Population heterogeneity and color stimulus heterogeneity in agent-based color categorization

Natalia L. Komarova [a,*], Kimberly A. Jameson [b]

[a] Department of Mathematics, University of California, Irvine, CA, USA
[b] Institute for Mathematical Behavioral Sciences, University of California, Irvine, CA, USA

ABSTRACT

Investigating the interactions between universal and culturally specific influences on color categorization across individuals and cultures has proven to be a challenge for human color categorization and naming research. The present article simulates the evolution of color lexicons to evaluate the role of two realistic constraints found in the human phenomenon: (i) heterogeneous observer populations and (ii) heterogeneous color stimuli. Such constraints, idealized and implemented as agent categorization and communication games, produce interesting and unexpected consequences for stable categorization solutions evolved and shared by agent populations. We find that the presence of a small fraction of color deficient agents in a population, or the presence of a "region of increased salience" in the color stimulus space, break rotational symmetry in population categorization solutions, and confine color category boundaries to a subset of available locations. Further, these heterogeneities, each in a different, predictable, way, might lead to a change of category number and size. In addition, the concurrent presence of both types of heterogeneity gives rise to novel constrained solutions which optimize the success rate of categorization and communication games. Implications of these agent-based results for psychological theories of color categorization and the evolution of color naming systems in human societies are discussed.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Human color categorization has a long history of empirical investigation by psychological scientists (see Hardin and Maffi, 1997 for a survey of perspectives), and is a research area that continues to attract active empirical interest (e.g., Davidoff et al., 1999; Roberson et al., 2000, 2005; Kay and Regier, 2003; Jameson, 2005a; Regier et al., 2005, 2007; Sayim et al., 2005; Lindsey and Brown, 2006). The empirical literature suggests that, on one hand, there is a good deal of universality in color categorization across cultures, whereas, on the other hand, a considerable amount of variation is also observed. A long-standing debate in the field has been whether specific universal tendencies exist in the ways different human linguistic societies categorize and name perceptual color experiences, and if so, to what factors (e.g., physical environment, human biology, perception, social features) might such tendencies be attributed. A major challenge for the area continues to be the development of a theory that would account for universal patterns seen in color categorization data while

explaining any differences observed. This is a fascinating problem that has recently attracted the investigative interests of computational scientists (Belpaeme and Bleys, 2005; Steels and Belpaeme, 2005; Dowman, 2007; Komarova et al., 2007a; Puglisi et al., 2007).

1.1. Two perspectives on human color categorization

One widely held perspective is that the commonalities of color categorization across individuals and cultures are largely explained by human perceptual color processing, and universal features of individual psychological processing believed to underlie perceptual color experience. This "*universalist*" view implicitly places the causes for observed color categorization systems *within* the individual observer. The universalist view has historically been successful in characterizing color categorization similarities across a number of ethnolinguistic groups (Kay and Regier, 2003; Kay, 2005; Kuehni, 2005a, b; Regier et al., 2005; Lindsey and Brown, 2006), but less successful identifying the sources of the considerable variation in color categorization and naming that is seen across ethnolinguistic groups.

A different view is that socio-cultural factors contribute substantially to the ways color appearance is categorized and

* Corresponding author. Tel.: +1 949 824 1268; fax: +1 949 824 7993.
E-mail addresses: komarova@math.uci.edu (N.L. Komarova), kjameson@uci.edu (K.A. Jameson).

named by different ethnolinguistic groups (Davidoff et al., 1999; Roberson et al., 2000, 2005). This "*culturally relative*" view suggests that to the degree that two societies share similar environmental, pragmatic and social circumstances, two such societies are more likely to exhibit color categorization systems that resemble one another, unlike systems from two different societies which do not share any environmental, pragmatic or social circumstances. Thus, in contrast to the universalist perspective, this socio-cultural perspective places a large emphasis on factors outside the individual observer such as demands from social practices, environment, and so on. Also implicit in this latter view is the suggestion that at least some of the features that are shared among individuals in a given ethnolinguistic group, stem from commonalities *specific to human society* as it interacts with color in the world, rather than shared due to a common human individual perceptual processing basis.

Although historically these two views[1] have been put forth as somewhat mutually exclusive positions, scenarios necessarily exist in which the universal perceptual processing view and the socio-cultural view both play substantial roles during the development and use of color categorization systems in everyday communications among observers.

1.2. Some challenges in color categorization research

A specific challenge confronting human color categorization research is to model the phenomenon in such a way that both universal individual features and socio-cultural features can contribute in tractable ways to color naming and categorization behaviors in a given ethnolinguistic group. Building one such empirical theory, general enough to serve as a model across ethnolinguistic groups while convincingly demonstrating how the different features trade off in the process of naming and categorization, has historically proven to be a difficult task.

A different, equally important, challenge is to find methods that, essentially, *look across time*, permitting the investigation of the evolutionary dynamics that were present during the development of a society's color categorization system. For this an evolutionary approach seems needed to study dynamically imposed pressures and how those pressures are responded to by a system and by the individuals using that system. Historically, diachronic analyses have been used to address these issues (e.g., Kay, 1975), but such investigations are difficult as they depend greatly on the accuracy of very sparse, subjective historical sources of information.

These two challenges (the empirically based modeling of the interplay between universalist and relativist influences, and the need to examine the evolutionary dynamics of color categorization) are substantial obstacles to understanding human color naming and categorization phenomena even now. This article presents an investigative approach that addresses these challenges by idealizing two features inherent in human color categorization phenomena which have raised recent interest in the empirical literature: namely (i) color categorization under variation in individual color perception and (ii) color categorization under color space variation.

1.3. Idealizing two realistic influences on color categorization

The first feature, namely, potential variation in individual color perception and its influence on color categorization systems,

originates from *within* the individual observer, representing what is generally viewed as a universalist type influence. Individual variation in color perception, while generally addressed in color vision science, has only recently played a role in the empirical color categorization literature (Sayim et al., 2005; Jameson, 2005d; Kuehni, 2005a, b), and the views on the impact of such variation are mixed. Jameson (2005a–d) recently discussed possible scenarios in which individual observer color perception variation might trade off with aspects of communication pragmatics arising outside observers in social situations of communicating about color. The present article explores such scenarios by formalizing simplified models of individual color perception variation and incorporating those models into individual category learning and population communication games. This provides a systematic way to assess the impact of idealized color perception variation on color category solutions.

The second feature idealized in our study, namely, the heterogeneity of color utility, potentially involves both cultural and environmental components, and therefore represents a relativistic influence. Realistically, several different sources may produce variation in color stimuli. Two such sources are: (a) environmental differences in color category exemplar frequency (i.e., frequency of green category exemplars may be greater in the rain forest compared to that for violet category exemplars), potentially expressed as a nonrandom sampling, or a heterogeneous "color diet"; (b) differences in pragmatically defined category ranges across the available color category exemplars (e.g., a nutritionally valuable reddish-orange exemplar may establish a trend for a more refined categorization, or a region of increased salience (RIS), compared to less valuable bluish category). Type (a) color stimulus variations were studied by Steels and Belpaeme (2005), Komarova et al. (2007a, b) and Dowman (2007). In the present article we introduce a novel investigation of type (b) heterogeneity and its effects on categorization.

The formal approach used here is based on the recent work of Komarova et al. (2007a, b), where our agent-based, evolutionary game-theoretic model of human color categorization was developed. We consider idealized populations of agents with different color perception abilities. Here individuals learn to categorize simulated colors by playing "discrimination-similarity games," and also learn to communicate the meaning of categories to each other. Population heterogeneity is introduced by varying, across agents, the "psychophysical transform" between the physical space of color stimuli and agents' perceptual spaces. The form of color stimulus heterogeneity used here is implemented by defining a RIS in the color stimulus domain.

1.4. Aims of the article

The present article aims to address several goals. First, using highly idealized models of real world phenomena, to develop investigative methods for examining possible trade offs in scenarios in which universal perceptual processing and socio-cultural factors both play substantial roles in the development and in the use of color categorization systems. Second, to examine the suggestion of Jameson (2005d) that nonrandom population heterogeneity will have a substantial impact on agent-based color category learning and on communicating population color category solutions, as well as figure prominently in the communication pragmatics (Jameson, 2005a, b). And third, to directly investigate the suggestion (Jameson, 2005a) that color salience (whether imposed by environmental scene statistics or through culturally imposed color sampling or salience) represents an important constraint on the development and use of a color

---

[1] Which typically appear in the literature as polarized extremes of the debate, but which are more akin to distant points in a continuum of possible explanations that involve a variety of potentially influential factors.

categorization system that is shared by individuals in a color observing population. Finally, although these investigations focus on agent-based categorization, we intend this approach to be used to complement the empirical study of human color categorization by providing a means to examine the theoretical plausibility of competing explanatory theories and hypotheses arising from the empirical literature.

The rest of this paper is organized as follows: In Section 2 we briefly review the phenomenology of population and color stimulus heterogeneity in ethnolinguistic societies. Section 3 describes the mathematical framework for modeling color categorization and explains how the two types of heterogeneity are incorporated. Section 4 examines the role of population heterogeneity in color categorization in our simulated environment, while Section 5 studies color stimulus heterogeneities. In Section 6 we investigate the interaction of the two types of heterogeneity, and provide conclusions and a discussion in Section 7.

2. Population and color space heterogeneities: a brief phenomenological description

This section gives the rationale underlying the idealized individual color perception variation used in our agent populations, and the rationale for the color stimulus variation we implement.

2.1. Observer heterogeneity and color categorization

Human trichromat color processing is, in general, the norm; although observer heterogeneity exists, seen as varying degrees of color vision anomaly across different populations. Normal trichromacy is initiated by the absorption of photons by three classes of cones in the human retina which have peak sensitivities in the long-wavelength (L), middle-wavelength (M), and short-wavelength (S) regions of the electromagnetic spectrum (see Cornsweet, 1970 for details). As a result, any real world color stimulus can be specified by three numbers representing the cone responses, and all colors visible to the color-normal observer are contained in a three-dimensional color space. While normal trichromat color processing arises from this trivariate neural signal that originates in the human retina, some forms of human color processing deficiency arise from an impoverished neural signal. The most common forms of impaired color vision processing[2] occur when, for example, an individual lacks a full complement of the three photoreceptor classes typically found in the normal observer; such individuals are called *dichromats* (Nathans et al., 1986a). Several forms of dichromacy exist, each lacking one of the retinal photopigments: That is, the L-photopigment type in protanopes, the M-type in deuteranopes, and the S-type in tritanopes (Nathans et al., 1986b). For these dichromatic observers any color stimulus essentially initiates only two cone responses, and all colors that they can discriminate are captured in a two-dimensional color space. Compared with trichromatic vision, dichromatic vision entails a loss of discrimination and often a reduced color gamut. For the modeling of observer heterogeneity in the present article, such deficiencies are thought of as resembling a loss of a dimension of color space compared to normal trichromatic processing.

Color deficiencies affect a substantial number of people world wide, although exact proportions vary among groups of different racial ancestry. In Caucasian groups, for example, it occurs in about 8% of males and only about 0.4% of females (Sharpe et al., 1999, p. 30). Communities that are isolated, or have a restricted gene pool, sometimes produce high proportions of color blindness. Examples include rural Finland, Hungary, and some of the Scottish islands. In the United States, about 7% of the male population—something on the order of 20 million men—and 0.4% of the female population either cannot distinguish red from green, or see red and green differently (Sharpe et al., 1999, p. 30). It has been found that more than 95% of all variations in human color vision involve receptors in males with peak sensitivities in spectral regions associated with reds and greens. Due, in part, to a difference in the inheritance mechanisms, it is very rare to find deficiencies in males or females for the blue end of the spectrum.[3]

It is important to note that while throughout the present investigations we refer to *dichromat* agents, the model we assume here is an extreme simplification of the discrimination model found for human dichromats under realistic circumstances. Nevertheless, as a starting point for introducing population heterogeneity into the present investigations, it is a very useful simplification, because (i) it is more realistic than implementing a form of random observer variation (Steels and Belpaeme, 2005, p. 517),[4] (ii) it is in part justified by simplified models of human behavior (Cornsweet, 1970, p. 184) and (iii) it is consistent with empirical results for dichromats (e.g., Shepard and Cooper, 1992). As the present results show, our method proves to be an appropriate investigative tool for understanding the ways population heterogeneity plays out in simulations of individual and population color categorization behaviors. We emphasize that our methods of evolving categorization systems do not intend to model how humans learn to categorize colors or how human color categorization systems came about. Nevertheless, the present results are informative with regard to experimental results concerning color categorization and theoretical explanations given for the frequently seen similarities between separately evolved human systems of color naming.

2.2. Color space heterogeneity and color categorization

Perceptually, colors vary continuously, and are generally ordered in three dimensions, with hue and saturation as polar coordinates of a color circle, and brightness as the third dimension. Human color naming reflects this perceptual structure in various ways. For example, in human color naming, each name describes a portion of color space that is a *connected set*. In other words, in the majority of empirically observed cases, from reliable and consistent participants, the meaning of a color name will almost never be a set of colors that can be partitioned into disconnected subsets. (Possible exceptions to this connected-set tendency are puzzling cases found in data from nonindustrialized, pre-literate, informants, for which the observation may be due to some form of random participant uncertainty regarding empirical task instructions, experimental error, actual disconnected sets of denotata, or some combination of these.) In the great majority of cases, colors that are perceptually very similar—that is, colors that

---

[2] Setting aside for this discussion, rarer forms such as acquired or congenital achromatopsia.

[3] The genes for the short-wavelength-sensitive photopigments arise through autosomal inheritance, linked to chromosome 7, and is not an X-chromosome-linked trait (see Sharpe et al., 1999, p. 44).

[4] Compared to Steels and Belpaeme (2005), the definition of an agent in the present investigation does not involve adopting the rich set of human color perceptual features that is typically used in investigations of color learning in simulated individuals. Rather our agents are only endowed with a primitive ability to carry out color discriminations. Endowing our agents with only the simplest of discrimination abilities permits a more direct assessment of the influences of reinforcement learning and between agent communication on the formation of any category structures observed.
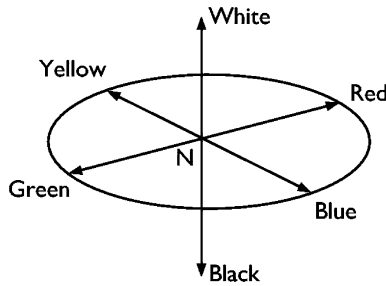
Fig. 1. Dimensions of a three-dimensional color solid consisting of color samples (or *chips*) arranged by approximately uniform color differences along each of the three Munsell dimensions: an angular dimension of *hue* for representing nonachromatic colors (with two axes—yellow-to-blue and green-to-red—given to orient the color circle); a vertically oriented linear dimension of *value* (lightness or brightness); a radial linear dimension of *chroma* (saturation) on which maximum vividness of color is achieved on the hue circle, and colors decrease in vividness in the direction of the achromatic axis, to become devoid of color at a neutral point *N*.

are very "close" to one another in the perceptual color space—will almost always be described by the same name, while different names will generally denote dissimilar colors.

Fig. 1 schematically depicts the dimensions (hue, value and chroma) widely accepted for the representation of human color experience. Such dimensions when configured as a color stimulus solid reflect the standard form of the stimulus spaces used in human color categorization research (e.g., *Munsell color order system*, Nickerson, 1976). Despite being of potential pragmatic importance, essentially no stimulus space used in human color categorization research incorporates any additional dimensions, such as texture, shimmer, gloss, etc.

When talking about color order systems, a useful concept is "perceptual uniformity". A system is said to be perceptually uniform if the distances separating each stimulus on a given color order system dimension is linearly related to the *cognitive* distances separating each stimulus in an observer's internal color appearance similarity space (Nickerson, 1979). While almost all physical color order systems like the Munsell system aim to model perceptually uniform color differences along the three dimensions illustrated, they only approximately achieve such perceptual uniformity, and in all perceptual uniformity is only dimension-specific, and is not achieved when dimensions interact.[5]

Variations in color appearance spaces generally do not involve omissions or extensions of any of the features mentioned above, but rather involve differences in how the three-dimensional color space is sampled or perceived. As suggested earlier, color space heterogeneity can manifest itself in several ways: it can arise from an environmental color bias (e.g., a bias for green foliage in rainforests; in the literature this kind of sampling is exemplified by "visual diet" discussions), or from an observer's visual mechanisms (e.g., a built in bias that does not distinguish green foliage from red foliage), or from some nonuniformity in pragmatic salience (e.g., a pragmatically tuned bias that benefits the successful identification of ripe nutritious fruit).

In the present investigations we constrain our study of the Fig. 1 space to the circular dimension of hue. The hue circle is a good choice as the initial categorization domain studied because it is a natural one-dimensional reduction of the entire color space, being the outcome of dimension reduction of spectral functions having any possible spectral form to three dimensions (Kuehni,

2003; Kuehni and Schwarz, 2008). We also manipulate this hue dimension stimulus sampling by examining cases where different regions of the color space carry a different pragmatic weight for our simulated observers compared to the rest of the color space. This obviously idealizes the third type of environmental color bias mentioned above. Thus, in a portion of the present investigations we define a RIS in the hue circle, where the pragmatic importance of the colors distinguishes it from the remaining colors on the hue circle.

Clearly, our choice of a hue circle as the stimulus domain on which categories are formed is a minimalization of what is typically considered the color appearance space of normal human observers. In the present paper, we use the hue circle stimulus in order to investigate if color categorization is influenced by increased complexity of communication scenarios, and to study the trade offs between stimulus variability and simulated observer variation. A higher-dimensional stimulus space is a subject of our further investigations.

3. Mathematical framework for modeling categorization

Here we present a mathematical formulation of the key notions such as "perceptual space", "color categorization", "optimal categorization", etc., and define a dynamical process of color categorization in a population of simulated observers. Many of these concepts have been first developed in our recent paper (Komarova et al., 2007a, b) and are briefly reviewed here. In addition, a novel mathematical concept (presented in the next subsection) incorporated here is a transformation from the physical space of stimuli to the perceptual space, which allows us to introduce observer heterogeneity in the model.[6]

3.1. Physical and perceptual spaces

The stimuli (color chips) are taken from a physical space which is the same for all the viewers. The physical space can be modeled for instance as a (discrete) circle. Each chip is characterized by a coordinate along the circle, with numbers from 1 through n, such that chip number n is next to chip number 1. As a model for such a circle we can for instance take a numbered physical set of chips from the Munsell set.

A viewer observes the color chips. As a result, the stimulus is represented in the viewer's "perceptual" space, which is individual for each viewer. This is modeled as a transformation from the physical space to the perceptual space which represents the individual's subjective impression of the color chip,

$$i \to \mathscr{F}(i), \quad 1 \neq i \leqslant n,$$

where $\mathscr{F}$ is the psychophysical function that maps stimuli to perception.

For example, in our modeling of the perceptual space of a dichromat, *reddish* and *greenish* chips generally appear the same. In the simplest model used here, the transformation between the physical space to the perceptual space of a dichromat collapses a hue circle to a line with coordinates 1 through $n/2 + 1$, such that

$$i \to \mathscr{F}(i) = \begin{cases} i, & i \leqslant n/2 + 1, \\ n + 2 - i, & i > n/2 + 1. \end{cases} \tag{1}$$

In a more general model we assume that a subset of size l chips collapses onto another subset of the same length, opposite to it. Namely, define the confusable regions to be two connected sets of
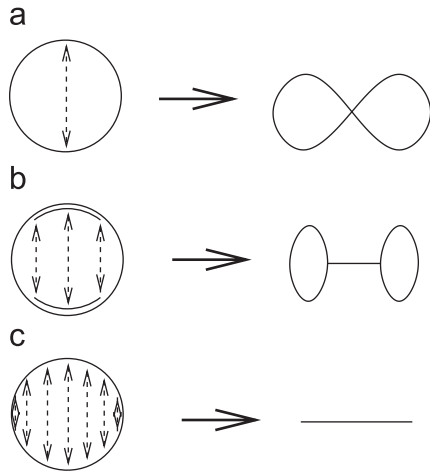

Fig. 2. Examples of the psychophysical transformation from the physical space to the perceptual space which represents the individual's subjective impression of the color chip, formula (2). (a) $l = 1$, (b) $1 < l < n/2 - 1$ and (c) $l = n/2 - 1$.

length l centered around $n/4 + 1$ and $3n/4 + 1$:

$$\frac{n}{4} - \frac{l-1}{2} + 1 \leqslant i \leqslant \frac{n}{4} + \frac{l-1}{2} + 1,$$
$$\frac{3n}{4} - \frac{l-1}{2} + 1 \leqslant i \leqslant \frac{3n}{4} + \frac{l-1}{2} + 1.$$

We set the following transformation:

$$i \to \mathscr{F}(i) = \begin{cases} n + 2 - i, & \frac{3n}{4} - \frac{l-1}{2} + 1 \leqslant i \leqslant \frac{3n}{4} + \frac{l-1}{2} + 1, \\ i, & \text{otherwise,} \end{cases} \quad (2)$$

see Fig. 2. In general, the parameter l varies between 1 and $n/2 - 1$, and defines the degree of the agents' color perception impairment (with low values of l corresponding to mild impairment, and high values of l corresponding to a large degree of impairment). The previous simple model (Eq. (1)) is obtained from the more general model by setting $l = n/2 - 1$. We can say that in case $l = n/2 - 1$ the circle collapses onto a horizontal line (Fig. 2(c)), and for $1 \leqslant l < n/2 - 1$ it transforms into an ∞-shaped object (Fig. 2(b)), with only one confusable pair of chips in the extreme case of $l = 1$ (Fig. 2(a)).

Note that the transformation from the physical space to the perceptual space of an agent with unimpaired color perception, or a "normal observer," is simply the identity transformation:

$$i \to \mathscr{F}(i) = i, \quad 1 \leqslant i \leqslant n \quad (3)$$

(as long as the physical stimulus space being considered is one that approximates a perceptual ordering as, for example, the Munsell color stimulus space aims to normatively achieve).

All transformations defined above, as well as all meaningful transformations in the context of this study, possess an invariant set, $I_d = \{i_1, \ldots, i_{n_d}\}$, such that $\mathscr{F}(I_d) = I_d$. The integer n_d is the size of the perceptual space of dichromats. In general, we have $n_d \leqslant n$.

The dichromat perceptual space described, and the psychophysical transform modeled by formula (2) are idealizations of actual dichromat observers. For the points made in this article these idealizations are appropriate and sufficient. Investigations of dichromat models which incorporate more realistic features are employed in a forthcoming article (Jameson and Komarova, 2008).

3.2. Color categorization

The color categorization of each agent is a vector function from the individual, perceptual space of each viewer (defined as the viewer's invariant set) to $[0, 1]^m$, where m is the total number of color categories used. We can represent it as n_d vectors, $\mathbf{f}_1, \ldots, \mathbf{f}_{n_d}$, where each $\mathbf{f}_k$ is a point in an m-dimensional simplex. In other words, the m entries of each of the vectors are nonnegative and sum up to one. The meaning of such vectors is that for each perceived signal, they give the probabilities that the viewer will categorize it into each of the m color categories.

A categorization is called a *deterministic categorization* if each vector $\mathbf{f}_k$, $1 \leqslant k \leqslant n_d$, contains a component equal to one. This means that the corresponding chip is categorized to a color category with certainty. We will see that both deterministic and probabilistic categorization solutions are meaningful in the context of heterogeneous populations.

3.3. Discrimination-similarity game and the similarity range

Our categorizing algorithms are based on the following idea: Colors that are highly similar perceptually to one another are highly likely to belong to the same category. More specifically it is based on the following three principles: (i) categorization is important; (ii) to be useful, categorization should attempt to minimize ambiguity and (iii) when color is a salient or meaningful cue for categorization, two randomly chosen objects that have similar color appearances are more likely to be categorized together than are two objects that have dissimilar colors. These three principles are captured by a similarity range parameter called *k-similarity* (defined below), used in individual and population discrimination-similarity games.

In general, the discrimination-similarity game proceeds as follows: Two chips are randomly drawn from a (common) physical stimulus space. For every game, the psychophysical transformation that maps the stimulus chips to that agent's perceptual space is performed. Then, an agent probabilistically assigns a color category to each chip, by using its individual categorization function.

Next, we need to assess whether the game is a success or a failure. For the purposes of this analysis, it is assumed that we are in a situation where color is used as the basis for pragmatically categorizing stimuli. This is done by means of the pragmatic quantity k_{sim}. This quantity applies to chips in the physical space. If the two chips are less than k_{sim} units apart then they probably have similar properties (and it pays off to categorize them in the same class). If they are farther than k_{sim} apart then they are probably very different and a viewer is better off putting them into different classes.

It is important to emphasize that the range k_{sim} is not another perceptual version of "just noticeable difference" (or *jnd*). Colors that differ by a perceptual jnd are well within the similarity range, as are adjacent chips in the Munsell color system described earlier. In general, some colors within the k_{sim} range are easily distinguished perceptually by any agents. The notion of k_{sim} intends to capture the *importance* of categorizing two chips as being "different" rather than "the same."

By applying the k_{sim}-test to the two physical chips we can say whether they are "close" or "far" apart from a pragmatic categories standpoint. A game is successful if (a) two chips that are pragmatically close are assigned to the same color category by the agent or (b) two chips that are pragmatically far apart are assigned to two different categories by the agent. The game is a failure otherwise.

By using the rules above, we can define the success rate, S, for a given color categorization. Given a categorization, let us play N rounds of the discrimination-similarity game. Then

$$S = \lim_{N \to \infty} \frac{\text{Number of successful games}}{N}.$$

In other words, S is the probability for a given color categorization that a game is successful.

3.4. The optimal color categorization

Our earlier studies (Komarova et al., 2007a, b) have identified the class of color categorizations that maximize the success rate of the discrimination-similarity game for agents modeled as normal observers (transformation (3)), with a homogeneous stimulus space. Let us suppose that the sampling procedure for the chip stimuli obeys the uniform distribution. Define the optimal number of categories, m_{opt}, as an integer number close to the value

$$m = n[2k_{sim}(k_{sim} + 1)]^{-1/2}. \qquad (4)$$

Then if the chips are points on an interval, then the following categorization maximizes the success rate: chips

$$\frac{(i - 1)n}{m_{opt}} + 1, \quad \frac{(i - 1)n}{m_{opt}} + 2, \ldots, \frac{in}{m_{opt}}$$

belong to color category v_i, where $1 \leqslant i \leqslant m_{opt}$ and for all $i \neq j$, $v_i \neq v_j$, see Fig. 3(a). If the chips are arranged on a circle, then the following categorization maximizes the success rate: the circle is divided into regions

$$\left[1, \frac{n}{m_{opt}}\right], \quad \left[\frac{n}{m_{opt}} + 1, \frac{2n}{m_{opt}}\right], \ldots, \left[\frac{(m_{opt} - 1)n}{m_{opt}} + 1, n\right],$$

all the vectors $\mathbf{f}_i$ with i inside the same region are identical to each other and have only one nonzero component. Any vectors $\mathbf{f}_i$ and $\mathbf{f}_j$ with i and j belonging to different regions are different from each other. Also, any shift of this pattern along the circle also yields an optimal categorization, see Fig. 3(b). There is no other categorization that has the success rate equal or bigger than those achieved by the above categorizations.

3.5. Individual dynamics with the reinforcement learning algorithm

The optimal categorizations described above have been found theoretically. However these results require the assumptions of (i) population homogeneity which forces the transformation $\mathscr{F}$ to be an identity function, (ii) unbiased sampling of the color stimulus space in discrimination-similarity games and (iii) homogeneity of color space with regard to pragmatic importance (or k_{sim}). Therefore it is important to design a numerical method which finds the optimal color categorization in different (i.e., less symmetrical) systems.

To achieve this goal, we have come up with the rules for individual and group learning dynamics which converge to a categorization close to optimal. Here we describe the individual learning dynamics with the reinforcement learning algorithm.
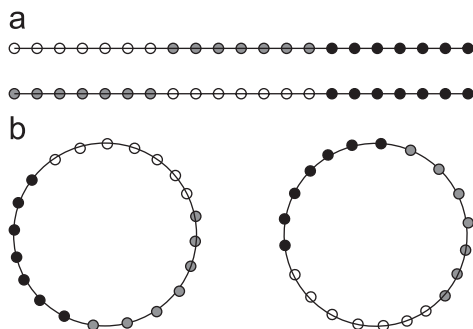

Fig. 3. Some optimal categorizations for (a) an interval and (b) a circle. White, gray and black circles represent the $m = 3$-color categories and $n = 21$.

Suppose many rounds of the discrimination-similarity game are played with an individual agent equipped with some initial color categorization. For successful games the agent earns a successful update of its color categorization, and for cases of categorization failures an unsuccessful update is earned. For such situations we adopt rules for a "uniform" reinforcement learner, briefly described below (see Komarova et al., 2007a, b for details).

Each color chip i is associated with a vector $\mathbf{X}_i$ whose m components are integer numbers that add up to some constant, L: $\sum_{j=1}^{m} [\mathbf{X}_i]_j = L$ for all $1 \leqslant i \leqslant n$. Then the categorization components are defined as normalized entries of these vectors:

$$[\mathbf{f}_i]_k = [\mathbf{X}_i]_k / L.$$

Let us suppose a learner plays a game with chips i and j, and assigns them to categories v_i and v_j, respectively. In case of a successful game, a successful update is applied to the color categorization of both chips. This means that for each chip i and j, the successful category is strengthened by a certain amount, Δ^+, and at the same time all other nonzero categories are reduced by a net amount equal to Δ^+. If a categorization fails, then for both chips, the value of the category associated with the chip decreases by an amount Δ^-, and all other categories are enhanced by the amount $\Delta^- / (m - 1)$.

Here is a more precise definition of the algorithm. If the game is a success, the following operation is performed regarding the component i of the categorization:

$$[\mathbf{X}_i]_{v_i} \to [\mathbf{X}_i]_{v_i} + \Delta^+, \quad [\mathbf{X}_i]_k \to [\mathbf{X}_i]_k - \delta_k^-, \quad k \neq v_i,$$

where

$$\Delta^+ = \sum_{k \neq v_i} \delta_k^-$$

and quantities δ_k^- are defined as follows:

$$\delta_k^- = \begin{cases} 1/(m - 1), & [\mathbf{X}_i]_k \geqslant 1/(m - 1), \\ [\mathbf{X}_i]_k, & \text{otherwise}. \end{cases}$$

In case of a failure, the categorization of chip i is updated as follows:

$$[\mathbf{X}_i]_{v_i} \to [\mathbf{X}_i]_{v_i} - \Delta^-, \quad [\mathbf{X}_i]_k \to [\mathbf{X}_i]_k + \delta^+, \quad k \neq v_i,$$

where we define

$$\Delta^- = \begin{cases} 1, & [\mathbf{X}_i]_{v_i} \geqslant 1, \\ [\mathbf{X}_i]_{v_i}, & \text{otherwise} \end{cases}$$

and $\delta_k^+ = \Delta^- / (m - 1)$.

Similar operations are performed regarding chip j.

3.6. Population dynamics

In order to model interactions in a population of individuals, we use the following rules for population dynamics. At each round of the game we draw two chips and also pick two observers. Both observers play the game. Three outcomes are possible.

If both players are unsuccessful, they both perform the reinforcement update for the failure scenario described above.

If one of the players is successful and the other one is not, the successful viewer is assigned the role of a teacher, and the other one—the role of a learner. The learner updates his color categorization for the stimuli in question in accordance with that of the teacher, which represents the process of learning of one agent from the other. This is done according to the following rule:

$$[\mathbf{X}_i]_{v_1^{(T)}} \to [\mathbf{X}_i]_{v_1^{(T)}} + \delta_{learn}, \quad [\mathbf{X}_i]_{v_1^{(L)}} \to [\mathbf{X}_i]_{v_1^{(L)}} - \delta_{learn},$$

where $v_1^{(L)}$ is the category chosen by the learner for chip i, $v_1^{(T)}$ is the category chosen by the teacher for the same chip, and we

define $\delta_{learn} = 1$ if $[\mathbf{X}_i]_{v_1^{(L)}} > 1$, and $\delta_{learn} = [\mathbf{X}_i]_{v_1^{(L)}}$ otherwise. Similarly, the learner updates the component corresponding to chip j. The teacher performs the usual successful update described earlier.

Finally, if both players are successful, then the teacher is picked at random. Again, the teacher performs a successful update while the other player "learns" from the teacher.

3.7. Numerical simulations

As our main tool we used computer simulations of population dynamics, using the rules presented above. Typically, we start from a random categorization for all players, and simulate rounds of the discrimination-similarity game for randomly chosen pairs of agents. Eventually, in many circumstances, a common (i.e., shared by most of the agents) and relatively stable categorization appears. Below we will use the following terminology: a "round" is one discrimination-similarity game played by one pair of agents. A "run" is a sequence of rounds terminated according to some rule (e.g., after a given number of rounds, or once a stable shared categorization is achieved).

A population of agents can be heterogeneous. This means that a given number of players are dichromats, whereas the rest are normals. All the parameters of simulation are defined in Table 1.

Table 1
Definitions for all the parameters of the model

N	Total number of agents
N_d	The number of agents that are dichromats
n	Total number of chips
m	The maximum number of color categories
k_{sim}	Parameter defining the pragmatic utility of chips
L	Parameter of the reinforcement learner algorithm
l	Confusable region length for dichromats

For illustration we present one example of a run, where a population of $N = 100$ individuals with $N_d = 2$ dichromats performed 10^8 rounds of the game, see Fig. 4. The outcome consists of N categorization functions (one for each of the N agents in the population). In Fig. 4, the bar charts show the color categorization of three individual agents. The horizontal axis represents the 50 chips that form a hue circle. The color bars correspond to the probabilities $[\mathbf{f}_i]_k$, that is, the probabilities with which a given chip is assigned to a given color category. In Fig. 4(a) we show a typical "random" color categorization near the beginning of a simulation. Each chip has a similar chance to be categorized into any of the 6-color categories. Note that the color code used in the charts is arbitrary and does not correspond to any actual color categories of humans.

Fig. 4(b) shows a typical color categorization of a normal individual after 10^8 runs. We can see that the first five chips are almost always classified as "blue," the next group of about six chips—as "purple," etc. This categorization is close to optimal, that is, it subdivides the hue circle into a number of nearly equal parts which are all deterministically categorized into different categories. Note that some nondeterministic categories are observed near the color boundaries. We believe that this is a realistic outcome, as there is a lot of variation in color naming at category boundaries even among normal observers. Color category assignments have been empirically shown to change in normal individuals depending on both internal and external circumstances, such as varying the surrounding context in which perceptually identical colors are judged and named (Sayim et al., 2005). Fig. 4(c) shows a color categorization of a dichromat and is discussed in Section 4.1.

3.8. Population results: the "winning category"

It is clear that a collection of 100 individual color categorizations contains a lot of information and is difficult to visualize for
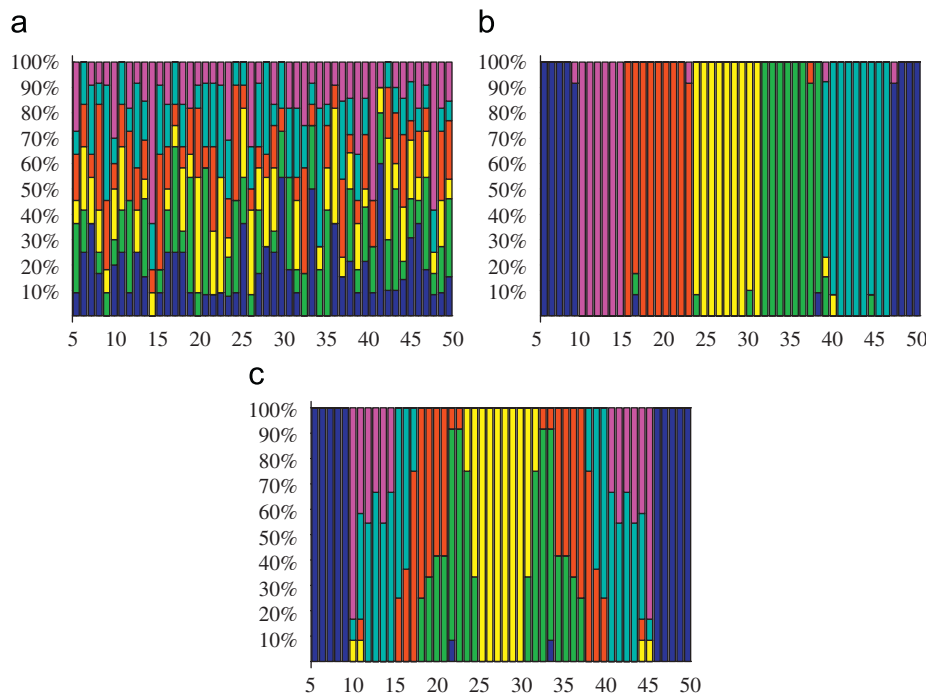

Fig. 4. Examples of color categorizations of individuals. The horizontal axis represents the 50 chips, and the vertical bars of each color correspond to the values $[\mathbf{f}_i]_k$, that is, the probabilities to use a given color category for a given chip. The color assignment of the bars is arbitrary and does not correspond to any real color categories. (a) A random color categorization of an agent, near the beginning of a simulation. (b) The color categorization of a normal person, after 10^8 runs. (c) The color categorization of a dichromat, after 10^8 runs. The parameters are $m = 6$, $L = 12$, $k_{sim} = 6$, $N = 100$, $N_d = 2$.

all agents. Therefore, we need a different way to represent and interpret population categorization results. It is useful to introduce the following quantities.

For each player, for each color chip, we define the "most likely" color category, the "second most likely" color category, etc. The most likely color category is the category that the player most frequently assigns to the chip, the second most likely category is the second most frequent category (for the particular agent), etc. As a result, the color categories are ranked for this chip for each player, from 1 to m.

Then, for each chip, we define the *winning category*. This can be done in a variety of ways. Here we will present the following general method. Suppose that for a given chip, color category k is the most likely category for $n_1^{(k)}$ agents, it is the second most likely category for $n_2^{(k)}$ agents, etc. If an agent does not use a certain category at all, we disregard such category and do not include it in the count. For this chip, we assign a numerical "score" to all categories, $1 \leqslant k \leqslant m$, according to formula:

$$s_k = \sum_{i=1}^{m} \xi_i n_i^{(k)}, \quad 1 \leqslant k \leqslant m,$$

where the values ξ_i are numerical "weights", and the summation is over all ranks of categories, from 1 to m. The winning category for this chip is the one with the highest score.

The result of the operation of identifying the winning category will of course depend on the choice of the coefficients. Here we concentrate on two examples:

- If we take $\xi_1 = 1$ and $\xi_i = 0$ for all $i > 1$, then the winning category is the "most popular" color category. That is, the winner is the color category that for the plurality (but not necessarily the majority) of the population is the most likely (as defined above) color for a given chip. This is similar to the simple plurality voting system. The corresponding score will be referred to as the *plurality score*, see Saari (2008).
- Alternatively, we choose $\xi_i = m - i$, and obtain the so-called Borda count (with truncated ballots). The Borda count method with truncated ballots is a standard procedure used in choice and decision theory and voting literature, see Saari (2008) for an overview of methods from the area. The corresponding score will be referred to as the *Borda score*.

Finally, we can calculate the "degree of agreement", which is measured as the proportion of the population using the winning (as defined above) color for each chip. This varies between $1/m$ and 1, and the closer it is to 1, the better the agreement.

Both the plurality method and the Borda method have advantages and disadvantages. The plurality method will give the best degree of agreement, as it is defined here. However, it disregards a larger amount of information about individual categorization functions, which is an important consideration here. For example, it may happen that even though 45% of the agents use category 1 as their most likely category, the majority of the population (the other 55% of the agents) never use it and use categories 2 or 3 instead. The Borda method seeks to find a category which is in some sense a compromise, a category which the largest majority will find to be least disagreeable, even if not their favorite.

To be more concrete, let us consider the following example. Suppose 45 out of 100 agents categorize a given chip as G (for "green") with probability 50%, as B (for "blue") with probability 30%, and as P (for "purple") with probability 20%. In other words, these agents have categorization vector $(50, 30, 20)$. Next, suppose that 35 agents have categorization $(5, 70, 25)$ and the remaining 20 agents—categorization $(5, 25, 70)$. We can calculate the values

Table 2

A hypothetical example of color categorization counts that illustrates the winning category concept

	1st	2nd	3rd	Borda score
G	$n_1^{(G)} = 45$	$n_2^{(G)} = 0$	$n_3^{(G)} = 55$	$s_G = 90$
B	$n_1^{(B)} = 35$	$n_2^{(B)} = 65$	$n_3^{(B)} = 0$	$s_B = 135$
P	$n_1^{(P)} = 20$	$n_2^{(P)} = 35$	$n_3^{(P)} = 45$	$s_P = 75$

n_k for the color categories G, B and P; they are presented in Table 2. The plurality score is given by the "1st" column in Table 2; it is the largest for the G category, so the green category is "the most popular color" for the chip in question. The Borda score is presented in the rightmost column of the table; it is the largest for the B category, so B is the winner in the Borda sense. Thus, by this example, different winning categories are produced by the plurality score method compared to the Borda score method. This occurs, largely, because the Borda score computation considers frequency information from ranks other than just the top-most ranked category, the latter being the basis for the plurality score.

Fig. 5 illustrates the two types of count by presenting the plurality and the Borda winning categories for $n = 50$ chips after $10,000$, $20,000$, etc. rounds of games. We can see that at the beginning, when the degree of agreement in the population is low, the two scores disagree, and as the degree of agreement becomes high, they tend to give very similar results. After less than 10^5 rounds, the population converges to a nearly optimal categorization (of six nearly equal color categories in this particular case), the degree of agreement is very high (not shown) and the two scores give nearly identical results (except perhaps some color chips right at the category boundaries).

This result is general, because as a nearly optimal solution gets established in the population, the categorization functions of individual agents become nearly deterministic, which means that only one value of the categorization function is nonzero, and the two counts (the plurality and the Borda counts) yield identical results. Because the emphasis in this article is interpreting stabilized, near-optimal solutions, below we present the simpler plurality analyses which, for the cases we consider here, produces the same results as the Borda score.

To summarize, instead of N bar-charts of the type presented in Fig. 4, only two graphs are sufficient to represent results of a population run: a line graph showing the winning color for each chip, and another graph showing the degree of population agreement for each chip. An example is presented in Fig. 6. This figure shows the color categorization assignments for each chip determined by a population of agents (as in Fig. 4), calculated according to the plurality score, with a smaller inset graph showing the degree of agreement among the agents.

4. Heterogeneous populations

4.1. Modeling dichromat and normal observers as agents

As described earlier, in real human populations individual categorization functions of normal observers and color deficient observers differ substantially. Similarly, in the present simulations the individual categorization results of normal and dichromat agents are different. Figs. 4(b) and (c) show the categorization of a normal observer and a dichromat, respectively, after 10^8 rounds of the game. In this example, two out of 100 agents are dichromats with psychophysical transformation (1). The color categorization of the normal observer is close to an optimal categorization with $m = 6$. The categorization of the dichromat agent is such that

Number of rounds	The winning categories:	plurality Borda
10,000	2 6 2 1 4 4 5 4 3 4 6 3 4 1 1 1 5 5 6 4 2 3 1 3 4 3 6 2 2 3 1 1 1 5 5 2 3 3 1 6 3 1 4 6 1 2 4 3 3 1 5 6 2 1 4 4 5 4 3 1 6 3 4 4 1 1 5 5 6 4 2 3 1 3 4 3 6 2 2 3 2 1 1 6 5 2 3 3 2 6 1 4 4 6 1 4 4 1 5 1	
20,000	5 1 4 1 4 4 4 4 2 4 6 1 4 1 1 6 5 1 6 3 2 3 2 2 3 3 6 2 2 3 2 2 2 3 5 2 6 3 6 6 6 4 4 6 1 5 4 1 5 1 5 4 4 1 4 2 4 4 2 4 6 1 4 1 1 6 5 1 6 3 2 5 2 2 3 3 5 2 2 2 2 2 2 2 5 2 6 6 6 6 6 4 4 6 1 5 5 1 5 1	
30,000	5 5 4 4 4 4 4 4 4 4 6 1 4 1 1 1 5 1 1 3 2 3 2 2 3 3 2 2 2 2 2 2 2 2 2 2 6 6 6 6 6 4 6 6 6 5 4 1 5 1 5 1 4 4 4 4 4 4 4 4 6 1 1 1 1 1 5 1 1 3 2 3 2 3 2 3 3 2 2 3 2 2 2 2 2 2 6 6 6 6 6 4 6 3 1 5 5 1 5 1	
40,000	5 5 4 4 4 4 4 4 4 4 1 1 4 1 1 1 1 1 1 3 3 3 2 2 3 3 2 2 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 1 5 4 5 5 4 4 4 4 4 4 4 4 1 1 4 1 1 1 5 1 1 3 3 3 3 2 3 3 3 3 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 1 5 5	
50,000	5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5 5 5 5 4 4 4 4 4 4 4 4 1 4 1 1 1 1 1 1 1 3 3 3 2 3 3 3 2 2 2 2 2 2 2 2 2 6 1 6 6 6 6 6 6 6 5 5 5 4 5	
60,000	5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5 5 5 4 5 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5	
70,000	5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5 5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 2 3 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5	
80,000	5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5 5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 3 2 3 2 2 2 2 2 6 6 6 6 6 6 6 6 5 5 5 5 5 5	
90,000	5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5 5 5 4 5 4 4 4 4 1 4 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 5 5 5 5 5 5	
100,000	5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5 5 5 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3 2 2 2 2 2 2 2 2 6 6 6 6 6 6 6 6 6 5 5 5 5 5	

Fig. 5. The winning color category for chips 1–50 along the horizontal axis, according to the plurality (top numerical sequence of each pair) and the Borda score (bottom numerical sequence of each pair), for different numbers of rounds. The parameters are the same as in Fig. 4, with the exception that $N = 10$ and $N_d = 1$.
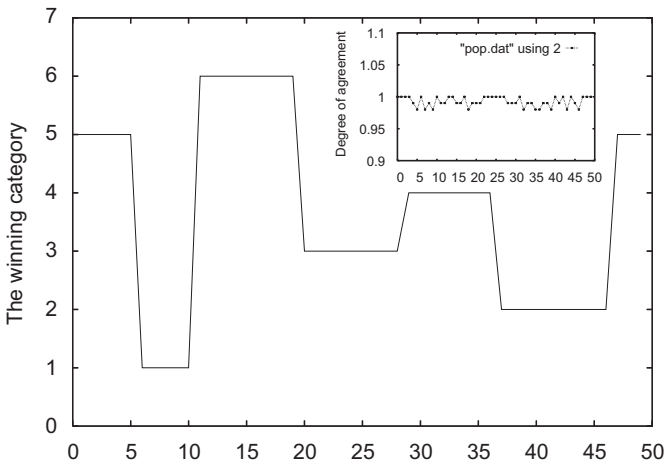

Fig. 6. A typical outcome of a heterogeneous population simulation, which is the winning color category for each of the $n = 50$ color chips. The inset shows the degree of agreement in the population. The parameters are the same as in Fig. 4; with the exception that the number of rounds is 10^7.

chips 1–5 and chips 45–50 are categorized as "blue", and chips 22–29 are classified as "yellow". These two categories coincide with those of a normal observer, being opposite each other on the hue circle, and correspond, given the model implemented, to a nonambiguous region of the color stimulus. On the other hand, the categorization of chips 6–12 and 39–44 oscillate between categories "cyan" and "purple", and the categorization of chips 13–21 and 30–38 oscillates between categories "purple, "red" and "green". Such nondeterministic categorization of these chips is attributable to the red–green confusion region specified by the psychophysical function $\mathscr{F}$, Eq. (1).

The fact that our model allows for nondeterministic categorization solutions becomes very important when simulating color deficient agents in the population. The color categorization of chips inside the confusable region is necessarily probabilistic for our simulated dichromats (that is, it contains nontrivial prob-abilities for more than one color term for chips in that region), which is a realistic model for such a color deficient observer. Also, our model allows for disconnected color categories. For example, in Fig. 4(c) we can see that regions around chips 20 and 30 are almost deterministically classified as "green", whereas a region around chip 25 is "yellow". Such disconnected color regions are not observed for our simulated normal agents from homogeneous populations, and they are not part of the theoretically obtained optimal solution class for normal agents. However, such categorizations are realistic for dichromats, because of a different topology of the color deficient's stimulus space, where disconnected regions in the hue circle become identified with one another, forming collapsed shape resembling an "8" digit (rather than a circular contour resembling a zero).

The outcome of a given run depends on the proportion of dichromats present in the population. We have found that as long as the fraction of dichromats is less or about 20%, the resulting categorization (that is, the structure of the winning color terms) is close to the optimal one found for a homogeneous population of normal agents (see Fig. 6). If the number of dichromats is much greater than that then the population converges to an impover-ished categorization system where the "confusable" (for dichro-mats) colors correspond to the same color term.

Even though for small percentages of dichromats the resulting common categorization is close to optimal, population hetero-geneity still plays a subtle but important role in population color category solutions.[7] To illustrate this, we consider the boundaries of color category regions for each run. Category region boundaries are defined as any color chip location for which the winning category of the chip to the immediate right is different from that for the given chip. For example, in Fig. 6 the boundaries of color

---

[7] It is important to emphasize that in categorization studies, uniformly assuming the exact same standard observer model for all individuals, whether in a simulated agent population or a actual observer population, will have regularizing and constraining influences on categorization results that may distort the agreement that would be found if normal variation was assumed. Thus, examining even subtle population heterogeneity using simulation can give insights into the regularizing influences that may occur when modeling data or simulation results.

terms are at 5, 10, 19, 28, 36 and 46. Given this definition, we can now examine how color boundaries are influenced by population heterogeneity.

4.2. Symmetry breaking

Let us first perform a number of runs with zero dichromats in the population (a homogeneous normal population). For each run, a sufficient number of rounds of the game is played such that the population converges to a shared, near-optimal categorization. We record the positions of the color boundaries for each run. It turns out that for the parameters chosen (see the legend of Fig. 7), the population always converges to one of three categorization types, containing either four, five or six categories. Homogeneous normal population results for 500 runs are presented in Fig. 7(a) as three histograms, each corresponding to one of the three categorization types. The histograms show that the frequencies of color boundaries are unsystematically, more or less uniformly, distributed among the chips. In other words, even though for every random simulation, the evolutionary process leads to an optimal categorization (with color categories of approximately equal length), the actual boundaries of the categories in homogeneous normal populations can be located anywhere around the circle, reflecting a rotational invariance (or rotational symmetry) feature previously noted by Komarova et al. (2007a, b).

Next, we include one, two,..., ten dichromats in the total population of $N = 100$ agents, $N_d = 1, 2, 3, 5, 10$, Figs. 7(b)–(f). Like in Fig. 4, in this experiment we will use the most extreme case of dichromats, which is characterized by the largest length of confusable region, $l = n/2 - 1$, see transformation (2) and Fig. 2(c). We can see from the graphs in Fig. 7 that the presence of dichromats removes the rotational invariance from the situation. Now, the boundaries of color categories are not uniformly distributed. Even 1% of dichromats in the population makes a clear difference in color boundary distributions, especially for outcomes with four categories, Fig. 7(b), the left column. And, generally, the larger the percentage of dichromats, the stronger their influence. For example, in Fig. 7(e), with 5% of dichromats, we can see that the distribution of boundaries in the presence of five categories has four clear bumps. Similarly, as the percentage of dichromats grows, 6-category outcomes also acquire well-defined boundary preferences.

The graphs in Fig. 7 suggest that the presence of dichromats leads to the phenomenon of symmetry breaking. In a homogeneous population of normals, category boundaries are placed with equal probability anywhere along the circle (note that the size of categories, i.e., the relative positions of boundaries with respect to one another, remain roughly the same). This is seen from the empirically obtained, uniform (modulo the noise) distribution of the boundary positions in Fig. 7(a). This distribution becomes very nonuniform as the percentage of dichromats increases. In Fig. 7(f) for 4-term categories, the boundaries occur more or less at only four fixed locations. The other figures also suggest that the presence of dichromats leads to established, preferred, boundary locations. We say that the heterogeneous populations' solutions are less symmetrical than those by homogeneous populations. This is because an arbitrary rotation of any categorization from a homogeneous population of normals would yield another, equally likely and equally "optimal" solution, while an arbitrary rotation of a typical categorization in Fig. 7(f) will lead to a categorization which happens with a very low probability (unless the angle of rotation is carefully chosen, e.g., equals $n\pi/2$ with an integer n in 7(f), 4-category solutions), and is suboptimal, i.e., it has a lower success rate.

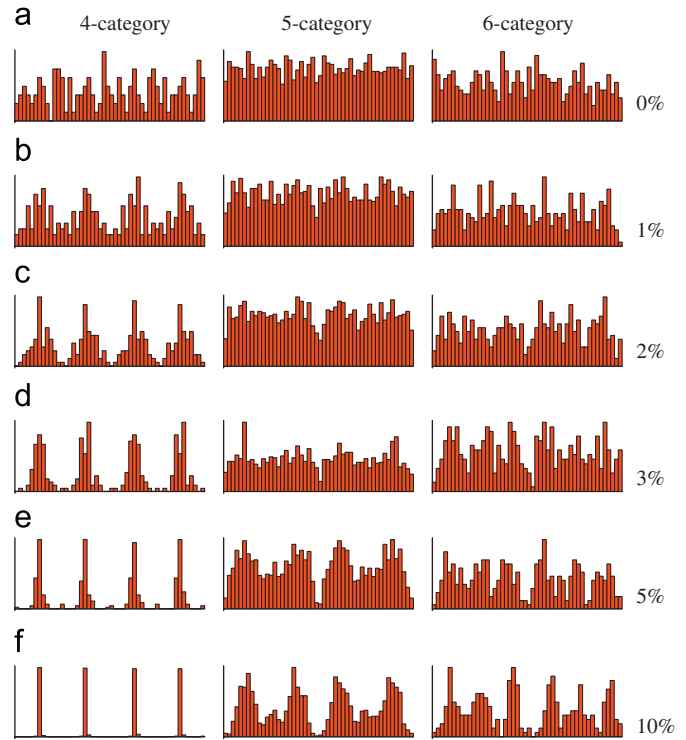The phenomenon observed here can be characterized as *explicit symmetry breaking* (as opposed to spontaneous symmetry

Fig. 7. The role of dichromats in population color categorization. Each row (a)–(f) represents a fixed percentage of dichromats, from 0% to 10% (the fractions of dichromats are marked on the right-hand side of the figure). Each histogram represents the positions of color categories for 4, 5- and 6-category solutions (columns 1–3, respectively). The horizontal axes of the histograms are the 50 hue circle chips, and the vertical axes shows how frequently a color boundary was established at a given chip. The vertical axes are all scaled to one to represent relative frequencies. The parameters are: $n = 50$, $m = 6$, $N = 100$, $k_{sim} = 6$, $L = 12$, the number of rounds is 10^7 for each run; the number of runs performed for each figure (a)–(f) is 500.

breaking observed often in pattern formation studies). This is because the "laws" underlying the system's behavior are no longer symmetrical for heterogeneous populations. While all points along the color circle are in some sense equivalent for normal agents, dichromat agents have regions of confusion, or ambiguity, in addition to non-ambiguous regions in the stimulus space. These both contribute to the regularization of boundary locations in a heterogeneous population consisting of normal and dichromat agents. This is discussed further in the next section where we look at the exact patterns of symmetry breaking.

Before we analyze the solutions found by heterogeneous populations in detail, we briefly report on some results pertaining to the statistics of mixed population categorization. In Fig. 7, the frequencies (see the vertical axes of each histogram) were scaled to one. In fact, the relative frequencies of 4-, 5- and 6-category outcomes change as the percentage of dichromats changes, see Fig. 8. While the number of 6-term categorizations decreases, the number of 4-term categorizations more than doubles, as the fraction of dichromats increases from 0 to 10%. This explains why the data in the rightmost column of Fig. 7 is very unsystematic. This also demonstrates that another effect of dichromats in a population is a changed likelihood for various types of categorization.

4.3. Categorization solutions from heterogeneous populations

Throughout this article, "solution" refers to an end-state color categorization scheme established by a population simulation. Such solutions may be of different "types". For instance, they may contain a different number of color categories on the hue circle.

Or, they may have the same number of categories but a different arrangement of "long" and "short" color categories along the hue circle.

In the previous section we saw that the statistics of categorization solutions is affected by the presence of dichromats in the population. To explain the observed nonrandom distribution of color boundaries in mixed populations, we note that it is transformation (1) (or (2)) that breaks the rotational symmetry. This transformation identifies what we call an *axis of no confusion*, or ANC. This axis crosses the color stimulus space at $i = 1$ and $i = n/2 + 1$ (note that these are the centers of the two connected invariant sets in the stimulus space).
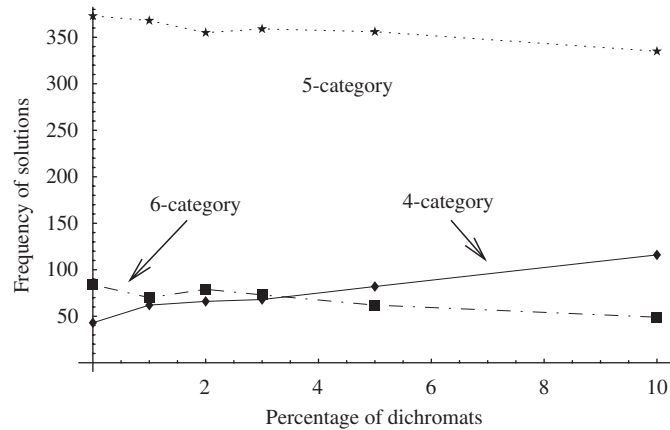
Fig. 8. The number of 4-, 5- and 6-term categorizations (seen across 500 runs) depending on the percentage of dichromats in the population. The parameters are as in Fig. 7.

Here is how the position of the ANC regularizes the positions of category boundaries in a population of normals and dichromats with $l = n/2 - 1$. If the percentage of dichromats is sufficiently large, an emerging shared categorization will tend to contain two color categories centered around the ANC (or, more precisely, around the points where the axis crosses the circle). These two color categories (we could think of them as "yellow" and "blue" categories) contain color chips that the dichromats can identify in a nonambiguous manner. These color categories are identically used by the dichromat and normal agents in the same population, see the "blue" and the "yellow" color categories in Figs. 4(b) and (c). Other categories (unless they are contained entirely within the invariant set of transformation (2)) will be different for the dichromats and the normals, Figs. 4(b) and (c).

In what follows, we present a case-study of color boundaries for a population of $N = 10$ individuals containing $N_d = 1$ dichromats with $l = n/2 - 1$, $n = 50$, $L = 12$, $m = 6$ and $k_{sim} = 6$, such that $m_{opt} \approx 5.5$. It turns out after 10^7 runs, a mixed population always comes up with one of four different solution types, as shown in Figs. 9 and 10.

Two of the solutions contain an even (4 or 6) number of equally sized categories, Figs. 9(a) and (b) and 10. In both cases, the solutions are oriented in such a way that there are two categories centered around the ANC.

Interesting results are observed when the optimal number of categories in a heterogeneous population is an odd number. There is no obvious way for the population solution to arrange the categories such that two opposite categories correspond to the chips that the dichromat can classify without confusion. In this case, the population comes up with two different solution types. In one group of solutions, there are five non-equal categories, such that the categories on the left of the input line (Fig. 9(c)) are
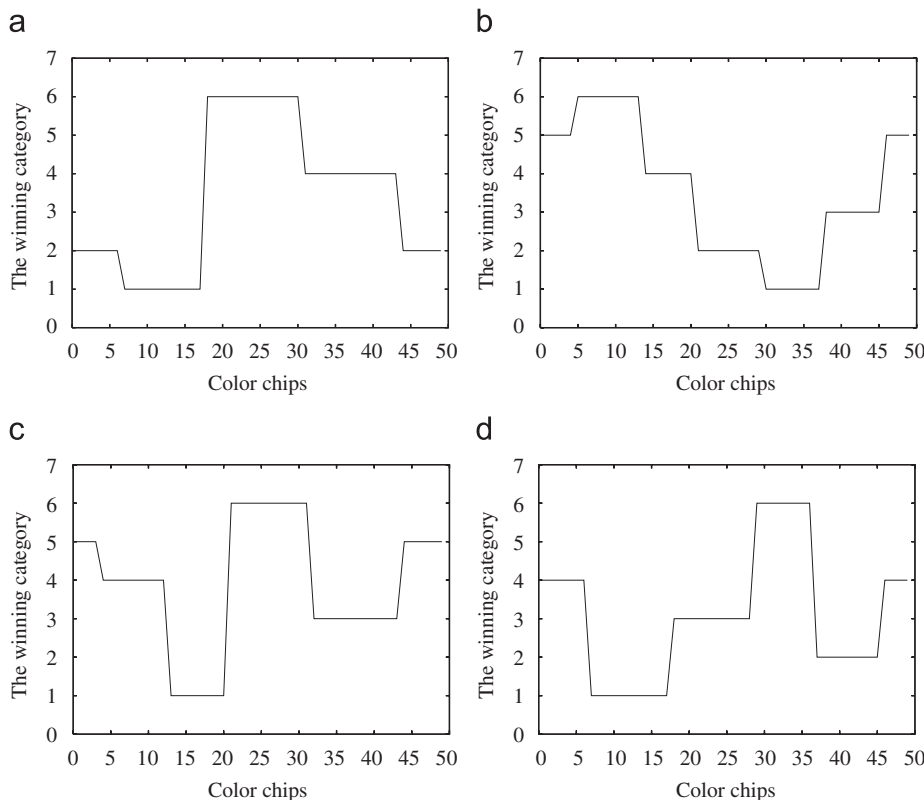

Fig. 9. Simulation results for a heterogeneous population (normal and dichromat agents), and a homogeneous color utility. The four types of solutions are: (a) the 4-category solution, (b) the 6-category solution, (c) the 5-category "short–long" solution, (d) the 5-category "long–short" solution. The parameters are $m = 6$, $L = 12$, $k_{sim} = 6$, $N = 10$, $N_d = 1$, $l = n/2 - 1$, $n = 50$, the number of rounds is 10^7.
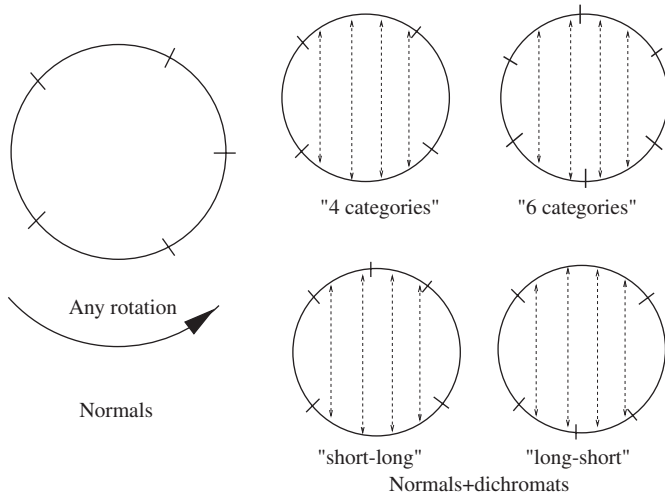
Fig. 10. Heterogeneous population (normal and dichromat agents), and a homogeneous color utility. The four types of solutions, a schematic. The parameters are as in Fig. 9.

shorter than those on the right.[8] The second group of solution is similar, except the shorter categories are on the right and the longer ones are on the left (Fig. 9(d)). We can see that in both of these solutions, there are two color categories centered around the points where the horizontal ANC crosses the color circle (the axis is not shown in the figure). In both cases this is achieved by shortening two of the five categories.

4.4. Varying the degree of agents' color perception impairment

So far we have only examined the extreme case of dichromats, namely, those with transformation (2) which collapses a circular stimulus space onto a line, with only two (diametrically opposite) nonconfusable color chips, see Fig. 2(c). Next let us explore how the degree of color deficiency, modeled by the parameter l, the length of the confusable region, affects the population categorization solutions.

In Fig. 11 we demonstrate how different values of l change the results, by running simulations with (a) $l = 2$ (a transformation similar to that in Fig. 2(a)), (b) $l = 8$ (Fig. 2(b)) and (c) $l = 18$ (close to Fig. 2(c)). For simplicity, only 4-category solutions are presented in Fig. 11. In Fig. 11(c), just like in the extreme case of dichromats studies in the previous section, the ANC plays an important symmetry breaking role. The psychophysical transformation (2) with large values of l splits the stimulus space into a wide "ambiguous" part and two short "nonambiguous" regions around $i = 1$ and $i = n/2 + 1$. This is where the ANC crosses the color circle, see Fig. 11(c), the dashed horizontal line at the bottom of the figure. These special points tend to be placed in the middle of color categories, which defines the observed pattern in Fig. 11(c).

On the other hand, Fig. 11(a) shows a different type of solution. There, transformation (2) defines a large non-ambiguous set (most of the circle) and two short ambiguous regions, around chips 13 and 38, or the "12 o'clock" and the "6 o'clock" in the diagram at the bottom of Fig. 11(a). Thus, instead of the ANC, we can talk about the "axis of confusion" (AC), which is denoted by a dashed vertical line. The points where the AC crosses the circle tend to contain color boundaries, which defines the observed pattern in Fig. 11(a).

---

[8] Note that we are still considering a circular stimulus space, and the terms "left" and "right" are used simply for reference to the horizontal axes of Fig. 9.

The general tendency is that confusable regions repel boundaries, and non-confusable regions attract boundaries. As is seen from the examples of Figs. 11(a) and (c), these tendencies may be opposing. However, in certain cases one of them is stronger than the other. For example, a very narrow confusable region (Fig. 11(a)) anchors color boundaries to well-defined locations, while large nonconfusable regions cannot impose much of a constraint onto category centers. Similarly, a very narrow nonconfusable region (Fig. 11 (c)) pins down color category centers, while large confusable regions cannot anchor boundaries to a certain location. Fig. 11(b) shows a case when both countervailing tendencies seem equally strong. In this case, we observe a lot of irregularity in the system, and also the boundary location distribution has a characteristic shape of four fuzzy double-humps.

In the next Section we will illustrate our discussion with strongly impaired dichromats, where l is not too far from $l = n/2 - 1$. By using similar methods, the results can be extended to small values of l.

To summarize our findings for heterogeneous populations, we note that the presence of dichromat observers leads to certain changes in the shared population categorization. In particular (a) it may lead to category number changes, (b) it may affect the uniformity of category sizes and (c) it influences category statistics, resulting in category boundaries acquiring preferred positions, and different solution types changing their likelihood.

Finally, although here we focus on the learning and evolution of color categories among artificial agents—and do not investigate human categorization phenomena—these results have implications for human color naming research. The present results support those shown by Komarova et al. (2007a) by showing through mathematical analyses and simulations that the simplest forms of discrimination and communications are sufficient for the evolution of color naming systems using simple learning algorithms.

5. Heterogeneous color utility

Besides investigating categorization solutions under idealized variations of observer populations, we also investigate idealized forms of color heterogeneity. To explain the notion of heterogeneous color utility, we envisage a situation where differences in shades of some reddish fruits convey a large qualitative difference for the observers (e.g., some shades of red may signal "edible" while others signal "poisonous"), whereas any other region of the stimulus space does not carry a similar qualitative difference. In this case, categorizing the reddish region of the stimulus space into different classes will pay off, or has a different color utility. At the same time, distinctions of purplish shades are not important, or salient, for any pragmatic reason, and thus all exemplars in that region can be considered as one category. We approximate such environmental color distribution variation here by implementing biases, or different levels of importance, in our simulations for a portion of the hue circle. Through such implementations it is possible to investigate the ways population heterogeneity (described above) interacts with biases that may exist in the way color stimulus categories are marked and differentially processed by our simulated observers.

5.1. Regions of increased salience in the color stimuli

We investigate the effect of color space biases on categorization solutions by including RIS in the stimulus space. A RIS is simply a differentiated stimulus region, which idealizes realistic circumstances where an especially salient portion of the color
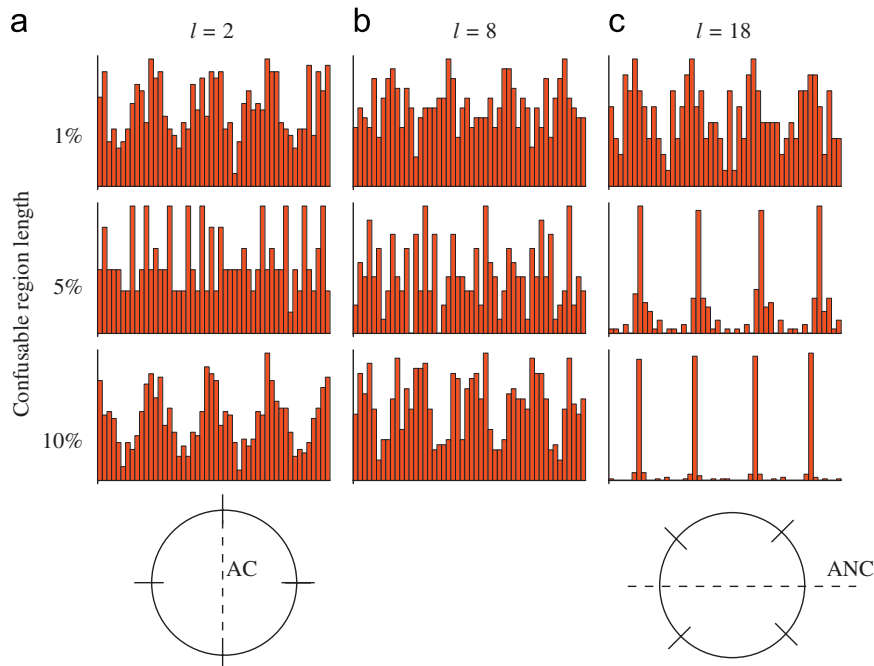
Fig. 11. Different types of dichromats and their role in population color categorization. Each row represents a fixed percentage of dichromats, from 1% to 10% (the fractions of dichromats are marked on the left-hand side of the figure). Each histogram represents the positions of color categories for 4-term categorizations only, for different parameters in transformation (2): (a) $l = 2$, (b) $l = 8$ and (c) $l = 18$. The diagrams below show the position of the AC for the least impaired dichromats and the position of the ANC for the most impaired dichromats. The axes and simulation parameters are the same as in Fig. 7.

space might demand a different pragmatic classification scheme. The pragmatic notion of k_{sim}, the similarity range defined above, provides a way to introduce color stimulus heterogeneities in this framework. Namely, instead of using one universal (for all color chips) measure of similarity, k_{sim}, we can simulate a world where k_{sim} is different in different regions of the stimulus space.

In general, the quantity k_{sim} may vary from chip to chip, such that we have a vector with n components, $k_{sim}^{(i)}$, $1 \leqslant i \leqslant n$. In particular, a RIS can be created by varying the parameter k_{sim} in a specific region of the hue circle, while a different value of k_{sim} is uniformly applied to stimuli in the rest of the hue circle:

$$k_{sim}^{(i)} = \begin{cases} k_{RIS}, & i_L \leqslant i \leqslant i_R, \\ k_{norm}, & \text{otherwise}, \end{cases}$$

that is, in a certain region of the color space (between chips i_L and i_R), we have $k_{sim} = k_{RIS}$ and everywhere else we have $k_{sim} = k_{norm}$, where

$$k_{RIS} < k_{norm}.$$

The region $i_L \leqslant i \leqslant i_R$ is the RIS of the color stimulus space.

In practice, when we sample the stimuli for use in discrimination-similarity games, as long as at least one of the chips sampled belongs to the RIS, we will use k_{RIS}. Otherwise we use k_{norm}. As as result, the optimal color categorization changes, and we find that the color categories around the RIS tend to be smaller than those outside the RIS.

5.2. The effect of the RIS on the color categorization of a homogeneous population

Simulation results show that in homogeneous populations, the presence of a RIS removes the rotational invariance of the solution. As a result, it may fix the boundaries of the color categories. In Fig. 12 we show a typical solution for a given parameter set where $n = 40$ (differing from $n = 50$ shown
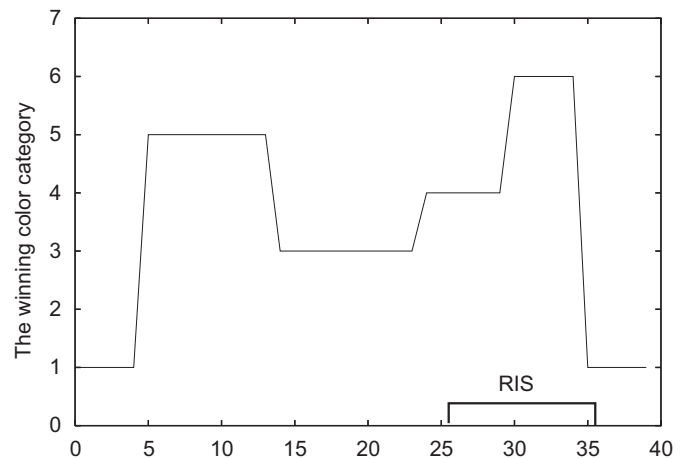

Fig. 12. A typical solution for a homogeneous (normal) population and a RIS between $i_L = 26$ and $i_R = 36$ and parameters $n = 40, N = 10, L = 12, k_{nor} = 6, k_{RIS} = 3$ and 10^7 runs.

earlier). Most of the time (about 70% of runs), the population comes up with a categorization which contains $[n - (i_R - i_L)]/\sqrt{2k_{norm}(k_{norm} + 1)} \approx 3$ color categories outside the RIS and $(i_R - i_L)/\sqrt{2k_{RIS}(k_{RIS} + 1)} \approx 2$ categories inside the RIS; the categories inside the RIS contain fewer color chips compared to nonRIS categories. In a smaller (but significant, about 25%) fraction of runs, we see solutions with three very short color categories squeezed inside the RIS. Finally, there are rare occasions where there are only two color categories outside the RIS and three inside; four categories outside and two inside, and so on. The exact configuration and the likelihood of solutions vary depending on the parameters. For instance, if we take the same parameters as before except $n = 42$ (a slightly larger stimulus space), the frequency of the second (6-category)
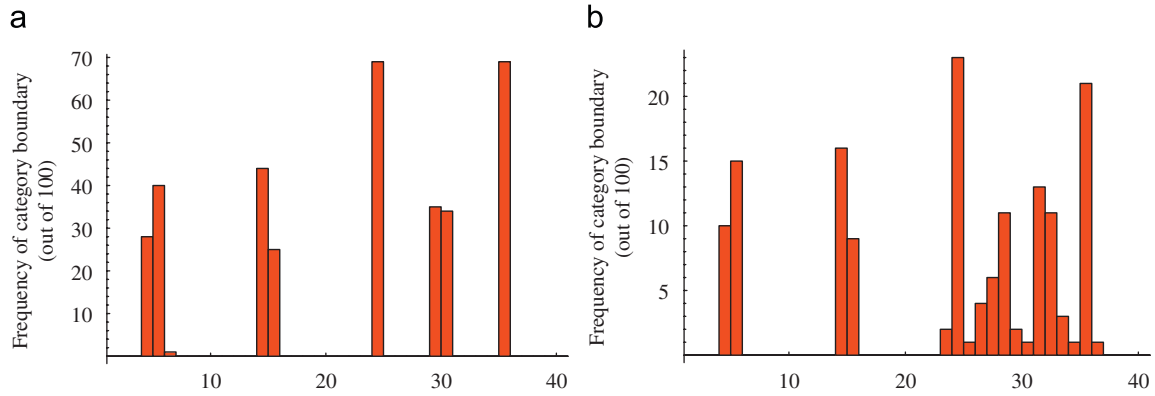
Fig. 13. Histograms showing the positions of category boundaries with a homogeneous population, in the presence of a RIS. We performed 100 independent runs. Parameter values are the same as in Fig. 12. (a) 5-category solutions and (b) 6-category solutions.
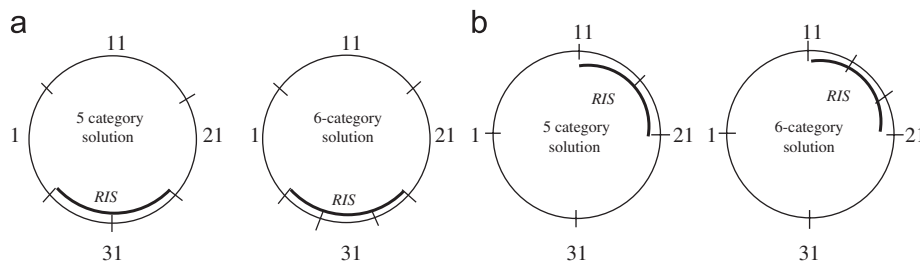

Fig. 14. A schematic showing the 5-category and the 6-category solutions of a homogeneous population in the presence of a RIS. (a) The RIS is between chips 26 and 36. (b) The RIS is between chips 11 and 21. Other parameters are as in Fig. 12.

solution increases and the first (5-category) decreases, such that one becomes almost as frequent as the other.

The results of Fig. 12 are also presented in Fig. 13 where we show histograms of positions of category boundaries for the more common 5- and 6-category solutions described above (see (a) and (b), respectively). We can see that the position of the RIS defines the position of color boundaries very precisely, with only a small amount of "spread". For example, out of 69 runs which resulted in a 5-color categorization, Figs. 12 and 13(a), 100% had color boundaries at positions 25 and 36. The position of the color boundary around the middle of the RIS which separates the two short color categories, varies between 30 and 31. This is because the number of chips inside the RIS is odd (11), so there are two ways to split them into two nearly equal parts (5 and 6, or 6 and 5).

The main conclusion that we draw from these experiments is that the presence of a RIS removes rotational invariance that would otherwise be seen in the system. The resulting solution(s) will have category boundaries dictated by the position of the RIS, and the number of color categories inside and outside of the RIS is determined by values of k_{sim}: in particular, k_{RIS} inside the RIS and k_{norm} everywhere else on the hue circle. The two main types of solutions for the particular parameter set case-studied here are given in Fig. 14. The positions of the smaller categories are completely defined by the position of the RIS.

6. The two types of heterogeneities combined

In this section we examine situations where the two types of heterogeneities described above are combined. In other words, we consider the stimulus space heterogeneity modeled with a RIS and study the color categorization of populations in the presence of a fraction of dichromats.
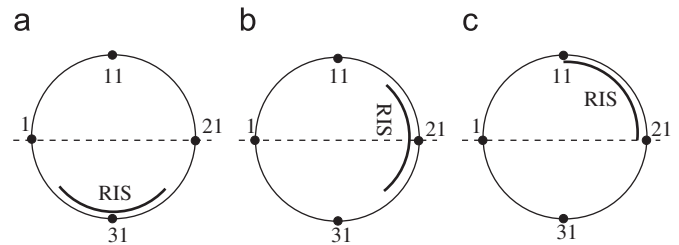

Fig. 15. RIS and dichromats. The three configurations (a–c), in terms of the position of the RIS with respect to the ANC. The ANC is marked by a dashed horizontal line; the confusable regions are centered around chips 11 and 31.

Generally, both the presence of a RIS and the presence of dichromats break the rotational symmetry of the system, but they do it in a different way. A RIS fixes a "special" region of the circular stimulus space using a lower value of k_{sim}, whereas dichromat agents in a population provide an anchor in the form of a special color direction, such as the ANC or/and the AC. Depending on the position of these axes with respect to the RIS, the results of the interaction of the RIS with population heterogeneities will be different.

Let us consider the example of the previous section, Figs. 12–14. For these parameter values, a homogeneous population of normals categorizes a color space with a RIS into five or six regions. The 2- or 3- color categories around the RIS are about two or three times smaller than the other three categories.

We will now add dichromats to the population and consider the following three qualitatively different cases:

(a) The ANC lies outside the RIS, Fig. 15(a).
(b) The ANC crosses the RIS in the center, Fig. 15(b).
(c) The ANC is adjacent to the RIS, Fig. 15(c).

In the simulations that follow we will keep the ANC fixed (it is defined by chips 1 and $n/2 + 1 = 21$), whereas we will vary the location of the RIS. To illustrate the three qualitatively different cases given above we define RIS that span three different regions of chips: namely, (a) 26–36, (b) 16–26 and (c) 11–21.

6.1. The RIS inside the confusion region

We first consider case (a) above where the RIS coincides with an idealized region of dichromat confusion (see Fig. 15(a)). In this case, the solutions of a heterogeneous population, comprised of both normal and dichromat agents, are identical to solutions seen from a homogeneous population of the normal agents. We have performed simulations with varying length of the confusable region for the dichromats, l. In our development here we kept l relatively large, corresponding to a large perceptual impairment of dichromats; in this case the ANC imposes a dominant constraint on the resulting categorization. Small values of l (low degrees of impairment) are characterized by the dominance of the AC and can be studied by similar methods. In all experiments with $11 \leqslant l \leqslant 19$, the heterogeneous population comes up with a 5-category solution about 70% of the time, and with a 6-category solution most of the rest of the time. These solutions are similar to those of Figs. 13 and 14(a); the shorter categories are located over the RIS.

In this example the presence of dichromats does not change the solution from that seen in the homogeneous normal simulation where the solution was made rotationally invariant by a RIS in this particular location. In other words, in situations where a RIS coincides with a confusion region relevant to a small proportion of the population, the solution that stabilizes is the one that optimizes the category solution for the unimpaired agents, as if the small contingent of impaired dichromats was not involved in the discrimination-similarity communication games.

6.2. The ANC crosses the RIS in the center

Next, consider a variation on the example just described, where a RIS is centered on the ANC, or Fig. 15(b). The solution found by a homogeneous population of normal agents is similar to that shown in Figs. 12–14, except that the categories are rotated with the small color categories shifted to this new RIS position (i.e., chip range 16–26).

Adding dichromat agents to the population changes the solution. In Fig. 16 we present histograms of category boundary locations for solutions found by such heterogeneous populations. To examine the interaction between these features we investigated three cases of dichromats with confusable regions of lengths $l = 11$, $l = 15$ and $l = n/2 - 1 = 19$, see Fig. 17. The results are presented in Fig. 16, top to bottom. We can see that most of the time, heterogeneous populations come up with two types of solutions, those with 5- and 6-color categories (similar to what was found for populations with only normal agents). In 6–7% of runs we also found 7-category solutions (not shown). First note that the shape of the 5-category solutions also changes as l grows. For the $l = 11$, the solutions are not distinguishable from those found by normal populations. However, as l increases, the color boundary near the center of the RIS gets pushed either to the left or to the right of the center. As a result, no category boundaries are placed in the immediate proximity of the ANC. This effect becomes more noticeable as the length of the confusable region, l, increases.

Another trend that we observed is that the fraction of 5-category solutions becomes smaller as l grows: recall that for homogeneous populations this fraction was 70%, and it drops down to only 12% for simulations that include the most impaired dichromat model, $l = 19$. Accordingly, the fraction of 6-category solutions grows from 30% to 82%.

These observations can be explained. The 5-category solution found by homogeneous populations has a color boundary at or near the ANC. Therefore, the frequency of this type of solution decreases as a result of the presence of dichromats, and also the middle category boundary inside the RIS becomes shifted away from the ANC. The 6-category solution has no conflict with the ANC and thus becomes more frequent in the presence of dichromat agents.

The exact location of the ANC with respect to the RIS is important. Fig. 18 presents simulations identical to those of the bottom row of Fig. 16, except the RIS is shifted by 1 position to the left, such that the ANC is not exactly in the center. Interestingly in this case the symmetry of solutions is broken. Now, the 5-category solution always has the boundary inside the RIS shifted to the left (and never to the right). The 6-category solution is also very asymmetrical, with the rightmost of the three categories inside the RIS being much smaller than the other two.

6.3. The ANC is adjacent to the RIS

Finally, consider a RIS positioned immediately adjacent to the ANC, Fig. 15(c). This configuration results in the largest degree of conflict between the basic homogeneous RIS solutions (Fig. 14(b)) and the position of the ANC. In our first set of experiments, we considered $N_d = 1$ dichromats in a population of $N = 10$ with $l = 15$. Later, we increased l to $l = 19$, the maximum value considered here.

The results for the $l = 15$ case are presented in Fig. 19. We can see that in this case, the heterogeneous population comes up with five different solutions. Two solutions are variants of the original 5-category solution (Fig. 19, top row), and three are variants of the 6-category solution (Fig. 19, bottom row). The more frequent 5-category solution is identical to that found by the homogeneous populations, Fig. 14(b). It has category boundaries exactly at both points where the ANC crosses the circle. The second 5-category solution is different in that it avoids placing category boundaries at the ANC. There are four instead of three shorter categories around the RIS, and only two categories in the remaining region.

The three 6-category solutions are presented in the bottom row of Fig. 19. Again, the most frequent one coincides exactly with the one found by homogeneous populations (see Fig. 14(b)), and the other two are its variants avoiding the overlapping of category boundaries with the ANC. The one encountered in 8% of the runs shortens one of the larger categories outside the RIS region to move a boundary away from the ANC at chip $i = 1$. The 3% solution contains four short color categories around the RIS and only two outside the RIS.

Now we increase the confusion length to $l = 19$. The results are presented in Fig. 20. The original, unmodified 5- and 6-category solutions do not arise anymore. Instead, the heterogeneous population comes up with their variants, all of which push the color boundaries away from the ANC.

These findings are summarized in schematic Fig. 21. There we depict the transformations of the basic 5- and 6-category solutions by means of changing the category number, by rotation and by changing category size. The solutions from left to right range from more common to less common.

Qualitatively the three kinds of transformations we observed in the solutions presented above (i.e., (i) adding or removing categories, (ii) rotating a solution and (iii) compressing or expanding category size) all work in concert to achieve a common goal: namely, to form a solution that maximizes communication
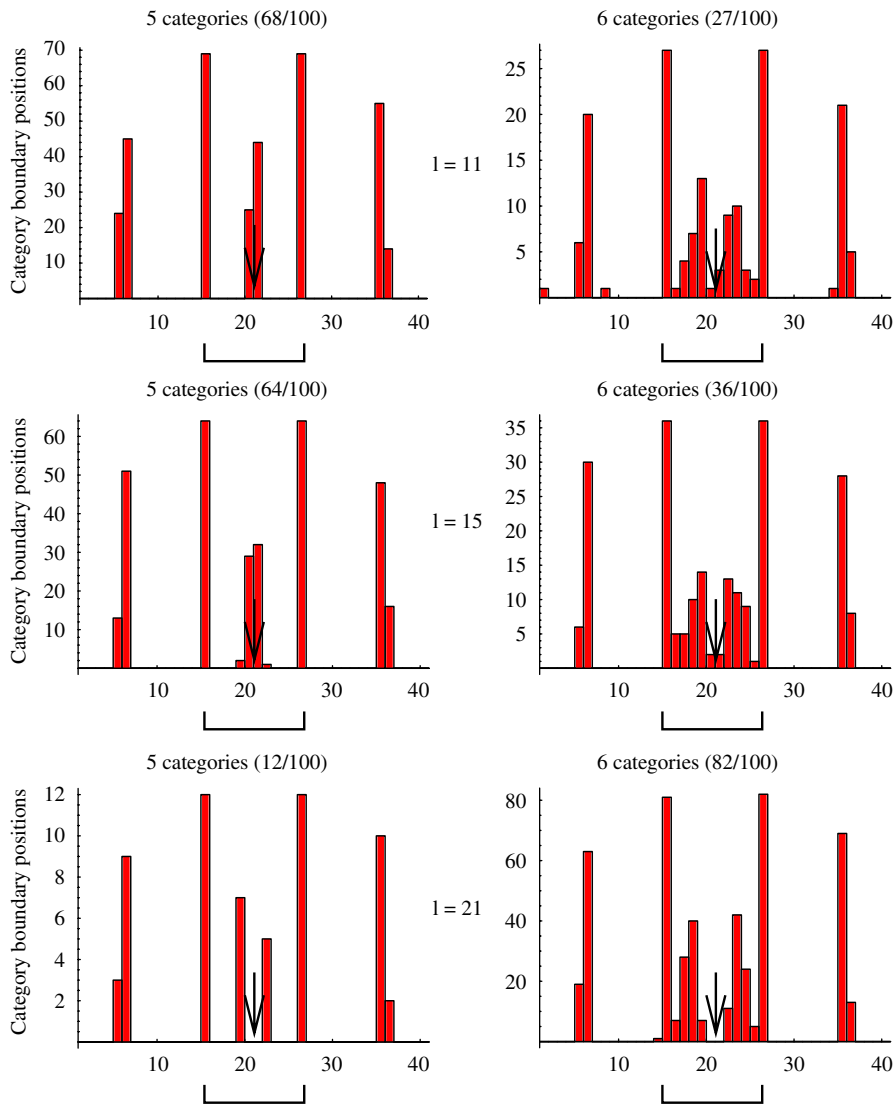
Fig. 16. RIS and dichromats, the ANC crosses the RIS in the center. Shown are histograms of color category boundary locations, out of 100 independent runs. The position of the RIS is indicated; it is between $i_L = 16$ and $i_R = 26$. The point in the RIS region that is coincident with the ANC is marked by an arrow (chip 21, exactly in the middle of the RIS). There are $N_d = 1$ dichromats out of $N = 10$ agents. Upper row: the confusable region of dichromats has length $l = 11$. The middle row: $l = 15$. The bottom row: $l = n/2 - 1 = 19$. Both categorizations into five and six categories are shown, and percentages of each are indicated above the graphs. The 7-category solution was found in 7 out of 100 runs for $l = 19$ and 6/100 for $l = 11$ (not shown). Parameters are $n = 40$, $k_{RIS} = 3$, $k_{nor} = 6$, 10^7 rounds are played for each run.
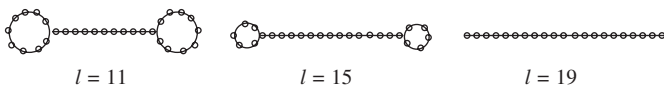

Fig. 17. A schematic representing the perceptual space of dichromats with varying lengths of the confusable region: $l = 11$, $l = 15$ and $l = 19$.

game success rates in the total population relative to a reliable anchor of nonconfusable stimuli for a portion of the population, while simultaneously taking advantage of a potential for richer partitioning (i.e., the k_{RIS} difference) for a majority segment of the population.

In general, the ways the present RIS conditions are shown to trade off with the population heterogeneity conditions suggests that, for the simulated agent populations investigated here, both factors contribute to how color category solutions take shape and stabilize. Whether these conditions similarly trade off in realistic

circumstances seems plausible, and is demonstrated as at least possible in view of the present results.

6.4. Can we predict the categorization?

The present investigations allow a formal description of the most likely structure of the optimal color categorization solution. The optimal "length" of a category is given by

$$l_{opt} = \sqrt{k_{sim}(k_{sim} + 1)}.$$

6.4.1. No RIS constraint in a homogeneous normal population

Under conditions in which there is no added constraint of a RIS, using a homogeneous normal population, the solution will consist of m nearly equal categories, where m is an integer approximation to the solution of
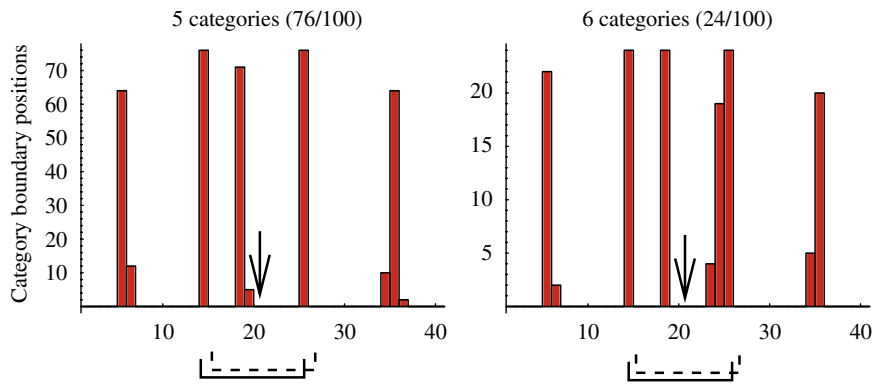
$$ml_{opt} = n.$$

Fig. 18. RIS and dichromats, the ANC crosses the RIS around its center. Shown are histograms of color category boundary locations, out of 100 independent runs. The confusable region has length $l = 19$. Other parameters are as in Fig. 16 except the RIS is shifted by 1 chip and is located between $i_L = 15$ and $i_R = 25$.
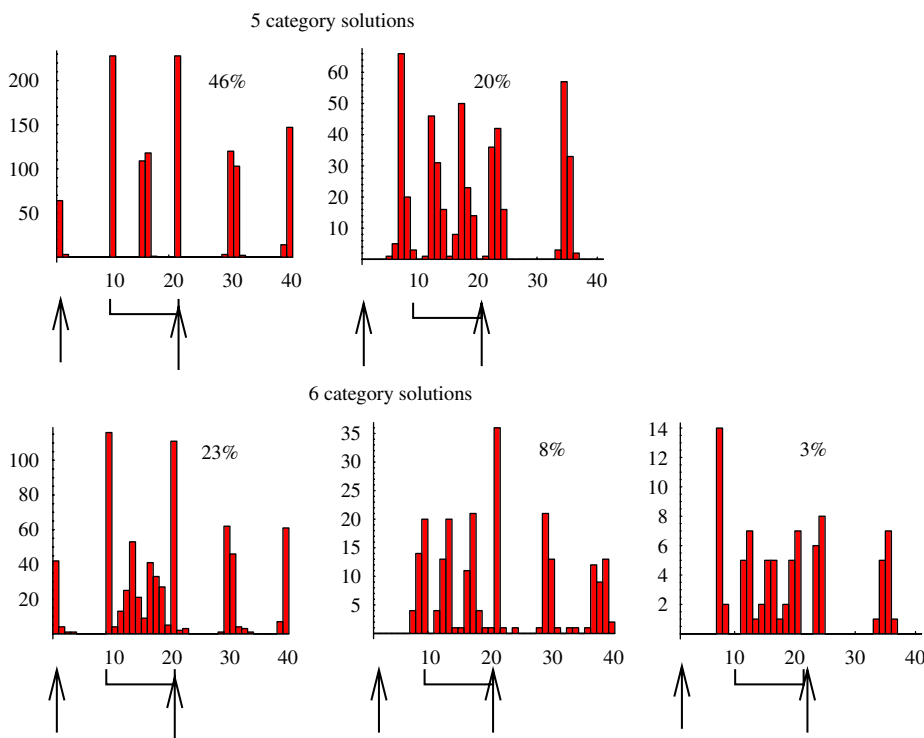

Fig. 19. Same as in Fig. 16 except the ANC is adjacent to the RIS (between $i_L = 11$ and $i_R = 21$), the confusable region has length $l = 15$ and frequencies are out of 500 independent runs.

If n/l_{opt} is not close to an integer, then the population will converge to one of two different solutions, one with the number of categories given by the largest integer $< n/l_{opt}$ and the other one with the number of categories given by the smallest integer $> n/l_{opt}$.

In all cases, the full solution class consists of all possible rotations of such solutions.

6.4.2. No RIS constraint in a heterogeneous population

Under conditions in which there is no added constraint of a RIS, if we consider both dichromat and normal agents in a population, the solution may change. In particular, if the optimal solution contains an even number of categories, and the dichromats are assumed to be close to the extreme case of $l = n/2 - 1$, the population will always converge to an optimal solution where the two least confusable colors are in the middle of two opposing

color categories. Therefore, the presence of dichromats in this case will not change the structure of the solution, but it will remove the rotational invariance, fixing (in a probabilistic sense) the boundaries of color categories.

If the optimal solution consists of an odd number of color categories, then the presence of dichromat agents will change the solution structure. For convenience let us assume that for the dichromats, l is close to $n/2 - 1$, and chips 1 and $n/2 + 1$ are the least confusable ones. With an odd optimal number of categories, m_{opt}, one of the following can happen:

- Unequal categories, same number of categories. In this solution, the optimal number of categories is maintained, but the boundaries are shifted. Namely, there is a category centered at chip 1 and a category centered at chip $n/2 + 1$, with $(m_{opt} - 3)/2$ categories and $(m_{opt} - 1)/2$ categories on
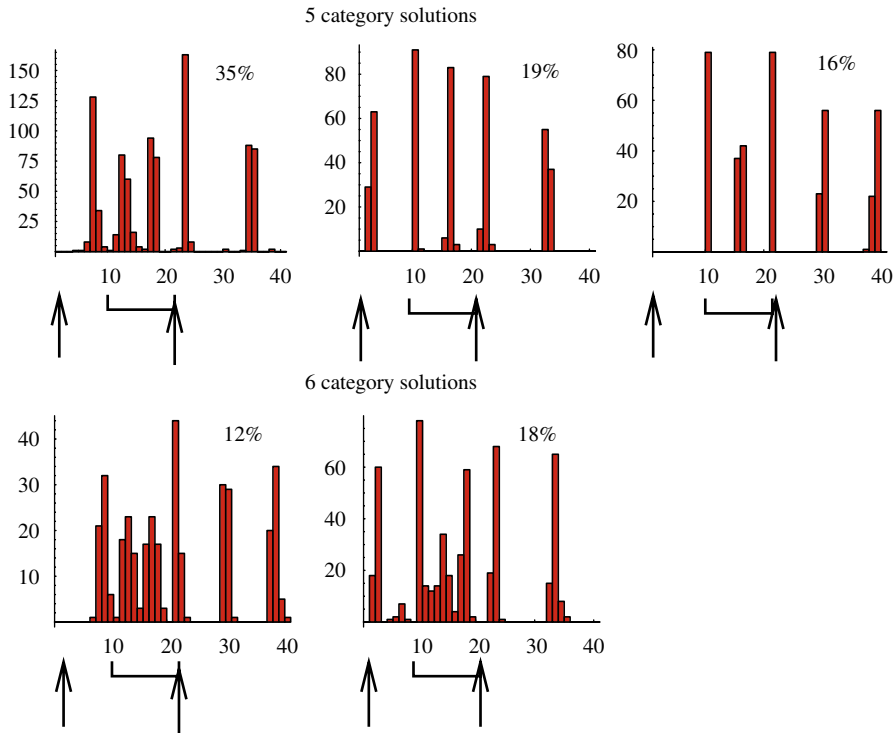
5 category solutions

6 category solutions

Fig. 20. Same as in Fig. 19, except the confusable region has length $l = 19$.
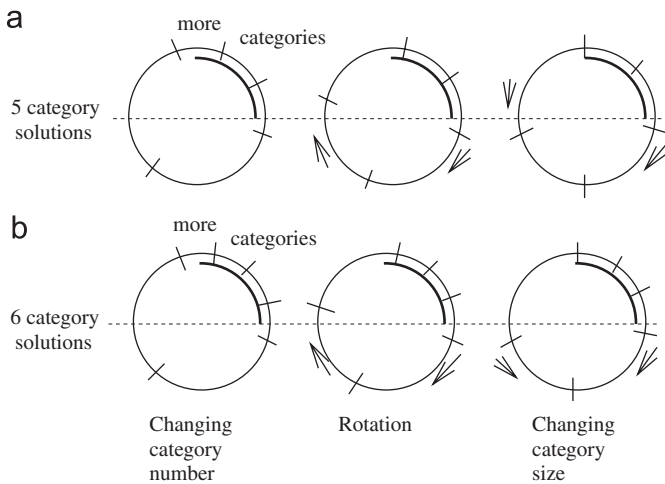

Fig. 21. The schematic show how the 5- and 6-category solutions get modified by the presence of dichromats, when the RIS is adjacent to the ANC.

each side. This way, the categories on one side of chips $n/2 + 1$ are on average shorter than those on the other side.

- Equal categories, suboptimal number of categories. This solution has an even number of color categories (smaller or larger). They are equally sized.

6.4.3. Presence of a RIS constraint in a homogeneous normal population

A RIS constraint (or different pragmatic category values of k_{sim}), in a homogeneous normal population has the following consequences for solution structures. There are categories of two sizes,

$$l_{opt}^{norm} = \sqrt{k_{norm}(k_{norm} + 1)}, \quad l_{opt}^{RIS} = \sqrt{k_{RIS}(k_{RIS} + 1)},$$

where $l_{opt}^{norm} > l_{opt}^{RIS}$. Suppose the length of the RIS is n_{RIS}. Then the system will have p long color categories and q short ones, the short ones located around the RIS. The integers p and q are found from the solutions of

$$p l_{opt}^{norm} = n - n_{RIS}, \quad q l_{opt}^{RIS} = n_{RIS}.$$

The solutions in such cases are not rotationally invariant. Often the solution is not unique: a couple of integers may be almost equally close to the solution of these equations.

6.4.4. Presence of a RIS constraint in a heterogeneous population

Finally, under conditions involving a RIS constraint, the presence of both dichromat and normal agents in a population may change the solution described just above. In particular, the presence of dichromats with a large perceptual impairment will force a relatively larger color category to be centered around chips 1 and $n/2 + 1$ (the nonambiguous chips). This will produce a tendency for the categories around the nonambiguous chips to shift away, or repel, from the ANC, making room for a relatively larger category centered around the ANC. As a result there can be a compensatory category rotation, with, in some cases, a change in category number and size.

7. Discussion

In this article we considered some very simple idealizations of situations where a signaling/categorization system evolves for objects from a continuous domain. We report findings from color categorization simulations, with a focus on two realistic features of categorization phenomena: population heterogeneity, and heterogeneity of color utility. The two are found to trade off in a dynamic fashion during the process of a population of agents forming and stabilizing an optimal color categorization solution. This solution aims to maximize successful communications

among agents about a circle of hue stimuli, in a context of the discrimination-similarity game.

These investigations reveal some features of the phenomena that are relevant to color categorization in general, and which may be further understood through formal analyses. Our framework creates a systematic approach to studying the dynamics between compensatory features (including trade offs that are possible and likely) in realistic situations where the demands on color categorization behaviors can be specified, regardless of how well their influences are understood.

Below we summarize our findings for the cases involving (I) populations of agents learning a shared category system under differences in population heterogeneity, (II) the introduction of a salient color region in the stimulus space and (III) the influences of combining population variation and variation in color utility on category solutions.

I. Populations of communicating agents and population heterogeneity.
   (a) Individual agents acquire color categorization systems in ways dependent on the psychophysical discrimination model they are endowed with (as well as the utility properties of the stimulus space, discussed below).
   (b) Communicating agents in a population can individually acquire a color categorization system that is shared with other agents in their society, and the properties of such shared categorization systems (e.g., where the category boundaries occur) depend (in part) on the discrimination model given to the agents.
   (c) The categorization solution achieved by a population for each region of the stimulus space is provided by the "winning color categories" based on population communication games. The shared, quasi-stable population solution depends (in part) on the composition of a population—that is, whether it is a society comprised of homogeneously modeled agents or whether the society includes some degree of nonrandom agent heterogeneity. When a population includes certain proportions of two different agent types (i.e., one type with a perceptual deficiency, the other modeled as normal observers), then the collective categorization solutions adjust to accommodate the constraints specific to agents from the population minority.
   (d) Solutions achieved by heterogeneous populations differ from those found by homogeneous populations. Namely, in some cases the presence of dichromat observers leads to changes in the number of categories achieved, and it can produce categories of nonuniform size. It always influences color category statistics, resulting in category boundaries acquiring (probabilistically) preferred positions. The latter effect can be described as a symmetry breaking phenomenon; in other words, the presence of color impaired observers breaks the rotational invariance found in the systems with homogeneous populations.
   (e) The exact nature of symmetry breaking depends on the form of psychophysical transformation used to model color vision impaired agents. In particular, in this study we employed a simple transformation which collapses two opposite regions of the stimulus space (the confusable regions) onto each other. For the case of large confusable regions (a high degree of impairment) the population solutions tend to center two opposing color categories around the (short) nonconfusable regions. For the case of short confusable regions (a mild impairment) the population solution tends to place category boundaries near the (short) confusable regions.

II. Heterogeneous color utility.
   (f) When color stimulus salience includes an inhomogeneity in the stimulus space—in the form of a region(s) of the space with different pragmatic utility compared to the rest of the space—then both individual learners and populations of learners acquire a color categorization without rotational invariance.
   (g) In this paper we concentrated on an explicit model, where one RIS was included in the stimulus space. Inside this region, the k_{sim} parameter (which correlates with the average size of color categories in the optimal solution) was assumed smaller than outside the RIS. Consequently, the population solutions did not possess a rotational invariance, and the category boundaries were (statistically) determined by the boundaries of the RIS. The category size and number were also affected by this constraint, placing shorter categories inside the RIS and larger categories outside.

And finally,

III. Both forms of heterogeneity.
   (h) When both forms of heterogeneity are combined in a population simulation, we observed several interesting trade offs between features of population heterogeneity and color utility heterogeneity. Briefly, these included trade offs between (i) the relative sizes of the population majority and minority, (ii) the relative degrees of color deficiency used to model population heterogeneity, (iii) the degree with which defined RIS areas coincided with varying regions of confusion and nonconfusion in the deficient agent models investigated.

All the results observed had the net effect of demonstrating that stable color category solutions are flexible and respond in predictable ways to countervailing demands that may be imposed by population and pragmatic constraints.

The remaining portion of the discussion presents several aspects that address the extension, or generalizability, of these findings, and relates them to noteworthy results and advances in the existing color categorization and game theory literatures.

7.1. Implications for color naming theory

Our results bear directly on an issue hotly debated over the last decade, and provide a strong counter example to the assertion that there are no nontrivial constraints on color categorization (e.g., Saunders and van Brakel, 1997). We found that by varying observer discrimination constraints, even at the idealized level needed for the present investigations, we could systematically (and in nontrivial ways) influence the stabilized color categorization systems observed. Also, introducing minimal color salience constraints into the investigations systematically changed the manner by which category partitions were sized and distributed. We also showed that such constraints interact, or trade off, in ways one would expect such constraints to trade off in realistic circumstances. The present study does not directly investigate the many potential realistic constraints that exist. However the two forms of constraints we did include in this study were shown to substantially affect the agent-based category systems, and, given the wide prevalence of such constraints in the real world and their potential relevance to the behavioral phenomena, we expect they may also serve as limiting conditions that influence the formation of human color categorization systems.

It is important to re-emphasize that the agent simulations we present are not intended to model human color category learning

or interactions between human categorizers. They are instead intended to demonstrate what can be achieved using only the most rudimentary forms of color observation and communication together with an elementary evolutionary dynamics. The evolutionary dynamics implemented here follow the ideas that (i) color naming should in part be based on pragmatic concerns, (ii) in general, perceptually similar colors are likely to be given the same name and (iii) perceptually different colors may be given different names. This kind of simulation has potential value for clarifying open issues in the empirical literature concerning the basis for human color naming. On the one hand, it can provide counter examples which show that various features of naming systems can evolve without making additional assumptions involving physiological processing, cognitive strategies, or socio-cultural methods of transmission. On the other hand, by having explicit evolutionary models and algorithms we may be able to demonstrate the feasibility of certain evolutionary theories presented in the literature.

Moreover, it should be noted that the *symmetry breaking* results presented here (Section 4.2) are not directly in conflict with the *maximized wellformedness* (also a form of symmetry breaking) results found by Regier et al. (2007). Indeed, in light of our *symmetry breaking* findings, Regier et al.s' observations that color vision processing could constrain how a stimulus domain is categorized are not surprising, although it is quite surprising that the influences of communication under stimulus heterogeneity and observer heterogeneity would alone be substantial enough to (i) give rise to a stable shared category solutions in the absence of perceptual model, and (ii) would trade off in the interesting ways that we show in the present paper. The interesting next step (which we are in the process of taking) is to show specifically how all the various constraining factors (e.g., perception, cognition, communication, etc.) trade off and interact.

Results presented here are important for evaluating how pragmatic constraints on color naming might influence the evolution and maintenance of a color signaling system in both artificial and nonartificial agents. They also give a good impression of how systematic variation in observer-type heterogeneity could influence convergent category solutions (Jameson, 2005d)—a topic of recent discussion in the color categorization literature (Steels and Belpaeme, 2005). Finally, both results accord with the organizational framework for human color categorization described by Jameson (2005a, pp. 316–325), as well as with suggestions regarding the possible influence of observer variation on color categorization systems (Jameson, 2005a, b, d).

7.2. The gap between the universalist and culturally relative views

Earlier it was noted that the universalist perspective is the most widely accepted explanatory framework for findings in human color categorization, and has generally been successful characterizing similarities seen across different ethnolinguistic groups, but somewhat silent on the issue of identifying the sources of variation in color categorization and naming that is seen across ethnolinguistic groups.

The present investigations, with those of Komarova et al. (2007a, b), give an approach that permits evaluation of both similarities and differences across groups. That is, despite the simplifications needed to conduct these investigations, our results show that it is theoretically possible for universal constraints (such as identifiable universal color vision processing features shared among humans) to impose important limiting conditions on the development and stabilization of shared color category solutions. This finding is at odds with any view suggesting that

human color categorization research lacks empirically tractable constraints. In addition, the present investigations show that such universal constraints can and will interact with other constraints that necessarily differ across local circumstances—or, generalizing to the domain of realistic phenomena, across culturally specific circumstances.

This is an important demonstration for the empirical study of human color categorization because although there have been hundreds of empirical reports evidencing culturally varying influences on human color categorization, the challenge of how to make universalist explanations compatible with culturally relative variation is largely unaddressed in the literature (although see Jameson 2005a–e), and certainly has not been adequately addressed through empirical study. This challenge is, in part, why universalist views remain largely unreconciled with culturally relativist views (despite a general awareness regarding the importance of both types of influences), and why mainstream pure universalist explanations give little assistance for understanding aspects of color naming phenomena that are so clearly linked with cultural variables (see discussion in Davidoff et al., 1999; Roberson et al., 2000, 2005).

7.3. Future research

The present study addresses a gap in the literature and in the mainstream theory regarding the testing of universal color vision processing factors in conjunction with factors specific to ethnolinguistic groups. Our approach is meant to complement empirical studies of human color categorization. Forthcoming research will further investigate how these qualitatively different sorts of factors trade off, and how the simulation findings compare with the empirically observed human phenomena (Jameson and Komarova, 2008). In particular, it will be interesting to study, in a quantitative way, how much population variation is observed in societies with little or no heterogeneity, compared to those with a large percentage of impaired observers. Also, questions of convergence speed to a stabilized shared solution with and without heterogeneity constraints will be addressed.

Jameson (2005a, b) previously suggested that real world pragmatics quite naturally allow for heterogeneous populations in which color categorization subcultures exist and reflect a specialized color expertise, or, perhaps, a marked pragmatic color salience that is only seen in a small portion of the population. It was also suggested that one possible model for communications between such population subcultures may be based on bilingual *code-switching* interactions generally seen among communicating individuals who are proficient in two languages. In such situations, a sender may trade off between modes of communicating as appropriate for linguistic variation in the receiver in any given communication game. Thus, for the case of Fig. 15(a), when pragmatic color utility for one portion of the population (e.g., normal agents) marks a distinctive RIS, but that region in turn is not part of the perceptual repertoire allowed by the psychophysical transform of a different portion of the population (e.g., dichromat agents), then communications about RIS colors between those two agent types unfolds in a way that differs from the mode of communication between, say, two normal agents who can equally perceive and discuss the RIS region (see related discussion in Jameson 2005a, b).[9] The present investigations, however, concentrate on capturing less complex scenarios,

---

[9] And the discussion surrounding the empirical differences in the use of color semantics in bilingual Vietnamese–English individuals in Jameson and Alvarado (2003a, b) and Alvarado and Jameson (2005).

and do not examine the kinds of code-switching circumstances just described. Forthcoming research explores such population communication variation in a more systematic way (Jameson and Komarova, 2008).

7.4. Conclusions

The main result of this study is that both forms of heterogeneity that were considered (population and color utility) served as process anchors and fix rotationally invariant category solutions when independently implemented (i.e., population heterogeneity implemented without color utility heterogeneity; or color utility heterogeneity implemented without population heterogeneity). Also, when the two types of heterogeneity are both incorporated in a simulation, they were shown to systematically trade off in ways that always optimize the communication game success rate for the population. And in all cases we considered, some of which pitted the two forms of heterogeneity against each other from a communication game standpoint, the category solutions that were reached were stable, coherent, and representative of the kinds of categorization schemes one would expect from realistic tests of population categorization of a hue circle continuum. Forthcoming research considers how such idealized categorization factors trade off in comparison with analogous situations from the human empirical color naming literature.

Acknowledgments

References

Alvarado, N., Jameson, K.A., 2005. Confidence judgments and color category best exemplar salience. Cross Cultural: Res. J. Comparative Soc. Sci. 39 (2), 134–158.

Belpaeme, T., Bleys, J., 2005. Explaining universal color categories through a constrained acquisition process. Adaptive Behav. 13, 293–310.

Cornsweet, T.N., 1970. Visual perception. Academic Press, New York.

Davidoff, J., Davies, I., Roberson, D., 1999. Color categories in a stone age tribe. Nature 398, 203–204.

Dowman, M., 2007. Explaining color term typology with an evolutionary model. Cognitive Sci. 31, 99–132.

Hardin, C.L., Maffi, L., 1997. Color Categories in Thought and Language. Cambridge University Press, England.

Jameson, K.A., 2005a. Culture and cognition: what is universal about the representation of color experience? J. Cognit. Cult. 5, 293–347.

Jameson, K.A., 2005b. Semantic and perceptual representations of color. In: Monahan, J.S., Sheffert, S.M., Townsend, J.T. (Eds.), Fechner Day 2005: Proceedings of the 21st Annual Meeting of the International Society for Psychophysics. Central Michigan University Printing Services, Mt. Pleasant, pp. 125–132.

Jameson, K.A., 2005c. The role of culture in color naming research. Cross Cultural Res. J. Comparative Soc. Sci. 39, 88–106.

Jameson, K.A., 2005d. Sharing a perceptually grounded categories in uniform and nonuniform populations. Commentary on steels, L. & Belpaeme, T. (Target Article). Coordinating perceptually grounded categories through language. A case study for colour. Behav. Brain Sci. 28, 501–502.

Jameson, K.A., 2005e. Why GRUE? An interpoint–distance model analysis of composite color categories. Cross Cultural Res. 39, 159–194.

Jameson, K.A., 2007. What is the role of computer modeling and evolutionary game theory in cross-cultural color categorization research? Presented at Invited Panel: "A Discussion of New Interdisciplinary Research on Color Naming and Categorization Within and Across Ethnolinguistic Groups." At The Biennial Meeting of the Society for Psychological Anthropology. May 10, 2007. Manhattan Beach, CA, USA.

Jameson, K.A., Alvarado, N., 2003a. The relational correspondence between category exemplars and naming. Philos. Psychol. 16 (1), 26–49.

Jameson, K.A., Alvarado, N., 2003b. Differences in Color Naming and Color Salience in Vietnamese and English. Color Res. Appl. 28 (2), 113–138.

Jameson, K.A., Komarova, N.L., 2008. Evolutionary Models of Color Categorization: Investigations Based on Realistic Population Heterogeneity. Technical Report Series MBS 08-04. Institute for Mathematical Behavioral Sciences University of California at Irvine, Irvine, CA, USA.

Kay, P., 1975. Synchronic variability and diachronic change in basic color terms. Lang. Soc. 4, 257–270.

Kay, P., 2005. Color categories are not arbitrary. Cross Cultural Res. 39, 39–55.

Kay, P., Regier, T., 2003. Resolving the question of color naming universals. Proc. Natl. Acad. Sci. 100, 9085–9089.

Komarova, N.L., 2007. Modeling the evolution of human language. At the Conference on the Evolution of Punishment. Sponsored by the Institute for Mathematical Behavioral Sciences (IMBS) and the Interdisciplinary Program in History and Philosophy of Science. February 10, 2007.

Komarova, N.L., Jameson, K.A., Narens, L., 2007a. Evolutionary models of color categorization based on discrimination. J. Math. Psychol. 51, 359–382.

Komarova, N.L., Jameson, K.A., Narens, L., Steingrimsson, R., 2007b. Evolutionary Models of Color Categorization based on Discrimination. Invited Presentation. UCI Human Sciences and Social Complexity Colloquium. Video Stream interaction with UCI, UCSD, UCR, UCSB and UCLA. May 4, 2007.

Kuehni, R.G., 2003. Color spaces. Workshop on Colour Ontology and Colour Science. University of British Columbia, Vancouver, October 2003. ⟨http://www4.ncsu.edu/~rgkuehni/PDFs/ColSp.pdf⟩. Retrieved on 09/23/2003.

Kuehni, R.G., 2005a. Unique hue stimulus choice: a constraint on hue category formation. J. Cognit. Cult. 5, 387–408.

Kuehni, R.G., 2005b. Focal color variability and unique hue stimulus variability. J. Cognit. Cult. 5, 409–426.

Kuehni, R.G., Schwarz, A., 2008. Color Ordered: A Survey of Color Systems from Antiquity to the Present. Oxford University Press, NY.

Lindsey, D.T., Brown, A.M., 2006. Universality of color names. Proc. Natl Acad. Sci. 102, 16608–16613.

Nathans, J., Thomas, D., Hogness, D.S., 1986a. Molecular genetics of human color vision: the genes encoding blue, green and red pigments. Science 232, 193–202.

Nathans, J., Piantanida, T.P., Eddy, R.L., Shows, T.B., Hogness, D.S., 1986b. Molecular genetics of inherited variation in human color vision. Science 232, 203–210.

Nickerson, D., 1976. History of the Munsell color system, company, and foundation. Color Res. Appl. 1, 7–10, 69–77, 121–130.

Nickerson, D., 1979. Optical society of America (OSA) uniform color scale samples. Leonardo 12, 206–212.

Puglisi, A., Baronchelli, A., Loreto, V., 2007. Cultural route to the emergence of linguistic categories. Available at: ⟨http://arxiv.org/abs/physics/0703164⟩.

Regier, T., Kay, P., Cook, R.S., 2005. Focal colors are universal after all. Proc. Natl Acad. Sci. 102, 8386–8391.

Regier, T., Kay, P., Khetarpal, N., 2007. Color naming reflects optimal partitions of color space. Proc. Natl. Acad. Sci. 104, 1436–1441.

Roberson, D., Davies, I., Davidoff, J., 2000. Color categories are not universal: replications and new evidence from a stone-age culture. J. Exp. Psychol. General 129, 369–398.

Roberson, D., Davidoff, J., Davies, I., Shapiro, L.R., 2005. Color categories: evidence for the cultural relativity hypothesis. Cognitive Psychol. 50, 378–411.

Saari, D., 2008. Disposing dictators; demystifying voting paradoxes. Cambridge University Press, New York.

Saunders, B.A.C., van Brakel, J., 1997. Are there nontrivial constraints on colour categorization? Behav. Brain Sci. 20, 167–228.

Sayim, B., Jameson, K.A., Alvarado, N., Szeszel, M.K., 2005. Semantic and perceptual representations of color: evidence of a shared color–naming function. J. Cognit. Cult. 5 (3–4), 427–486.

Sharpe, L.T., Stockman, A., Jägle, H., Nathans, J., 1999. Opsin genes, cone photopigments, color vision, and color blindness. In: Gegenfurtner, K.R., Sharpe, L.T. (Eds.), Color vision: from genes to perception. Cambridge University Press, Cambridge, pp. 3–51.

Shepard, R.N., Cooper, L.A., 1992. Representation of colors in the blind, color blind, and normally sighted. Psychol. Sci. 3, 97–104.

Steels, L., Belpaeme, T., 2005. Coordinating perceptually grounded categories: a case study for colour. Behav. Brain Sci. 28, 469–529.

Webster, M.A., Kay, P., 2007. Individual and Population Differences in Focal Colors. In: MacLaury, R.L., Paramei, G., Dedrick, D. (Eds.), The Anthropology of Color. John Benjamins: Amsterdam, The Netherlands.