Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 5805–5808, September 1985
Evolution

Color categorization: A possible concordance between genes and culture

(evolution/psychology/socialization/ethnography/gene–culture coevolution)

CHARLES J. LUMSDEN

Department of Medicine, University of Toronto, Toronto, Ontario, Canada

Communicated by Edward O. Wilson, April 11, 1985

ABSTRACT **A marked correspondence is found to relate the categories of infant wavelength discrimination and the ethnographic distribution of adult color terms. The structure of the infant category system accounts for at least 75% of the variance in the world ethnographic categories ($P < 0.001$). Such a correspondence is predicted to be favored by natural selection when perceptual mechanisms interact with their associated cultural traits.**

Although a reciprocal interaction between genetic and cultural evolution has been widely discussed (1–8), the precise nature of the linkage remains uncertain. Gene–culture theories (1, 3, 4, 6, 9, 10) suggest that in circumstances involving adaptive cultural traits of sufficient stability, a strong effect of genes on culture will be favored. In this mode of relationship innate constraints on behavioral development significantly influence not only the direction but the statistical distribution of cultural diversity as well. The cultural patterns of information and learned behavior in turn form an environment in which natural selection acts on genes, and a marked correlation between biology and culture is expected to exist (3). The effects are predicted to be especially noticeable in the case of traits that reflect the more automatic processes (11) leading from sensory filtering to perception. The organizing effects that these processes exert on the development of behavior have been termed primary epigenetic rules, and possible instances involving the different sensory modalities have been documented (3).

To investigate the existence of such a concordance between an ethnographic distribution and a primary epigenetic rule, the case of color terms was considered. The categorization and naming of colors is a cultural universal (12–14). Although lexicon size varies, all human societies so considered have terms for colors or degrees of light and dark (13, 14). The explanation of the ethnographic distribution of color terms and their evolution is considered an important anthropological problem (12–15). The neurophysiology of color perception has also been extensively studied using human subjects and primate models (16, 17). Discussion of ethnographic distributions of color naming has therefore taken on an increasingly biological tone (13–15, 18, 19). Moreover, the study of color categorization has recently been extended to very young infants prior to the onset of language acquisition (20–23). It is now possible to compare the performance of humans who have not yet begun color term socialization through language with that of fully enculturated adults. Qualitative correspondences between the color categories perceived by young infants and by the adult populations examined by psychologists have been previously noted (3, 24).

METHODS

The present study, which presents quantitative measures of the degree of relationship between infant and ethnographic categories, used the specifications of infant color categorization reported by Bornstein and his co-workers (20–26). Sixteen-week-old infants were found to respond to variation in wavelength as though four basic spectral color categories were being discriminated. Conventional color terms for these categories would be *red*, *yellow*, *green*, and *blue*. The categorization was detected by measuring the span of the infants' attention to monochromatic lights ordered across the visible spectrum. Within a short time the infants habituated to the repeated presentation of the stimulus light. Recovery from habituation was strong only when changes in wavelength crossed certain wavelength values, which were judged to be the boundary regions between perceptual categories of color. Boundaries between the categories were mapped, together with the perceptual responses near the category centers.

To quantify the relationship between the infant data and the ethnographic observations, boundary values between infant categories were identified: *red–yellow* at 600 nm, *yellow–green* at 560 nm, and *green–blue* at 480 nm. The values given correspond to the peaks of maximum wavelength discriminability in the boundary regions (21) and to the midrange crossover points between the respective categories as measured by Bornstein *et al.* (24, 25). The width of the boundary regions is in all cases narrow: 20 nm for *red–yellow*, 10 nm for *yellow–green*, and 20 nm for *green–blue* (24, 25).

The ethnographic observations were drawn from the data on color naming developed by Berlin and Kay (12). In the Berlin–Kay (BK) study, native speakers of 20 languages were shown arrays of patches ordered in color and brightness by the Munsell system. A list of basic color terms (terms operationally defined to include such characteristics as monolexemic structure and broad applicability: see ref. 12, p. 6) was elicited from each subject. Subjects were then asked to place each basic color term of their native language on the two-dimensional color chart by picking the color patch or patches that best represented each term. Except for the 40-member Tzeltal group, the subjects were bilingual in their native language and American English. The Tzeltal subjects varied from monolingual fluency in Tzeltal to Tzeltal–Spanish bilingualism. The close accord observed between the response patterns of the Tzeltal speakers and those of the other informants suggested that biases arising from bilingual ties to American English (27) were not substantial (12). The center of gravity of the best exemplars was plotted for each basic color term in each language. The resulting map (ref. 12, figure 3) was the starting point for the ethnographic tabulations used in the present study.

The BK map was analyzed in the following way. The Munsell brightness axis holds the dominant wavelength of the

---

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: BK, Berlin–Kay.

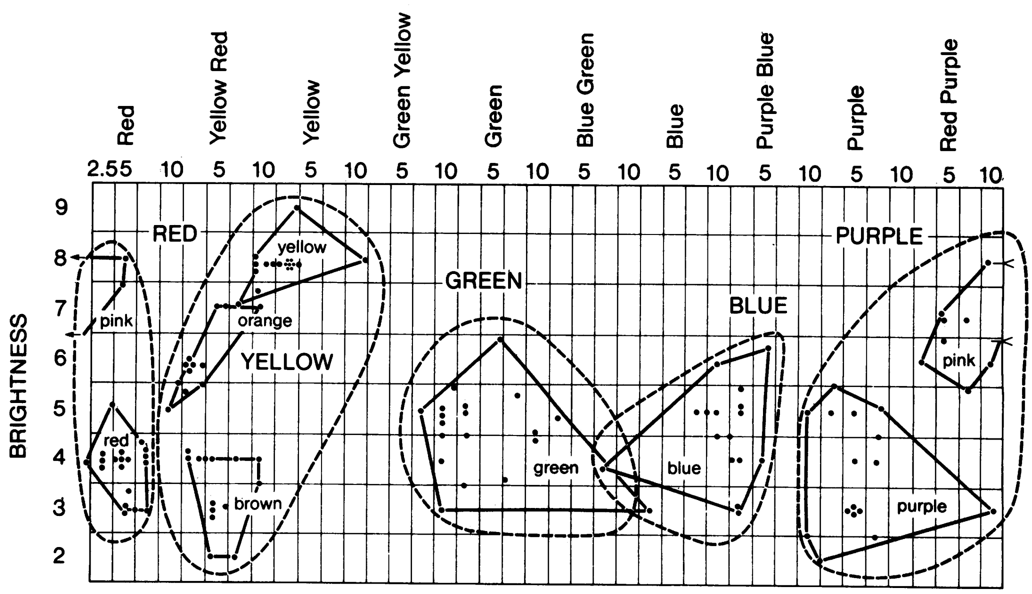

FIG. 1. Clustering of the BK color-term groups into four major spectral hue categories: RED, YELLOW, GREEN, and BLUE. PURPLE corresponds to a nonspectral purple and was omitted from the analysis (see text). Solid polygons, original groups according to Berlin and Kay (12). Dashed lines, proposed clusterings.

color patches approximately constant (Table 1). To allow tests of correspondence with the four infant spectral categories, the BK color groups were clustered on the basis of vertical order to bring together all sample points in a given wavelength range. This procedure led to BK red forming RED (R), BK brown, orange, and yellow forming YELLOW (Y), BK green forming GREEN (G), and BK blue forming BLUE (B) (Fig. 1). Lateral boundaries for the clusters were those reported by Berlin and Kay for their original color groups (12). The BK group purple and all of the points in BK pink except two were omitted from the analysis for two reasons. First, the infant category data were restricted to the spectral categories *red*, *yellow*, *green*, and *blue*, involving the *red–yellow*, *yellow–green*, *green–blue* boundary values (20–26). A category of spectral violet did not appear to be clearly distinguished from spectral blue by these young subjects. Second, purple in the Munsell system does not appear to be a unitary phenomenon corresponding to spectral violet. It resembles a nonspectral purple that uses an assortment of dominant wavelengths ranging across all four basic spectral color categories: *red*, *yellow*, *green*, and *blue*. Two data points from BK pink fell in the Munsell patch (R5, 8), vertically aligned with BK red, and were included in RED.

Charts (28) of the dominant wavelength and excitation purity (saturation of the color patch emission by the dominant wavelength) for the Munsell color notations were kindly provided by Munsell Color (Baltimore, MD). The brightness and hue coordinates were read from the BK map for each data point observable on it. Chroma values were not available for the data. Maximum chroma, or saturation of the Munsell patch stimulus by the dominant wavelength, was assumed. With few exceptions, to be noted below, dominant wavelength values changed by <10% along the chroma axis of the Munsell system at the brightness and hue values of the BK data. Dominant wavelength and excitation purity were recorded for each point with the Munsell charts being used to look up the corresponding coordinates of hue, brightness, and chroma. The results are shown in Table 1.

RESULTS

Now transformed into wavelength equivalents, the BK data can be directly compared to psychological studies of infant categorization. The qualitative correspondence between the infant category structure and the BK ethnographic distribution proves to be quite marked (Fig. 2). The boundaries of the ethnographic clusters align closely with those of the infant categories. In Fig. 2 the histograms of sample points represent distributions of wavelengths that speakers of different languages judge to be best examples of particular colors. Thus, Fig. 2 presents wavelength ranges that infants group as similar in hue against wavelength ranges that adults from different linguistic communities each label the same way.

To quantify the relationship, each of the data points on the BK map was assigned to an infant category on the basis of its dominant wavelength compared to the values selected for the boundaries between the infant categories. The results were organized into a 4 × 4 cross-tabulation of the infant categories and the categories of the BK ethnographic distribution as shown below (the actual sample counts are shown in boldface; sample counts are given as row percents in brackets and as column percents in parentheses).

Cross-cultural

	R	Y	G	B
red	**11** [100] (64.7)			
yellow	**2** [4.3] (11.8)	**44** [95.7] (100.1)		
green	**4** [18.2] (23.5)		**18** [81.8] (94.7)	
blue			**1** [5.9] (5.3)	**16** [94.1] (100.0)

In this form the data show a striking diagonal order, with most of the ethnographic BLUE points falling in the cell corresponding to infant *blue* and similarly through GREEN, YELLOW, and RED. The strength of this association is highly significant: $\chi^2 = 246.5$, degrees of freedom = 9, $P < 0.001$; Spearman's $\rho = 0.88$; Kendall's tau*b* = 0.86. The measures of category association indicate that >75% of the

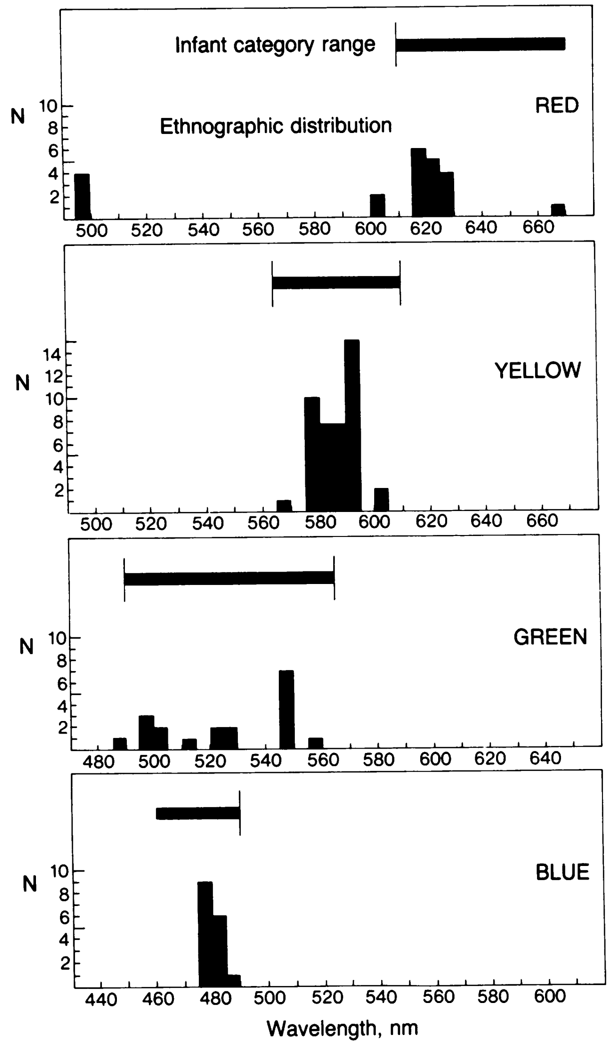

FIG. 2. Alignment of the infant color categories (24, 25) and the BK (12) ethnographic distribution of basic color terms.

variance in the membership in the ethnographic color categories is accounted for solely by the prelinguistic infant color categories.

The assumption of maximum chroma at each Munsell hue and brightness value led to the four sample points at Munsell (R2.5, 4) being assigned a dominant wavelength of 493 nm (*green*). In contrast, the dominant wavelength values at lower chroma values for (R2.5, 4) all lie in the spectral region 625–690 nm. These points were the only data in the entire sample at which such a large discrepancy across chroma was noted and at which the usefulness of the maximum-chroma assumption is questionable. In order to test the effects of moving the (R2.5, 4) points into the RED group, which would be the case if the original (R2.5, 4) data did not correspond to maximum chroma, the dominant wavelength value of 493 nm was replaced by a value in the range 625–690 nm. The strong concordance between the infant color categories and the ethnographic clusters was observed to further increase: χ^2 = 278.9, degrees of freedom = 9, $P < 0.001$; Spearman's ρ = 0.98; Kendall's tau$b$ = 0.98. With this partitioning the residual ethnographic variance unaccounted for by the infant hue categories is ≈4%.

DISCUSSION

The results establish the existence of a marked concordance of the type predicted by gene–culture theories to relate primary epigenetic rules and their associated cultural traits.

Table 1. Dominant wavelength and excitation purities evaluated for the BK color-term data

Group	Munsell coordinates (hue, brightness)	Number of points	Dominant wavelength, nm[*]	Maximum excitation purity, %
		RED		
BK red	R2.5, 5	1	665 (45)	46.5
	R2.5, 4	4	493 (200)[†]	57
	R5, 4	5	624 (16)	68
	R7.5, 4	4	615 (14)	86.5
	R5, 3	4	629 (19)	60.5
	R7.5, 3	2	618 (16)	81
BK pink	R5, 8	2	603 (1)	19.5
		YELLOW		
BK yellow	Y2.5, 9	1	578 (1)	43.5
	Y2.5, 8	9	579 (2)	93.5
	YR10, 8	5	582 (3)	86
	YR10, 7	1	582 (2)	91
	YR7.5, 7	1	585 (2)	94
	GY2.5, 8	1	569.5 (0.5)	81
BK orange	YR10, 7	1	582 (2)	91
	YR5, 7	2	588 (3)	84.5
	YR2.5, 6	5	593 (4)	91.5
	YR2.5, 5	2	594 (5)	91.5
	R10, 5	2	600.5 (7.5)	92
BK brown	YR2.5, 4	3	594 (3)	82
	YR5, 4	2	589.5 (2.5)	75
	YR7.5, 4	1	586 (1.5)	80
	YR10, 4	1	583 (2)	83.5
	YR5, 3	4	590 (2)	70
	YR7.5, 3	1	587 (2)	74.5
	YR5, 2	1	591.5 (1.5)	57.5
	YR7.5, 2	1	587 (1)	64
		GREEN		
BK green	G5, 6	1	512 (5)	22
	G2.5, 5	2	525.5 (9.5)	37.5
	G7.5, 5	1	504 (3)	29.5
	G10, 5	1	499 (1)	33.5
	G2.5, 4	1	523.5 (8.5)	36
	G7.5, 4	1	503 (2.5)	36
	G10, 4	1	499 (1)	41
	G2.5, 3	1	520 (11)	39.5
	GY7.5, 5	1	558 (2)	72
	GY10, 5	4	548 (4)	66
	GY10, 4	2	548.5 (3.5)	50
	GY10, 3	1	547 (3)	39.5
	BG2.5, 5	1	496 (1)	39
	BG10, 3	1	486 (1)	63
		BLUE		
BK blue	BG7.5, 4	1	489 (1)	49
	PB5, 6	1	476 (1)	44
	PB2.5, 5	3	478 (3)	62
	PB2.5, 4	2	478 (1)	61.5
	PB5, 4	1	475 (1)	58.5
	PB2.5, 3	2	477 (1.5)	73.5
	B10, 6	1	481 (1)	49.5
	B7.5, 5	2	482 (2)	58
	B10, 5	1	480 (2)	65.5
	B10, 4	2	480 (1.5)	65
		OTHER ("PURPLE")		
BK pink	RP10, 8	1	650 (5)	12
	RP7.5, 7	1	495 (1)	20
	RP5, 7	3	498 (0)	26
	RP2.5, 6	1	510 (2)	37
	RP10, 6	1	493 (0)[‡]	33
	RP5, 5	1	499 (0)	43
BK purple	PB10, 5	1	455 (3)	39
	PB10, 2	2	444 (3)	60.5
	P2.5, 5	2	566 (0)	36
	P5, 5	1	561 (1)	36.5
	P7.5, 5	1	551 (1)	37
	P5, 4	1	561 (0)	48.5
	P7.5, 4	2	553 (0)	50
	P5, 3	4	561 (0)	52.5
	P7.5, 2	1	554 (0)	43
	PR10, 3	1	493 (2)	54.5

[*] Parentheses list the variation in nm over the chroma range at the given hue and brightness.

[†] Maximum chroma occurred at 493 nm; the dominant wavelength values for lower chroma values of this patch varied between 625 and 690 nm.

[‡] Lowest chroma values have dominant wavelength values of 675, 690, and 700 nm at purities of 2.8%, 4.5%, and 9.5%. These wavelengths were omitted from the cited range of variation.

They suggest that the wavelength clustering of the BK ethnographic color-term distribution can be parsimoniously explained on the basis of a primary epigenetic rule (3) serving color category development. This epigenetic rule involves procedures of color discrimination present since early infancy. Other findings support such a conclusion. Bornstein has related the centers of the BK clusters to response peaks of neurons in the noncortical visual systems serving color processing (18). Similar correspondences have been documented at length by Kay and McDaniel (14). Investigation of color categorization has also been extended to nonhuman primates. The macaque, whose visual system appears neurophysiologically similar to the human case, categorizes the spectrum into similar basic categories of *red, yellow, green*, and *blue* (26). Whether these properties are shared on the basis of parallel adaptation or descent from a common ancestor remains to be determined.

The fact that the basic hue categories of infants, which match those of the macaque, account for most of the variance in the basic color categories of the ethnographic distribution does not imply that culture is irrelevant or plays no role. Color terms themselves are part of the language system and are transmitted culturally. Human societies differ greatly in the complexity of their color lexica and in the ways in which color terms are used in social exchange (12, 13, 15). The primary epigenetic rule does not address the determinants of this complexity and the associated systems of conventions. The social evolution of color-term systems is also an involved process whose regularities might be explained by a combination of cultural and biological mechanisms (13–15, 18, 19). However, a primary epigenetic rule for basic color categories would indicate that these important cultural processes have been occurring in a manner consistent with innate constraints operating since early infancy.

I thank Christine Littlefield for discussions on statistics and the referees for points of stylistic clarification. The manuscript was prepared with care and skill by Anne Hansen-Johnston. This work was supported by the Natural Sciences and Engineering Research Council of Canada. The author is a Scholar of the Medical Research Council of Canada.

1. Richerson, P. R. & Boyd, R. (1978) *J. Soc. Biol. Struct.* **1**, 127–154.
2. Pulliam, H. R. & Dunford, C. (1979) *Programmed to Learn: An Essay on the Evolution of Culture* (Columbia Univ. Press, New York).
3. Lumsden, C. J. & Wilson, E. O. (1981) *Genes, Mind and Culture: The Coevolutionary Process* (Harvard Univ. Press, Cambridge, MA).
4. Durham, W. H. (1982) *Am. J. Phys. Anthropol.* **57**, 183 (abstr.).
5. Freeman, F. (1983) *Margaret Mead and Samoa: The Making and Unmaking of an Anthropological Myth* (Harvard Univ. Press, Cambridge, MA) pp. 295–302.
6. Lumsden, C. J. & Wilson, E. O. (1983) *Promethean Fire: Reflections on the Origin of Mind* (Harvard Univ. Press, Cambridge, MA).
7. Swanson, C. P. (1983) *Ever-Expanding Horizons: The Dual Information Sources of Human Evolution* (Univ. of Massachusetts Press, Amherst, MA).
8. Lopreato, J. (1984) *Human Nature and Biocultural Evolution* (Allen and Unwin, Boston, MA).
9. Feldman, M. W. & Cavalli-Sforza, L. L. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1604–1607.
10. Lumsden, C. J. (1984) *Ethol. Sociobiol.* **5**, 111–124.
11. Fodor, J. A. (1983) *The Modularity of Mind* (MIT Press, Cambridge, MA).
12. Berlin, B. & Kay, P. (1969) *Basic Color Terms* (Univ. of California Press, Berkeley, CA).
13. Witkowski, S. R. & Brown, C. H. (1977) *Am. Anthropol.* **79**, 50–57.
14. Kay, P. & McDaniel, C. K. (1978) *Language* **54**, 610–646.
15. Ratliff, F. (1976) *Proc. Am. Philos. Soc.* **120**, 311–330.
16. Mollon, J. D. (1982) *Annu. Rev. Psychol.* **33**, 41–85.
17. Lennie, P. (1984) *Trends Neurosci.* **7**, 243–248.
18. Bornstein, M. H. (1973) *Psychol. Bull.* **80**, 257–285.
19. Whitfield, T. W. A. (1981) *Percept. Psychophys.* **29**, 87–90.
20. Bornstein, M. H. (1975) *J. Exp. Child Psychol.* **19**, 401–419.
21. Bornstein, M. H. (1979) in *Psychological Development from Infancy: Image to Intention*, eds. Bornstein, M. H. & Kessen, W. (Lawrence Erlbaum, Hillsdale, NJ), pp. 37–81.
22. Bornstein, M. H. (1981) *Infant Behav. Dev.* **4**, 119–150.
23. Bornstein, M. H. & Marks, L. E. (1982) *Psychol. Today* **16**, 64–73.
24. Bornstein, M. H., Kessen, W. & Weiskopf, S. (1976) *Science* **191**, 201–202.
25. Bornstein, M. H., Kessen, W. & Weiskopf, S. (1976) *J. Exp. Psychol. Human Percept. Perform.* **2**, 115–129.
26. Sandell, J. H., Gross, C. G. & Bornstein, M. H. (1979) *J. Comp. Physiol. Psychol.* **93**, 626–635.
27. Conklin, C. H. C. (1973) *Am. Anthropol.* **75**, 931–942.
28. Anonymous (1976) *Dominant Wavelength and Excitation Purity for Designated Munsell Color Notations* (Munsell Color Company, Baltimore, MD).