 1 bit of information per territory, since it only sees red or green. Therefore, because there are three territories, it sees a total of 3 bits of information. *Truth* sees $\log_2(m)$ bits of information per territory, which for $m = 100$ is approximately 20 bits. There are also, of course, energy costs for decision, not just for perception. But for simplicity of analysis we ignore these here.

We now plot, in Fig. 3, how c_e and the red/green boundary used by *simple* affect the asymptotic performance of each strategy. White is where *simple* wins, black where *truth* wins, and gray where the two strategies stably coexist.

As seen from Fig. 3, *simple* drives *truth* to extinction for most boundaries. Even if *simple* has a poor boundary (e.g., 0–10 or 90–100), *truth* wins only for small c_e . This suggests that as the cost of information increases, simpler perceptual strategies can enhance survival.

6. Environmental complexity

The evolutionary game we have studied so far is quite simple. Now, to better understand the strengths and weaknesses of each strategy, we increase the complexity of the external world in three ways: we vary the number of territories, the number of resources per territory, and the correlations between resources. One might expect that additional complexity would favor *truth* but, as we show in this section, *simple* can still drive *truth* to extinction.

6.1. Correlated resources

We first modify our simple game by adding a single resource (e.g., water) to each territory, and then study the effects of varying the correlation between food and water. Since *truth* sees every aspect of the world, each additional resource will increase the number of bits it sees and the energy it expends. *Simple* only sees food and therefore expends no energy on perceiving water. If the correlation between food and water is high and *simple* happens to pick a territory with a large food value, then it probably also gets a high water value; *simple* effectively gains information about water without expending more energy. However, it gains little if correlations are low. In nature, of course, both high and low correlations are found between various resources.

As before, the quantities of food and water have uniform distributions on $[1, 2, \dots, m]$; we now assume they have correlation ρ . We studied the effect of correlation on the evolutionary stable strategies by running Monte Carlo simulations of the competition between *truth* and *simple*. Each input to the payoff matrix was computed as the average of 10^6 interactions, which was sufficient to provide stable values.

Fig. 4 shows that as correlation increases, *truth* does steadily worse when *simple* has a central boundary, but does better when *simple* has a boundary at either extreme. The range of boundaries for which *simple* drives *truth* to extinction also increases with increased correlation. The increased correlation effectively gives *simple* free access to more information, allowing it to compete more effectively.

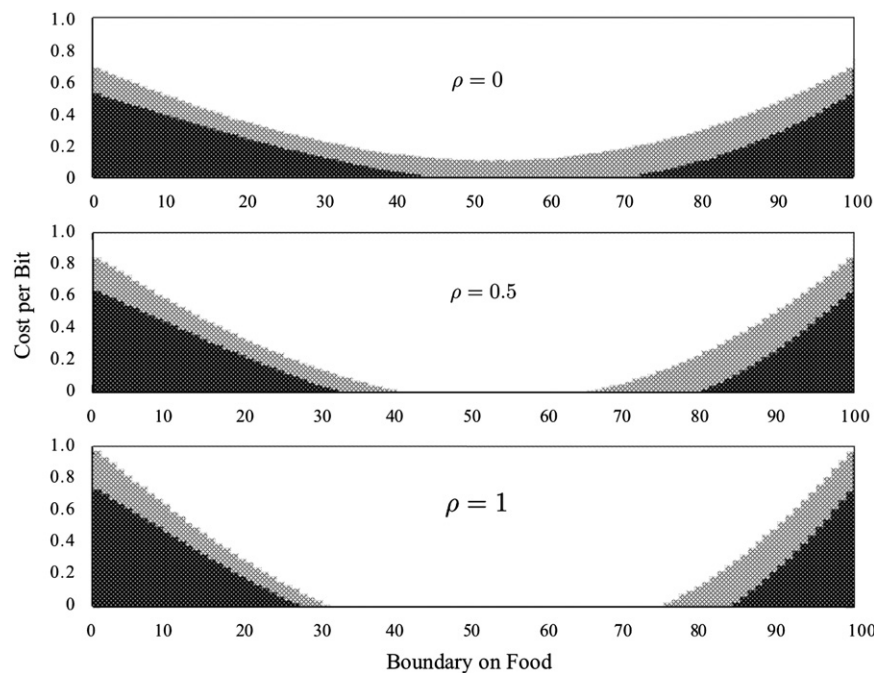


Fig. 4. Evolutionarily stable strategies with increasing correlation. When $\rho = 0$, *truth* can survive against an optimal *simple* agent so long as the cost per bit is about < 0.2 , which corresponds to an energy cost of about 6.8% of the expected payout to *truth*.

It is useful now to define the *optimal boundary* for *simple* to be the boundary that maximizes its expected payoff against *truth*, i.e., that maximizes b in the payoff matrix of Table 1. Fig. 5 shows that, as the correlation between resources increases, a *simple* agent using an optimal boundary does better against *truth*. In this figure, we quantify the performance of *truth* by the maximum cost per bit of information at which *truth* drives the optimal *simple* agent to extinction. This figure assumes seven territories and four resources per territory which, we shall see, is favorable for *truth*; the effects of adding more territories and resources are discussed in more detail in the next section.

6.2. Additional resources

We now increase the complexity of the game by adding more resources to each territory. Fig. 6 shows that adding resources improves the fitness of *truth* up to a point, but then adding more resources beyond this slowly reduces the fitness of *truth*. The fitness of *truth* increases initially because, with each additional resource, *simple* sees a smaller proportion of the information in

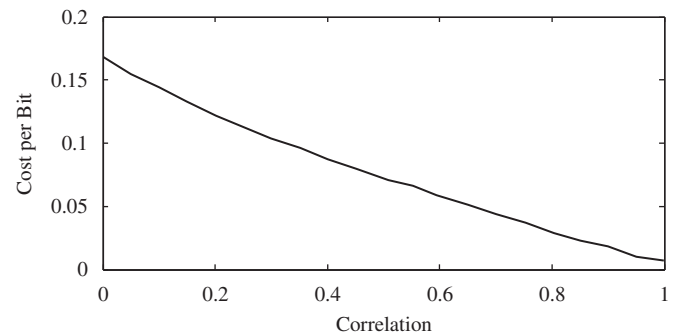


Fig. 5. Effect of correlation on the maximum cost per bit at which *truth* drives to extinction an optimal *simple* agent in an environment with seven territories and four resources. When $\rho = 0$ this maximum cost corresponds to about 11.6% of the expected payout to *truth*, and when $\rho = 1$ it corresponds to about 1%.

the environment and, as a result, is less likely to choose the best territory. This allows *truth* to pick the best territory more often.

As we continue to add resources, though, the territories become more homogenous, i.e., the difference between the best and worst territories becomes smaller. Therefore, while *truth* is more likely to choose a better territory than *simple*, the difference in their payouts becomes less significant. At the same time, *truth's* cost increases with each additional resource while *simple's* cost remains constant. This combination of higher costs and more similar territories leads to the asymptotic decrease in *truth's* fitness and, in consequence, an asymptotic increase in the fitness of *simple*.

6.3. Additional territories

We now return to one resource per territory and increase the number of territories. The additional territories provide a slightly more complex environment in which we compare the two strategies. Both *simple* and *truth* are charged as before for each additional bit of information they use to see the extra territories in the environment. But while both strategies see increases in costs, *truth's* cost increases significantly faster with the increases in information.

Fig. 7 shows that while the initial addition of territories aids *truth*, the costs of seeing more than eight territories overcomes the benefits. To understand this effect, we revisit how the expected maximum value of n territories changes with increasing n . Recall from Eq. (18a) that the expected maximum value for three territories is approximately 75. If we add an additional territory, that value jumps to about 80. As we continue to add more territories; however, the increases in the expected

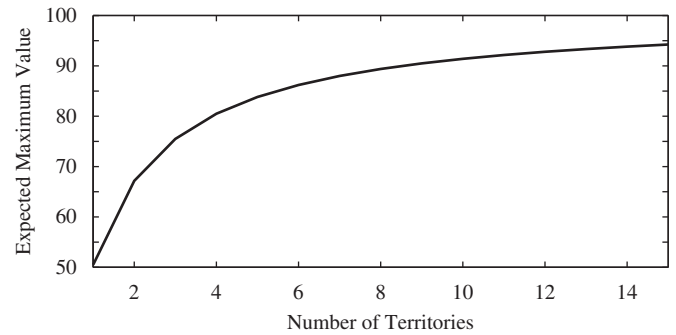


Fig. 8. Expected maximum value for n territories.

maximum territory level off as shown in Fig. 8. For *truth*, this means that for more than eight territories, the increase in cost required to see the extra territories exceeds the increase in its expected maximum payout. However, for the optimal *simple*, the increase in cost required to see each extra territory (i.e., one bit) does not exceed the increase in its expected maximum payout. This payout grows quickly for *simple* because, for any fixed boundary < 100 , as the number of territories increases the number of green territories will also increase. Therefore the optimal *simple* uses a higher boundary and receives a greater payout. As a result, we see that *simple* maintains its evolutionarily advantage as complexity increases.

7. Optimality in a complex environment

The previous section focused on the effects of increasing environmental complexity in the competition between *truth* and the optimal *simple* agent. The simulations show that, even in more complex environments, the optimal *simple* agent still drives *truth* to extinction for small cost-per-bit values. We now study how the optimal boundary for the *simple* agent changes with increasing environmental complexity. We also study strategies that use multiple, optimally placed boundaries, rather than the single boundary used by *simple*.

7.1. Optimal boundary placement

As we increase complexity by adding more territories, the boundary for the optimal *simple* agent shifts. With more territories available, the probability increases that one of them has a large food value. To increase the probability that it will choose a more valuable territory, the optimal *simple* agent must have a higher boundary, so that only territories with high food values are labeled green. In other words, it is useful for *simple* to become more discriminating in its selection of territories, as shown in Fig. 9.

7.2. Multiple category strategies

We now study how a *simple* agent performs if, instead of two categories, it has n categories; we will call this an n Cat agent. For example, a 3Cat agent might have boundaries at 70 and 40 and would label territories > 70 “green,” territories > 40 but ≤ 70 “yellow,” and territories ≤ 40 “red.” The three perceptual categories are ordered, e.g., *green* $>$ *yellow* $>$ *red*. In addition to this perceptual order, the agent has another order describing its decision strategy, such as *yellow* $>$ *green* $>$ *red*, i.e., choose yellow territories if available, then green, then red. The n Cat agent is a hybrid of *simple* and *truth* (e.g., at 100 categories the n Cat agent is identical to *truth*). This allows us to study how additional

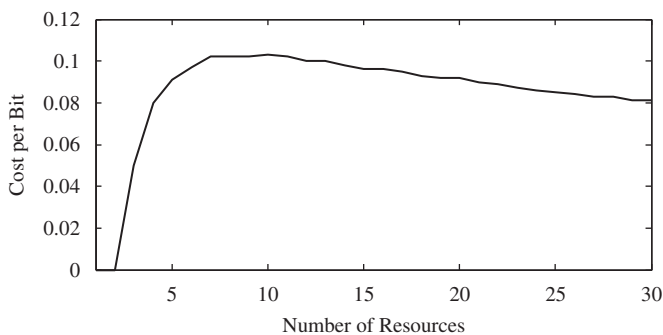


Fig. 6. Effect of the number of resources on the maximum cost per bit at which *truth* drives to extinction an optimal *simple* agent. In an environment with 10 resources per territory, this maximum cost corresponds to about 3.7% of the expected payout to *truth*, and for 30 resources it is about 3.0%.

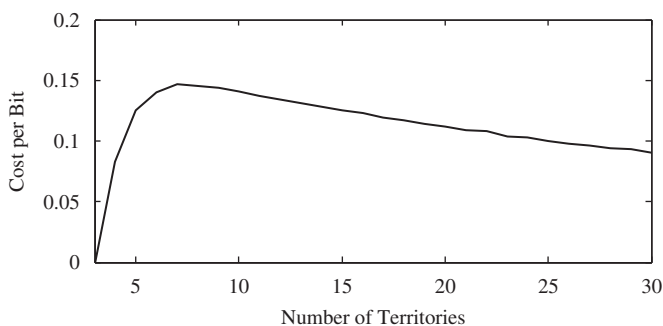


Fig. 7. Effect of the number of territories on the maximum cost per bit at which *truth* drives to extinction an optimal *simple* agent. In an environment with eight territories, this maximum cost corresponds to about 10% of the expected payout to *truth*, and for 30 territories it is about 18.8%.

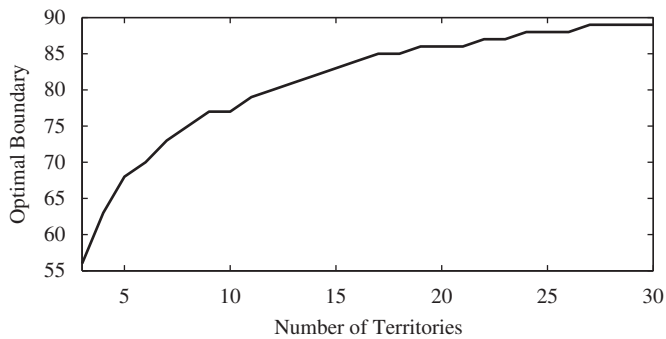


Fig. 9. Increase in the optimal boundary as environmental complexity increases.

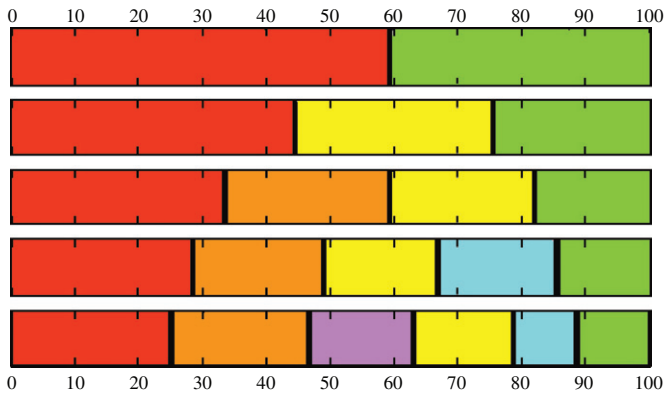


Fig. 10. Optimal boundary placement in a simple game (3-territories, 2-resources, and no correlation). Each strategy has a preference order that is identical to its perceptual order (e.g., green > yellow > red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

perceptual information about the environment affects the evolutionary fitness of a strategy.

Fig. 10 shows the optimal placement for one through five boundaries in a game with three territories and two uncorrelated resources.

As the number of categories increases, the optimal *nCat* agent receives more information about the environment. If the *nCat* agent wished to glean the most information about its environment, it would space its $n-1$ boundaries evenly because the distribution of resources in this environment is uniform. Instead, the lower part of the scale gets relatively few boundaries while the higher end becomes crowded with additional boundaries. This mirrors results from Komarova and Jameson (2008) showing that optimal boundary placement is governed to a large degree by utility. The optimal *nCat* agent, in this case, has more boundaries near the high end of the resource scale, allowing it to better distinguish among those territories with high utility. In this manner, the *nCat* agent maximizes its fitness.

7.3. Truth seeking strategies

As we saw in the previous section, the optimal *nCat* agent has categories with unevenly spaced boundaries, allowing it to be in tune with the utility of its environment. However, since the resources in the environment have uniform distributions, if the *nCat* had categories with evenly spaced boundaries, this would allow it to be better in tune with the truth of its environment (for instance, in the sense of a least squares estimate of the true value of a resource). The categories best tuned to utility need not be the categories best tuned to truth.

In this section, we study the categories best tuned to truth, i.e., we study the *nCat* agent with evenly spaced boundaries. As the number of boundaries increases, this agent sees a more refined representation of the truth of its environment, and will approximate the *truth* strategy as the number of boundaries approaches 100.

Fig. 11 shows the expected payoff for the *nCat* agent as the number of evenly spaced boundaries increase. For the “cost per bit=0” curve, there is no increase in payoff for additional boundaries beyond 10 boundaries.

Fig. 11 also shows that if the cost of information is greater than 0, then there is an optimal number of categories that maximizes the expected payoff. As the cost of information increases, this optimal number of categories decreases. Adding further categories actually decreases the expected payoff and, in consequence, will probably decrease fitness.

One might object that Fig. 11 assumes that the cost of information grows linearly with the number of bits, thus ignoring the possibility that there could be an economy of scale. To investigate this objection we ran new simulations, shown in Fig. 12, in which the cost for information only grows logarithmically with the number of bits. Again we see that there is no increase in payoff for additional categories beyond about 16 categories.

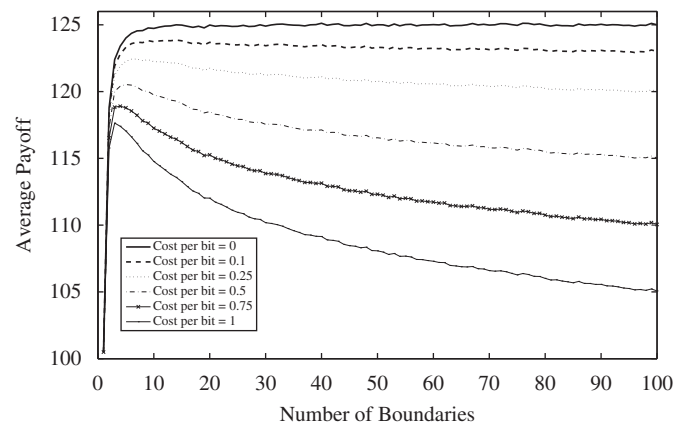


Fig. 11. Effect of additional boundaries on the expected payout of the *nCat* agent with evenly spaced boundaries. Environment consists of three territories each with two uncorrelated resources and an expected payout of approximately 125 for the best territory.

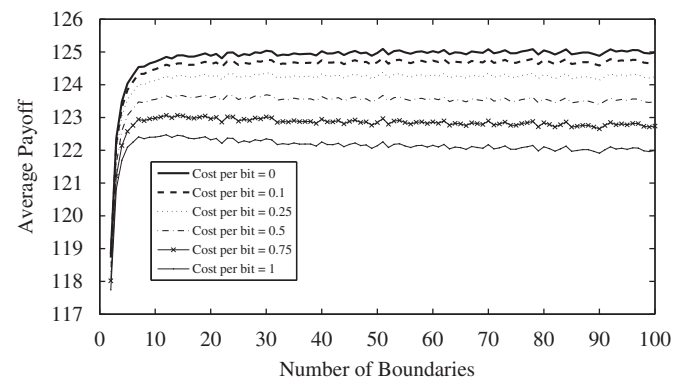


Fig. 12. Effect of additional categories on the expected payout of the *nCat* agent with evenly spaced boundaries when cost increases logarithmically with the number of bits. Environment consists of three territories each with two uncorrelated resources and an expected payout of approximately 125 for the best territory.

8. Interface strategies

One might argue that the simulations discussed so far suggest that perceptions need not resemble reality to be fit. After all, the *simple* strategy sees only red or green, and not the reality of its simulated environment, viz., food and other resources with values ranging from 1 to 100. Moreover, red and green are arbitrary representations and could just as easily be replaced by blue and yellow, smooth and rough, pen and pencil, or even nonsense syllables such as dax and vox; none of these resemble food, water, or 1–100 in any way. And yet, in many simulations, *simple* drives *truth* to extinction.

However, one could counter that *simple*'s perceptions still preserve a vestige of truth. Even though green and red do not resemble the truth, if *simple* orders its perceptual categories green > red then this reflects a true feature of the environment: every food value that leads to a green perception is greater than every food value that leads to a red perception. (It also happens, in this special case, that green reflects greater utility than red; but in general the order on perceptual categories need not be the same as the utility order.) This homomorphism between the order in the environment and the order on perceptual categories is why we, in Section 3, call *simple* a critical-realist strategy. Thus, one could argue that the simulations so far show only that some critical-realist strategies, which see limited but true information about the environment, can be fit.

This raises the question of the fitness of interface strategies: If we eliminate the last vestige of truth in the perceptions of an agent, i.e., if we eliminate any homomorphism between environment and perceptions, can a perceptual strategy still be fit? Here we show that the answer is yes. Interface strategies can outcompete critical realist strategies, and can continue to do so when naive realist strategies are also present. But first, we briefly discuss nonlinear utilities.

8.1. Gaussian utilities

Our simulations, so far, posit a specific linear relationship between utility and quantity of resources, in which, e.g., 10 units of food yield twice the payout of 5. Of course utilities in nature need not fit this model. For a thirsty mouse, a little dew might have more utility than a large lake. For a hungry shark, a small tuna might have more utility than a large whale. The utility of eating spices depends on local annual temperature (Billing and Sherman, 1998) and the utility of geophagy, the deliberate eating of dirt, depends on the phenolic content of local plants (Krishnamani and Mahaney, 2000).

The spice and geophagy examples illustrate that utilities of resources can interact with each other and with other aspects of the environment. The mouse and shark examples illustrate that, even if such interactions are ignored, utilities can vary nonlinearly with the quantity of resources. For now, we will omit interactions between utilities, and simply introduce a truncated Gaussian utility function independently on each resource in the environment. In this case, agents receive the highest payout by choosing territories with resources falling inside an optimal midrange, rather than by finding the greatest quantity of resources.

This poses a new problem, because perceptual information about resource quantities is not enough to determine the payout of each territory. In order to pick the best territory, a strategy must know the payout associated with each possible resource quantity. We must now charge each strategy for both seeing the quantity and knowing the utility of each resource, and its cost becomes

$$c_e[\text{trlog}_2(q)] + c_k[rqn_b], \quad (20)$$

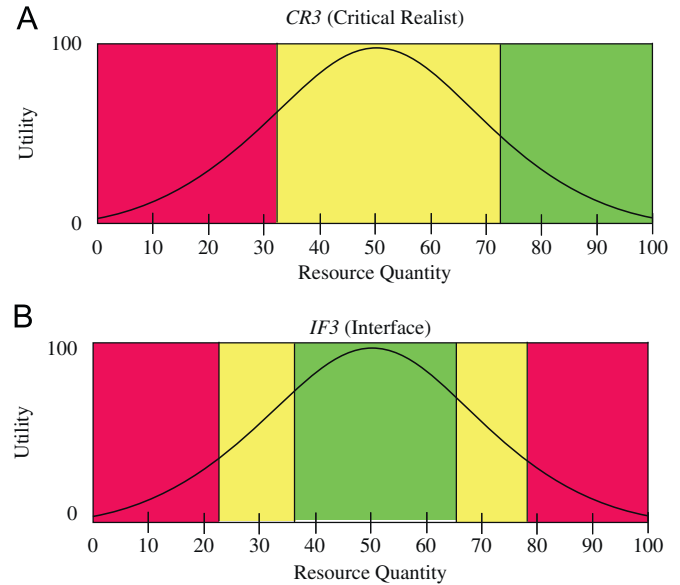


Fig. 13. (A) Optimal boundary placement on a Gaussian utility structure for a 3*Cat* critical realist. (B) A 3*Cat* interface strategy for a Gaussian utility structure. Because the perceptual map of an interface strategy is not required to be a homomorphism of the structure of reality, it can instead be tuned to the utilities that are critical to evolutionary fitness.

where c_e is again the cost per bit of information, c_k is the cost per bit of knowledge about utility values, t is the number of territories, r is number of resources, q is the number of perceptual categories for that strategy, and n_b is the number of bits used to represent the utility of a resource quantity. In the case of *truth*, we take q to be 100.

In this situation, the optimal *simple* agent has its boundary just below the peak of the Gaussian curve, so that its green category contains this maximum value.

An optimal critical realist with three perceptual categories (red, yellow, green) and perceptual order red < yellow < green is illustrated in Fig. 13(A); we will call it CR 3. Its decision rule would be to prefer yellow to green, and green to red. Notice that the order governing its decision rule is red < green < yellow which differs from its perceptual order. Even though perceptual orders and decision rules can differ, together they coevolve.

An interface strategy is illustrated in Fig. 13(B). It has four boundaries, but only three perceptual categories (red, yellow, green), so we call it IF 3. As is clear in the figure, the resource values that get mapped to the perceptual category yellow are not all contiguous, nor the resource values that get mapped to red. Thus, this mapping is not a homomorphism. However, it is better tuned to what really matters in evolution, namely utility. This kind of mapping is not possible for the critical realist because its perceptions must be homomorphic to relationships in reality. This restriction limits the flexibility of the critical realist to be tuned to utilities.

8.2. Three strategy games

We now have the framework to directly compare the three perceptual strategies described at the start of this paper. Taking *truth* to be a naive realist strategy, CR 3 to be a critical realist strategy, and IF 3 to be an interface strategy, then, by including all three strategies in one population, we can evaluate the fitness of each strategy relative to the others by simulating their evolutionary dynamics through successive generations.

In our games involving three strategies, the agents are repeatedly paired at random to compete. This gives a 3×3 payoff matrix and, unlike in two-player games, we cannot immediately calculate which strategy survives and which goes extinct. Instead we compute the time derivatives of the frequencies of each strategy using the replicator equation (Hofbauer and Sigmund, 1998; Nowak, 2006; Taylor and Jonker, 1978),

$$\dot{X}_i = X_i[f_i(\vec{X}) - \phi(\vec{X})], \tag{21}$$

where $\phi(\vec{X})$ is the average fitness and, for each strategy i , X_i denotes its frequency, \dot{X}_i its time derivative and $f_i(\vec{X})$ its fitness. Here, $f_i(\vec{X}) = \sum_{j=1}^n P_{ij}x_j$ where P_{ij} is the payoff matrix, and $\phi(\vec{X}) = \sum_i f_i(\vec{X})x_i$.

We plot the resulting dynamics as a flow field on a 2-simplex (see Fig. 14), where each vertex i represents a population composed entirely of strategy i , each point on the edge opposite to vertex i represents a population devoid of strategy i , and each interior point represents a population containing all three strategies.

In Fig. 14, the vertices represent the naive-realist (*truth*), critical-realist (*CR 3*), and interface strategies (*IF 3*), competing in an environment with a Gaussian utility structure.

We assume that the *truth* agent assigns a utility to each of the 100 possible resource values, so that n_b in (20) is $\log_2(100)$; since *CR 3* and *IF 3* only need to order their three categories, their n_b in

(20) is $\log_2(3)$. We also assume that $c_k = c_e/10$, making energy costs of perception and knowledge of utility roughly on a par. Instead of varying over the cost per bit, c_e , as we have done in previous figures, we compute the cost to *truth* as a percentage of *truth's* expected payout averaged over competitions with all three strategies. The cost to *CR 3* and *IF 3* is then computed from (20) and *truth's* cost. In this way, the cost for each strategy can be directly compared to the expected payoff for *truth*.

As shown in Fig. 14, for low costs *truth* dominates and for higher costs *IF 3* dominates; in fact, *IF 3* dominates for costs $> 4.25\%$, *truth* dominates for costs $< 4.29\%$, and in between they coexist. For these competitions, *CR 3* is driven to extinction so long as the initial population contains *IF 3* agents. In populations with no *IF 3* agents, *CR 3* dominates *truth* for costs $> 9.3\%$, is dominated by *truth* for costs $< 9.9\%$, and in between they coexist.

In order to show that *IF 3* does not drive *truth* and *CR 3* to extinction only in this isolated case, we ran simulations with additional territories and resources. We first increased the number of territories. In an environment with 30 territories and one resource per territory, *IF 3* drives *CR 3* and *truth* to extinction for costs $> 2.8\%$ of *truth's* expected payoff.

We then increased the number of resources. In an environment with three territories and 30 resources per territory, *IF 3* and *CR 3* each have three categories per resource but use $n_b = \log_2(100)$, so that they can adequately represent the expected utility of each category. In this case *IF 3* drives *CR 3* and *truth* to extinction for costs $> 0.5\%$ of *truth's* expected payoff. These simulations show that, due to the large costs required for *truth* to see more complex environments, *IF 3* is in fact more fit than *truth* in these environments with increased complexity.

In summary, these competitions between naive realist, critical realist and interface strategies show that natural selection does not always favor naive realism or critical realism, and that in many scenarios only the interface strategy survives.

9. Discussion

Perceptual researchers typically assume that “it is clearly desirable (say, from an evolutionary point of view) for an organism to achieve veridical percepts of the world” (Feldman, 2009, p. 875). They assume, that is, that truer perceptions are ipso facto more fit. Although they acknowledge that natural selection can shape perception to be a “bag of tricks” or heuristics (Ramachandran, 1990), they assume that these tricks or heuristics are short cuts to the truth.

We tested this assumption using standard tools of evolutionary game theory. We found that truer perceptions need not be more fit: Natural selection can send perfectly, or partially, true perceptions to extinction when they compete with perceptions that use niche-specific interfaces which hide the truth in order to better represent utility. Fitness and access to truth are logically distinct properties. More truthful perceptions do not entail greater fitness.

One key insight here is that perceptual information is not free. For every bit of information gleaned by perception there are typically costs in time and energy. These costs can of course be defrayed. To reduce costs in time, for instance, neurons can process information in parallel. But each additional neuron in a parallel architecture requires additional energy. Costs in time and energy can trade off, but not be eliminated.

A second key insight is that perceptual information is shaped by natural selection to reflect utility, not to depict reality. Utility depends idiosyncratically on the biological needs of each particular organism. Idiosyncracies in utility will necessarily lead to concomitant differences in perception, even for organisms that

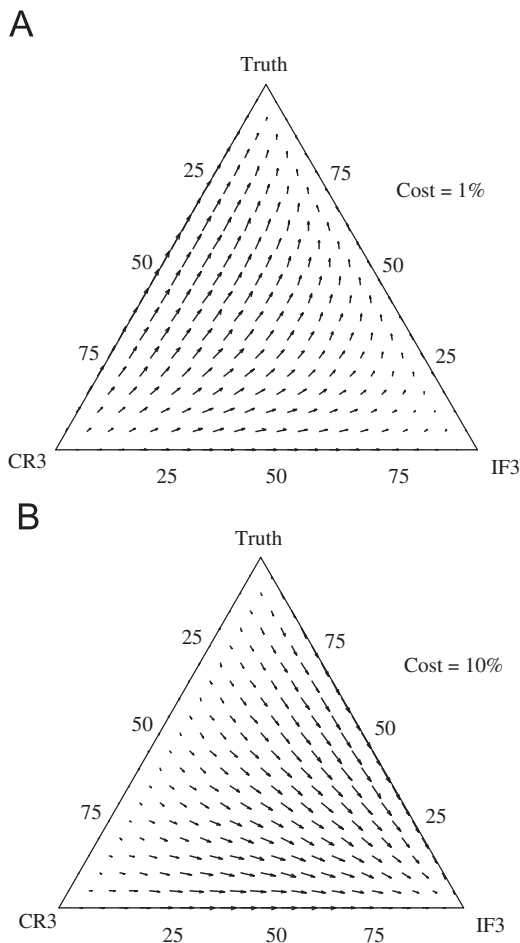


Fig. 14. Competition between *truth*, *CR 3*, and *IF 3* in a game with three territories and one resource per territory. Utility is given by a Gaussian with mean 50 and standard deviation 20. Arrows indicate evolutionary flows. In (A) the cost to *truth* is 1% of its expected payout averaged over competitions with all three strategies (cost per bit = 0.0074). In (B) the cost to *truth* is 10% (cost per bit = 0.074). Payoff matrices for these competitions can be found in Appendix B.

inhabit precisely the same environment and use precisely the same number of bits of perceptual information.

We do not wish to overstate our findings. Our simulations do not find that natural selection *always* drives truth to extinction. They show instead that natural selection *can* drive truth to extinction. However, this still calls into question the standard assumption of perceptual researchers that natural selection favors veridical perception. In consequence, perceptual researchers now face a question. Are there empirically plausible conditions in which natural selection favors true or partially true perceptions? Among the conditions of interest are those concerning environments, competing perceptual strategies, and costs for time and energy.

For the environments and perceptual strategies used in our simulations, we find that naive realist and critical realist strategies go extinct except for certain cases in which the costs for perceptual information are small (as the figures in the previous section illustrate). This finding might reflect a general and important feature of perceptual evolution, or it might be an artifact of our particular simulations. To decide which is the case, a much wider range of simulations is required. For instance, in our simulations the environment is probabilistically invariant over time, but in nature the environment is constantly in flux. Our simulations assume infinite populations and complete mixing, but in nature the populations are finite and competitions are constrained by the spatial locations of potential competitors. Our simulations involve only a rudimentary form of foraging, whereas in nature a wide variety of foraging strategies are possible (Goldstone et al., 2005; Roberts and Goldstone, 2006; Sernland et al., 2003). Our simulations assume that a perceptual strategy is a deterministic map from the world to perceptual experiences; but in nature this map is, in general, probabilistic rather than deterministic. Our simulations assume that perceptual categories have precise boundaries, and thus precise conditions for membership; but psychologists find that membership in a category can depend on degree of similarity to a prototype (Rosch, 1983) or by means of a similarity range parameter (Jameson and Komarova, 2009). Our simulations assume that a perceptual strategy is fixed for the lifespan of an organism, but in nature many organisms are equipped with learning mechanisms and conspecific interactions that allow their perceptual strategies, within certain endogenous limits, to change over their lifespan (Barrett, 2009; Hutteger, 2007; Komarova et al., 2007; Skyrms, 2004). Our simulations use classical computations and classical perceptual strategies, but biological systems can apparently use quantum computations, e.g., in photosynthesis (Engel et al., 2007), and evolution might exploit quantum computation to sift through quantum and classical strategies (McFadden, 2000). Our simulations have involved *truth* and *simple* strategies in many of the competitions, but the real interest for future simulations is in competitions between critical-realist and interface strategies.

Nevertheless, the evolutionary simulations presented here demonstrate that naive-realist and critical-realist strategies can be driven to extinction when they compete with interface strategies that hide the truth. More simulations, in the directions outlined above, are needed to determine if there are plausible environments in which critical-realist perceptions can survive.

Acknowledgments

For helpful discussions and comments on previous drafts, we thank Pete Foley, Simon Hutteger, Kimberly Jameson, Don Macleod, Julia Mossbridge, Louis Narens, V.S. Ramachandran, Don Saari, John Serences, Brian Skyrms, Kyle Stephens, Mark Steyvers, Carol Trabing, Nicole Winter and Julian Yarkony. We thank Elaine Ku for the rabbit images used in Fig. 2.

Appendix A. Truth vs simple payoffs

The expected payoffs for *truth* when competing against *simple* are shown in their analytical form below, as calculated from Eq. (14b). These equations are used in Eq. (15d), to calculate the expected payoffs for *truth* associated with the events G_0 , G_1 , G_2 , and G_3 .

$$E_{3:3}(v|R) = \frac{3b^2 + 2b - 1}{4b},$$

$$E_{2:3}(v|R) = \frac{b + 1}{2},$$

$$E_{2:2}(v|R) = \frac{4b^2 + 3b - 1}{6b},$$

$$E_{3:3}(v|G) = \frac{3(m-b)^2 + 2(m-b) - 1}{4(m-b)} + b,$$

$$E_{2:3}(v|G) = \frac{m-b+1}{2} + b,$$

$$E_{2:2}(v|G) = \frac{4(m-b)^2 + 3(m-b) - 1}{6(m-b)} + b,$$

$$E_{1:2}(v|G) = \frac{(m-b+1)(2m-2b+1)}{6(m-b)} + b.$$

Appendix B. Payoff matrices for three-strategy games

The payoff matrices for Fig. 14 were computed from Table 3 minus a cost of 1% or 10% of *truth's* expected payout averaged over competitions with all three strategies. Each entry to the payoff matrix was computed as the average of 100,000,000 interactions. The exact payoff matrices used for Fig. 14 are displayed in Tables 4 and 5, which clearly show that *truth's* payoffs decrease the most with the increase in cost. Costs are computed using Eq. (20).

Table 3
Payoff matrix (cost = 0).

	CR 3	IF 3	Truth
CR 3	60.15	59.05	58.11
IF 3	63.19	61.77	60.83
Truth	65.72	64.46	63.43

Table 4
Payoff matrix (cost = 1%).

	CR 3	IF 3	Truth
CR 3	60.12	58.02	58.08
IF 3	63.15	61.73	60.80
Truth	65.08	63.82	62.78

Table 5
Payoff matrix (cost = 10%).

	CR 3	IF 3	Truth
CR 3	59.76	58.66	57.73
IF 3	62.80	61.38	60.44
Truth	59.27	58.01	56.97

References

- Antal, T., Nowak, M., Traulsen, A., 2009. Strategy abundance in 2×2 games for arbitrary mutation rates. *Journal of Theoretical Biology* 257, 340–344.
- Barrett, J.A., 2009. The evolution of coding in signaling games. *Theory and Decision* 67, 223–237.
- Billing, J., Sherman, P.W., 1998. Antimicrobial functions of spices: why some like it hot. *The Quarterly Review of Biology* 73, 3–49.
- Calik, S., Gungor, M., 2004. On the expected values of sample maximum of order statistics from a discrete uniform distribution. *Applied Mathematics Computation* 157, 695–700.
- Engel, G.S., Tessa, R.C., Read, E.L., Ahn, T., Mancal, T., Cheng, Y., Blankenship, R.E., Fleming, G.R., 2007. Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature* 446, 782–786.
- Feldman, J., 2009. Bayes and the simplicity principle in perception. *Psychological Review* 116, 875–887.
- Fudenberg, D., Tirole, J., 1998. *Game Theory*, sixth ed MIT Press, Cambridge, MA.
- Geisler, W.S., Diehl, R.L., 2003. A Bayesian approach to the evolution of perceptual and cognitive systems. *Cognitive Science* 27, 379–402.
- Goldstone, R.L., Ashpole, B.C., Roberts, M.E., 2005. Knowledge of resources and competitors in human foraging. *Psychonomic Bulletin & Review* 12, 81–87.
- Heyer, D., Mausfeld, R. (Eds.), 2002. *Perception and the Physical World: Psychological and Philosophical Issues in Perception*. Wiley, New York.
- Hofbauer, J., Sigmund, K., 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge.
- Hoffman, D.D., 1998. *Visual Intelligence: How We Create What We See*. W.W. Norton, New York.
- Hoffman, D.D., 2009. The interface theory of perception. In: Dickinson, S., Tarr, M., Leonardi, A., Schiele, B. (Eds.), *Object Categorization: Computer and Human Vision Perspectives*. Cambridge University Press, Cambridge.
- Huttenger, S.M., 2007. Evolution and the explanation of meaning. *Philosophy of Science* 74, 1–27.
- Jameson, K.A., Komarova, N.L., 2009. Evolutionary models of color categorization. *Journal of the Optical Society of America A* 26 (6), 1414–1436.
- Knill, D., Richards, W. (Eds.), 1996. *Perception as Bayesian Inference*. Cambridge University Press, Cambridge.
- Komarova, N.L., Jameson, K.A., Narens, L., 2007. Evolutionary models of color categorization based on discrimination. *Journal of Mathematical Psychology* 51, 359–382.
- Komarova, N.L., Jameson, K.A., 2008. Population heterogeneity and color stimulus heterogeneity in agent-based color categorization. *Journal of Theoretical Biology* 253, 680–700.
- Krishnamani, R., Mahaney, W., 2000. Geophagy among primates: adaptive significance and ecological consequences. *Animal Behavior* 59, 899–915.
- Lehar, S., 2003. Gestalt isomorphism and the primacy of subjective conscious experience: a Gestalt Bubble model. *Behavioral and Brain Sciences* 26, 375–444.
- Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- Maynard Smith, J., Price, G.R., 1973. The logic of animal conflict. *Nature* 246, 15–18.
- McFadden, J., 2000. *Quantum Evolution*. W.W. Norton, New York.
- Nowak, M.A., 2006. *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Harvard University Press, Cambridge, MA.
- Nowak, M.A., Sigmund, K., 2004. Evolutionary dynamics of biological games. *Science* 303, 793–799.
- Palmer, S.E., 1999. *Vision Science: Photons to Phenomenology*. MIT Press, Cambridge, MA.
- Radnitzky, G., Bartley, W.W. (Eds.), 1987. *Evolutionary Epistemology, Theory of Rationality, and the Sociology of Knowledge*. Open Court, La Salle, Illinois.
- Ramachandran, V.S., 1990. Interactions between motion, depth, color and form: the utilitarian theory of perception. In: Blakemore, C. (Ed.), *Vision: Coding and Efficiency*. Cambridge University Press, Cambridge.
- Roberts, M.E., Goldstone, R.L., 2006. EPICURE: Spatial and knowledge limitations in group foraging. *Adaptive Behavior* 14, 291–313.
- Rosch, E., 1983. Prototype classification and logical classification: the two systems. In: Scholnick, E.K. (Ed.), *New Trends in Conceptual Representation: Challenges to Piaget's Theory?*. Lawrence Erlbaum Associates Hillsdale.
- Samuelson, L., 1997. *Evolutionary Games and Equilibrium Selection*. MIT Press, Cambridge, MA.
- Sandholm, W.H., 2007. *Population Games and Evolutionary Dynamics*. MIT Press, Cambridge, MA.
- Schwartz, R., 2004. *Perception*. Blackwell Publishing, Malden, MA.
- Sernland, E., Olsson, O., Holmgren, N.M.A., 2003. Does information sharing promote group foraging? *Proceedings of the Royal Society of London* 270, 1137–1141.
- Skyrms, B., 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press, Cambridge.
- Szabó, G., Fáth, G., 2007. Evolutionary games on graphs. *Physics Reports* 446, 97–216.
- Taylor, P., Jonker, L., 1978. Evolutionary stable strategies and game dynamics. *Mathematical Bioscience* 40, 145–156.
- Weibull, J., 1995. *Evolutionary Game Theory*. MIT Press, Cambridge, MA.
- Yuille, A., Bülhoff, H., 1996. Bayesian decision theory and psychophysics. In: Knill, D., Richards, W. (Eds.), *Perception as Bayesian Inference*. Cambridge University Press, Cambridge.