Universality of Color Categorization[1]

Paul Kay
U.C. Berkeley (Linguistics), Stanford U. (Linguistics)

To appear in *Handbook of Color Psychology.* Andrew J. Elliot & Mark D. Fairchild (eds.)
Cambridge University Press.

Introduction

In 1965 two freshly minted linguistic anthropology Ph. Ds were chatting about their recent respective fieldwork experiences, one in Chiapas, Mexico, among speakers of the Mayan language Tzeltal the other in Tahiti. No two languages could be less related to each other or to the languages of Europe, Asia, and the Mediterranean with long written histories. Both researchers had learned that the Sapir-Whorf hypothesis is established fact and that color vocabulary is the *locus classicus* of its empirical demonstration – as reflected in confident statements like the following.

> . . . there is no such thing as a natural division of the spectrum. Each culture has taken the spectral continuum and has divided it upon a basis which is quite arbitrary (Ray 1952).

As graduate students heading for the field each had expected that the meanings of the color words of the language he encountered would be hard to learn, since every language (or language family) segments the spectrum in its own arbitrary way. But as the conversation progressed they discovered that they had both had the same disconcerting experience in the field. Each one had found to his surprise that the color words of the language he encountered were easy to learn. In fact it turned out that the major color words in both Tzeltal and Tahitian corresponded closely in

---

[1] I am grateful to Anna Franklin for helpful comments on this ms. The usual disclaimers apply.

[2] The chapter will not discuss attempts that have been made to *explain* these findings in terms of

meaning to the basic color words of English (or French, or Mandarin, …) – with a single exception. And in the one respect in which each of the two unwritten languages differed from English and other familiar languages they differed in the same way: Tzeltal and Tahitian each had words that readily translates as black, (*noir, hēi sè,* ...), white, red, and yellow, and both Tzeltal and Tahitian had a single word that covered both green and blue. That surprising set of facts motivated some research that eventually led to a small study in which it two hypotheses were offered.

> [1] … although different languages encode in their vocabularies different *numbers* of basic color categories, a total inventory of exactly eleven basic color categories exists from which the eleven or fewer basic color categories of any given language are always drawn,

and

> [2] If a language encodes fewer than eleven basic color categories, then there are strict limitations on which categories it may encode (Berlin & Kay 1969 [B&K]: 2).

The details of these two broad findings that were proposed in 1969 have been considerably modified by further research, as this chapter will discuss. The broad findings themselves have, however, stood up over the years. This chapter will outline both the modifications that have been made to these claims in the last near half-century and the relevant evidence.[2]

---

[2] The chapter will not discuss attempts that have been made to *explain* these findings in terms of human vision, the color statistics of the environment, the exigencies of human communication, the vagaries of cultural transmission, or any combination of these or other factors.

Basic Color Terms (1969)

Berlin and Kay and members of a graduate seminar interviewed native speakers of twenty different languages resident in the San Francisco Bay area using a set of stimuli approximated by the palette in Figure 1.

INSERT FIG 1

For each of the major ("basic") color terms of his or her language every participant was asked to point out the best example(s) of the term ("focal color") and to indicate the term's extent. The color-naming data obtained in this way were supplemented by seventy-eight reports of color naming systems from the literature. Based on these data, B&K proposed that all the color terms in the ninety-eight languages considered could be described in terms of eleven basic categories, black, white, red, yellow, green, blue, brown, purple, pink, orange, and gray and their combinations. They tacitly assumed that every language has a set of basic color terms that partitions the perceptual color space, an assumption that later had to be modified. They proposed that if a language has only two basic color terms, these terms are black and white; if it has three they are black, white, and red, if four either black, white, red, and green or black, white, red, and yellow, and so on as indicated in Figure 2.

INSERT FIG 2

Figure 2. The B&K hypothesis regarding possible color naming systems (Source B&K, p.4)

There is an implied contradiction implied in the last two sentences, which was obscured by an equivocation in B&K. If every language has a set of terms that partitions the color space – as stated in the first sentence – and some languages have, for example, only terms for black, white and red – as indicated in the second sentence – the contradiction lies in the fact that *black, white,* and *red* do not partition the color space. B&K equivocated by sometimes talking about black, white, red, etc. and sometimes about BLACK, WHITE, RED, etc., the latter assumed to be composed of more colors than just black, or white, or red, ... Color words written in capitals were intended to denote *composite* categories. Composite categories are composed of a set of adjacent simple categories, the names of which were written in lower case. The sequence in Figure 2 was written in lower case but was in fact discussed using a mix of composite categories – which should have been written in caps in Figure 2 – and simple categories.

The encoding sequence post-1969

The situation was subsequently clarified as follows: BLACK was defined as comprising black, green, and blue; WHITE was defined as comprising white, red, and yellow; RED was defined as comprising red and yellow; and GRUE (a neologism) was defined as comprising green and blue (Kay 1975).[3] The sequence in Figure 3 was proposed to make clear that the temporal progression

---

[3] B&K had written that the original BLACK and WHITE named all dark colors and all light colors, respectively. Based on the work of Eleanor Rosch (Heider 1972a,b), Kay (1975) proposed the definitions in the text, which have remained apparently valid, although no two-term system other than the Dugum Dani reported by Rosch (Heider 1972a,b) has ever been experimentally documented. Several three-term systems have however been carefully attested, with terms for (1) black, or green, or blue, or other ‘cool’ colors, (2) red, or yellow, or other ‘warm’ colors, and (3) white and some light shades of chromatic colors.

in color-naming systems involves the successive refinements of an original partition of the color space rather than the progressive naming of previously unnamed regions.[4]

INSERT FIG 3

Not long after, the capital letter names were abandoned in favor transparent disjunctions of primaries, as in Figure 4.

INSERT FIG 4

The World Color Survey

As reported in Cook, Kay & Regier (2005: 224):

> The B&K results were immediately challenged … on the grounds that the sample of experimental languages was too small, too few collaborators per language were questioned (sometimes only one), all native collaborators also spoke English, the data were collected in the San Francisco Bay area rather than in the homelands of the target languages, certain regions of the world and language families were underrepresented or overrepresented in the sample of 20, and the sample of 20 had too few unwritten languages of low-technology cultures [Hickerson (1971), Durbin (1972), Collier (1973), Conklin (1973)]. The results were nevertheless supported by various ethnographic and experimental studies conducted after 1979 [Footnote: For example, Berlin and Berlin (1975), Dougherty (1975, 1977), Hage and Hawkes (1975), Harkness (1973), Heider (1972a, 1972b), Heider and Olivier (1972), Heinrich

[4] Although, as mentioned above and discussed below, this general tendency was later discovered to not necessarily repreent the initial state of the history of every color lexicon.

(1972), Kuschel and Monberg (1974), MacLaury (1986, 1987, 1997), Maffi (1990b), Monberg (1971), Senft (1987), Snow (1971), and Turton (1978, 1980).and were largely accepted by psychologists and vision researchers [e.g., Brown (1976), Miller and Johnson-Laird (1976), Ratliff (1976). See also Boynton (1997, p. 133 ff), Kaiser and Boynton (1996, p. 498 ff).]

In response to these criticisms The World Color Survey (WCS) was undertaken in the late 1970s, with results to be discussed below. Color naming data were collected from 110 unwritten languages using stimuli closely approximating those of Figure 1 at locations where linguistic missionaries of the Summer Institute of Linguistics (now SIL International) were located. These languages are primarily spoken in regions of comparatively low technological development.[5]

On the basis of the WCS data, a final adjustment to the encoding sequence was proposed. Ever since B&K had noted the occasional "premature" occurrence of gray terms, exceptions had accumulated to the proposition that not all the unary categories (black, white, red, yellow, green, and blue) receive separate basic terms before any of the binary categories (pink, purple, orange, and gray). Robert MacLaury, based on his extensive study of the color terms of the native languages of Mesoamerica (MacLaury1986, 1987), was the first to propose two mostly disjoint but occasionally overlapping developmental sequences for (i) the composite-to-unary progression of the unary categories and (ii) the binary categories. Kay, Berlin, & Merrifield

[5] A major function of SIL linguistic missionaries is to study unwritten languages with the goal of translating the Bible into them. See Cook et al. (2005), Kay et al. (2009 Section 2), and Kay & Cook (forthcoming) for further details regarding the WCS.

(1991) observed that in the WCS data brown or purple or both occasionally receive basic terms before grue is divided into separate green and blue terms. MacLaury had also observed that the encoding sequence of the binary categories is less regular than the composite-to-unary sequence and this held also for the just-noted observation regarding brown and purple. Since 1991 the seven evolutionary stages of B&K, still recognized in (Kay & McDaniel 1978; See Figure 4) have been reduced to five. The binary categories have been dropped from consideration. The current version of the evolutionary typology of basic color term systems based on the WCS data is given in Figure 5.

INSERT FIG 5

In Figure 5 each box represents a type of color terminology and the arrows represent systems that were scored as being transitional between two types. Of the 110 languages in the WCS, 91 or 83% were found to belong to the "main line" of color term evolution, represented by the middle row of Figure 5. Figure 6 displays the types of the main line with the number of WCS languages assignable to each type in the top row.[6, 7]

---

[6] There are no WCS languages of type I. Two-term languages are known only from Rosch's experimental assessment of the Dugum Dani, several ethnographic reports in the pre-B&K literature that were probably not based on systematic elicitation with controlled color materials, and a personal communication from K.F Koch; these are discussed in B&K (§ 2.3.1). All the known languages judged to present stage I (two-term) color systems appear to be in or related to the Dani (Non-Austronesian) language group and from the same region of Irian Jaya.

[7] The assignment of languages to stages was not based on an explicit count but on the subjective assessment of a number of explicit indices presented by the WCS data. For details of the procedure see Kay et al. (2009, Section 2). The same work gives a detailed discussion of each

INSERT FIG 6

The Emergence Hypothesis

As mentioned above, B&K's tacit assumption that every language has a color lexicon that partitions color space has been questioned and abandoned. The hypothesis that not all languages contain color lexicons adequate to name all colors has been referred to as the Emergence Hypothesis. The first to suggest it in print was Lyons (1999). Kay (1999) proposed the name Emergence Hypothesis and cited unpublished work of Luisa Maffi dating from 1990 as embodying the idea. The emergence hypothesis has been supported by WCS data in Kay et al. (1997), and Kay et al. (2009). It had been strongly supported by Levinson's field experiments with speakers of Yélî dnye (Levinson 2000).

Evidence from the WCS for universal constraints on color naming

Studies supporting the claim that there are universal constraints on color naming can be divided into two groups, those that deal statistically with the entirety of the WCS online data set[8] and those that deal with specific languages that have been claimed to present counterexamples to this claim.

Studies dealing statistically with the entire WCS database

The first to use the WCS data to establish the universality of constraints on color naming was the late Robert E. MacLaury (1997). Before the online WCS database was established MacLaury hand tabulated every focal (best example) choice in the WCS data archive. His description of his Figure 1, reproduced in part as our Figure 7, is as follows:

---

language's color terminology with details on how decisions were made regarding which were the basic terms and what stage (if any) to assign to the language.

[8] Permanently available at http://www1.icsi.berkeley.edu/wcs/data.html

> Part (b) displays 15,186 color-term foci placed by 2,476 speakers of 107 of the 110 WCS languages. (Data from three languages pend a check of their preliminary processing.) …The darkened cells mark the plurality peaks of foci in white (2,105), black (1,983), red (472), yellow (537), green (194), and blue (169), the 6 densest, noncontiguous clusters on single chips. Part (c) represents the frequency of foci distributed across the hue columns in the ethnographic array. There are four steep ascents to an apex in the single columns that intersect the purest examples of red, yellow, green, and blue (MacLaury 1997: 200).

INSERT FIG 7

The first study to use the newly created WCS online database was Kay & Regier (2003). This study examined two aspects of the WCS naming data: (1) Is there greater clustering in color naming across the WCS languages than one would expect by chance? (2) Are the color categories of the WCS languages, which are all unwritten and spoken in mostly technologically simple societies, more similar than chance expectation to those of the B&K study, which are mostly written and spoken in technologically advanced societies? To test hypothesis (1) the WCS Munsell stimulus palette was translated into CIE L*a*b* (hereafter LAB) perceptual color space in order to make available psychologically meaningful inter-point distances. Each term in each language was represented by the centroid[9] of all the points in LAB space corresponding to a naming choice of an individual participant. In order to measure the degree to which the naming centroids of different languages cluster it was convenient to measure the opposite of clustering: *dispersion*, the degree to which points are scattered, failing to form clusters, and focus on low values. Low values of dispersion correspond to high values of clustering. The dispersion D in the WCS dataset as a whole was taken to be the sum of the distances in LAB space from each term

[9] The centroid of a set of points in a space is the point location in the space at minimum total distance from the points in the set.

centroid to the nearest term centroid in another language. In order to test whether D was smaller than might be expected by chance, the following Monte Carlo test was devised. In creating pseudo-random values of D it was considered important to maintain the shape of the actual WCS terminologies, since any naming system is likely to display a certain amount of dispersion for reasons of effective communication. Consequently, in each language the set of observed term centroids was rotated in the a*b* hue plane a different random amount and the result coerced back into the Munsell coordinates. This rotation produced a hypothetical set of WCS languages, each presenting the same distance among centroids as the true WCS languages but located at a random remove in the hue plane of LAB space from the centroids in the actual WCS. The procedure was repeated 1,000 times to create a distribution of D values calculated for hypothetical WCS corpora that respect the relative spacing of the actual observations but locate them randomly in the hue dimension. The value of actual D was found to lie well below the entire set of hypothetical Ds, as shown in Figure 8a. According to this test, the WCS data show a degree of clustering greater than chance with $P < .001$. To answer question (2) a similar Monte Carlo test was conducted with an analogous rotation routine except that this time the statistic was not D but the sum S (for *separation*) of the distances from each WCS naming centroid to the closest naming centroid in the B&K data. Analogously to the inverse relation between dispersion and clustering, the lower the value of separation, S, the greater the similarity of WCS to B&K. Again actual S fell well below any of the hypothetical Ss, as shown in Figure 8b, again with $P < .001$, establishing that there is no important difference between the naming behavior of the 110 unwritten languages of the WCS study and the 20 experimentally assessed languages of the B&K study (seventeen of them written).[10]

[10] The same test was carried out comparing the WCS languages to the seventeen written B&K

INSERT FIG 8

The tendency for WCS naming responses to be more clustered than chance and more similar to B&K naming responses than chance having been established, Regier et al. (2005) considered the question whether focal (best example) responses clustered even more tightly than naming responses. If focal responses cluster more tightly than naming responses it would suggest that universals in color naming stem from there being privileged places in the mental representation of color that serve as anchors for color categories and would counter a claim by Roberson et al. (2000) that ‘‘color categories are formed from boundary demarcation based predominantly on language.’’[11] In LAB space for both naming centroids and modal focal choices for each term in every WCS language the distance from the closest term in any B&K language was observed. Summing these two sets of LAB distances produced two numbers, one measuring the Focus Separation of WCS naming responses to B&K naming responses and the other measuring the Naming Separation of WCS focal responses to B&K focal responses. A paired *t* test showed the Focus Separation to be significantly less than the Naming Separation ($P < .0001$). It is unlikely that if focal colors were epiphenomena of categories boundaries determined by language that the focal responses of the WCS would approximate those of B&K even more closely than the

---

languages with the same qualitative result, $P < .001$.

[11] The hypothesis that color categories derive from color foci does not entail that the foci represent biases of the visual system. Such foci might also arise from adaptation to the color statistics of the visual environment or to other factors not heretofore considered. The hypothesis does contradict the claim of Roberson et al. that languages can set the boundaries of color categories capriciously, subject only to the weak constraint that color categories must carve out unbroken regions of color space: “grouping by similarity” (Roberson et. al 2000).

naming responses do. Regier et al. also expanded MacLaury's hand-counted observations regarding clustering of WCS focal responses by producing the contour plot in shown in Figure 9, summarizing the focal responses in the WCS corpus and comparing them to the foci for English reported in B&K. Regier et al. found that the two chips that received the largest number of focal choices across the entire WCS database were the chips A0 (2,084 hits) and J0 (1,988) of Figure 1, the whitest and blackest chips, respectively. In Figure 9, restricted to the chromatic palette, the outer contour of each cluster marks 100 focal choices and each inner contour another 100 focal choices. The black dots represent the English focal choices reported by B&K.

INSERT FIG 9

Lindsey and Brown (2006) adopted an approach to the WCS dataset distinct from the foregoing. Instead of first creating a summary of the responses for a language, such as the centroid of all the language's naming choices or the modal focal choices, and then comparing these summary statistics across languages, Lindsey and Brown analyzed individual participants' naming responses in the following way. For each participant *p*, for each term *t* that *p* used, and for each chromatic chip *c*, a 320-place binary naming vector was created which recorded (+ or –) whether *p* assigned *t* to *c*. There were 14,236 such vectors. Lindsey and Brown performed a *k*-means cluster analysis of these naming vectors based on the Pearson correlation coefficients. Among their results were the following:

> When **K**, the number of **k**-means clusters, varied from 2 to 10, we found that (**i**) the average color-naming patterns of the clusters all glossed easily to single or composite English patterns, and (**ii**) the structures of the **k**-means clusters unfolded in a hierarchical way that was reminiscent of the Berlin and Kay sequence of color category evolution (Lindsey & Brown 2006: 16,608).

Jäger (2012) performed a principal components analysis of the WCS naming responses. He reports that his "results largely confirm the generalizations that [Kay & Maffi (1999) and Kay et al. (1991)] achieved with non-statistical techniques, even though not all the proposed universals could be confirmed" (p. 525). The biggest difference Jäger found between his PCA analysis and the non-statistical analysis of Kay & Maffi (1999) was that he found yellow more frequently associated with white than they did (p. 533). In a follow-up analysis to their 2006 paper, Lindsey and Brown (2009) found that although there is significant variation across speakers of the same language,

> The color naming idiolects of 2,367 WCS informants fall into three to six ''motifs,'' where each motif is a different color-naming system based on a subset of a universal glossary of 11 color terms. These motifs are universal in that they occur worldwide, with some individual variation, in completely unrelated languages. Strikingly, these few motifs are distributed across the WCS informants in such a way that multiple motifs occur in most languages (Lindsey & Brown 2009: 19785).

They further stated:

> Berlin and Kay … proposed two conjectures: (**i**) there exists a limited set of ''universal'' categories from which all languages draw their color lexicons, and (**ii**) languages ''evolve'' by adding color names in a relatively fixed sequence. There is now overwhelming empirical support for the first conjecture (Lindsey & Brown 2009: 19785).

Studies devoted to specific languages

Two of the languages that have been advanced as exceptions to the hypothesis of universal constraints on color naming are Berinmo [Ethnologue: Berinomo], a Sepik language of Papua New Guinea [PNG] (Roberson, Davies, & Davidoff 2000) and Yélî dnye, an unclassified language and probably a language isolate of Rossel Isand in the Louisade Archipelago, PNG (Levinson 2000). MacLaury (1997: 467), based on inspection of the original B&K data, the WCS data, and his Mesoamerican Color Survey data, proposed four chips in the 320 chip B&K Munsell palette as representing the "elemental hues", red, yellow, green, and blue. Kay (2005) calculated the naming centroids for Yélî dnye from the data published in Levinson (2000) and from unpublished Berinmo naming data generously provided by Roberson and compared the naming centroids for each of the three well-established chromatic terms of these languages[12] with MacLaury's proposed elemental hues. Levinson reports red, yellow, green, and blue categories for Yélî dnye, for each of which he records several expressions. He reports the naming responses for all except blue, apparently because the blue category is the least distinctly lexicalized. The hue terms reported for Berinmo are red, yellow and grue; those of Yélî dnye are red, yellow, and green. The results are presented in Table 1. The naming centroid for the Berinmo grue terms is precisely halfway between MacLaury's elemental green and elemental blue chips. The remaining red, yellow, and green centroids in both languages are either on or adjacent in the 320-chip palette to the chip named by MacLaury as representing the corresponding "elemental hue."

[12] Comparable data were not available for achromatic terms, but it is clear from the descriptions of these authors that both languages have terms for black and white.

MacLaury's Elemental Hues	*Red* 2.5R/4	*Yellow* 2.5Y/8	*Green* 2.5G/5	*Blue* 2.5PB/5
Berinmo	*mehi* (red) 2.5R/5	*wor* (yellow) 5Y/7	*nol* (grue) 7.5BG/5	
Yélî dnye	red terms 2.5R/4	yellow terms 5Y/7	green terms 2.5G/4	

Table 1. Berinmo and Yélî dnye naming centroids and MacLaury's elemental hues, in Munsell notation (Source: Kay 2005: 48, Table 1)[13]

Kay & Regier (2007) considered the boundaries of Berinmo color terms. Roberson et al. (2000) write, "[C]olor categories [are] a function of cultural experience and only, at most, loosely constrained by the default neural organization …[We] propose that color categories are formed from boundary demarcation based predominantly on language. Thus in a substantial way we present evidence in favor of linguistic relativity." Roberson (2005) specifies the loose constraint as follows, "No language has ever been reported to have a category that includes two areas of the color space (e.g., yellow and blue) but excludes an area between them (green)….Grouping always follows principles of similarity… and *the only free parameter appears to be the placement of boundaries between categories.*" (italics added).

Think of a color terminology as determining a map on the surface of color space. If the placement of boundaries is a free parameter, then any transformation of a terminology in which every point is moved the same distance in the same direction is as likely as any other. Kay and

[13] Expressions like "2.5R/4" refer to the Munsell Hue and Value (lightness) coordinates of colors in the stimulus palette, which was described in Figure 1. The notation preceding the slash, e.g., 2.5R, denotes the Munsell Hue coordinate; the notation following the slash, e.g., 4, denotes the Munsell Value coordinate.

Regier tested the Roberson et al. hypothesis as follows. Nineteen hypothetical Berinmo color lexicons were created by rotation of the actual Berinmo map by 2,4, etc. hue columns. Any two color terminology maps can be compared for similarity of boundaries by simply counting how many boundary chips they have in common. If the Roberson hypothesis is correct, actual Berinmo should not be specially privileged over the nineteen hypothetical Berinmos in a comparison with the WCS languages in general. The test consisted simply in calculating for each version of Berinmo – actual or hypothetical – the average number of shared boundaries between it and each of the languages in the WCS. The result is shown in Figure 10, in which it can be seen that not only do the boundaries of actual Berinmo match more closely to those of the languages of the WCS, but the closer a hypothetical Berinmo is to actual Berinmo the better it matches.

INSERT FIG 10

An argument against the usual data

All the studies reviewed above have analyzed data on color naming. However, an argument has been made against universal tendencies in color naming that is not based on the analysis of color naming data but on the claim that what have been taken for color-naming data are not color-naming data.[14] Wierzbicka (2008 and earlier works cited there) claims that there cannot be any

[14] There have been challenges to universal tendencies in color naming based on critiques of analyses of color naming data and also responses to those challenges. Full discussion of these matters might require a text as long as the present chapter and will not be undertaken here. See, e.g., Lucy (1997) and Saunders and van Brakel (1997) for critiques; Kay & Berlin (1997) and Kay (2006) for responses.

color-naming universals because there are languages that have no word for 'color'.[15] And it's not just a few languages: Wierzbicka claims that no language that lacks a word meaning 'color' can have color words. This is a major claim, because probably most – certainly a large fraction – of the languages of the world have no word for 'color'. Wierzbicka does not maintain – in fact she denies – that lacking a word for 'temperature' prevents a language from having words for 'hot' or 'cold', or that lacking a word for 'size' prevents a language from having words for 'big' or 'small' (pc from Wierzbicka to Kay cited in Kay and Kuehni 2008). Indeed, BIG and SMALL figure among the thirty-seven universal semantic primitives of Wierzbicka and Goddard's Natural Semantic Metalanguage (Goddard 1994: 22) and not all languages have a word meaning 'size'. Other universal primitives according to Wierzbicka and associates include ONE, TWO, VERY, and ALL, although many languages, perhaps most, lack abstract terms for 'extent', 'number', or 'quantity'. Wierzbicka does not hold that *in general* a language cannot have words for members of a set of related properties (e.g., {big, small,…}, {hot, cold,…}, {one, two,…},{ red, green,…}etc.) if it lacks a more abstract word denoting the semantic field they occupy. But she does hold that if a language lacks a word for 'color' it cannot have words for 'red', 'green', etc. She does not explain why the domain of color alone is subject to this stricture while no other domain, such as size, temperature, quantity, etc., is. This unsupported assumption makes the argument against color-naming uiversals from absence of a word for 'color' difficult to accept.

Conclusion

[15] The fact that many languages which have words that appear to be words for colors have no word for 'color' has been general knowledge at least since the 1960s.

The broad lines of the Berlin and Kay hypotheses were that (1) different languages' color lexicons are more similar than they would be if each language partitioned the color space idiosyncratically and (2) that within these limits there is further order in the way languages acquire new color terms. The original statements of B&K have been significantly modified and loosened, but these broad lines persist. For example, there are languages that fail to distinguish green and blue and there are languages that fail to distinguish red and yellow, but there are no languages that distinguish green and blue and fail to distinguish red and yellow.

References

Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley & Los Angeles: University of California Press.

Berlin, B. & Berlin, E.A. (1975). Aguaruna color categories. *American Ethnologist, 2*, 61-87.

Boynton, R. M. (1997). Insights gained from naming the OSA colors. In C.L. Hardin and L. Maffi (eds.) *Color Categories in Thought and Language*. Cambridge: Cambridge University Press, 135-150.

Brown, R. W. (1976). Reference. *Cognition, 4,* 125-153.

Collier, G. A. (1973). Review of Basic Color Terms. *Language, 49,* 245-248

Conklin, H. C. (1973). Color categorization: Review of Basic Color Terms, by Brent Berlin and Paul Kay. *Language,* 75, 931-942.

Cook, R. S., Kay, P. & Regier, T. (2005). The world color survey database: History and use. In *Handbook of Categorisation in the Cognitive Sciences*. Cohen, H. & Lefebvre, C. (eds.), Amsterdam: Elsevier.

Durbin, M. (1972). Review of Basic Color Terms. *Semiotica, 6*, 257-278.

Goddard, C. & Wierzbicka, A. (1994). *Semantic and lexical universals: Theory and empirical findings. Vol. 25.* Amsterdam: John Benjamins.

Hickerson, N. (1971). Review of Berlin and Kay (1969). *International Journal of American Linguistics,* 37, 257-270.

Hage, P. & Hawkes, K. (1975). Binumarin color categories. *Ethnology, 24,* 287-300.

Dougherty, J. W. D. (1975). *A universalist analysis of variation and change in color semantics*. Unpublished Ph. Dissertation, University of California, Berkeley.

Dougherty, J. W.D. (1977). Color categorization in West Futunese: Variability and change. In *Sociocultural Dimensions of Language Change*. Blount, B.G. & Sanchez, M. (eds.) New York: Academic.

Harkness, S. (1973). Universal aspects of learning color codes: A study in two cultures. *Ethos, 1*(2), 175–200.

Heider, E. R. (1972a). Universals in color naming and memory. *Journal of Experimental Psychology, 93,* 10-20.

Heider, E. R. (I972b). Probabilities, sampling and ethnographic method: The case of Dani colour names, *Man, 7,* 448-66.

Heinrich, A. C. (1972). A non-European system of color classification. *Anthropological Linguistics, 14,* 220-227.

Jäger, G. (2012). Using statistics for cross-linguistic semantics: a quantitative investigation of the typology of colour naming systems. *Journal of Semantics, 29*(4*),* 521-544.

Kaiser, P. K. and Boynton, R. M. (1996). *Human Color Vision*, Washington D.C.: Optical Society of America.

Kay, P. (1999).The emergence of basic color lexicons hypothesis. In *The Language of Color in the Mediterranean*, Alexander Borg (ed.). Stolkholm: Almqvist and Wiksell International.

Kay, P. (2005). Color categories are not arbitrary. *Cross Cultural Research,* 39, 72-80.

Kay, P. (2006). Methodological Issues in Cross-Language Color Naming. *Language, Culture and Society*, Christine Jourdan and Kevin Tuite (eds.), Cambridge: Cambridge University Press, 115-134.

Kay, P. (1975). Synchronic variability and diachronic change in basic color terms. Language in Society, *4*(03), 257-270.

Kay, P., & Berlin, B. (1997). Science ≠ imperialism: There are nontrivial constraints on color naming. *Behavioral and Brain Sciences*, *20*(02), 196-201.

Kay, P., Berlin, B., Maffi, M., & Merrifield, W.R. (1997). Color naming across languages. In *Color Categories in Thought and Language.* Hardin, C.L. & Maffi, L. (eds.). Cambridge: Cambridge University Press.

Kay, P., Berlin, B., Maffi, L., Merrifield, W.R., & Cook, R.S. (2009). *The World Color Survey*. Stanford, CA: CSLI Press.

Kay, P., Berlin, B., & Merrifield, W.R. (1991). Biocutural implications of systems of color naming. *Journal of Linguistic Anthropology, 1*(1), 12-25.

Kay, P. & Kuehni, R. G. (2008). Why colour words are really … colour words. *Journal of the Royal Anthropological Insitute, 14*, 886-887.

Kay, P. & Maffi, L. (1999). Color appearance and the emergence and evolution of basic color lexicons. *American Anthropologist, 101*(4), 743-760.

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, *54*(3), 610-646.

Kay, P. & Cook, R.S. (forthcoming). The World Color Survey. In *Encyclopedia of Color Science and Technology*. Ronier Luo (ed.). New York/Heidelberg: Springer.

Kay, P. & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(15), 9085-9089.

Kay, P. & Regier, T. (2007) Color naming universals: the case of Berinmo. *Cognition. 102*(2), 289-98.

Kuschel, R. & Monberg, T. (1974) 'We don't talk much about colour here': A study of colour semantics on Bellona Island. *Man, 9,* 213-242.

Levinson, S. C. (2000).. Yélî dnye and the theory of basic color terms *Journal of Linguistic Anthropology, 10*(1), 3-55.

Lindsey, D. T. & Brown A.G. 2006. Universality of color names. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(44), 16609-16613.

Lindsey, D. T., & Brown, A.G. (2009). World Color Survey color naming reveals universal motifs and their within-language diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(47), 19785-19790.

Lucy, J. M. (1997). The linguistics of color. In *Color Categories in Thought and Language*. Hardin, C.L. & Maffi, L. (eds.) Cambridge: Cambridge University Press, 135-150.

Lyons, J. (1999). The vocabulary of colour with particular reference to Ancient Greek and Classical Latin. In *The Language of Colour in the Mediterranean*. Borg, A (ed.) Wiesbaden: Otto Harrassowitz, 76-90.

MacLaury, R. E., (1986). *Color in Meso-America: Vol I. A theory of composite categorization*. Unpublished Ph.D. thesis. University of California, Berkeley. (Later revised and published as MacLaury 1997)

MacLaury, R. E. (1987). Color-category evolution and Shuswap yellow-with-green. *American Anthropologist, 89,* 107-124.

MacLaury, R. E. (1997). *Color and Cognition in Mesoamerica*. Austin: University of Texas.

Maffi, L. (1990). Somali color term evolution: Grammatical and semantic evidence. *Anthropological Linguistics, 32,* 316–334.

Miller, G. A. & Johnson-Laird, P. (1976). *Language and Perception.* Cambridge, Massachusetts: Harvard University Press.

Monberg, T. (1971). Tikopia color classification. *Ethnology, 10,* 349-358.

Ratliff, F. (1976). On the psychophysiological bases of universal color terms. *Proceedings of the American Philosophical Society, 120*(5)*,* 311-330.

Roberson, D. (2005). Color categories are culturally diverse in cognition as well as in language. *Cross-Cultural Research, 39*(1), 56-71.

Ray, V. F. (1952). Techniques and problems in the study of human color perception. *Southwestern Journal of Anthropology, 8,* 251-259.

Roberson, D., Davies, I., & Davidoff, J. (2000). Colour categories are not universal: Replications and new evidence from a stone age culture. *Journal of Experimental Psychology: General, 129,* 369-398.

Saunders, B. A., & van Brakel, J. (1997). Colour: An exosomatic organ? *Behavioral and Brain Sciences*, *20*, 167-228.

Senft, G. (1987). Kilivila color terms. *Studies in Language, 11,* 313–46.

Snow, D.L. (1971). Samoan color terminology: A note on the universality and evolutionary ordering of color terms. *Anthropological Linguistics, 13,* 385-390.

Regier, T., Kay, P., & Cook, R.S. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(23), 8386-8391.

Turton, D. (1978). La catégorization de la couleur en Mursi. In Voir et nommer les couleurs. Tornay, S. (ed.) Nanterre, Laboratoire d'Éthnologie et de Sociologie comparative (diffusé par Klincksieck)("Recherches thématiques" 2).

Turton, D. (1980). There's no such beast: Cattle and colour naming among the Mursi, *Man, 15,* 320-338.

Wierzbicka, Anna (2008) Why there are no 'colour universals' in language and thought. *Journal of the Royal Anthropological Institute (N.S.), 14,* 407-25.

Figure Captions

Figure 1. The B&K stimulus palette (approximation). Rows correspond to Munsell Value (lightness); top to bottom 9.5, 9.0, 8.0, 7.0, 6.0, 5.0, 4.0, 3.0, 2.0, 0.5. Columns correspond to equally spaced Munsell Hues, left to right 2.5R, 5R, …, 10RP. All colors are at maximal available Chroma (saturation) for the Value-Hue combination depicted.

Figure 2. The B&K hypothesis regarding possible color naming systems (Source B&K, p.4)

Figure 3. Revision of the encoding sequence in (Kay 1975).

Figure 4. The encoding sequence as portrayed in (Kay & McDaniel 1978: 639, Fig. 13).

Figure 5. Current WCS typology and encoding sequence (Source Kay et al. 2009 p. 30, Fig. 1)

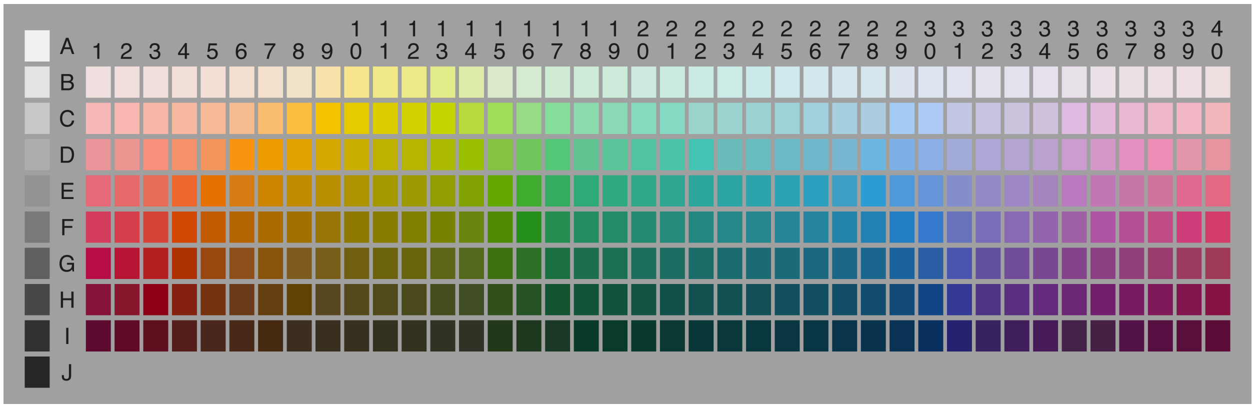
Figure 6. The man line of basic color term evolution with number of languages of each type (Source Kay et al. 2009, p. 30, Fig.1)

Figure 7. 15,186 WCS best example choices. (Source: MacLaury 1977, p. 200, Figure 1)

Figure 8. a. Dispersion of real (arrow) and hypothetical (histogram) WCS naming centroids. b. Separation of actual (arrow) and hypothetical (histogram) WCS and B&K naming centroids. (Source: Kay & Regier 2003: 9088, Figure 3.)

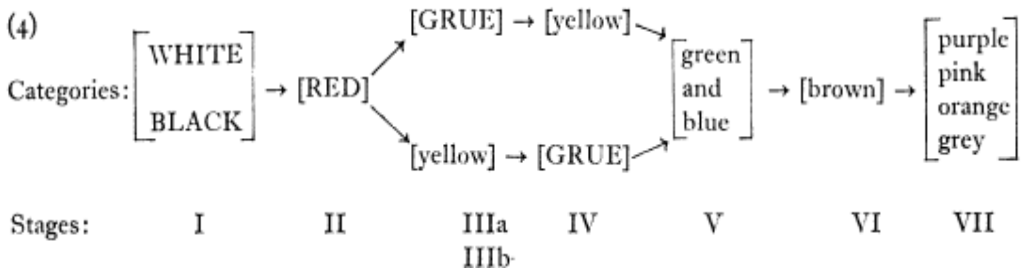
Figure 9. Contour plot of WCS best example choices showing English data from B&K. (Source: Regier et al. 2005: 8387, Figure 2)

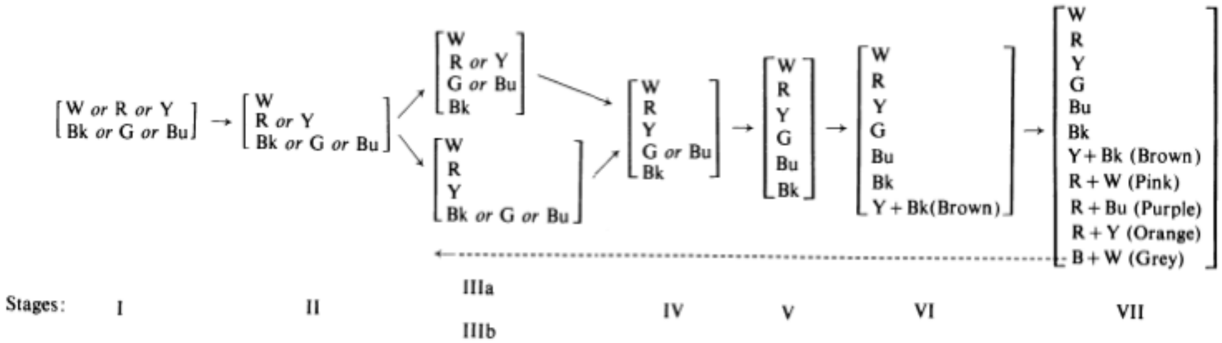
Figure 10. Boundary matches of real and hypothetical Berinmo to the WCS as a whole (Source Kay & Regier 2007)

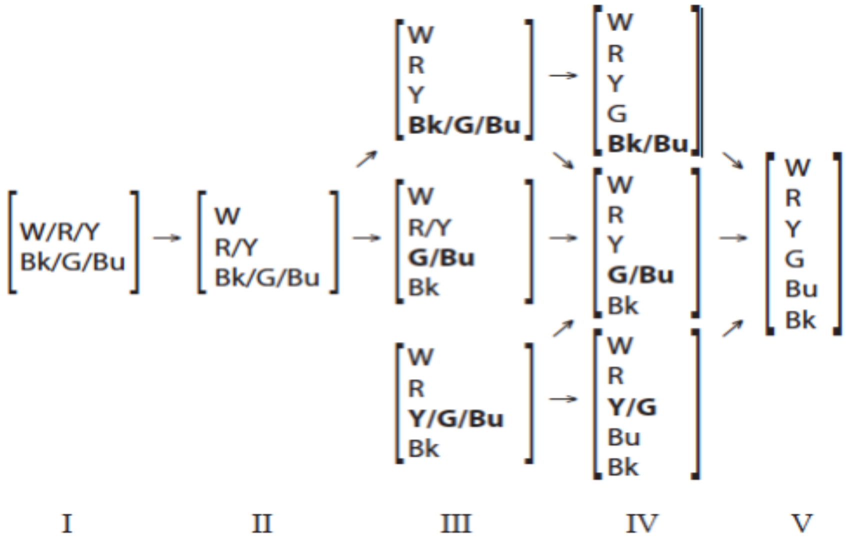

A
B
C
D
E
F
G
H
I
J
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40

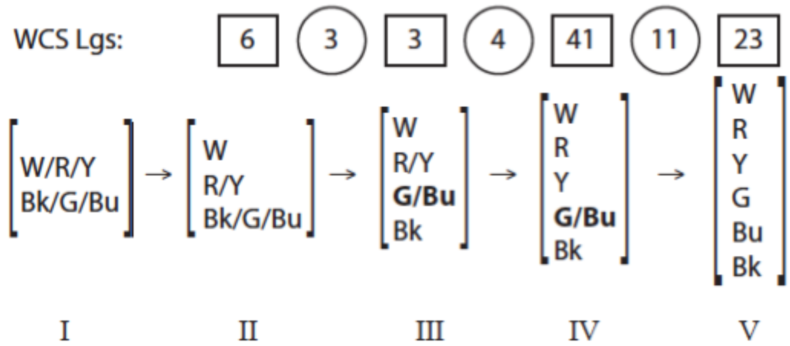
$$\begin{bmatrix}\text{white}\\ \text{black}\end{bmatrix} < [\text{red}] < \begin{bmatrix}\text{green}\\ \text{yellow}\end{bmatrix} < \begin{bmatrix}\text{yellow}\\ \text{green}\end{bmatrix} < [\text{blue}] < [\text{brown}] < \begin{bmatrix}\text{purple}\\ \text{pink}\\ \text{orange}\\ \text{gray}\end{bmatrix}$$

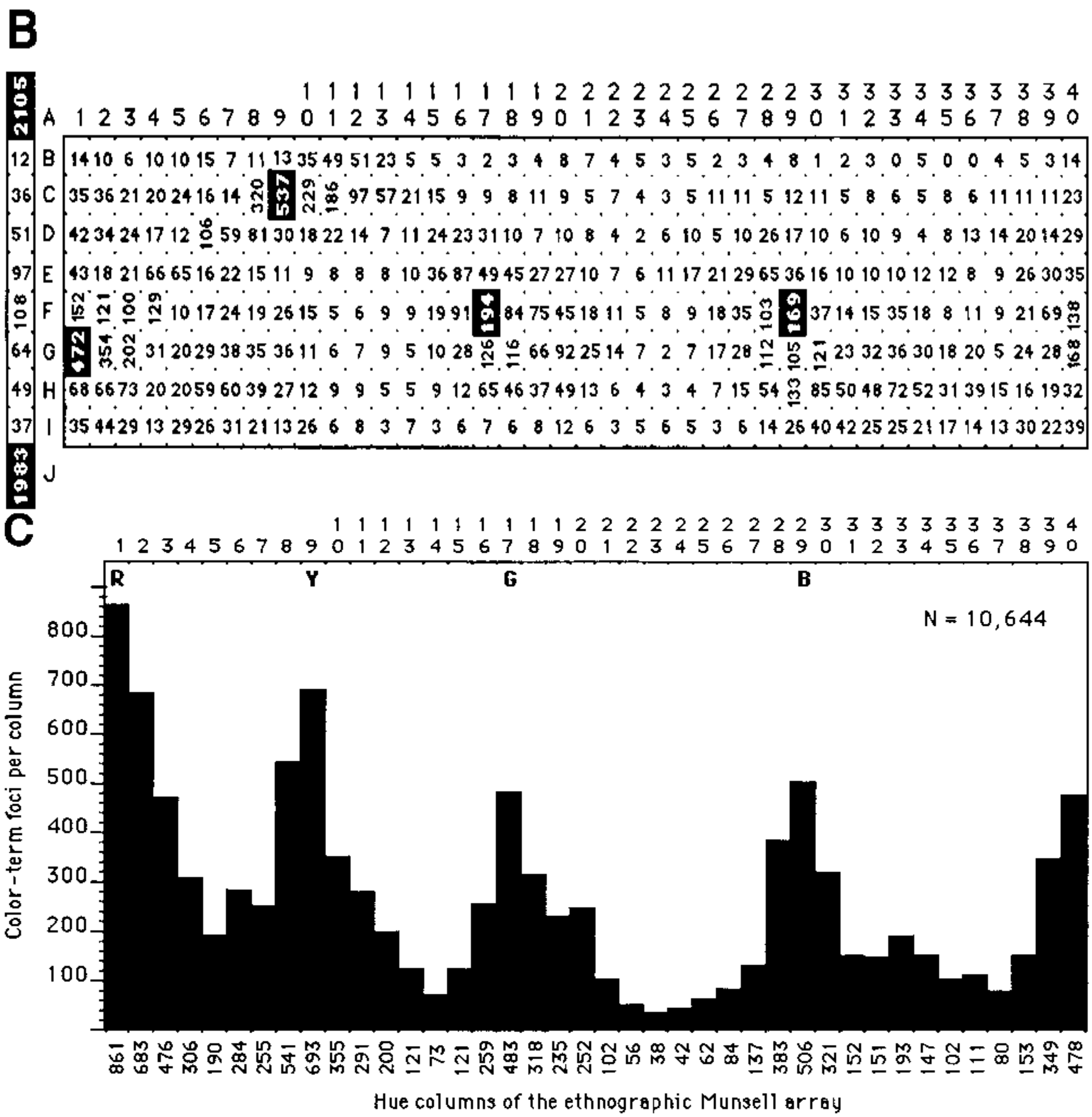
(4)


Categories:
WHITE
BLACK
[RED]
[GRUE] → [yellow]
[yellow] → [GRUE]
green
and
blue
[brown]
purple
pink
orange
grey
Stages:
I
II
IIIa
IIIb
IV
V
VI
VII

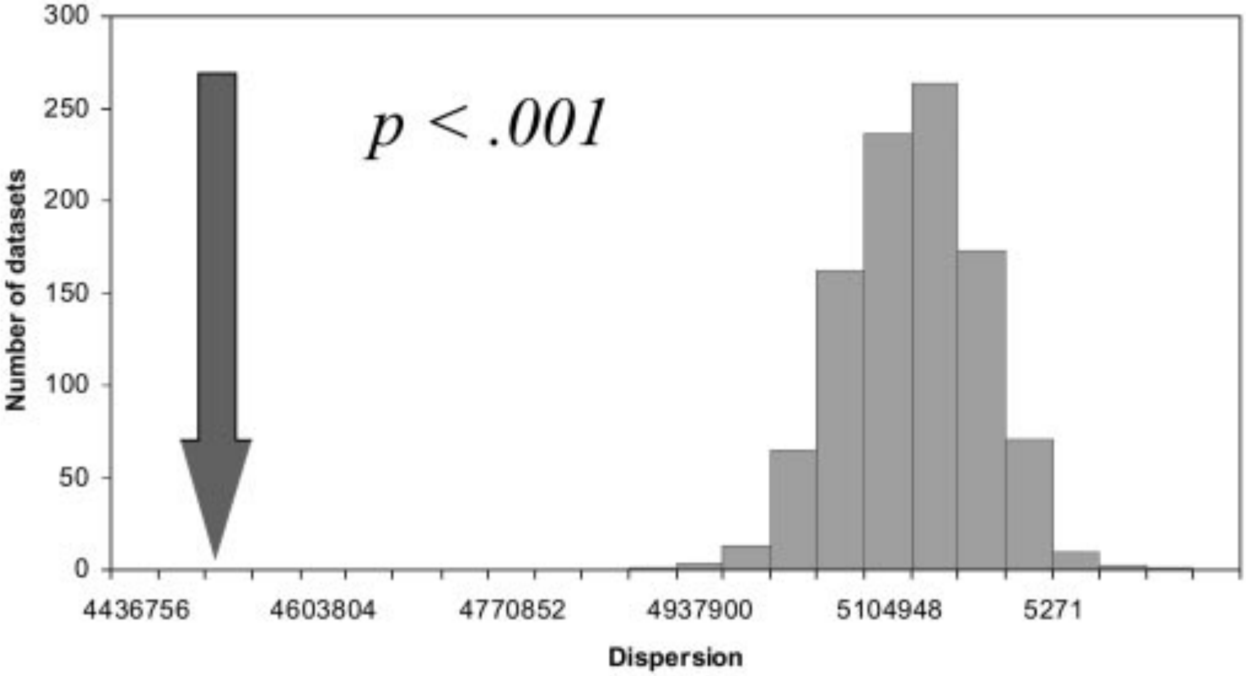

W or R or Y
Bk or G or Bu
W
R or Y
Bk or G or Bu
W
R or Y
G or Bu
Bk
W
R
Y
Bk or G or Bu
W
R
Y
G or Bu
Bk
W
R
Y
G
Bu
Bk
W
R
Y
G
Bu
Bk
Y + Bk(Brown)
W
R
Y
G
Bu
Bk
Y + Bk (Brown)
R + W (Pink)
R + Bu (Purple)
R + Y (Orange)
B + W (Grey)
Stages:
I
II
IIIa
IIIb
IV
V
VI
VII


W/R/Y
Bk/G/Bu
W
R/Y
Bk/G/Bu
W
R
Y
Bk/G/Bu
W
R/Y
G/Bu
Bk
W
R
Y/G/Bu
Bk
W
R
Y
G
Bk/Bu
W
R
Y
G/Bu
Bk
W
R
Y/G
Bu
Bk
W
R
Y
G
Bu
Bk
I
II
III
IV
V

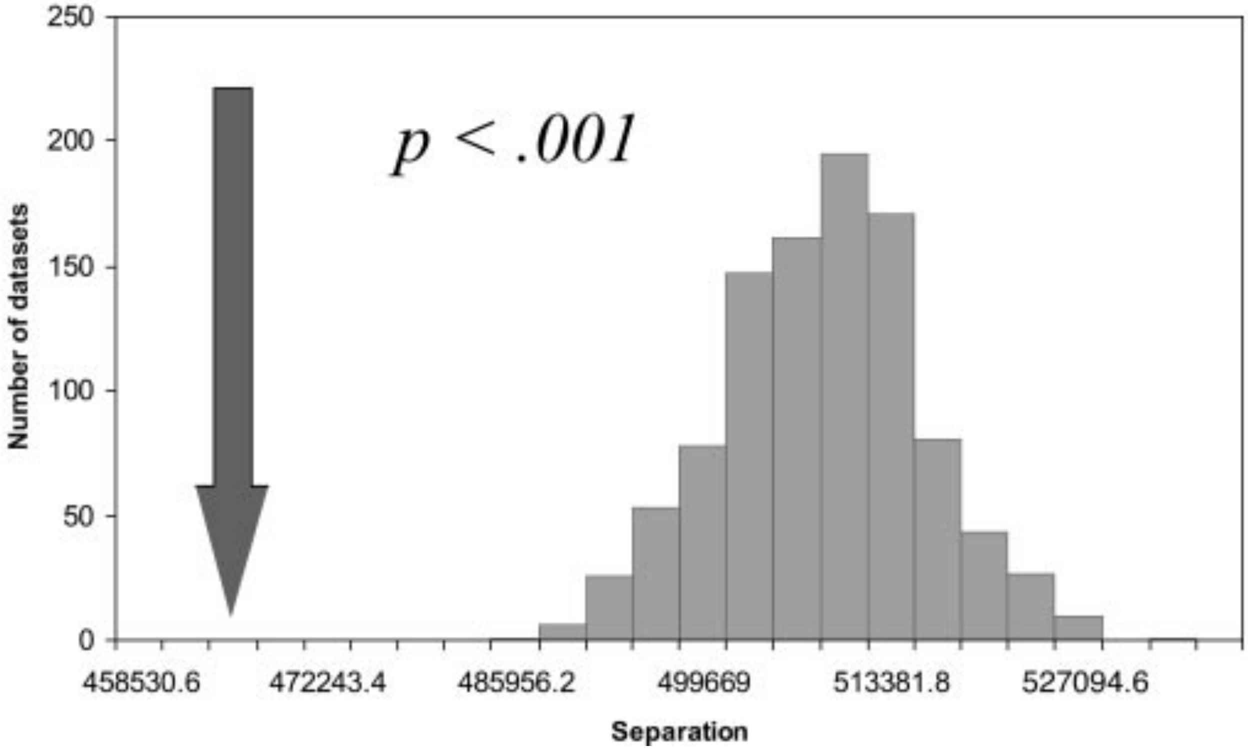

WCS Lgs:
6
3
3
4
41
11
23
W/R/Y
Bk/G/Bu
W
R/Y
Bk/G/Bu
W
R/Y
G/Bu
Bk
W
R
Y
G/Bu
Bk
W
R
Y
G
Bu
Bk
I
II
III
IV
V

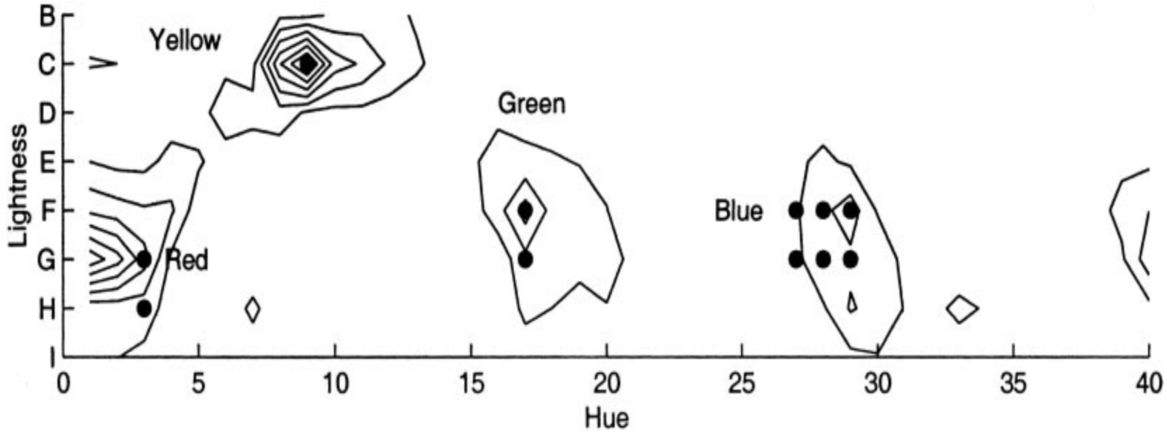

B
C
R
Y
G
B
N = 10,644
Color-term foci per column
Hue columns of the ethnographic Munsell array
861 683 476 306 190 284 255 541 693 355 291 200 121 73 121 259 483 318 235 252 102 56 38 42 62 84 137 383 506 321 152 151 193 147 102 111 80 153 349 478

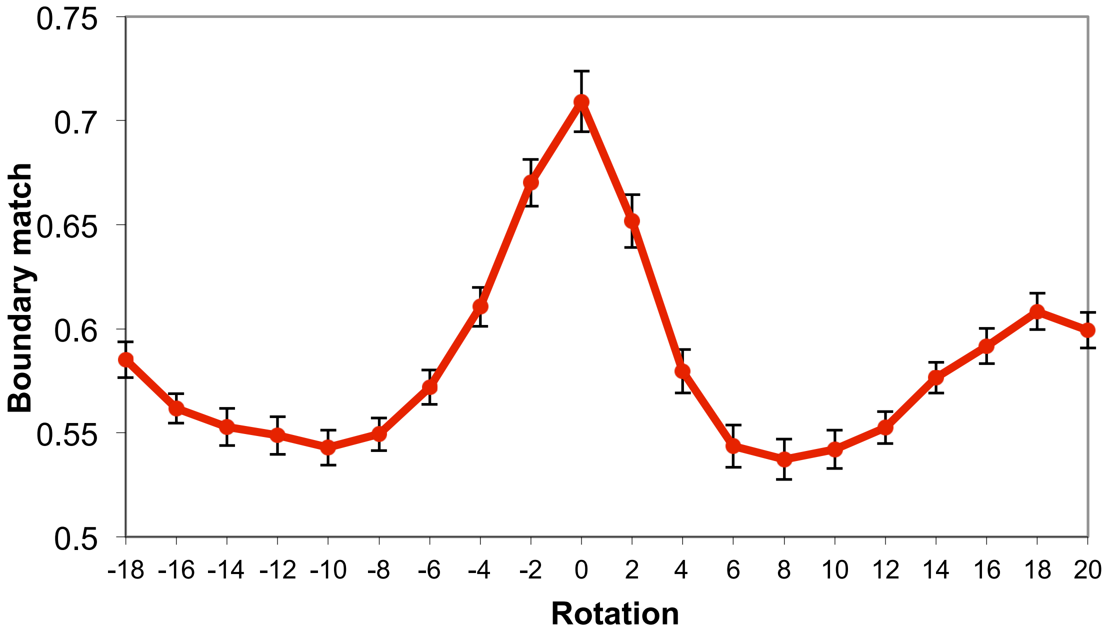
a


300
250
200
150
100
50
0
Number of datasets
p < .001
4436756
4603804
4770852
4937900
5104948
5271
Dispersion

b


250
200
150
100
50
0
Number of datasets
p < .001
458530.6
472243.4
485956.2
499669
513381.8
527094.6
Separation


Yellow
Green
Blue
Red
Lightness
B
C
D
E
F
G
H
I
Hue
0
5
10
15
20
25
30
35
40



Boundary match
0.75
0.7
0.65
0.6
0.55
0.5
-18 -16 -14 -12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12 14 16 18 20
Rotation