Evolutionary models of color categorization. II. Realistic observer models and population heterogeneity

Kimberly A. Jameson[1,3] and Natalia L. Komarova[2,4]

[1]*Institute for Mathematical Behavioral Sciences, University of California, Irvine, Social Science Plaza, Irvine, California 92697-5100, USA*
[2]*Department of Mathematics, University of California, Irvine, Rowland Hall, Irvine, California 92697-3975, USA*
[3] *kjameson@uci.edu*
[4] *komarova@math.uci.edu*

Received April 10, 2009; accepted April 10, 2009;
posted April 16, 2009 (Doc. ID 110009); published May 22, 2009

The evolution of color categorization is investigated using computer simulations of agent population categorization games. Various realistic observer types are implemented based on Farnsworth–Munsell 100 Hue Test human performance data from normal and anomalous trichromats, dichromats, and humans with four retinal photopigments. Results show that (i) a small percentage of realistically modeled deficient agents greatly affects the shared categorization solutions of the entire population in terms of color category boundary locations; (ii) for realistically modeled populations, dichromats have the strongest influence on the color categorization; their characteristic forms of color confusion affect (i.e., attract or repel) color boundary locations and accord with our findings for homogeneous dichromat populations [J. Opt. Soc. Am. A **26**, 1414–1423 (2009)]; (iii) adding anomalous trichromats or trichromat "experts" does not destabilize the solutions or substantially alter solution structure. The results provide insights regarding the mechanisms that may constrain universal tendencies in human color categorization systems.

OCIS codes: 330.1690, 330.5020, 330.4060.

1. INTRODUCTION

An extensive empirical literature exists on human color perception and categorization (e.g., [1–6]) and color categorization as it manifests and varies across different language groups (e.g., [7–13]). Recently, novel computational approaches have given insight into population color category system evolution for both empirically observed and simulated population categorization scenarios [14–20]. A particularly informative computational approach investigates population color categorization behaviors from an evolutionary game theory perspective [21–23].

The present article explores the impact of visual processing constraints on the evolution of shared color category systems using a computer simulation methodology similar to that used in our companion paper [23]. The present investigations examine greater realism with regard to both color vision observer type variation and population heterogeneity. Here the range of observers modeled is extended to include anomalous trichromats and observers with four retinal photopigments, in addition to the forms of normal and dichromat observers previously modeled [23]. As a result, the present paper (i) extends our color categorization studies by including a wider variety of realistic observer types and (ii) increases the realism of our modeling by investivating the effects of population diversity, which in the present study is modeled to resemble that found in some human populations. These realistic population scenarios permit identification of mechanisms that play an instrumental role in simulated category solution variation. They also provide for comparisons with mechanisms identified under less realistic situations to evaluate their usefulness as constructs for modeling color category evolution in a wide range of human populations.

Our earlier research considered either varied populations of agents with different but simply idealized color perception features [21,22] or unrealistically uniform populations of agents with realistically modeled color perception features [23]. By comparison, the present research investigates the effects of both realistic observer variation and realistic population heterogeneity on the formation of stable population category structures. The types of observer variation examined are based on human performance data from (1) normal color perception variations, (2) anomalous trichromat variations, (3) dichromats (both protanopes and deuteranopes), and (4) nonnormative trichromats with color processing expertise.

Below is a review of the modeling and methods used, including: the stimulus continuum on which all our agent models are based, the parameterization methods, and some of our existing agent models [23]. We also introduce new agent models, present investigations that use increasingly realistic population models, and report their results. The last section of the article discusses the research findings and their relevance to human categorization behaviors.

2. COMPONENTS MODELED

To investigate the effect of specific realistic constraints on color categorization and inter-individual communication,

we use an evolutionary game theory modeling framework in which individual agents learn to (i) categorize simulated colors through reinforcement learning by playing "discrimination–similarity games," and (ii) communicate the meaning of categories to each other [21,22]. The main components of this approach are (*a*) the stimulus space, (*b*) the observer models used, and (*c*) the evolutionary game. Component (*c*) is described in detail in recent research [21–23]. All three components are detailed in the companion paper (this issue) [23], where they were used to examine the effect of limited observer variation on constrained population categorization solutions.

A. Stimulus Space Used

Similar to previous investigations [21–23], in this article a hue circle stimulus space is employed. A hue circle is a natural subspace of the color appearance space and is a structure that can be used to distinguish normal observers from color-deficient observers by means of hue circle similarity relations (see [23] for a discussion). The hue circle stimulus model used is the Farnsworth–Munsell 100 Hue Test, abbreviated FM100 [24]. The FM100 contains 85 color "caps" designed to form a perceptually smooth gradient of hue, ostensibly at a fixed level of brightness and a fixed level of saturation [25]. Advantages of using the FM100 stimulus set include that it is well understood colorimetrically and is in common use as a diagnostic for human color vision deficiencies. Also, because the FM100's 85 hues form an approximate continuum of uniform Munsell *Chroma* and comparatively uniform Munsell *Value* levels [[25], p. 2239, Fig. 1] , it provides a standard for comparisons between our simulated population category solutions and existing human color category partitioning results on a Munsell Book of Color stimulus set (e.g., [4]). In the investigations described here, populations of simulated agent observers engage in color communication games using FM100 stimuli as the sampled color space domain. Using the FM100 constrains the present investigations since it provides only a subspace of full color space, and the resulting observer models reflect only that subset of human color discrimination behaviors. However, the FM100 stimulus provides much of the perceptual color space variation needed to capture the primary factors of any hue-based color categorization system.

B. Realistic Models of Population Heterogeneity

A major goal here is to compare variation in color categorization systems found under realistic population heterogeneity (i.e., populations in which individual agent models differ) with that previously observed under population homogeneity (i.e., populations in which all agents possess the same discrimination model). In addition to models of normal trichromats and dichromats used earlier [23], here we include several forms of *normal* trichromat (including some "expert" observers), and also *anomalous trichromat* agents. While the entire hue circle stimulus is employed in our simulations, our models of color perception deficiency primarily reflect confusions among hue stimuli along the reddish-to-greenish continuum of color appearance [26], attributable to human X-chromosome recessive inheritance (i.e., protan and deutan forms of deficiency).

Table 1 shows the frequency with which such deficiencies are found in, for example, populations of European ancestry. Table 1's values are compiled from multiple sources ([26], p. 98; [27], p. 667; [28], p. 217; [29], p. 464; [30], p. 30) and give the ranges of the observed population frequencies approximated in our simulations. These data guide primarily the modeling of relative frequencies in the artificial agent populations investigated here, since we make no attempt to replicate real-world population demographics underlying human color category evolution. The reasons for this are: (*a*) frequencies of deficiency vary across different human populations and different empirical surveys, and (*b*) the population size used throughout these investigations is, for technical reasons, held constant at 100 agents. Observer frequencies were adjusted in one or two instances in our simulations (e.g., Table 1's values close to a 1.5% observed frequency were approximated as a 2% simulated population frequency) to permit evolutionarily important peer-to-peer interactions in population communication games. Minor adjustments in population frequencies were well investigated and generally found to have no, or negligible, effect on the present findings.

Table 1 distinguishes deficiencies that involve a shift of photopigment sensitivity from forms caused by a lack of a photopigment class; these important variations are investigated and compared here. Other important features constrain color categorization dynamics in human populations, but not all such constraints are incorporated in the agent-based scenarios simulated here. For example, in the present investigations cognitive factors do not contribute to color categorization—e.g., factors involving individual color memory, or personal color salience. Omitting complex cognitive factors permits a focus on clarifying the specific influences of color perception and discrimination

Table 1. Observed Frequency Ranges of Human Color Deficiency Used to Model Heterogeneous Agent Populations

Type of Color Deficiency	Number of Photopigments	Percent Frequency Ranges in Populations of European Descent
Protanopia	2: No long-wavelength-sensitive photopigment	1.01–1.60
Protanomalous Trichromacy	3: Abnormal long-wavelength-sensitive photopigment	1.03–1.27
Deuteranopia	2: No middle-wavelength-sensitive photopigment	1.01–1.97
Deuteranomalous Trichromacy	3: Abnormal middle-wavelength-sensitive photopigment	4.01–5.35

on categorization. In contrast to recent investigations [21,22], here no assumptions are made about the salience of color space with respect to (a) variation in environmental color distribution or (b) variation in the putative importance, or utility, of some color appearances compared with others (see [6]). Thus, here regions of increased salience and environmental color hot spots are not investigated [21,22], nor are there any investigations at the individual observer level—e.g., specific color appearances that are universally salient [2].

C. Modeling Different Forms of Observer Variation

Here we simulate both normal and deficient observer groups using human data as described in [23]. "Normal" agents include those modeled on normal trichromat FM100 performance, with trichromacy confirmed by pseudoisochromatic plate assessment [5]. "Anomalous trichromat" agents are modeled using FM100 performance as reported by Bimler and colleagues [[31], Table 1, p. 166]. "Expert" agents are modeled using FM100 performance reported by Jameson and colleagues [32,33]. And "deficient" agents are modeled using protanope and deuteranope performances from the FM100 diagnostic database [34].

To realistically model human observers, we use a probabilistic observer model [23]. The hue circle to be categorized by all agents is assumed to be identical (i.e., the FM100 stimulus). Two colors are chosen randomly from the circle and presented to the agent. An appropriate color confusion transformation reflects the probabilities that an agent perceives the two chosen color stimuli as the same or different in color. In general, the transformation is described as an 85×85 matrix, $\{C_{ij}\}$, whose entries C_{ij} give the probability that stimulus i is perceived as stimulus j. This confusion matrix varies across individual agents. Table 2 lists modeling and simulation notation used.

In modeling our agents we use several types of discrimination-based confusions among FM100 stimuli. The first one is a form of unsystematic noise in an agent's categorization that resembles the *sorting confusion* errors seen when human observers sort the FM100 stimulus continuum. Normal color vision observers typically exhibit random transposition errors or confusions that occur between adjacent FM100 stimuli. Such confusions tend to be 2-cap exchange errors that are random seldom reoccur on retest, and are not bunched in any one region of the hue circle [35]. This normal unbiased sorting confusion is modeled here by a matrix $\{C_{ij}^{sorting}\}$, which assumes that with probability p, each color cap can be confused with either of its neighboring caps on the FM100 stimulus (Eq. 1, [23]).

Sorting confusions are included in all agent models used here. Those for normal dichromat agents were defined earlier [23]. Appendix A derives sorting confusion parameter values for the anomalous trichromat agents introduced here.

A second form of confusion is referred to as *local confusions*, or confused stimuli adjacent to one another on the FM100 array. Local confusion regions may occur within a FM100 tray or span two trays their locations depend on the severity type of observer deficiency. The matrix of an agent's local confusions is denoted as $\{C_{ij}^{local}\}$ (Eq. 2, [23]). FM100 performance data used here suggest that local confusion regions are delimited by cap pairs, denoted as I_1 and I_2. A parameter that also varies across observer models characterizes the *range of confusion*, or the number of adjacent caps confused within a local confusion region denoted as w.

Normal agent models do not include local confusions. Dichromat agent local confusions were defined earlier [23]. Appendix A details the local confusion calculations for the anomalous trichromat agent models introduced here. Figure 1 presents the anomalous trichromat average local confusion regions derived in Appendix A. These provide Table 3's anomalous trichromat parameter values.

Third, we also model *global confusions* or confused FM100 stimulus pairs, connected by a line of confusion, opposite on the hue circle ([26], pp. 65–69). In previous simulations such a confusion line was referred to as an *axis of confusion* [22]. Here global confusions are used exclusively in dichromat models (similar to [23]).

Dichromat FM100 global confusion features were included in our realistic models of dichromats after considering the appropriateness of global confusions in the everyday color experience of human dichromats. Here, as in [23], we assume that dichromat global confusion pairs exist in practice when similarity is important, as in the FM100 task. Existing data [36] define three global confusion pairs for protanope agent models separately from those for deuteranope agents. The global confusion matrix, denoted as $\{C_{ij}^{global}\}$, was defined by Eq. 3 in [23].

Table 2. Simulation Notation

Symbol	Definition	Value
N	Total number of agents	100
n	Total number of hue circle stimuli (FM100 samples)	85
k_{sim}	Parameter defining the pragmatic utility of colors	11
p	Probability of sorting confusion	Table 3
w	Confusion range operating in a local confusion region	Table 3
C_{ij}	Probability that FM100 stimulus i is confused with j defined by $C_{ij}^{sorting}$, C_{ij}^{local}, and C_{ij}^{global} for different observer models.	Table 3

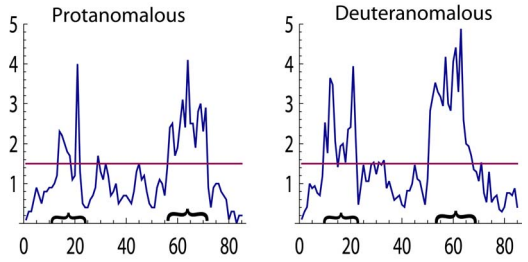

Fig. 1. Average FM100 sorting score for protanomalous ($n=10$) and deuteranomalous ($n=17$) individuals. Horizontal axes are the 85 caps of the FM100, and the vertical axes are the observed average error at each cap for the two samples considered. Bracketed regions illustrate continuous regions of confusion arising from confusion patterns in the data, with the value 1.5 obtained for each anomalous type.

Table 3. Agent Confusion Matrix $\{C_{ij}\}$ Parameter Values Derived from FM100 Observer Data

	Sorting ($C_{ij}^{sorting}$)	Local (C_{ij}^{local})			Global (C_{ij}^{global})		
Observer Model	p	$I_1=[i,j]$	$I_2=[i,j]$	w	$(x_1,\bar{x}_1)$	$(x_2,\bar{x}_2)$	$(x_3,\bar{x}_3)$
Normal	8% or 0%	—	—	—	—	—	—
Protanope	8%	(12, 27)	(59, 70)	10	(9, 29)	(1, 40)	(74, 53)
Deuteranope	8%	(9, 25)	(51, 63)	10	(11, 21)	(2, 31)	(79, 42)
Protanomalous	26%	(14, 21)	(57, 71)	5	—	—	—
Deuteranomalous	24%	(10, 22)	(51, 67)	5	—	—	—
Expert	22%	(21, 28)	(43, 47)	3	—	—	—

Summary of the agent models used. The following agent models are used in Study 1 and Study 2 simulations:

- Normal trichromat agents: $C_{ij}=C^{sorting}$ [23].
- Anomalous trichromat agents: $C_{ij}=C^{sorting}C^{local}$, where confusion parameter matrices are based on observed FM100 data and where local confusion regions (Fig. 1) accord with Farnsworth's [35] major axes for anomalous subjects. (Details of parameter derivation for anomalous trichromat agent variants are given in Appendix A.)
- Deficient dichromat agents: $C_{ij}=C^{sorting}C^{local}C^{global}$ [23].

Study 3 simulations additionally use "expert" agents:

- Expert trichromat agents: $C_{ij}=C^{sorting}C^{local}$, where confusion parameter matrices are based on observed FM100 data [32,33]. (Details for "expert" agent modeling are presented later, in Section 3.)

Table 3 presents parameter values for all observer models investigated. Parameter values depend on the data used to model agent types. If different data are used, parameter values may also vary. However, during model development we varied the underlying data to evaluate the effect on categorization solutions. Such tests found that the main findings presented here were not altered by parameter changes that would be expected from typical variation in observed human data.

3. POPULATION HUE CATEGORIZATION INVESTIGATIONS

Homogeneous populations were investigated in the companion article [23]. These are compared with heterogeneous population investigations (Studies 1, 2, and 3) presented here. Study 1 investigates heterogeneous population categorization solutions, and the effect of local and global confusions, by comparing population solutions under protan or deutan biases. Study 2 approximates realistic observer populations with realistic agent models occurring at frequencies that approximate some human populations. Study 3 examines whether (i) agent models based on nonnormative FM100 performance data from individuals with otherwise unimpaired color perception further constrain population solutions and (ii) whether the presence of such agents disrupts the stabilization of optimal population solutions. Sample populations investigated in Studies 1–3 are summarized in Table 4.

In all three studies population color naming is achieved by discrimination–similarity communication games between agents involving the 85 FM100 caps. We previously detailed communication games using the FM100 stimulus (see [23], Section 3). Briefly, two agents are randomly chosen and shown two randomly selected FM100 stimuli. Based on the agents' individual $\{C_{ij}\}$ transforms, and on their individual categorization system, each agent names these two stimuli. A pragmatic category width parameter, k_{sim}, is used to determine whether two stimuli compared should belong to the same category or different categories. Agent responses that accord with this k_{sim} determination are reinforced. Each agent participates in thousands of these communication games, gradually optimizing the success rate of the games. Through reinforcement each agent eventually learns and stabilizes a category solution for all 85 FM100 stimuli. The population solution is evolved by this process, and is the conglomerate of all individual agent solutions.

In general, the value of k_{sim} can be made to vary across the hue circle if some colors have more pragmatic or social utility than others [22]. Here k_{sim} is held constant. The same approach extended to populations with ideal heterogeneous color observers with fixed and variable k_{sim} was

Table 4. Agent Observer Type Percentages in the Populations Investigated

	Study 1		Study 2		Study 3	
Observer Model	A_1	B_1	A_2	B_2	A_3	B_3
Normal	97	93	97	90	75	0
Protanope	1	0	1	1	1	0
Deuteranope	0	2	2	2	2	0
Protanomalous	2	0	0	2	2	0
Deuteranomalous	0	5	0	5	5	0
Expert	0	0	0	0	15	100

recently reported [22]. In all these articles, the discrimination–similarity game evolves population naming behavior that is near theoretically optimal ([21] details optimal naming behavior).

A. Study 1: Population Heterogeneity under Varying Protan and Deutan Perceptual Biases

Study 1 introduces agent models that approximate anomalous trichromat performance (Fig. 1) and compares categorization solution variation across populations that include different forms and varying degrees of perceptual biases. Heterogeneous populations of agents were constructed using normal, anomalous, and dichromat models at frequencies approximating those found in some real populations while restricting the form of a population's bias to either a protan or deutan type. Note that the constraint of either protan or deutan forms of deficiency in Study 1's populations does not resemble the manner in which deficiencies exist in human populations.

Table 4 shows the composition of Study 1's heterogeneous groups with population A_1, which involves protan biases only, and population B_1 involving deutan biases only. Simulation results for these populations are shown in Figs. 2(a) and 2(b), respectively. Results are summarized as follows:

1. Categorization solutions from populations A_1 and B_1 agree with optimal solutions found previously [21] in that they consist of disjoint category regions of roughly equal length. In other words, the presence of a small percentage of dichromat agents with global confusion pairs does not lead to a prevalence of heterogeneous population defects (as found in [23] for homogeneous dichromats with global confusion pairs). Therefore, in contrast to results from [23]'s homogeneous dichromat population results, for this situation, the notion of color boundaries is well defined.

2. Figure 2 shows histograms of category boundary locations for stabilized population solutions over 5,000 independent simulations. It is clear that some locations are more preferred as category boundaries than others, as indicated by the pronounced histogram peaks observed. This phenomenon of *symmetry breaking* is similar to that observed for populations with 100% deficient agents [23], except that boundary histogram peaks from populations of mostly normal observers (e.g., populations A_1 and B_1) are less pronounced than peaks found for homogeneous dichromat populations [23]. Accordingly, 5,000 simulations were required to obtain comparable Fig. 2 histograms from mostly normal A_1 and B_1 populations compared with the 1,000 simulations needed for the histograms based on 100% deficient populations [[23], Figs. 4 and 5].

3. B_1 populations boundary peaks appear more pronounced than with those from A_1 populations (Fig. 2). This result is due to the higher total percentage of agents with perceptual bias in B_1 (at 7%) than that in A_1 populations (at 3%). That is, investigations of A_1 populations that include 7% agents with some protan bias (not shown) produced boundary distributions as pronounced as that shown for B_1 populations [Fig. 2(b)].

4. Despite relatively low percentages of individuals with deficiencies, influences are clearly observed in Fig. 2 for both forms of deficiency. We also performed similar runs without anomalous trichromat agents. That is, populations of 1% protanopes and 99% normals as well as 2% deuteranopes and 98% normals were considered. In these latter cases similarly clear peaks were also observed in the color boundary histograms (not shown). The observed symmetry breaking in a population with as few as 1% dichromat agents underscores the importance of modeling population heterogeneity, since even infrequent perceptual biases have significant effects on population categorization solutions.

5. The locations of category boundary peaks for population A_1 [Fig. 2(a)] are similar to analogous observed solutions from homogeneous populations of protanopes ([23] Fig. 4), whereas the category boundary locations for population B_1 [Fig. 2(b)] resemble those from previously observed homogeneous populations of deuteranopes ([23], Fig. 5). Such resemblances underscore the importance of our earlier investigations of unrealistic homogeneous dichromat populations [23]. Earlier investigations showed that boundary peak positions from homogeneous populations were based entirely on local and global confusions [23]. The same explanations generalize to Study 1's A_1 and B_1 populations, where dichromat and anomalous trichromat agents comprise only a small percentage of the populations studied.

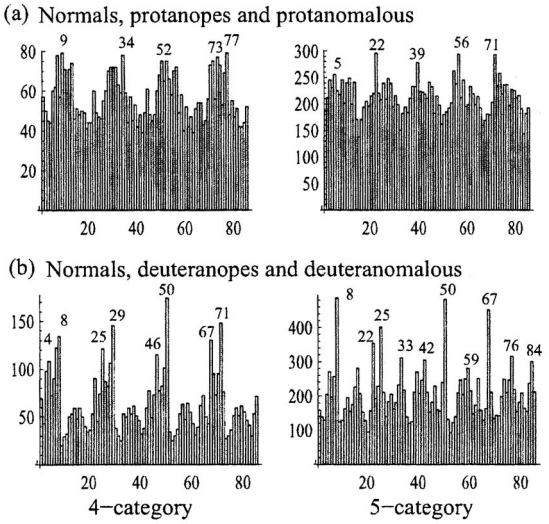

Fig. 2. Symmetry breaking in heterogeneous population categorization solutions for population (a) A_1, protan biases (i.e., 1% protanope, 2% protanomalous trichromat, and 97% normal agents) and (b) population B_1, deutan biases (i.e., 2% deuteranope, 5% deuteranomalous trichromat, and 93% normal agents). Solutions shown are: four-category (left) and five-category (right). Horizontal axes show the FM100's 85 caps. Vertical axes show color boundary frequency at a given cap (over 5,000 independent simulations).

B. Study 2: Realistic Population Heterogeneity Comparisons

Study 1 introduced anomalous trichromat models to increase population heterogeneity for populations with *either* protan *or* deutan biases. However, while Study 1's populations A_1 and B_1 served to increase heterogeneity

compared with earlier investigations [23], they do not achieve the observer type diversity found in real populations that contain both forms of deficiency. Study 2 increases population heterogeneity with such realism in mind, based on the rationale that the agent variety in a population may influence the category learning in discrimination–similarity communication games.

Color communications between two individuals with highly similar color perception (say, monozygotic twins with the same photopigment opsin genes, producing similar trichromat phenotypes) are likely to experience fewer perceptually based discordances of color identification and naming (see [37]), compared with the interactions between two individuals with different color perception bases (e.g., dizygotic twins [38] or siblings who share only half of their opsin genes) or even compared with unrelated individuals with different forms of normal variation (for example, a normal trichromat compared with an anomalous trichromat). Perceptually based discrepancies in hue naming have been widely examined in the literature and are well understood. For example, as Pokorny and Smith [39] described hue estimations by five deuteranomalous trichromats using restricted response categories "red," "green," "yellow," "blue," and "white," they explain:

…the response category "blue" was the same for normals and deuteranomals… [While] the deuteranomals showed greater intersubject variability than normals in using response categories "red," "yellow," and "green," a phenomenon that does not occur in normal trichromats…. The deuteranomalous trichromats did not always need three independent color names to describe the spectrum above 510 nm ([39], pp. 1202-1203).

It is plausible, then, that color perception differences that separate normal trichromats from both forms of anomalous trichromats should have influences on the shared color categorization system formed by a population. Study 2 examines this possibility and investigates how solutions vary from Study 1's partially heterogeneous population solutions.

Two heterogeneous populations investigated in Study 2 are shown in Table 4, namely, populations A_2 and B_2. The results are shown in Figs. 3(a) and 3(b) and are summarized as follows:

1. Both populations produced close-to-optimal solutions (as in Study 1) and exhibited a tendency to a high degree of symmetry breaking, with certain stimulus locations preferentially chosen as color boundaries.

2. When compared with Study 1's Fig. 2, the boundary location histograms in Fig. 3 more closely resemble those obtained for Study 1's-population B_1 population results than those obtained for A_1 populations. Thus, a higher total percentage of deutan-biased types (7%) in a population, in the company of fewer protan-biased types (3%), is found to exert a stronger influence on the boundary locations in Study 2 solutions [see Study 1, item (3)].

3. Variation in the normal agent parameter values shown in Table 3 is illustrated in the Fig. 3(a) results that assume zero sorting errors ($p=0$). To evaluate the role of sorting confusions in population A_2 simulations, additional investigations were conducted with values of p increased to higher and more realistic values. Such increases did not make any tangible difference in population A_2 observed outcomes (not shown).

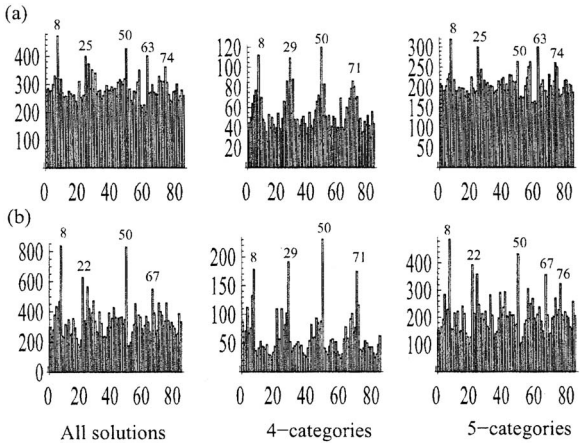

Fig. 3. Color category boundary locations from two heterogeneous populations. (a) Population A_2 consists of 1% protanope, 2% deuteranope, and 97% normal agents (noise at the $p=0\%$ level), and (b) population B_2 consists of 1% protanope, 2% protanomalous trichromat, 2% deuteranope, 5% deuteranomalous trichromat, and 90% probabilistic normal agents with $p=8\%$ noise. Solutions shown: All (left), four-category (middle), and five-category (right). Each histogram shows the distribution of boundary locations resulting from 5000 population simulations.

4. Adding realistic percentages of anomalous trichromat agents to a population [Fig. 3(b)] makes no difference regarding the stabilization of near-optimal solutions compared with populations that differed only by the absence of anomalous trichromats [Fig. 3(a)].

5. Figures 3(a) and 3(b) show that the presence of anomalous trichromat agents makes symmetry-breaking effects stronger and refines boundary location distributions as a result of overall increased percentages of deficiency in the population (i.e., 3% versus 10% of the total population with either anomaly or deficiency).

6. Solution robustness was investigated in simulations where local confusion regions for all protanomalous agents were varied slightly or were made identical, and where for all deuteranomalous agents were varied slightly or made identical. The results indicated that slight modifications of this kind did not noticeably change the resulting boundary histograms for obtained categorization solutions (not shown) from those shown in Fig. 3.

C. Study 3: Modeling Agents from Observers with Diverse Photopigment Genotypes

While Studies 1 and 2 modeled agents of known perceptual variation linked to color vision deficiency, Study 3 additionally models individuals with diverse photopigment opsin genotypes that are linked to a normal form of color perception variation with some uncertainty.

1. Observer Groups Defined by Photopigment Opsin Gene Variation

The protan and deutan agent models investigated above parallel human perceptual groups attributable to long- and medium-wavelength-sensitive photopigment opsin gene variations. Photopigment genes are located in a tan-

dem array on the X-chromosome and share 96% of their exon coding regions [40,41]. The long- and medium-wavelength-sensitive genes differ at only 15 sites of their nucleotide sequences. Single nucleotide substitutions at three particular sites (codons 180, 277, and 285, exon 3) produce substantial shifts in photopigment spectral sensitivity [30], and sensitivity shifts increase monotonically with substitutions [42,43].

Some human females have different long-wavelength-sensitive opsin genes on each X-chromosome and, as a result, the genetic potential to express more than the usual three retinal photopigment classes. These heterozygous females are *putative retinal tetrachromats* [33] and may express (in addition to rods) four retinal cone classes, each with a different spectral sensitivity distribution. Such individuals have the potential to experience tetrachromatic vision [44]. Frequency estimates of females who are potential tetrachromats range between 15% and 47%, whereas less is known about the actual frequency of expressing four retinal cone classes.

2. Perceptual Differences Associated with Retinal Tetrachromacy

Expressing four retinal cone classes is accepted, but functional human tetrachromacy is debated. Theory limits humans to no more than a trivariant color signal. Thus, four retinal cone classes may be a necessary (but not a sufficient) condition for tetrachromatic color perception, since, for true tetrachromacy, four channels of cortical color signal processing also seem to be needed. It has been suggested that female carriers of photopigment opsin gene variants show signs of impaired chromatic discrimination (e.g., [45]). However, carriers of protan deficiencies are not always chromatically impaired, and some need significantly less red/green contrast to detect the chromatic modulation in a grating ([46], p. 2898). This suggests that carriers of protan deficiencies can have lower red/green chromatic discrimination thresholds than normal trichromats.

Jameson and colleagues [33] investigated the FM100 performance of females with heterozygous (Serine-180-Alanine) long-wavelength-sensitive opsin genes. Some of these experienced no color vision impairment or weakness and exhibited increased sensitivity for detecting chromatic bands in a diffracted spectrum task [33]. Jameson and colleagues [32,33] suggest that the color perception of some female carriers of protan deficiencies can differ from that of female trichromat controls but not in a deficient way. Rather, such carriers detect significantly more chromatic contrast (e.g., [46]), producing categorical color differences in the chromatic banding task used [32]. Similarly, Birch [47] indicates that female compound mixed heterozygotes for protan and deutan color deficiency (e.g., among those studied by Jameson *et al.* [32,33]) are usually reported to have normal, not deficient, color vision.

Individual color similarity data are also relevant to these hue circle categorization studies. In a twin study of perceptual color space, Paramei *et al.* showed that similarity mappings of color circle stimuli are nearly identical across individual same-sex monozygotic twins possessing the same underlying photopigment opsin genotypes; but color circle mapping spatial relations can vary substantially across other normal trichromat twin pairs owing, apparently, to differing opsin genotypes [[37], p. 308, Fig. 2]. Thus, normal trichromat observers with slightly different photopigment genotypes can exhibit different similarity relations among hue circle stimuli. Although such differences seem to have little effect in many perceptual categorization studies these differences, modeled by FM100 data, may affect the evolutionary dynamics underlying population categorization solutions.

Study 3 uses FM100 data from opsin gene heterozygotes to model a variation on normal trichromat agents. Heterozygote-based agent models are referred to as "experts" here, based on findings that they outperform female control subjects on some color perception tasks [32,33,46], despite (in some cases) nonconformance to FM100 norms. We explore the possibility that the FM100 error patterns of these otherwise normal color vision individuals might influence evolutionary game theory dynamics of population color categorization solutions. Used as the basis for expert agents, such individuals' FM100 error patterns may be uninterpretable, and may have no discernable effect on inter-agent color communications or population categorization solutions. Study 3 evaluates these possibilities.

Modeling expert agents. Modeling expert agents on FM100 performance accentuates the diagnostic's theory and expectations. One might assume that actual color experts perform like ideal normals on the FM100, with zero deficiency-based sorting transpositions and no probabilistic sorting error. Modeled this way, experts would be expected to produce color continua that duplicate the "correct" sequential ordering of the FM100 caps, as prescribed by the FM100 diagnostic. In fact, this scenario does not characterize the human performance data used to model our expert agents (see Fig. 4).

Alternatively, a model of expert FM100 performance that is compatible with the actual performance data supposes that experts require a personal "correct" ordering that does not exactly follow the FM100 stimulus sequence. In this scenario, an expert may exhibit transpositions in the sorting task that disagree with the diagnostic's standard sequence. This personal ordering scenario is possible if in some cases FM100 cap transpositions reflect an unimpaired individual's nonnormative just-

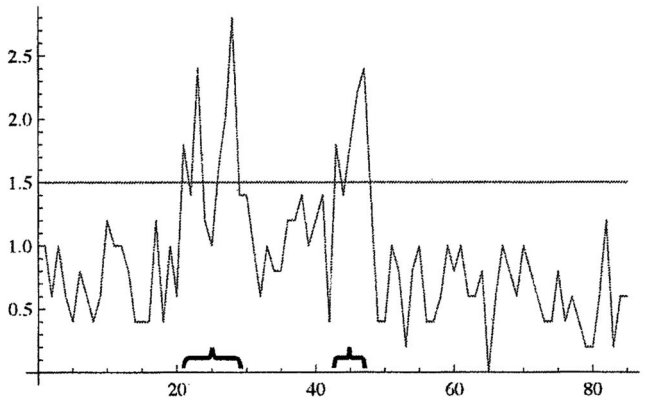

Fig. 4. Confusion patterns from FM100 performance of five expert observers [33]. Axes as defined in Fig. 1 caption. Used in modeling expert agents in both population A_3 and B_3.

noticeable-difference (jnd) variation along one or more color space dimensions (rather than sorting errors due to poor color sensitivity). If patterns of *expert* transpositions are sufficiently structured, they may present a new symmetry-breaking potential, similar to Studies 1 and 2 above. Incorporating such variation in agent modeling may also affect categorization by varying FM100 jnd mappings across groups of agents and varying the stimulus continuum relative to parameter k_{sim}.

Expert agents were modeled using existing FM100 data from five female individuals with diverse photopigment opsin genotypes [48]. All five were heterozygous for both X-chromosome linked opsin genes [32,33]. Despite otherwise excellent color perception, compression parameter analyses showed that all individuals' patterns of FM100 confusion were displaced in a direction corresponding to a 15° axis in a polar coordinate compression parameter space ([33], Fig. 2(a), p. 12). Their FM100 performance differs from normative age-adjusted performance by an average Z value of 1.31 standard deviations (individuals either failed or performed very poorly), and were described diagnosed as "… false-positive deficient when their color perception is otherwise unimpaired and their color sense is generally regarded as excellent" ([33], p. 15).

Figure 4 shows the expert-agent model average FM100 performance, with local confusion regions (i.e., 21–28 and 43–47) computed as described for anomalous trichromats, with confusion range $w=3$. Similar to anomalous trichromats (Fig. 1), expert observers exhibited diffuse color confusion, with greater than normal sorting confusions outside their local confusion regions and a probability of sorting errors (derived from an average transposition calculation described in Appendix A) $p=22\%$ per cap.

Simulations included expert agents at frequencies (a) likely to influence obtained categorization solutions and (b) in accord with Table 1's populations, for which at least 15% of women are carriers of X-chromosome-linked color deficiencies (see [46]). Study 3 permits comparisons between solutions from populations that include expert agents with those that do not include experts, shown in Table 4 as populations A_3 and B_2, respectively. Parameter values for expert agents are shown in Table 3. We also consider homogeneous expert populations (B_3), similar to the spirit of [23]. Study 3's aim is to examine whether FM100 error patterns from these otherwise normal color observers (i) disrupt or alter population solution stabilization, (ii) introduce new symmetry breaking influences, or (iii) affect category boundary robustness. Observing any of these three would provide new information for human population color category evolution.

Populations including expert agents. Results are presented in Fig. 5 and summarized below.

1. Categorization solutions for both the A_3 population including experts and the B_3 population of 100% experts achieve forms of near-optimal solutions seen previously [21].

2. The presence of 15% experts in a realistically heterogeneous population [Fig. 5(a)] does not change the boundary location histograms dramatically, compared with populations without experts, Fig. 3(b).

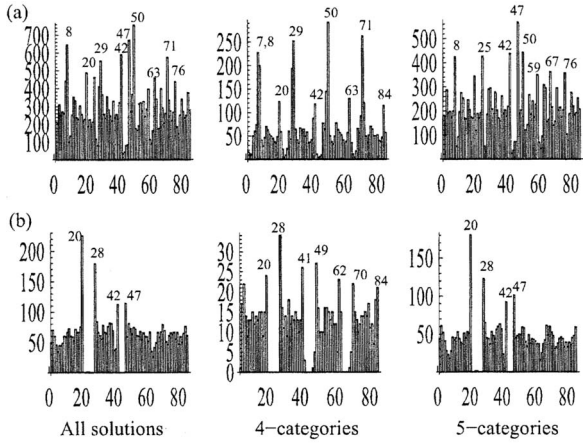

Fig. 5. Experts in population color categorizations. Results from two populations are shown: (a) population A_3 consists of 1% protanope, 2% protanomalous trichromat, 2% deuteranope, 5% deuteranomalous trichromat, 75% normal, and 15% expert agents; (b) population B_3 consists of 100% expert agents. All (left), four-category (middle), and five-category (right) solutions are shown. Parameters are as described earlier. Histogram shows the results of (a) 5000 runs, (b) 1,000 runs.

3. Aside from small shifts in boundary locations for five-category solutions, the only differences are new "gaps" in color boundary distributions, one between caps 42 and 47 and the other between caps 20 and 28. These gaps are consistent with our conjecture that local confusion regions repel color boundaries. Expert local confusion regions occur between caps 21–28 and 43–47: exactly where color boundaries are few. A more pronounced form of this is seen in Fig. 5(b) for homogeneous expert populations, where no boundaries occur in regions of local confusions. In addition, four-category solutions reflect a third gap between chips 62 and 70 due to the absence of color boundaries inside the local confusion regions [middle histogram, Fig. 5(b)]. The same tendency is observed in realistic heterogeneous populations that include experts [middle histogram, Fig. 5(a)].

4. Increasing experts to 25% of the population does not qualitatively change the resulting picture (results not shown).

Study 3's results suggest that nondeficient color perception observers (putatively possessing a form of genetically based color expertise) who perform poorly on the FM100 diagnostic can constitute a considerable proportion of a heterogeneous population without substantially disturbing the stablization or the form of a population's shared categorization of FM100 stimuli.

4. DISCUSSION

Realistic approximations of population heterogeneity were investigated to increase understanding of the effect of observer variation on simulated color category evolution. Similar to our previous research that used idealized or constrained realism, population simulations show that trade-offs exist between the features of various agent models and frequencies of different agent types in populations. These features balanced in a dynamic fashion during the formation and stabilization of near-optimal

color categorization solutions. Even under diverse population heterogeneity, observed solutions reach stabilization by maximizing successful color communications among agents. This was achieved in ways compatible with an optimal partitioning theory of color space that uses an Interpoint Distance heuristic [49–51].

A. Summary of Findings

The general findings from these realistically stratified populations, composed of approximated real observers, parallel those from our companion paper [23] and are summarized as follows:

1. Under realistic observer heterogeneity (e.g., populations B_2 and A_3), systematic observer biases exert significant symmetry-breaking effects on categorization solutions compared with unsystematic error variations or random sorting confusion variations, which have little effect on category solution stabilization. Varying the levels of observer-model sorting confusions (e.g., population A_2 and related analyses) primarily decreases population agreement at color category boundaries. These findings confirm similar results observed under less realistic simulation scenarios [21–23].
2. In realistically heterogeneous population solutions, the effect of anchoring color boundaries to a subset of possible locations is primarily defined by the presence of dichromats in the population. The resulting preferred boundary locations resemble those found for homogeneous dichromat populations.
3. The presence of as little as 1% dichromats in a population leads to a noticable bias in color category stabilization.
4. Global confusion pairs break symmetry in categorization solutions by anchoring category boundaries (seen in populations A_1, B_1, A_2, B_2, and A_3). Local confusion regions break symmetry in categorization solutions (e.g., population B_3) by repelling category boundaries.
5. Increasing the number of agents with similarly biased deficiencies in a heterogeneous population produces greater symmetry-breaking effects in population solutions (e.g., compare populations A_1 and B_1).
6. Increasing the proportion of anomalous trichromat agents in a population up to 25% has only a minor effect (beyond that imposed by dichromats) on the location of category boundaries in observed solutions (e.g., Study 2 results compared with [23]'s dichromat solutions). As discussed below, this may be due to similarities in the FM100's assessment of anomalous trichromats and dichromats.
7. Adding expert agents at realistic frequencies to already heterogeneous populations did not disrupt convergence or dramatically alter the form of optimal population solutions (e.g., four-category solutions of population B_2 compared with those of population A_3). Furthermore, increasing the proportion of expert agents in a population up to 25% did not diminish the stability of shared category solutions (e.g., population A_3 and related discussion).
8. Individual variation in agent observer features—for example, protan versus deutan biases—affects the form of population category solutions obtained (e.g., population A_1 and B_1, Fig. 2).
9. Realistic approximations of observer types and population frequencies (e.g., all populations considered here) confirm earlier symmetry-breaking and confusion axis results found with idealized populations [21,22]. To our knowledge, the present kind of modeling of color categorization dynamics based on realistic data has not been previously attempted. These findings underscore the importance of tracking population heterogeneity when modeling shared color category systems [52].
10. Generally, the category solutions shown here are shaped by the observer features employed, and it was previously shown that solutions obtained under such constraints are easily modified by simulating external factors, such as variable color utility (see [22]).

B. Effects of Systematic Individual Variation on Population Solutions

Jameson [53] suggested that adding nonrandom individual variation to a population by, for example, introducing individuals with some systematic X-linked photopigment deficiency, should influence population color categorization solutions. Webster and Kay suggested that to the extent that sources of individual variation within and across human populations (including physiologically based perceptual differences of the type simulated here) "… affect differences in color appearance, they are also important for understanding the similarities or universal tendencies in color naming" ([6], p. 34).

Studies 1, 2, and 3 support the suggestion that individual variation—especially nonrandom variation—influences population color categorization solutions. Results from [23] show that homogeneous populations of realistic normal observers produce category solutions that differ in important ways from those produced by the heterogeneous populations simulated here. Further differences were also seen when individual differences were made more frequent and more diverse to approximate realistic observer variation to greater degrees.

Despite the influences of individual differences, results from the present local and global confusion analyses suggest one possibly universal color-naming tendency that emerges: evolved systems tend to minimize the likelihood that colors perceptually confusable by *some individuals* in a population will tend to be classified by the entire population into different color categories [53].

Such tendencies in color naming solution dynamics were generally observed even when the underlying population response models tested had different perceptual biases. This undermines the suggestion that uniformities in color naming necessarily imply that the color relevant aspects of physiology are also uniform (see [6], p. 34).

C. Influences from Anomalous Trichromat Variation

Study 2 examined whether FM100 performance patterns that differentiate anomalous trichromats from normal trichromats were shown to affect color categorization solutions formed by a population containing both observer

types. Study 2 results suggest that (with the exception of a tendency to refine category boundary locations) anomalous trichromat agents did not have any unique, differentiable effect on population color categorization solutions. However, note that Study 2 examined results from populations that included both normal and deficient agents (population A_2) compared against results from populations that additionally included anomalous trichromat agent models (population B_2). In both kinds of populations it is likely that (i) strong dichromat biases likely dominate any potential for anomalous agent symmetry-breaking effects and (ii) finding an anomalous trichromat influence depends critically on the degree to which FM100 performance is capable of capturing performance differences between dichromats and anomalous trichromats—a limitation of the FM100 that has been previously noted [54]. The net result for the Study 2 simulations is that the presence of anomalous agents just makes dichromat symmetry-breaking effects stronger, because the anomalous agents' FM100 response patterns reinforce the dichromat biases that serve as the basis for boundary locations.

The present findings for anomalous trichromat agents underscore further the fact that the present modeling of observer types is necessarily limited by the FM100's ability to differentiate observer types assessed. To the degree that the FM100 test and analysis does not unequivocally differentiate, for example, an anomalous trichromat from a dichromat observer, this ambiguity is inherent in the data used to model our agent populations. It seems likely that if observer models were constructed on data that more decisively distinguished anomalous trichromats from dichromats, a more specific set of anomalous trichromat influences on categorization solutions could be observed. Thus, anomalous trichromats may be more influential to such category solutions than those revealed in FM100-based investigations.

Despite the strong ties between our simulated category solutions and the data used for observer modeling, the results for symmetry-breaking and local and global confusion influences on boundary locations are unavoidable consequences arising from the hue circle communication game described here. Whether such results are also found in categorization solutions on three-dimensional color spaces is of additional interest, especially in view of expected observer model complexities, such as variation in saturation discrimination across observer types [55,26].

D. Influences from Expert Agents' Nonnormative FM100 Performance

Generally, Study 3's expert observers exhibit FM100 performance similar to that of anomalous trichromats in that they have a large degree of diffuse color sorting error around the FM100 hue circle. However, the data used to construct expert agent models differ in important ways from the data used to construct our anomalous trichromat models.

First, as seen in Fig. 4, the FM100 confusion regions in expert-observer average performance plots do not coincide with the confusion regions for protan and deutan defects or with anomalous trichromat local confusion regions observed.

Second, the expert group confusions include five caps in Tray 3 but are more pronounced in Tray 2 (caps 22 to 42) of the FM100 test, in the greenish-yellow region of the test. These two regions may have significance beyond a simple model of deficient sorting. For example, in eigenvector analysis of cone responses to reflectance spectra measures of the 85 caps of the FM100, Mantere and colleagues [25] showed that when the stimuli are scaled in a two-dimensional opponent color space [56], departures from a smooth color circle contour can be clearly seen in the range from approximately cap 22 through cap 40 ([25], Fig. 2, p. 2239). Disruptions in this region of the color circle are similarly seen when spectral reflectance measures of the FM100 stimuli are plotted in a Munsell color metric space [57].

Thus, subtle disruptions in the smoothness of the perceptual spacing of a color circle probably exist ([35], p. 568). One could speculate that small perceptual nonuniformities in some regions of the FM100 sequence might be ignored by normal color vision observers but noticed by observers with heightened color experience who produce regions of nonnormative sorting on portions of the FM100 and normative FM100 sorting elsewhere. For observers with four retinal photopigments, then, the possibility exists for compensatory heightened discrimination in other zones (which may not be picked up by the FM100, because of floor effects). In those regions, a compensatory heightened discrimination might dictate a subjective ordering of the FM100 samples that is at odds with the "correct" ordering recommended by the scoring manual [33].

It is also possible that despite aims to maintain isobrightness of the test, lightness differences in the FM100 caps might provide sorting advantages (see [36], p. 283), especially in individuals with long-wavelength-sensitive opsin gene polymorphisms (e.g., the five experts used in Study 3). Also, some combination of both subtle hue and lightness variation might contribute to the nonnormative expert observers' FM100 sorting. The suggestion that some perceptual nonuniformities in the FM100 might be accentuated for some observers but not others is consistent with other research [33,45,58].

FM100 modeling limitations, similar to those discussed above for anomalous trichromats, also exist for expert agents because the FM100 was not designed to assess individuals expressing more than three photopigment classes. For this reason, the test's normative scoring may not be an appropriate method for evaluating the perceptual differences that might follow from individuals with weak (retinal) tetrachromacy.

Still, Study 3's results are generally very informative. The suggestion that such expert color observers may occur, at relatively large frequencies, in heterogeneous populations without disrupting the stablization and form of a populations' shared category solution and without driving it to an indeterminant oscillating solution or a nonoptimal solution is a very useful result. It provides a proof of concept for human population heterogeneity by implying that extensions of a trichromatic constraint would not destroy a population's ability to communicate optimally about color.

E. Identifying Universal Tendencies in Color Categorization System Evolution

In analyses of 110 human color categorization systems from the World Color Survey (WCS), Webster and Kay observe *constellations of similar color categories* across languages and conclude that "… group differences place some limits on universal tendencies in color naming," further suggesting that the factors underlying individual differences may have strong universal tendencies ([6], p. 50). Consistent with these observations, the present simulations reveal universal tendencies in the evolution of color category solutions produced by artificial agent populations. The present focus on endogenous observer model variation suggests that some universal tendencies found here include the following:

(i) Unsystematic perceptual variation (e.g., random sorting error) minimally affects categorization solutions.

(ii) Systematic perceptual variation affects substantial constraints on solutions.

(iii) Countervailing mechanisms (arising from perceptual variation) trade off in the process of categorization solution stabilization.

(iv) Trade-off demands increase as perceptual constraints (e.g., color confusion pairs and regions) become more varied or more frequent in a population or engage larger areas of color space.

(v) Even small proportions of a population's observers with strong systematic perceptual bias (e.g., 1% of dichromats in an otherwise normal population) can influence shared population categorization solutions by making some regions of the color space better suited as population color category boundary locations.

(vi) Regardless of observer variations or realistically modeled population diversity, evolved categorization solutions typically optimize communication among *all population members*, as opposed to evolving solutions based on majority rule or a marked population specialization.

It seems likely that the categorization solution tendencies observed here suggest mechanisms that are similar to those contributing to universal tendencies in human color categorization. As modeled here, such mechanisms primarily involve communication dynamics and perceptual constraints. Presumably any influence on population categorization solutions also depends on the role played by cultural and linguistic influences on categorization systems [9,10,12,13]. Our existing simulation research previously investigated idealizations of exogenous environmental and sociocultural influences and similarly demonstrated strong symmetry breaking in simulations from color regions of increased salience or environmental color hot spots [21,22]. Such influences, and others that can be generalized from pragmatic constraints found in human color-naming situations, are strong candidates for modeling cultural and linguistic influences that universally exist in color communication and categorization situations. The present simulation research supports the suggestion that a variety of constraints and sources of influence give rise to universal tendencies in color categorization and that the more common sources of influence may in turn underlie color-naming similarities across human languages.

APPENDIX A: ANOMALOUS TRICHROMAT PARAMETER ESTIMATION

Table 3 presents parameter values for "anomalous trichromat" agents' sorting confusions and local confusions. These parallel Eq. (1) and Eq. (2) described elsewhere [23]. All parameter calculations for anomalous trichromat models introduced here are based on the data of Bimler *et al.* ([31], Table 1, p. 166) from 10 protanomalous and 17 deuteranomalous individuals diagnosed as severe or extreme anomalous trichromats (confirmed by anomaloscope, FM100, and pseudoisochromatic plate assessment).

Sorting errors. The method for calculating sorting error parameter values, p, were described in [23]. It is based on counting the total number of sequence inversions (inversions of two adjacent caps) needed to recreate the perfect cap ordering of 1 to 85 out of the individual's sorting data. Anomalous trichromats (Fig. 1) exhibit patterns of diffuse color confusion (even outside local confusion regions) where they show considerable sorting difficulties compared with the normal observers. Thus, to calculate the probability of sorting errors for our anomalous agents, an average transposition calculation was performed similar to that for the normals, with the computation restricted to sequence inversions outside the anomalous local confusion regions. This was based on the observation that protanomalous and deuteranomalous individuals experience two types of confusion errors. One type arises from a systematic bias in color space (consistent with either a shift of the peak sensitivity of the L photopigment curve toward the M photopigment (protanomalous) or a shift of the peak sensitivity of the M photopigment curve toward the L photopigment (deuteranomalous), which produces higher confusion in local confusion regions similar to that found for dichromats. The second type of error arises from the general tendency for diffuse color confusions outside these local confusion regions.

Accordingly, for protanomalous individuals, the probability of inversion was estimated at 26%, and for deuteranomalous individuals it was 24% per cap.

Local confusion regions. Anomalous trichromat agents' local confusion regions were derived from patterns of local confusion found in observers' data. Suppose that for consecutive circle positions z_{i-1},z_i,z_{i+1}, observer (I) has placed (generally nonconsecutive) cap numbers c_1,c_2,c_3. Then the Munsell scored error for individual (I) at position i is defined as $x_i^{(I)}=|c_1-c_2|+|c_2-c_3|$, and the individual's FM100 performance can be represented as a vector of 85 numbers corresponding to an individual's ordering of the test's 85 color caps, $\{x_1^{(I)},\ldots,x_{85}^{(I)}\}$, where the superscript (I) is used to distinguish different individuals. From these vectors for groups of protanomalous and deuteranomalous individuals, the *averaged* count of FM100 transpositions in ordering were calculated as follows: Denote by S_{PN} (S_{DN}) the set of all protanomalous (deuteranomalous) individuals, and by N_{PN} (N_{DN}) the respective numbers of anomalous individuals. Then the average FM100 performance vectors are given by

$$\bar{x}_i^{PN}=N_{PN}^{-1}\sum_{I\in S_{PN}}x_i^{(I)},\qquad \bar{x}_i^{DN}=N_{DN}^{-1}\sum_{I\in S_{DN}}x_i^{(I)}. \tag{A1}$$

Average FM100 performance vectors for protanomalous and deuteranomalous individuals are shown in Fig. 1. Note that 2 has been subtracted from the y-axis values of Fig. 1 because usually a FM100 score of 2 is the lowest possible score that any cap can obtain, and represents perfect serial order [[36], p. 570]; here, for simplicity, "no error" equals zero. Two "bumps" in each Fig. 1 plot are easily detected; they correspond to regions of local confusion. Boundaries of these confusion regions are given by continua for which average cap error scores exceed 1.5 and show the boundary positions of the local confusion regions. For protanomalous individuals, the confusion regions obtained are the bracketed ranges 14–21 and 57–71; for deuteranomalous individuals they are 10–22 and 51–67.

The confusion range. A width parameter, w, was also defined for cap sequence inversions that occur inside local confusion regions (see Table 3). To compute w for anomalous trichromats, we find the quantity $q \equiv (1/N)\Sigma_I(1/n)\Sigma_i|x_i^I - i|$, where the inner summation is over all the caps inside the confusion regions, n is the number of caps inside the confusion region, and N is the total number of protanomalous or deuteranomalous individuals. Then the value of w is evaluated as $w \approx \pi q^2$, (using $q \approx (\int e^{-x^2/w}|x|dx)/(\int e^{-x^2/w}dx) = \sqrt{w/\pi}$). The result is $q = 1.2$ for protanomalous and $q = 1.4$ for deuteranomalous agents, which gives roughly the value $w = 5$.

ACKNOWLEDGMENTS

The authors thank Louis Narens and Ragnar Steingrimsson for helpful suggestions, and David Bimler for providing data from [31]. Partial support was provided by National Science Foundation (NSF) grant NSF 07724228 from the Methodology, Measurement, and Statistics (MMS) Program of the Division of Social and Economic Sciences (SES). N. Komarova gratefully acknowledges support of a Sloan Fellowship.

REFERENCES AND NOTES

1. K. Uchikawa and R. M. Boynton, "Categorical color perception of Japanese observers: comparison with that of Americans," Vision Res. **27**, 1825–1833 (1987).
2. M. A. Webster, S. M. Webster, S. Bharadwadj, R. Verma, J. Jaikumar, J. Madan, and E. Vaithilingam, "Variations in normal color vision: III. Unique hues in Indian and U.S. observers," J. Opt. Soc. Am. A **19**, 1957–1962 (2002).
3. P. Kay and T. Regier, "Resolving the question of color naming universals," Proc. Natl. Acad. Sci. U.S.A. **100**, 9085–9089 (2003).
4. T. Regier, P. Kay, and R. S. Cook, "Focal colors are universal after all," Proc. Natl. Acad. Sci. U.S.A. **102**, 8386–8391 (2005).
5. B. Sayim, K. A. Jameson, N. Alvarado, and M. K. Szeszel, "Semantic and perceptual representations of color: evidence of a shared color-naming function," J. Cogn. Culture **5**, 427–486 (2005).
6. M. A. Webster and P. Kay, "Individual and population differences in focal colors," in *Anthropology of Color: Interdisciplinary Multilevel Modeling*, R. E. MacLaury, G. V. Paramei, and D. Dedrick, eds. (Benjamins, 2007), pp. 29–53.
7. C. L. Hardin and L. Maffi, eds., *Color Categories in Thought and Language* (Cambridge U. Press, 1997).
8. C. C. Moore, A. K. Romney, and T. Hsia, "Shared cognitive representations of perceptual and semantic structures of basic colors in Chinese and English," Proc. Natl. Acad. Sci. U.S.A. **97**, 5007–5010 (1997).
9. J. Davidoff, I. Davies, and D. Roberson, "Colour categories in a stone-age tribe," Nature **398**, 203–204 (1999).
10. D. Roberson, I. R. L. Davies, and J. Davidoff, "Color categories are not universal: Replications and new evidence from a stone-age culture," J. Exp. Psychol. Gen. **129**, 369–398 (2000).
11. K. A. Jameson and N. Alvarado, "Differences in color naming and color salience in Vietnamese and English," Color Res. Appl. **28**, 113–138 (2003).
12. D. Roberson, J. Davidoff, I. R. L. Davies, and L. R. Shapiro, "The development of color categories in two languages: a longitudinal study," J. Exp. Psychol. **133**, 554–571 (2004).
13. D. Roberson and J. R. Hanley, "Color vision: Color categories vary with language after all," Curr. Biol. **17**, R605–R607 (2007).
14. T. Belpaeme and J. Bleys, "Explaining universal color categories through a constrained acquisition process," Adapt. Behav. **13**, 293–310 (2005).
15. L. Steels and T. Belpaeme, "Coordinating perceptually grounded categories: A case study for colour," Behav. Brain Sci. **28**, 469–529 (2005).
16. L. D. Griffin, "The basic colour categories are optimal for classification," J. R. Soc., Interface **3**, 71–85 (2006).
17. Delwin T. Lindsey and Angela M. Brown, "Universality of color names," Proc. Natl. Acad. Sci. U.S.A. **103**, 16608–16613 (2006).
18. T. Regier, P. Kay, and N. Khetarpal, "Color naming reflects optimal partitions of color space," Proc. Natl. Acad. Sci. U.S.A. **104**, 1436–1441 (2007).
19. M. Dowman, "Explaining color term typology with an evolutionary model," Cogn. Sci. **31**, 99–132 (2007).
20. A. Puglisi, A. Baronchelli, and V. Loreto, "Cultural route to the emergence of linguistic categories," Proc. Natl. Acad. Sci. U.S.A. **105**, 7936–7940 (2008).
21. N. L. Komarova, K. A. Jameson, and L. Narens, "Evolutionary models of color categorization based on discrimination," J. Math. Psychol. **51**, 359–382 (2007).
22. N. L. Komarova and K. A. Jameson, "Population heterogeneity and color stimulus heterogeneity in agent-based color categorization," J. Theor. Biol. **253**, 680–700 (2008).
23. K. A. Jameson and N. L. Komarova, "Evolutionary models of color categorization. I. Population categorization systems based on normal and dichromat observers," J. Opt. Soc. Am. A **26**, 1414–1423 (2009).
24. D. Farnsworth, *The Farnsworth–Munsell 100 Hue Test for the Examination of Color Vision* (Munsell Color Company, 1949/1957).
25. K. Mantere, J. Parkkinen, M. Mäntyjärvi, and T. Jaaskelainen, "Eigenvector interpretation of the Farnsworth–Munsell 100-hue test," J. Opt. Soc. Am. A **12**, 2237–2243 (1995).
26. J. Birch, *Diagnosis of Defective Colour Vision*, 2nd ed. (Butterworth-Heinemann, 2001).
27. J. H. Nelson, "Anomalous trichromatism and its relation to normal trichromatism," Proc. Physiol. Soc. **50**, 661–702 (1938).
28. J. Pokorny, V. C. Smith, G. Verriest, and A. J. L. G. Pinckers, *Congenital and Acquired Color Vision Defects* (Grune & Stratton, 1979).
29. G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed. (Wiley, 1982).
30. L. T. Sharpe, A. Stockman, H. Jägle, and J. Nathans, "Opsin genes, cone photopigments, color vision, and color blindness," in *Color Vision: From Genes to Perception*, K. R. Gegenfurtner and L. T. Sharpe, eds. (Cambridge U. Press, 1999), pp. 3–51.
31. D. Bimler, J. Kirkland, and R. Jacobs,"Colour-vision tests considered as a special case of multidimensional scaling," Color Res. Appl. **25**, 160–169 (2000).
32. K. A. Jameson, S. M. Highnote, and L. M. Wasserman, "Richer color experience in observers with multiple

photopigment opsin genes," Psychon. Bull. Rev. **8**, 244–261 (2001).
33. K. A. Jameson, D. Bimler, and L. M. Wasserman, "Reassessing perceptual diagnostics for observers with diverse retinal photopigment genotypes," in *Progress in Colour Studies 2: Cognition*, N. J. Pitchford and C. P. Biggam, eds. (Benjamins, 2006), pp. 13–33.
34. Farnsworth–Munsell Scaling Software, Version 2.1 (MacBeth Division of Kolmorgen Corporation, 1997).
35. D. Farnsworth, "The Farnsworth–Munsell 100-Hue and Dichotomous Tests for color vision," J. Opt. Soc. Am. **33**, 568–578 (1943).
36. S. J. Dain, "Clinical colour vision tests," Clin. Exp. Optom. **87**, 276–293 (2004).
37. G. V. Paramei, D. Bimler, and N. O. Mislavskaia, "Colour perception in twins: Individual variation beyond common genetic inheritance," Clin. Exp. Optom. **87**, 305–312 (2004).
38. G. V. Paramei, D. L. Bimler, and C. R. Cavonius, "Effect of luminance on color perception of protanopes," Vision Res. **38**, 3397–3401 (1998).
39. J. Pokorny and V. C. Smith, "Evaluation of single-pigment shift model of anomalous trichromacy," J. Opt. Soc. Am. **67**, 1196–1209 (1977).
40. J. Nathans, D. Thomas, and D. S. Hogness, "Molecular genetics of human color vision: the genes encoding blue, green and red pigments," Science **232**, 193–202 (1986).
41. J. Nathans, T. P. Piantanida, R. L. Eddy, T. B. Shows, and D. S. Hogness, "Molecular genetics of inherited variation in human color vision," Science **232**, 203–210 (1986).
42. A. B. Asenjo, J. Rim, and D. D. Oprian, "Molecular determinants of human red/green color discrimination," Neuron **12**, 1131–1138 (1994).
43. S. Yokoyama and F. B. Radlwimmer, "The 'Five-Sites' Rule and the evolution of red and green color vision in mammals," Mol. Biol. Evol. **15**, 560–567 (1998).
44. G. Jordan and J. D. Mollon, "A study of women hererozygous for colour deficiencies," Vision Res. **33**, 1495–1508 (1993).
45. D. Bimler and J. Kirkland, "Colour-space distortion in women who are heterozygous for colour deficiency," Vision Res. **49**, 536–543 (2009).
46. S. M. Hood, J. D. Mollon, L. Purves, and G. Jordan, "Color discrimination in carriers of color deficiency," Vision Res. **46**, 2894–2900 (2006).
47. J. Birch, "Extreme anomalous trichromatism," in *Normal and Defective Colour Vision*, J. D. Mollon, J. Pokorny, and K. Knoblauch, eds. (Oxford U. Press, 2003), pp. 364–369.
48. Subjects 27, 52, 58, 61, and 85 reported in [33]: ages 18–21 years, with above average chromatic banding behaviors [$\mu(medianbands)=10$ versus 7.9 for controls].
49. K. Jameson and R. G. D'Andrade, "It's not really red, green, yellow, blue: An inquiry into cognitive color space," in *Color Categories in Thought and Language*, C. L. Hardin and L. Maffi, eds. (Cambridge U. Press, 1997), pp. 295–319.
50. K. A. Jameson, "Culture and cognition: What is universal about the representation of color experience?" J. Cogn. Culture **5**, 293–347 (2005).
51. T. Regier, P. Kay, and N. Khetarpal, "Color naming reflects optimal partitions of color space," Proc. Natl. Acad. Sci. U.S.A. **104**, 1436–1441 (2007).
52. K. A. Jameson, "Sharing perceptually grounded categories in uniform and nonuniform populations," Behav. Brain Sci. **28**, 501–502 (2005).
53. Fewer confusion pairs make minimization more likely.
54. J. Birch, "Use of the Farnsworth–Munsell 100-hue test in the examination of congenital colour vision defects," Ophthalmol. Physiol. Opt. **9**, 156–162 (1989).
55. G. V. Paramei, "Color space of normally sighted and color-deficient observers reconstructed from color naming," Psychol. Sci. **7**, 311–317 (1996).
56. With axes corresponding to the second and third vectors of the Eigen solution, namely, a Y–B axis and a R–G+B axis, respectively.
57. A. K. Romney, "Relating reflectance spectra space to Munsell color appearance space," J. Opt. Soc. Am. A **25**, 658–666 (2008).
58. D. Bimler, J. Kirkland, and K. A. Jameson, "Quantifying variations in personal color spaces: Are there sex differences in color vision?" Color Res. Appl. **29**, 128–134 (2004).