

## Linguistic Society of America

---

The Linguistic Significance of the Meanings of Basic Color Terms

Author(s): Paul Kay and Chad K. McDaniel

Source: *Language*, Vol. 54, No. 3 (Sep., 1978), pp. 610-646

Published by: [Linguistic Society of America](#)

Stable URL: <http://www.jstor.org/stable/412789>

Accessed: 01/08/2011 17:55

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=lsa>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



Linguistic Society of America is collaborating with JSTOR to digitize, preserve and extend access to *Language*.

<http://www.jstor.org>

# THE LINGUISTIC SIGNIFICANCE OF THE MEANINGS OF BASIC COLOR TERMS

PAUL KAY and CHAD K. MCDANIEL

*University of California, Berkeley*

There are semantic universals in the domain of color; i.e. there are constraints on the types of possible basic color lexicons. These constraints arise from the structure and function of the visual system. Thus in the case of color at least, rather than language determining perception (cf. Sapir and Whorf), it is perception that determines language.

In deriving the semantic universals from properties of the visual system, one must employ a continuous rather than discrete mathematics, in particular the theory of fuzzy sets. The resulting model of color semantics thus conflicts with the discrete-feature concept of semantic primes shared by structuralists and generativists. It is argued on this basis that discrete-feature semantic theories are of limited accuracy.\*

## INTRODUCTION

1. Recent empirical research into the meanings of words for color provides evidence that contradicts two widely-held beliefs in linguistics and the philosophy of language. This paper presents a summary of this evidence, uses it as a basis to construct a general model of basic color-term semantics, and explores the implications of this model for general semantic theory.

The first belief against which we will present evidence is familiar to most linguists as the Sapir-Whorf hypothesis. There are various so-called strong and weak forms of this hypothesis (Fishman 1960); in all forms, the basic notion is that each language imposes on the individual's 'kaleidoscopic flux of impressions' its own idiosyncratic semantic structure. This doctrine emphasizes the relativity of semantic structures, and minimizes the role of linguistic universals. The lexical categorization of color has often been presented as a paradigmatic instance of this arbitrary, language-specific imposition of semantic structure. Thus, in his widely-used introductory textbook, Gleason claimed (1961:4), 'There is a continuous gradation of color from one end of the spectrum to the other. Yet an American describing it will list the hues as red, orange, yellow, green, blue, purple—or something of the kind. There is nothing inherent either in the spectrum or the human perception of it which would compel its division in this way.'

Studying 20 languages experimentally and investigating the literature on color-term semantics for 78 additional languages, Berlin & Kay 1969 (hereafter B & K) presented evidence indicating that, contrary to the claims of Gleason and others, all languages share a universal system of basic color categorization. B & K argued that there are universal basic color categories, and that the basic color-term inventories of most languages expand through time by lexicalizing these categories in a highly constrained, universal order. In addition, McDaniel has argued (1972, 1974, ms) that these universals are inherent in the human perception of color. The

\* We have benefited greatly regarding the matters discussed here from the ideas of John Atkins, Brent Berlin, Janet Dougherty, Craig Molgaard, Sue Thompson, Bill Wooten, and an anonymous reader for *LANGUAGE* named George Lakoff.

color perception of all peoples is the result of a common set of neurophysiological processes, and McDaniel suggests that these pan-human neurophysiological processes are the basis of the universal patterns in the meanings of basic color terms.

We argue, then, in direct opposition to Gleason and other relativists, that the human perception of color offers an explanation of why English speakers segment the visual spectrum as they do—and why, furthermore, speakers of other languages exhibit the limited and systematic set of alternative segmentations of the color space that they do. Working with a biologically based understanding of basic color-term semantics, we can show the natural relations that exist between the numerous color categories encoded in highly differentiated color terminologies, such as English, and the fewer categories encoded in languages with less differentiated and therefore superficially simpler terminologies. Thus, in extension of the arguments advanced by B & K, we present the lexical categorization of color as a paradigmatic example, not of the relativity of semantic structures, but of the existence of biologically based semantic universals.

A second widespread belief in linguistics and the philosophy of language, challenged by the data reviewed here, is the doctrine that there exist ultimate semantic primes which are DISCRETE entities. These units are called semantic 'features' by both European and American structuralist schools, and 'markers' or 'distinguishers' by Katz (1964, 1966) and those generative linguists who follow him. Although there are, of course, major differences between structuralist and standard generative theories as to how semantic primes are combined into the meanings of words or larger linguistic units, until recently there was nearly universal agreement that all semantic primes are discrete, i.e. non-continuous, entities. When we say that semantic primes have been considered to be discrete entities, we mean that they have been viewed as properties that are predicable of things in an all-or-none fashion. This tacit premise is directly reflected in the common plus/minus notation of semantic features. When [+round] or [-human] is written, it indicates that roundness and humanity are being taken as properties which are simply predicable or not predicable of something being talked about. These properties are not treated in the discrete-feature theory as predicable of something TO A DEGREE.<sup>1</sup>

In discussions of the discrete-feature theory, as in discussions of semantic relativity, the domain of color is often used as a paradigm example. In Katz's theory, one of the major relationships that may exist between discrete semantic features is that of antonymy. For Katz (1966: 195 ff.), two linguistic expressions are antonymous just if they have readings that are identical except that, where one reading has a semantic feature from a (previously specified) antonym set, the other reading has a different feature from that same set. Katz (1964: 532) exemplifies antonymous lexical items as follows:

'There are many special antonymy relations between words and expressions. One example is the relation of "sex-antonymy". A pair of lexical items is SEX-ANTONYMOUS just in case they have identical paths except that where one has the semantic marker (*Male*) the other has the semantic marker (*Female*). Some instances are: *woman and man*; *bride and groom*; *aunt and uncle*; *cow and bull*. The majority of antonymous lexical items are not sets of

<sup>1</sup> For recent challenges to this view, see Lakoff 1972, Fillmore 1975, Kay 1975b.

pairs but sets of  $n$ -tuples. For example, there are the species-antonymous lexical items, one example of a species-antonymous  $n$ -tuple being: *child, cub, puppy, kitten, cygnet*, and so on. Moreover, there are  $n$ -tuples of lexical items that are distinguisher-wise antonymous, e.g., the  $n$ -tuple of simple color adjectives (*blue, yellow, green, red, orange*). These form an antonymous  $n$ -tuple because the path associated with each is identical except for the distinguisher which differentiates that color adjective from the others.<sup>7</sup>

Katz here proposes that *yellow, green, and blue* each has assigned to it in the semantic component of the grammar a discrete feature, called a distinguisher, which distinguishes it from each of the others and from the remaining color terms. The inadequacy of such a treatment is apparent when one considers compound color terms such as *yellow-green* or *blue-green*. These terms are not self-contradictory, as one might deduce from Katz's treatment, and their meanings are relatively transparent: something which is 'blue-green' is blue to some degree and green to some degree. This and related observations developed below show that the meanings of basic color terms can not be accurately represented with discrete semantic features. We propose instead that color categories, like the neurophysiological processes that underlie them, are continuous functions; and that a non-discrete formalism, in this instance fuzzy set theory, provides the most concise and adequate description of the semantics of basic color terms.

Preliminary to this analysis and further discussion of the general linguistic significance of basic color-term semantics, it is necessary to review some of the anthropological and psychological research from which our understanding of color categorization has developed. In §2, the major findings of B & K regarding semantic and developmental universals in basic color-term vocabularies are summarized. This section also discusses certain inadequacies in the original analysis that can now be corrected. In §3, we offer a brief summary of current theory and research regarding the neurophysiological bases of color perception as they are relevant to understanding McDaniel's psychophysiological re-interpretation and explanation of B & K. Our current analysis is presented in §§4–5, where the work of B & K and of McDaniel is integrated with new data from recent field studies of basic color-term systems. This integration leads to a reformulation of B & K's universals of color categorization and color-category development. In this reformulation, fuzzy set theory is used to model the structure of individual color categories and to explicate the relations which develop between the various universal color categories which appear as basic color vocabularies expand.

#### ANTHROPOLOGICAL RESEARCH INTO BASIC COLOR-TERM SEMANTICS

2. B & K's experimental and library research into the systems of color categorization in 98 languages, representing a wide range of major language stocks, focused on the categorization as represented by the basic color terms of each language. Basic color terms were defined as those (a) which are monolexic (unlike *reddish-blue*); (b) whose signification is not included in that of any other term (unlike *crimson* and *vermillion*, both of which are kinds of *red*); (c) whose application is not restricted to a narrow class of objects (unlike *blond* and *roan*); and (d) which are relatively salient as evidenced in frequent and general use (unlike *puce* and *mauve*).<sup>2</sup>

<sup>2</sup> For further discussion of the concept 'basic color term', see B & K (p. 5 ff.) and Kay 1975a.



The experimental portion of B & K's research, dealing with twenty languages, began with the elicitation, in the absence of any particular colored stimuli, of the basic color terms in each informant's language. After an informant's basic color-term inventory had been determined in the elicitation interview, the informant was asked to perform two tasks that involved matching these basic color terms to color chips in a standardized stimulus array.<sup>3</sup> The informant's first task was to pick out the color chips which he felt were the best examples of each of the basic color terms in his language. The second task was to indicate the boundaries of each color category, i.e. to indicate ALL the chips in the array whose color might be denoted by a given color term.

Informants selected best examples with far greater ease, speed, and reliability than they determined boundaries. B & K found that these judgments of best example, indicating what they called the focus of each color category, were more useful than the judgments of boundaries in describing and comparing the meanings of basic color terms in the languages studied experimentally. As expressed in the original report, the basic experimental findings were that 'color categorization is not random and the foci of basic color terms are similar in all languages' (10). A total of eleven foci were identified, located in the color space where English speakers locate the best examples of black, white, red, orange, yellow, brown, green, blue, purple, pink, and grey. These foci were interpreted as the primary designata of a set of universal semantic categories.

Using these results as a guide to interpretation, B & K considered published data from 78 additional languages, and determined that there is a temporal order in which languages encode these universal categories. They concluded as follows (3-4):

'Although different languages encode in their vocabularies different numbers of basic color categories, a total universal inventory of exactly eleven basic color categories exists from which the eleven or fewer basic color terms of any given language are always drawn ... If a language encodes fewer than eleven basic color categories, then there are strict limitations on which categories it may encode. The distributional restrictions of color terms across languages are:

- '1. All languages contain terms for white and black.
- '2. If a language contains three terms, then it contains a term for red.
- '3. If a language contains four terms, then it contains a term for either green or yellow (but not both).
- '4. If a language contains five terms, then it contains terms for both green and yellow.
- '5. If a language contains six terms, then it contains a term for blue.
- '6. If a language contains seven terms, then it contains a term for brown.
- '7. If a language contains eight or more terms, then it contains a term for purple, pink, orange, grey, or some combination of these.

<sup>3</sup> The stimulus array consisted of 329 Munsell color chips, including (a) 320 chromatic chips representing a uniform sampling of the colors on the outer skin of the perceptual color solid, where the most highly saturated colors of every hue and brightness are found; and (b) nine neutral, achromatic chips ranging from black through grey to white. For the reader unfamiliar with the color-solid model of the psychological color space, the important point here is that the stimulus array represents a reasonable sampling of the estimated seven million distinguishable colors. For further discussion of the stimuli, see B & K (5; 160, fn. 3; 162, fn. 12).

Collier 1973 and Collier et al. 1976 have examined the possibility that restricting the selection of chromatic chips to the maximally saturated chips at each hue and brightness level may have biased B & K's results; but it was found that no such bias exists.

'These distributional facts are summarized in [Table 1], in which each row corresponds to an actually occurring type of basic color lexicon. The pattern displayed by the actual distribution is a tight one; of the 2,048 (that is, 2<sup>11</sup>) possible combinations of the eleven basic color terms, just twenty-two, about 1 per cent, are found to occur in fact.

Type	No. of basic color terms	Perceptual categories encoded in the basic color terms										
		white	black	red	green	yellow	blue	brown	pink	purple	orange	grey
1	2	+	+	-	-	-	-	-	-	-	-	-
2	3	+	+	+	-	-	-	-	-	-	-	-
3	4	+	+	+	+	-	-	-	-	-	-	-
4	4	+	+	+	-	+	-	-	-	-	-	-
5	5	+	+	+	+	+	-	-	-	-	-	-
6	6	+	+	+	+	+	+	-	-	-	-	-
7	7	+	+	+	+	+	+	+	-	-	-	-
8	8	+	+	+	+	+	+	+	+	-	-	-
9	8	+	+	+	+	+	+	+	-	+	-	-
10	8	+	+	+	+	+	+	+	-	-	+	-
11	8	+	+	+	+	+	+	+	-	-	-	+
12	9	+	+	+	+	+	+	+	+	+	-	-
13	9	+	+	+	+	+	+	+	+	-	+	-
14	9	+	+	+	+	+	+	+	+	-	-	+
15	9	+	+	+	+	+	+	+	-	+	+	-
16	9	+	+	+	+	+	+	+	-	+	-	+
17	9	+	+	+	+	+	+	+	-	-	+	+
18	10	+	+	+	+	+	+	+	+	+	+	-
19	10	+	+	+	+	+	+	+	+	+	-	+
20	10	+	+	+	+	+	+	+	+	-	+	+
21	10	+	+	+	+	+	+	+	-	+	+	+
22	11	+	+	+	+	+	+	+	+	+	+	+

NOTE: Only these twenty-two out of the logically possible 2,048 combinations of the eleven basic color categories are found.

TABLE 1. The twenty-two actually occurring types of basic color lexicon.

'Moreover, the twenty-two types which do occur are not unrelated but may be summarized by (or generated from) a rather simple rule [Figure 1]:

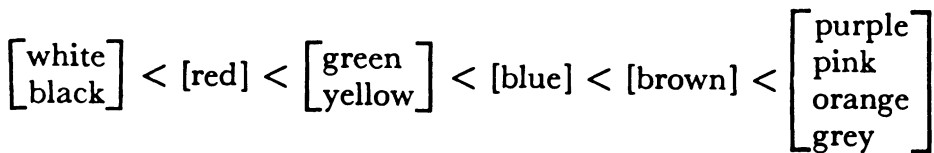


FIGURE 1.

where, for distinct color categories (*a*, *b*), the expression *a* < *b* signifies that *a* is present in every language in which *b* is present and also in some language in which *b* is not present. [Fig. 1] is thus a partial order on the set of basic color categories, the six bracketed sets being a series of six equivalence classes of this order ... [Fig. 1] represents not only a distributional statement for contemporary languages but also the chronological order of the lexical encoding of basic color categories in each language. The chronological order is in turn interpreted as a sequence of evolutionary stages.'

This distribution of color categories in the ethnographic present must reflect a sequence through which each language has to pass as it changes its number of basic color terms. If this were not the case, we would be forced to assume that languages existed in the past which did not conform to the observed distribution; e.g., a language with words for black, white, and blue, but none for red, green, or yellow. This would imply that the pattern displayed by the current synchronic sample of 98 languages is fortuitous with respect to languages in general. This is an unlikely conjecture for which no evidence exists. Moreover, considerable direct evidence has been adduced supporting the hypothesized sequence of temporal development in basic color-term systems (B & K, 36-41; Berlin & Berlin 1975, Dougherty 1975, Kuschel & Monberg 1974).

To continue our quotation from B & K (4-5):

'The logical, partial ordering of [Fig. 1] thus corresponds, according to our hypothesis to a temporal-evolutionary ordering, as follows [Figure 2]:

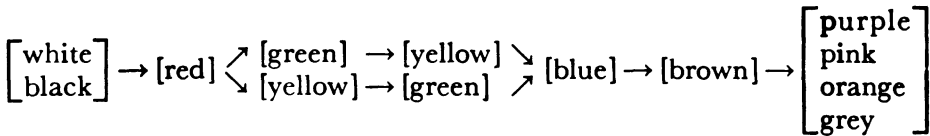


FIGURE 2.

where the arrow may be read 'is encoded before' ... In sum, our two major findings indicate that the referents for the basic color terms of all languages appear to be drawn from a set of eleven universal perceptual categories, and these categories become encoded in the history of a given language in a partially fixed order.'

Although subsequent research has substantiated B & K's basic findings, their 1969 report contained some errors of fact and a theoretical equivocation. The latter involved the conflicting treatments of category foci and category boundaries in the explanation of the successive stages of basic color-term development. B & K interpreted the developmental sequence primarily as the successive encoding of new foci. Stage I (i.e. two-term) systems were described as consisting of the categories BLACK and WHITE, where these categories included (on the one hand) black and all darker hues, and (on the other) white and all the lighter hues, with the foci of the categories being at pure black and pure white respectively. At Stage II, the warm hues were described as being accorded their own separate basic color term, RED, which focused at red. Thus the development at Stage II was thought essentially to consist of adding, to the categories focused at black and white, a category focused at red. But an equivocation enters at this point with respect to the meanings of the labels BLACK and WHITE. In Stage I, these labels referred to categories, focused at black and white, whose extensions (considered together) take in all other colors. At Stage II, though, the extensions of BLACK and WHITE have retracted, so that the warm colors are not included in either of them, but

instead are included in the extension of the new term RED, focused at red. Thus BLACK and WHITE mean one thing at Stage I, and something else at Stage II.

The same equivocation occurs with respect to all B & K's category labels BLACK, WHITE, RED, and GREEN. These labels were used sometimes to refer to a category having a particular focus, and at other times to a category having a particular extension. Moreover, the extensions referred to were not constant across occasions of use of the category label. These equivocations are eliminated in our present analysis of basic color-term semantics, where the non-discrete formalism of fuzzy set theory is shown to provide a unitary mechanism for describing the relationships between color-category foci, extensions, and boundaries.

Studies appearing after 1969, testing B & K's hypotheses, produced empirical results which highlighted the equivocation just noted and revealed certain factual errors in the original report. Heider 1972a,b studied the Stage I system of the Dugum Dani in detail. She found that the Dani's two color categories, *mola* and *mili*, were better labeled 'white-warm' and 'dark-cool', rather than simply WHITE and BLACK (as B & K had suggested for all Stage I systems), since *mola* contains not only white but all the warm colors (reds, oranges, yellows, pinks, and red-purples), while *mili* contains black and all the cool colors (blues and greens). Diagrammatically, the situation may be represented as in Figure 3.

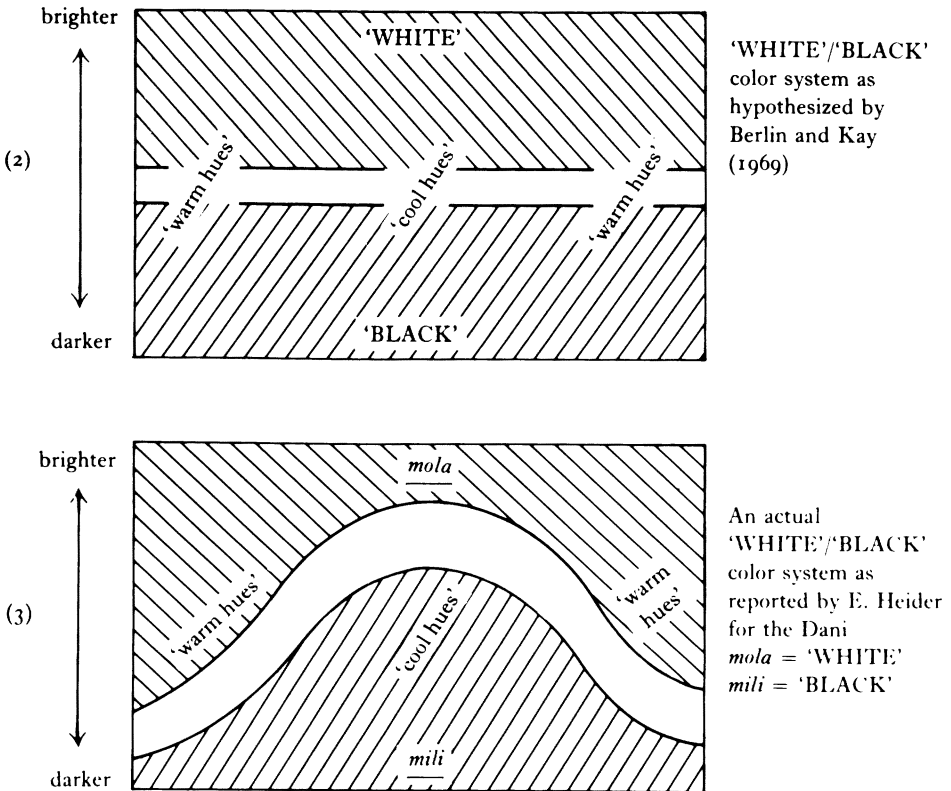


FIGURE 3. Source: Kay (1975a:259).

Significantly, Heider also found that the foci of these white-warm and dark-cool categories were variable across informants. *Mola* and *mili* were not always focused at only white and black. For example, 69% of her informants focused the white-warm category *mola* at English focal red. B & K's original analysis assumed that all basic color categories have a single focus. Other recent empirical studies of early-stage basic color-term systems, where monolingual informants were tested in their native environments, have corroborated Heider's finding that a language may have multiply-focused basic color terms.<sup>4</sup>

This and related factual emendations to B & K show that their treatment of the basic color-term encoding sequence as the simple successive encoding of single foci cannot be maintained. The re-interpretation of the encoding sequence presented here follows McDaniel 1974, MS, in treating the successive stages of basic color-term development as the progressive differentiation of color categories rather than as the successive encoding of foci. This re-interpretation, as our general analysis, involves a model of basic color-category formation that utilizes findings from recent researches into the neurophysiological bases of color perception. These findings are reviewed in §3.

#### THE NEUROPHYSIOLOGICAL BASES OF COLOR PERCEPTION

3. Research conducted in the past two decades has significantly increased our knowledge of the physiological processes which underlie the human perception of color. This research is concerned, for the most part, with discovering how differences in the wavelengths of light reaching the eye are transformed into response differences in the visual nervous system. It has been known for some time that color perception begins, neurophysiologically, at the retina, with the stimulation of color-sensitive cells called cones. There are three types of cone, and each type is distinguished by its own unique pattern of wavelength-dependent response. The light of each wavelength distinguishes itself neurally at the retinal level by evoking a unique set of neural responses from this three-cone system. Each distinct color is associated with a unique pattern of responses among the three types of cones.

The neural processing and coding of color continues, however, beyond this retinal level. Many recent studies of the visual processes have been concerned with the neural representation of color at some remove from the retina, in the neural pathways between the eye and brain. These studies, which have used microelectrodes to monitor single neurons, indicate that, by the time wavelength-governed neural impulses reach higher points in the visual pathway, the tripartite response of the retina's three-cone system is transformed into a set of opponent neural responses.

Two properties distinguish a cell as an opponent response cell. First, an opponent response cell has a spontaneous rate of firing—a basal response rate that it

<sup>4</sup> The studies encompass one Jivaroan language, one Mayan, one Papuan, one Eskimoan and three Austronesian (Berlin & Berlin 1975, Harkness 1973, Hage & Hawkes MS, Heinrich 1972, Dougherty 1974, 1975, Heider 1972a,b, and Kuschel & Monberg 1974). These data have been critical in the development of our new understanding of the encoding sequence. The results of these studies, as regards the substance of the encoding sequence, are summarized in Berlin & Berlin 1975, Dougherty 1975, Kay 1975a, and McDaniel 1974, MS.

maintains without external stimulation. Second, the cell shows an increased rate of firing in the presence of lights whose dominant wavelengths are from certain regions of the visual spectrum, while lights from the complementary spectral regions will decrease its rate of firing below its basal rate. The opposing effects of complementary regions of the visual spectrum on these cells gives rise to the term 'opponent response.'

Cells with opponent response characteristics have been identified by R. De Valois and his co-workers in the lateral geniculate nucleus (LGN) of the macaque (*Macaca irus*), an Old World monkey with a visual system similar to man's (De Valois et al. 1966, De Valois & Jacobs 1968). These scholars discovered four types of opponent cell. The mean response patterns that distinguish these four types are shown in Figure 4, where horizontal dotted lines indicate the mean spontaneous

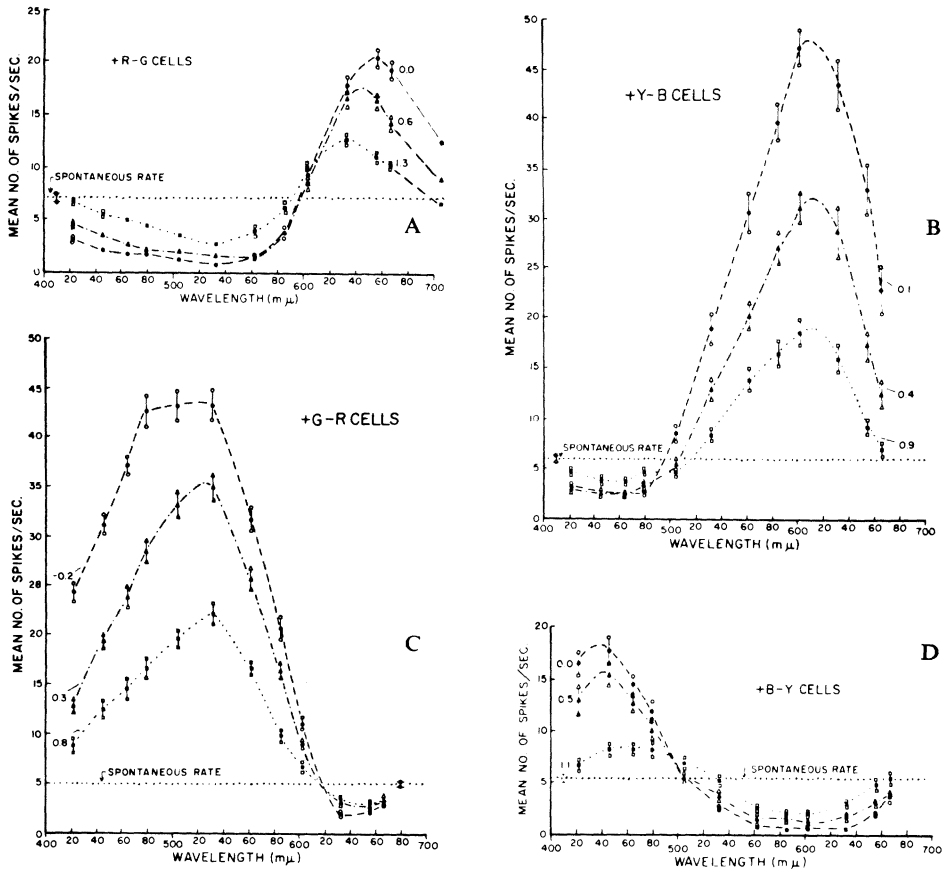


FIGURE 4. Source: De Valois et al. 1966:972-73; Figs. 9-12). Note: Horizontal dotted lines indicate mean basal response rate for each cell type.

firing rates of the cell, and the curvilinear dotted, dashed, and dotted-and-dashed lines indicate the observed response rates of the cell to lights of varying wavelengths at three different levels of brightness. The +R-G and +G-R cells shown in Figs. 4A and 4C have responses opposed between the red (far long) and green

(near short) regions of the spectrum. These cell types differ fundamentally only in the directions of the response deviations that represent the opposition between the red and green regions of the spectrum. The  $+Y-B$  and  $+B-Y$  cells illustrated in Figs. 4B and 4D have responses opposed between the yellow (near long) and the blue (far short) regions of the spectrum. These two types again differ primarily only in the response modes (increased vs. decreased firing) in which this opposition is expressed.

In their analysis of the translation of these response patterns into perceptual effects, De Valois et al. (1976) argue, with respect to the cell types that are opposed between the same regions of the spectrum, that 'it seems reasonable to say the excitation in one type [e.g.]  $(+R-G)$  carries the same information as inhibition in the other [e.g.]  $(-R+G)$ .' Following this argument, they sum the absolute values of the red, yellow, green and blue deviations of the  $+R-G$ ,  $-R+G$ ,  $+Y-B$ , and  $-Y+B$  opponent cells into unitary red, yellow, green, and blue response effects. They then take these response distributions as the neutral bases of color discriminations at the perceptual level.

Figure 5 presents these response distributions for human observers as estimated by Wooten 1970 from human psychophysical data. The direction of deviation above or below the X-axis signifies in this figure the chromatic value (i.e. red, yellow,

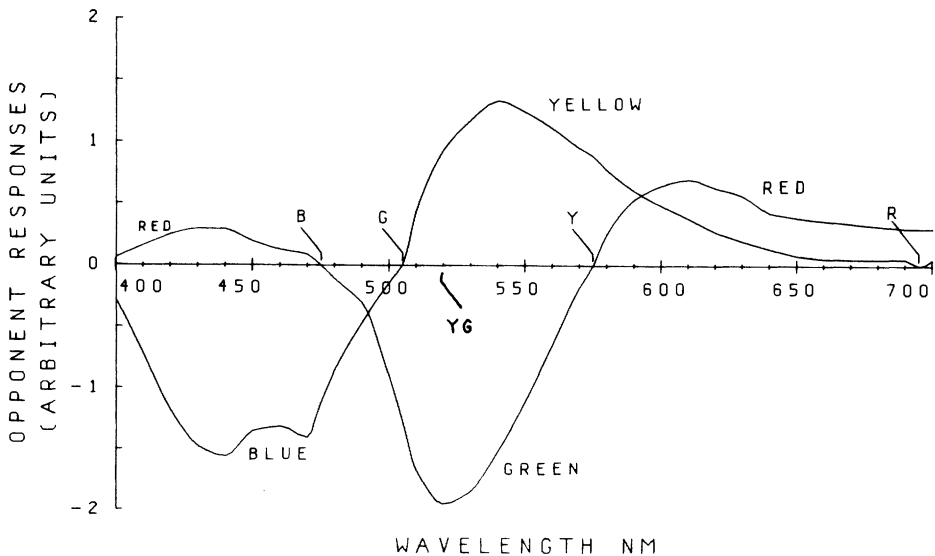


FIGURE 5. Note: Values are taken from Wooten 1970, who estimated them using methods outlined in Jameson & Hurvich 1968.

green, and blue) of the deviations in the combined RG and YB opponent systems (not the actual direction of response rate changes from a basal rate as in Fig. 4). The plus and minus values are used here simply to represent the opposition of red with green and of yellow with blue in their respective opponent systems. The absolute magnitudes of the red, yellow, green, and blue deviations represent the strengths of the response states at different wavelengths (regardless of whether the

neural representation of that response at some lower level is through inhibition or excitation).

In sum, the work of De Valois and his colleagues indicates that, while color is coded at the retinal level in the differential distributions of responses from the three types of cone, this code is transformed beyond the retina into one of opposed red and green and opposed yellow and blue neural response states distributed by wavelength as shown in Figure 6, below. It is the relative strengths of these states, then, that directly determines perceived hue. (In addition to the opponent cells just discussed, De Valois et al. also found in the macaque LGN two additional types of cells, both of which are non-opponent; i.e., they lack the characteristic of opposed response to complementary spectral regions that distinguishes opponent cells. The role of these cells in the definition of the fundamental categories black and white is discussed in §4.2, below. For the moment, however, we continue to restrict our discussion to the four hue response states basic to color perception.)

The above findings are particularly significant in the support they provide for an analysis of the phenomenal quality of color perception presented by Hering 1920, who argued that any color appears subjectively either as a pure, unique instance of one of four fundamental hues—red, yellow, green, and blue—or as a mixture of these primaries. All colors other than unique red, unique yellow, unique blue, or unique green, Hering claimed, could be seen (on introspection) to consist of the simultaneous perception of two of the primaries. His treatment regarded oranges, e.g., not as unitary percepts, but as having recognizable red and yellow components. Purples are composed of red and blue, while yellow-green (*chartreuse*) and blue-green (*turquoise*) show their true composite character in their descriptive names.<sup>5</sup>

While claiming red, yellow, green, and blue to be the only four fundamental hue sensations, Hering also noted certain antagonistic relations between red and green, and a corresponding antagonism between yellow and blue. He observed that green and blue may be perceived simultaneously (in what we call *greenish-blues* or *bluish-greens*) and similarly that red and yellow may be perceived simultaneously (in what we call, among other things, *orange*). Contrastingly, he pointed out, there is no such thing as a simultaneous perception of green and red, nor of blue and yellow. Perceptually, there are no yellowish-blues or bluish-yellows. From these and other observations, Hering suggested that the four primary hues are in fact two opposing hue pairs, red-green and yellow-blue, whose opposition indicates that these hue sensations are the product of antagonistic neural processes (as yet undiscovered in Hering's time).

The correspondence between the phenomenal analysis which Hering presented

<sup>5</sup> A number of recent studies have noted that a relation exists between the B & K semantic results, on the one hand, and, on the other, either (a) the results of the micro-electrode studies of De Valois and his co-workers, (b) the abstract opponent-process model of Hering, (c) other material in the physiological and psycho-physical vision literature, or (d) some combination of these. None of these studies traces in detail the line of argument developed here and, in any case, limitations of space preclude any attempt to summarize their contents. These studies include Bornstein 1973a,b, 1975; Cairo (ms); Faris 1974; Miller & Johnson-Laird (1976: 342-55); Ratliff 1976; Sahlins 1976; Stephenson 1973a,b; 1976; Witkowski & Brown 1977; and Zollinger 1972, 1973, 1976.



and the neurophysiological findings of De Valois and his associates is striking. Hering's observation of the phenomenal opposition of red to green and of blue to yellow anticipated De Valois' discovery of single cells which have opposing responses to these regions of the spectrum. Clearly, the R, G, Y, and B outputs of the opponent processes are the four fundamental hues described by Hering. The relative magnitude of the deviations in the RG and YB systems then indicate the relative strength of each fundamental hue in the total color sensation associated with light of a given wavelength. The phenomenal appearance of light of any given wavelength is thus the result of (1) the response state (R or G, Y or B) of each opponent system, and (2) the relative strengths of the responses in each state. Thus in Fig. 5, which shows Wooten's estimates of the human opponent process functions, light of 520 nm is shown to produce a yellow response in the YB system. It also produces a green response in the RG system, and this is of greater magnitude than the yellow response. The result is that light of 520 nm is seen as predominantly green with a 'veil' of yellow, i.e. as yellowish-green. The appearance of the entire visible spectrum can be accounted for in this manner, by evaluating at each wavelength the states and relative strengths of the two opponent response functions shown in Fig. 5.

The articulation of De Valois' findings with Hering's observations indicates that a particular structure is inherent in the human perception of color, a structure which is not deducible from the physical properties of light alone. The opponent process analysis identifies and describes four specific categories of neural response: the R (**red**), G (**green**), Y (**yellow**), and B (**blue**) response states.<sup>6</sup> As is shown below, the semantics of basic color terms in all languages directly reflect the existence of these pan-human neural response categories. Furthermore, when the universal basic color categories are treated non-discretely as fuzzy sets, their structure and formation from these prior perceptual categories can be given a precise representation.

#### COLOR CATEGORIES AS FUZZY SETS

##### 4.1. FUZZY SET MEMBERSHIP AND UNIVERSALS OF COLOR-CATEGORY STRUCTURE.

Fuzzy set theory has developed from several relatively straightforward modifications of the concepts of standard set theory.<sup>7</sup> The basic differences between the two stem from the fact that membership in standard set theory is categorically defined (an element simply IS or ISN'T a member of a given set), while degrees of membership are recognized in fuzzy set theory.

<sup>6</sup> The words **red**, **yellow**, **green**, and **blue** in boldface refer to response states of LGN cells in an organism that need not have color words at all. That is, they label neural responses which are shared by macaques, humans, and probably all higher primates. As such they have no NECESSARY relation to the corresponding English words or their meanings; but we will see that, empirically, there is a strict relation between these response states and the meanings of color words. Later, when we have established the identity of certain universal semantic categories with neural response categories, it will not always be necessary or desirable to maintain a distinction in notation.

<sup>7</sup> This and subsequent sections present, for the convenience of the non-mathematical reader, a non-rigorous description of the basic elements of fuzzy set theory. The mathematical reader is referred to Zadeh 1965, 1971.

Consider the classes denoted by the words *Congressman* and *gourmet*. The former seems to denote a set in the accepted sense; in particular, someone is either a Congressman or he is not—the Congress does not admit of degrees of membership. This is the only sort of set countenanced by standard set theories. However, *gourmet* (like many other words) seems to denote something very like a set, except that individuals appear to have different degrees of membership. Charles may be more of a gourmet than Harry, and less of a gourmet than Anne. Zadeh 1965, 1971 has constructed the notion of fuzzy set to formalize this sort of intuition.

A fuzzy set  $A$  is defined by a characteristic function  $f_A$  which assigns, to every individual  $x$  in the domain under consideration, a number  $f_A(x)$  between 0 and 1 inclusively, which is the degree of membership of  $x$  in  $A$ .<sup>8</sup> For example, letting  $f_G$  symbolize the characteristic function of the fuzzy set 'gourmet,' perhaps  $f_G(\text{Harry}) = .4$ ,  $f_G(\text{Charles}) = .7$ , and  $f_G(\text{Anne}) = .9$ . If so, the inequalities given above in words are satisfied:  $f_G(\text{Charles}) > f_G(\text{Harry})$ , and  $f_G(\text{Charles}) < f_G(\text{Anne})$ .

That color-category membership is a matter of degree in English is apparent from even casual consideration of some of the ways that color is talked about. We can speak of something as (a) *a good red*, (b) *an off red*, (c) *the best example of red*, (d) *sort of red*, (e) *slightly red*, (f) *yellowish-red*, (g) *blue-green*, (h) *light pink*, or (i) *dark blue*. All these constructions indicate the degree to which the color referred to approximates an ideal example of the root color term. A *good red* has a high degree of similarity to some norm for red. Something that is either *sort of red* or *slightly red* is, in a lesser degree, an approximation to this norm. Lakoff has presented a general treatment of hedges like those found in (a)–(e), where such expressions are analysed as modifying or qualifying the degree to which something is a member of a particular category. In (a)–(e), a color is denoted (1) by reference to some basic color category, and (2) by the use of a hedge which indicates how much the color actually named deviates from the norm for this basic category.

Constructions (f)–(i) also indicate degree of approximation to a norm, but they indicate the direction of variation from the norm as well. A *yellowish-red* fails to be a *good example of red* by virtue of having some degree of yellow. *Blue-greens* are neither *good blues* nor *good greens*; and they differ from *good greens*, e.g., by being blue to a significant degree. While (f)–(g) indicate variation in hue from a norm, constructions such as (h)–(i) indicate variation in brightness and/or saturation from a norm. *Light pinks* are pinks which are to some degree lighter or whiter than the pink norm. *Dark blues* are black to a greater degree than blues that are just blue. A phrase such as *slightly purplish blue* combines specification of both degree and direction of deviation from a norm, to denote a sensation that is only marginally a member of the class of purple things, but is a nearly perfect member of the class of blue things.<sup>9</sup>

Constructions of these sorts are found in all languages, showing that all speaker-hearers recognize and talk about degrees of color-category membership. Since color categorization is in general a matter of degree, color categories are best regarded

<sup>8</sup> Actually, the range of the function may be expressed more generally, but we need not be concerned with such mathematical niceties here.

<sup>9</sup> Rosch (1973, ms) has also extended, into semantic domains other than color, the notion that semantic categories can be defined in terms of approximation to prototypes.

as fuzzy sets. The members of the fuzzy set corresponding to each basic color category are chosen from the set of all possible color percepts. The degree to which each percept is a member of a particular category is specified as a value between zero and unity; each category is thus characterized by the function that assigns, to each color percept, a degree of membership in that category.

Adopting this framework, it can be seen that when B & K asked informants to pick *the best X* or *the best example of X*, they were in effect asking informants to indicate which colors had the highest degree of membership in the category *X*. It was in these judgments of focal colors that B & K found universal agreement. The regions in the color space where these universal basic color-category foci are found can thus be understood as regions where universal basic color-category membership functions reach their maxima, i.e. unity.

Similarly, the non-focal colors included in the extension of any given basic color term are colors with positive but non-maximal degrees of category membership. B & K's study and subsequent ones show that, in these non-focal regions, the basic color categories intergrade. Colors are found here which are members of more than one basic category. In English, this overlap of basic color-category memberships is obvious for colors such as *yellow-green*, *reddish-purple* etc.

It is the existence of these colors with positive degrees of membership in more than one of the basic color categories that produced the variability in boundary judgments reported by B & K. They collected boundary data by simply asking informants to indicate 'all' the colors included in each basic color term. This instruction required subjects to judge class membership categorically. Informants were not allowed to indicate degrees of membership, only 'membership'. An English-speaking informant asked to indicate 'all the reds' could reasonably pick out all the stimuli which are red to a greater degree than they are any other color—thereby excluding, e.g., red-oranges. Or the informant could indicate all the stimuli with even a trace of red, since an expression including *red* or *reddish* would be appropriate to a precise description of any of these stimuli. This ambiguity in the elicitation frame for category boundaries left each informant free to determine what degree of membership he felt was sufficient to permit inclusion of a color in a category. Wide variation in boundary judgments resulted.

McDaniel 1972 has attempted to take this ambiguity in the B & K data into account—arguing that, if absolute universal boundaries exist (as do universal foci), then all the boundary judgments that B & K collected for each category should fall between its focus and some specifiable absolute bounds for it. Examining the extensions of each color term across all B & K's informants for all languages, McDaniel has shown that (1) each universal category has well-defined limits in the color space beyond which the category is never extended, and (2) these absolute boundaries are the foci of the adjacent fundamental color categories. In the fuzzy set framework, these universal boundaries are points in the color space beyond which membership values for a given category are always zero. Once individual variations in boundary placement, as found by B & K, are attributed to individual differences in the selection of criteria for categorical category inclusion, it can be seen that the universal basic color categories have, in addition to universal foci, universal absolute boundaries.

In sum, each basic color category can be regarded as a fuzzy set where the elements in each set are chosen from the set of all color percepts. The degree to which each percept is a member of a particular basic category is specified as a value between zero and unity. Each category is thus distinguished by the set of color percepts which are assigned some positive degree of membership in it. The general structure of these categories, as shown in their pattern of membership assignments, is such that, as one moves through the perceptual color space from the focus of a category toward its boundaries, there is a continuous and gradual decline from unity to zero in the membership values of successive color percepts. In addition, the locations of the basic category foci and their absolute boundaries are universally fixed in the color space.

This formulation of basic color categories as fuzzy sets makes clear the inadequacies of the model of color categorization embodied in the 'focus' and 'boundary' language of earlier works. To talk simply of foci and boundaries entails (or at least invites the inference) that color categories admit of only three degrees of membership: focal member, non-focal member, and non-member. In this respect the focus/boundary model is but a minor variation of the two-degrees-of-membership, discrete-feature model questioned above. The arguments that can be made against that model also apply to the three-degrees-of-membership model implicit in the language of 'focus' and 'boundary'. The data discussed above and below indicate that color categories need to be modeled with continuously graded degrees of membership, not a small finite number such as two or three.

The formalism of fuzzy set theory is a natural device for expressing the continuity of basic color-category membership, allowing accurate description of the full range of memberships admitted. The special status formerly sought for foci and boundaries is nevertheless preserved in the fuzzy set representation, as these become characteristic points (maxima and minima) of the fuzzy membership functions that represent the basic categories. The result is a more accurate description of basic color categories. In addition, when basic color categories are given this non-discrete characterization, the membership functions that then represent the semantic structures of these categories can be derived directly from the neural response functions (discussed in §3, above) that make up the physiological code for color.

**4.2. FUNDAMENTAL NEURAL RESPONSES, FUZZY IDENTITY, AND THE PRIMARY BASIC COLOR CATEGORIES.** Fuzzy set representations of the universal basic color categories red, yellow, green, and blue can be developed directly from neural response functions inherent in the neural code for color. This code consists in the wavelength-dependent behaviors of the opponent neural states **red**, **yellow**, **green**, and **blue** (see §3). As shown in Fig. 5, at a given wavelength, at least one and usually two of these states are evoked. Consider, e.g., point YG in Fig. 5, which represents a (yellowish-green) light of 520 nm. At this wavelength, neither **red** nor **blue** is evoked. The absolute value of the **green** response is 1.8, and the absolute value of the **yellow** response is 0.9. These simultaneously evoked states compete in perceptual effect, and the net effect at this wavelength is the perception of a color zero per cent **red**, zero per cent **blue**, 67% **green**  $\left(\frac{1.8}{1.8 + .9}\right)$ , and 33% **yellow**  $\left(\frac{.9}{1.8 + .9}\right)$ . From computations such as this, we can calculate for each response

state its proportional contribution to the total chromatic response at each wavelength. The results of these calculations are represented by the four curves in Figure 6, which indicate for each wavelength THE PROPORTION OF THE TOTAL CHROMATIC RESPONSE CONTRIBUTED BY EACH OPPONENT STATE. (These proportion-of-total-chromatic-response functions are a standard means of expressing perceived color quality in the vision literature, where they are referred to as 'hue coefficients'.)

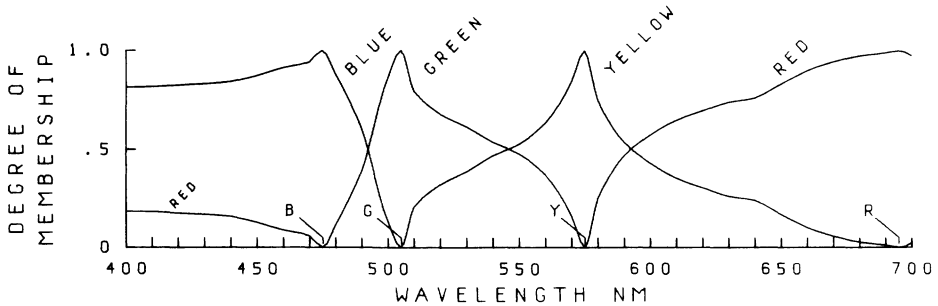


FIGURE 6.

Each of the four functions shown in Fig. 6 has the mathematical properties of a fuzzy set. Every individual in the domain, i.e. light of each wavelength, is assigned a number between zero and unity, inclusively, by each function. These net neural response curves are thus formally the graphs of four fuzzy sets, and substantively the graphs of the four fundamental neural response categories  $f_{\text{red}}$ ,  $f_{\text{yellow}}$ ,  $f_{\text{green}}$ , and  $f_{\text{blue}}$ . It is these four neural response categories that are encoded as the universal semantic categories  $f_{\text{red}}$ ,  $f_{\text{yellow}}$ ,  $f_{\text{green}}$ ,  $f_{\text{blue}}$ . The latter fuzzy categories are the designata of the English words *red*, *yellow*, *green*, and *blue*, and of the words that are exact translations of these in languages of Stages V, VI, and VII. After presenting the evidence for this claim, we will argue further that the semantic categories corresponding to all the basic color terms occurring in the world's languages are either (a) one of these four fundamental response categories; (b) one of two additional fundamental response categories,  $f_{\text{black}}$  and  $f_{\text{white}}$ , to be introduced below; (c) fuzzy unions among these six fuzzy categories; or (d) simple functions of fuzzy intersections among these six fuzzy neural response categories.

The evidence for the identity of the fundamental neural response categories  $f_{\text{red}}$ ,  $f_{\text{yellow}}$ ,  $f_{\text{green}}$ , and  $f_{\text{blue}}$  with the semantic categories  $f_{\text{red}}$ ,  $f_{\text{yellow}}$ ,  $f_{\text{green}}$ , and  $f_{\text{blue}}$  begins with the observation just made that the neural response categories and the semantic categories are both manifestly non-discrete and naturally representable as fuzzy sets.

Second, each of the neural response categories has a single point of maximal membership, as do the membership functions characterizing the corresponding semantic categories.

Third, the particular wavelengths at which the neural response functions reach their maxima—in opponent process terms, their unique hue points—coincide with the wavelengths at which the semantic categories reach their maxima, i.e. have their foci (McDaniel 1972). The points where  $f_{\text{red}}$ ,  $f_{\text{yellow}}$ ,  $f_{\text{green}}$ , and  $f_{\text{blue}}$  reach maximum (points R, Y, G, and B in Figs. 5–6) are referred to as unique hue points: at these points, one of the opponent channels is at its basal response rate and makes no

chromatic contribution, so that the net chromatic response is determined exclusively (and uniquely) by the response state of the other channel. In a replication of B & K's original experiment, McDaniel has shown experimentally that wavelength measures of these unique hue points correspond to wavelength measures of focal yellows, greens, and blues chosen by English-speaking subjects.<sup>10</sup> The wavelengths at which these unique hue/focal points occur are 695 nm for  $f_{\text{red}}$  and  $f_{\text{red}}$ ,<sup>11</sup> 575 nm for  $f_{\text{yellow}}$  and  $f_{\text{yellow}}$ , 510 nm for  $f_{\text{green}}$  and  $f_{\text{green}}$ , and 475 nm for  $f_{\text{blue}}$  and  $f_{\text{blue}}$ .

A fourth item of evidence for the identity of these neural and semantic categories is that both show a gradual decline in membership values as one considers wavelengths at increasing distances, in either direction, from their membership maxima.

Finally, for each pair of matched neural response and semantic categories, the declining membership values become zero at the same points in the color space (see §4.1, above). Membership values for both the neural and the semantic categories decline continuously, though remaining positive, as one moves away from the unique hue point/focus of each category, until the unique point/focus of an adjoining category is reached. At these points membership values become zero, and they remain zero across the remainder of the spectrum. For example, both the neural response category  $f_{\text{green}}$  (as psychophysically determined) and the semantic category  $f_{\text{green}}$  (as inferred from B & K, and from other semantic investigations) have positive values in the interval between 575 nm (unique yellow) and 480 nm (unique blue), but zero values elsewhere.

In sum, the distinctive properties of the semantic categories red, yellow, green, and blue as discovered in semantic investigations correspond precisely to the properties of fuzzy response functions derivable from the opponent process model of the neural mechanisms that underlie color vision. Each of these semantic categories thus bears the identity relation to one of the four fuzzy fundamental neural response categories. We have termed these identity-based semantic categories PRIMARY basic color categories.

The analysis of the full roster of universal semantic color categories to be presented below requires two further fundamental neural response categories,  $f_{\text{black}}$  and  $f_{\text{white}}$ . The existence, at some neural level of response, of categories corresponding to the sensations of black and white is supported not only by the evidence presented here but also by a wide range of psychophysical evidence. At the neurophysiological level, De Valois et al. found that the LGN of the macaque has two

<sup>10</sup> McDaniel did not investigate the relation of unique to focal red because of limitations in the equipment available to him (see fn. 10). There is no reason to suspect, however, that a test of that relation would have yielded a result different from that found for yellow, green, and blue.

<sup>11</sup> This is an oversimplification: actually, the purest sensation of red occurs with presentation to the visual system of no monochromatic light, but rather of a mixture of mostly long-wavelength and a little short-wavelength light (Dimmick & Hubbard 1939). That is, the subjectively purest red is in physical terms a kind of 'purple', a mixture of lights of long and short wavelengths. (McDaniel could not study red in his experiment, because only one source of single wavelength light was available.) This complication does not vitiate the present argument, because the phenomenon in question equally affects the neural response category  $f_{\text{red}}$  and the semantic category  $f_{\text{red}}$ . We may thus, without loss of generality, treat unique red as if it corresponded to a single wavelength near the long end of the visual spectrum.

further types of cells beyond those playing a role in the opponent processes. Both of these were non-opponent, in the sense that each cell was either excited or inhibited relative to its basal rate by light of every wavelength. These were called excitatory non-opponent cells and inhibitory non-opponent cells, respectively. By further experimentation and analysis that we will not report here, De Valois et al. concluded that 'the non-opponent excitatory cells carry luminosity information' (974), while the opponent cells do NOT carry luminosity information (975). (For technical reasons, the investigators were unable to make a comparable test on the non-opponent inhibitory cells.) De Valois et al. concluded: 'The brightness of a light is almost certainly encoded in the firing rate of the non-opponent cells; we have presented an analysis only of the non-opponent excitatory cells, but non-opponent inhibitors appear to give comparable information (this is clearly so for the squirrel monkey as shown by Jacobs [1964]).' That is, at the level of the LGN, beside the two opponent-process systems of cells that determine the four fundamental hue sensations, we also find a separate channel, consisting of brightness-sensitive and darkness-sensitive cells that inform us regarding the whiteness or blackness of a stimulus. On this basis we posit the two additional fundamental neural response color categories  $f_{\text{black}}$  and  $f_{\text{white}}$ .

Membership functions for  $f_{\text{black}}$  and  $f_{\text{white}}$  are given in Figure 7. As we have just suggested, the neural categories  $f_{\text{black}}$  and  $f_{\text{white}}$  do not have the opponent characteristic. They may be perceived in the same part of the visual field at the same

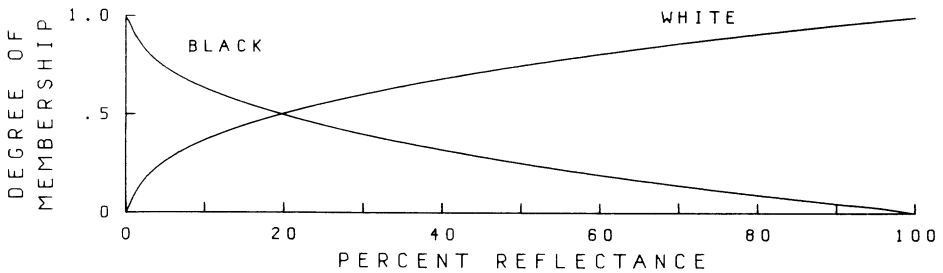


FIGURE 7. Source: Newhall et al. 1943.

time, unlike the opponent pairs red-green and yellow-blue. This is indicated in Fig. 7 by the fact that both curves have positive values throughout the domain. When black and white are perceived together, the resulting sensation is, of course, grey, the lightness (darkness) of the grey reflecting the proportional contribution of the  $f_{\text{white}}$  ( $f_{\text{black}}$ ) response. In Fig. 7, degree of membership in  $f_{\text{white}}$  is taken as the Munsell value (brightness) divided by ten, and degree of membership in  $f_{\text{black}}$  as unity minus this quantity. (The Munsell value scale runs from one to ten.) The abscissa represents the proportion of incident light that a surface reflects.

The membership functions of  $f_{\text{black}}$  and  $f_{\text{white}}$  are identical to those of the semantic categories  $f_{\text{black}}$  and  $f_{\text{white}}$  encoded by Stage V systems along with the four fundamental chromatic categories. Black and white are of course brightness categories, in contrast to the chromatics we have been considering up to now. Introduction of the brightness dimension requires a brief discussion of the fact that the dimensions of hue and brightness interact, and that both interact with the

dimension of saturation (to be discussed below) in determining perceptual color categories. These interactions are illustrated in the representation of the full color space of Figure 8. In this standard schematization of the color space, the dimensions

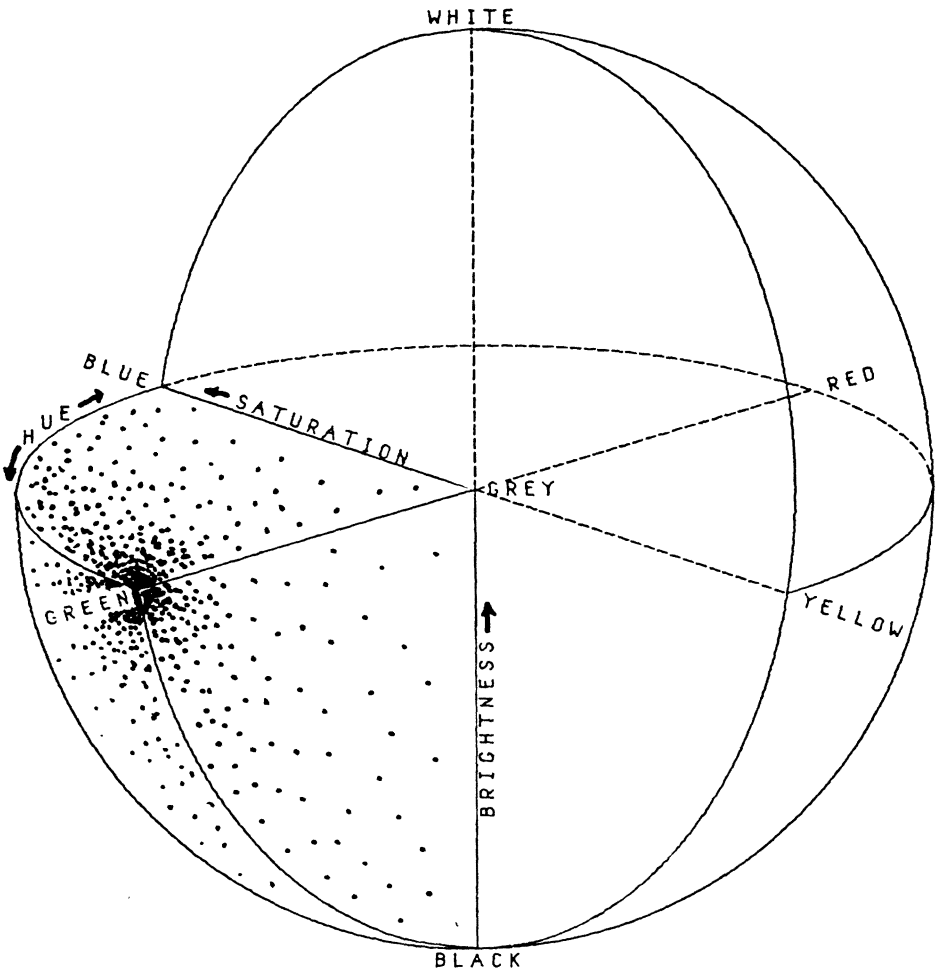


FIGURE 8. Note: Density of stippling indicates degree of membership in green.

of hue and brightness are orthogonal. The black/white (brightness) dimension is a polar axis marking the achromatic core of the color space, a line of greys that varies from black to white. Around this central axis the continuous dimension of hue circles in perpendicular planes. A color's distance out from the central axis toward the surface in a particular brightness-defined hue plane determines its chromatic purity or saturation. Chromatic purity is a measure of a color's vividness or chromaticity, an indication of how free a hue is from dilution by achromatic blacks, greys, or whites. Chromatic purity, i.e. saturation, is thus a function of the relative strengths of the chromatic (opponent) and achromatic (non-opponent)



neural responses. (Fig. 8 represents a cut-away view of the color sphere, with the upper blue-green quadrant removed, as well as both upper and lower green-yellow quadrants.)

Because hue, brightness, and saturation are all dimensions along which the neural responses that code color vary, complete membership specifications for the fundamental neural response categories (and hence the semantic categories defined in terms of them) would require that membership in each of these categories be expressed as a function of all three dimensions. Fig. 8 illustrates this situation for the category  $f_{\text{green}}$ . Membership in  $f_{\text{green}}$  declines, as was shown in Fig. 6, with movement away from focal green on the hue circuit. Membership in  $f_{\text{green}}$  also declines with movement from focal green toward the achromatic axis, since the greens along this path are becoming progressively less saturated. Perceived hue remains constant, but chromaticness decreases. This decline in green membership is matched by a corresponding increase in the membership of some achromatic category.

As the spheroidal shape of the color space indicates, the maximal saturation that green can have decreases, as greens darker and lighter than focal green are considered. Thus membership in  $f_{\text{green}}$  is also a function of brightness. In defining the fundamental chromatic categories, formal description of these three-dimensional membership functions is possible, but the arguments of this paper are not affected thereby; and exposition is facilitated if (as in the beginning of this section and in Fig. 6) brightness and saturation are tacitly held constant.

**4.3. FUZZY UNIONS, COMPOSITE CATEGORIES, AND EARLY STAGE BASIC COLOR-TERM SYSTEMS.** In standard set theory, the union of two sets A,B is the set that contains everything that is in A, or in B, or in both. Thus, if the set of people eligible for cheap tickets is the union of the set of registered students and the set of people under twelve years old, people who are either registered students or under twelve or both are eligible. The union of the fuzzy sets A,B, which we will denote as 'A OR B' is defined by a function which assigns to each individual  $x$  the larger of the two values  $f_A(x)$ ,  $f_B(x)$ . In symbols, we define the union of two fuzzy sets A,B by this equation:

$$(1) f_{A \text{ OR } B} = \text{Max} [f_A, f_B]$$

Let us suppose that we are interested in forming the union of the fuzzy sets 'competent basketball player', B, and 'competent landscape painter', P. (Perhaps we are composing a guest list for a potentate whose principal avocations are basketball and landscape painting.) Let us further suppose that Kareem Abdul-Jabbar has a degree of membership of .99 in B and .02 in P; while Joe Furge, who has played semi-pro basketball and sold a few watercolors, has degrees of membership in these fuzzy sets of .5 and .6, respectively. Intuitively, our potentate will be more interested in meeting Abdul-Jabbar than in meeting Furge; and so we are glad to note that our definition of (B OR P) gives a higher degree of membership to Abdul-Jabbar (.99) than to Furge (.6), despite the fact that the sum and product of Furge's degrees of membership both exceed those of Abdul-Jabbar. In standard set theory, an individual is in the union of two sets if it is in either set; in fuzzy set theory, an individual is in the union of two sets to the greatest degree that it is in either set (not to the degree that it is in both, whatever that might mean).

All basic color-term systems prior to Stage V have at least one term that encodes a fuzzy union of two or more of the six fundamental neural response categories. A category formed by such a union will be referred to as COMPOSITE, in the sense of being composed of all the colors that have any degree of positive membership in any of the fundamental response categories from which it is formed. The composite category found most often is GRUE, equivalent to the fuzzy union  $f_{\text{green OR blue}}$ ; cf. Figure 9. Gatschet 1879 described a grue category as characteristic of many American Indian languages; and the recent world surveys by B & K and by Bornstein 1973a,b have shown that many of the world's languages have a basic color term that means grue.

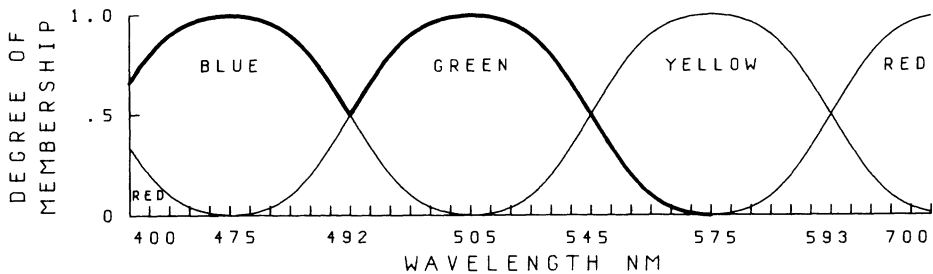


FIGURE 9. Note: heavy line indicates  $f_{\text{green OR blue}}$ .

The heavy line in Fig. 9 represents the membership values that result from evaluating  $f_{\text{green OR blue}}$ . (The wavelength scale in Fig. 9 has been transformed from the uniform scale of Fig. 6, to normalize the membership functions of the fundamental response categories.) An interesting feature of this membership function is that a blue-green-appearing stimulus of, say, 492 nm is assigned a lower degree of membership in grue than a stimulus near either the blue or green focal points, 475 nm and 510 nm, respectively. This corresponds to the claim that, in languages that encode grue, the colors intermediate between focal blue and focal green are relatively poorer members of grue than either focal blue or focal green, even though these intermediate colors are nearer the mathematical center of the category. This somewhat counter-intuitive prediction, a direct product of the treatment of grue as a fuzzy union, is supported by evidence on the distribution of grue focal choices from anthropological field studies (Kay 1975a). In the experimental studies that have been conducted subsequent to B & K's work, it has been frequently observed (as B & K found for Tzeltal) that early-stage categories, including grue, are often multiply-focused. In these studies, focal grue selections have often proved to be bimodal, being chosen from both the focal blue and focal green regions. But grue has never been found to be focused in the intermediate blue-green region. The absence of focal choices from this intermediate region is strong evidence that these colors have lower grue membership values, and that grue has the membership structure stipulated by the fuzzy union analysis.

Fifty-seven composite categories can be formed by taking all possible unions of two or more of the six fundamental neural response categories; but only three in addition to grue (cool) have actually been observed as basic color categories. These observed composites are 'warm' (red OR yellow), 'light-warm' (white OR red

OR yellow), and ‘dark-cool’ (black OR green OR blue) (Kay 1975a; McDaniel 1974, MS). Only Stage I systems, like that of the Dani, are made up entirely of composite categories. At Stage I, all the fundamental neural response categories are joined into the two composite categories light-warm and dark-cool. Systems at Stages II, III, and IV contain both primary and composite categories. The transitions between these stages take place through the partial or total decomposition of composite categories, with the separate encoding of the primary categories of which they were composed. Grue is the last of the composite categories that languages lose; hence its frequency. The transition to Stage V, where all basic color categories are primary, occurs when grue is replaced by its component primaries blue and green.

Beyond Stage V, the development of basic color-term vocabularies follows a different pattern. Stage transitions no longer occur through the addition of primary categories and the loss of composites. Instead, basic color-term lexicons expand by the addition of terms that refer to regions of the color space where the fundamental neural response categories overlap. These later, ‘derived’ categories—brown, orange, pink, purple, and grey—are related to the fuzzy intersections of the fundamental response categories.

**4.4. FUZZY INTERSECTION, DERIVED CATEGORIES, AND LATER STAGE BASIC COLOR-TERM SYSTEMS.** In standard set theory, the intersection of two sets A,B is the set that contains just those individuals that are members of A and also members of B. The intersection of fuzzy sets A,B—denoted here as ‘A AND B’—is defined by a function that assigns to each individual x the smaller of the two values  $f_A(x)$ ,  $f_B(x)$ . In symbols, the intersection of the fuzzy sets A,B is defined by this equation:

$$(2) f_{A \text{ AND } B} = \text{Min} [f_A, f_B]$$

If the category corresponding to the fuzzy set ‘green AND yellow’ is called *chartreuse* (as it is by some speakers of English), then *chartreuse* encodes a category whose membership function has the values indicated by the heavy line in Figure 10

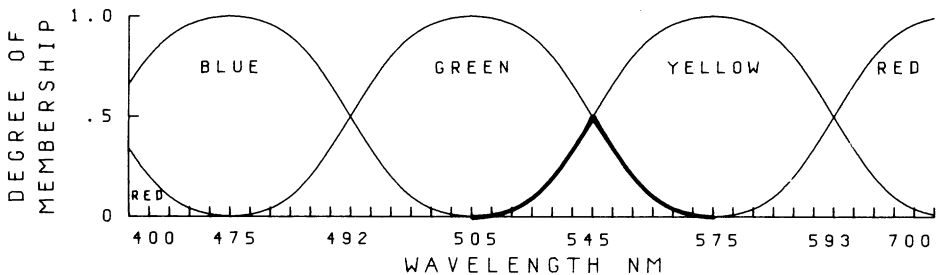


FIGURE 10. Note: heavy line indicates  $f_{\text{green AND yellow}}$ .

(or perhaps some monotone function thereof, a point to which we return below). Colors below 505 nm and above 575 nm have zero degree of membership in this category. As one advances to the right from unique green (505 nm) and to the left from unique yellow (575 nm), one initially finds quite poor examples of chartreuse; but as one continues from either end, the stimuli exemplify chartreuse increasingly well, with membership values reaching a maximum somewhere around 545 nm.

*Chartreuse* is a non-basic color term in English; yet the category it encodes shares an important property with the basic color categories that are added in the later stages of basic color-term development, i.e. brown, orange, pink, purple, and grey. Like *chartreuse*, these later-stage basic color categories have membership functions whose positive (non-zero) ranges are restricted to regions of the color space where two of the fundamental response states co-occur, i.e. to regions where two adjacent fundamental response categories overlap. Brown is found in the region where **yellow** and **black** overlap, pink where **red** and **white** overlap, purple where **red** and **blue** overlap, orange where **red** and **yellow** overlap, and gray where **black** and **white** overlap. Thus, as with *chartreuse*, the regions over which later-stage basic categories have positive values are regions where fuzzy intersections of certain fundamental response categories are positive.

Fuzzy intersections can thus be used to specify the regions where later-stage basic color categories have positive membership values; but the actual membership values which these categories have in these regions are not precisely the values that the simple fuzzy intersections yield. Consider orange: in Figure 11A, this category

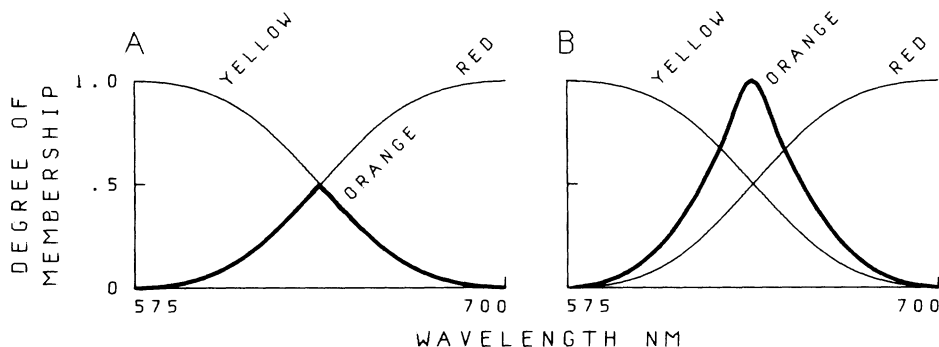


FIGURE 11. Note: B is the preferred formulation for reasons given in the text.

is modeled as identical to the fuzzy set  $f_{\text{yellow AND red}}$ . This model makes two claims that are contradicted by both casual and experimental observations. First, the model claims that there are no really good examples of orange. The maximum of the membership function for orange, when modeled this way, is well below the maxima for the previously encoded primary categories yellow and red; in fact it is precisely 0.5. This contradicts experimental evidence that subjects are essentially as confident about assigning good examples of orange to orange as they are about assigning good examples of red to red, or good examples of any other primary category to that category. Many subjects declare that for them orange is just as fundamental a category—just as distinctive a color sensation—as any of the six primaries (Sternheim & Boynton 1966). Second, this model implies, not only that there are no really good examples of orange in comparison to the primaries, but also that there is no hue sensation with a higher degree of membership in orange than in either red or yellow. Every hue point under the left half of the orange membership function in Figure 11A has a higher degree of membership in yellow than in orange, and every point under the right half has a higher degree of membership in red than in orange. As is shown by subjects' confidence about the existence

of good oranges, there is unquestionably a range of hues that speakers of English more readily label *orange* than either *yellow* or *red* in both experimental and natural contexts.

These observations require that we abandon the simple fuzzy intersection model of orange, illustrated in Fig. 11A, in favor of a model which allows some colors to have a higher degree of membership in orange than in either yellow or red, and which also has a membership function whose maximum is unity. Since orange encodes the simultaneous occurrence of yellow and red, the best examples of orange (i.e. the colors with the highest degrees of membership in orange) will be those that are most nearly equal in their yellowness and redness. In other words, as the absolute difference between a color's degree of yellowness and degree of redness decreases, its orangeness approaches unity. The fuzzy set 'orange' is therefore defined, over the region of the spectrum where  $f_{\text{red AND yellow}} > 0$ , by

$$(3) f_{\text{orange}}(x) = 1 - |f_{\text{yellow}}(x) - f_{\text{red}}(x)|$$

The membership function for orange that this equation yields is shown by the heavy line in Figure 11B. This function reaches its maximum of unity at the point where the functions  $f_{\text{yellow}}$  and  $f_{\text{red}}$  intersect. Also, at the yellow and red unique hue points, the value of the absolute difference term is unity; hence the degree of membership is zero, as desired, at these boundary points. Equations of the same general form can be constructed to derive membership functions for all of the later-stage categories. The basic categories whose membership functions are formed in this fashion are thus referred to as DERIVED basic color categories.

While Equation 3 is not the simple fuzzy intersection  $f_{\text{yellow AND red}}$ , it can be rewritten in a form that relates  $f_{\text{orange}}$  directly to the fuzzy intersection  $f_{\text{yellow AND red}}$ . It will be recalled that each of the four fundamental chromatic response categories reflects the PROPORTION of total chromatic response resulting from particular response state at each wavelength. Consequently, at each wavelength, the sum of the four fuzzy set functions  $f_{\text{red}}$ ,  $f_{\text{yellow}}$ ,  $f_{\text{green}}$ , and  $f_{\text{blue}}$  is unity. Furthermore, within the interval in which orange is defined, the functions  $f_{\text{blue}}$  and  $f_{\text{green}}$  both have the value zero; hence, in this interval, the functions  $f_{\text{red}}$  and  $f_{\text{yellow}}$  sum to unity. Recalling that Equation 2 defines  $f_{\text{red AND yellow}}$  as the minimum of the functions  $f_{\text{red}}$  and  $f_{\text{yellow}}$ , it follows that, in this interval, (a) the smaller of the two functions  $f_{\text{red}}$ ,  $f_{\text{yellow}}$  will always be equivalent to  $f_{\text{red AND yellow}}$ ; and (b) the larger of the two will always be equivalent to  $1 - f_{\text{red AND yellow}}$ . Thus Equation 3 may be rewritten

$$(3') f_{\text{orange}}(x) = 1 - (1 - f_{\text{red AND yellow}}(x) - f_{\text{red AND yellow}}(x)) \\ = 1 - 1 + 2f_{\text{red AND yellow}}(x)$$

$$(4) f_{\text{orange}}(x) = 2f_{\text{red AND yellow}}(x)$$

That is, membership in a derived category such as orange, though not equivalent to the fuzzy intersection of the membership functions of its two constituent categories, is equivalent to twice this intersection.<sup>12</sup>

While Equations 3–4 assign the same membership function to orange, they may represent two distinct cognitive processes in the formation of derived categories.

<sup>12</sup> We are indebted to John R. Atkins for pointing out to us the equivalence of Equations 3 and 4.

Assuming that the fuzzy set operations referred to in these equations correspond in some way to actual neural and cognitive events, 3 and 4 make different claims about the cognitive processes available for derived color-category formation. Equation 4 suggests that the only operations needed for the formation of a derived category are fuzzy intersection and scalar multiplication. Equation 3, however, suggests the existence of three entirely different cognitive operations. Membership in orange, according to 3, is determined by taking (a) the complement of, (b) the absolute value of, (c) the difference between two functions representing membership in fundamental response categories. Making a choice between 3 and 4 can thus have general cognitive implications. For example, if 3 is taken as the proper representation of orange, this encourages consideration of the role which complement operations play in fuzzy domains. Accepting 4 provides no such motive. In addition, the simple fact that there may be processually distinctive cognitive paths by which the same semantic structure can be formed suggests that semantic processes and semantic structures may sometimes be independent. For example, one might be able to examine experimentally the hypothesis that different individuals—or a given individual under varying conditions—would process the semantic category orange in ways that sometimes suggested 3, and sometimes 4.

There is indirect evidence indicating that these and the other equations used in this paper to relate semantic to neural structures are in fact representative of real neural and cognitive events, albeit at a grosser level. Taking the equations presented in this paper as models of actual cognitive processes, using either 3 or 4 as a model of the cognitive process which underlies derived categories would be to assert that these categories have more complex cognitive bases than do the primary categories based on the identity relation. A likely observable effect of this greater cognitive complexity would be an increase in the time needed by subjects to determine derived category memberships, over the time needed to determine primary category memberships. Heider found just this effect in a study of English focal color naming, even though the 'focal' colors used in the study were probably not optimal for a proper test of this hypothesis. Her finding was that the 'primary focal colors black, white, blue, green, yellow, red were named significantly more rapidly than the non-primary focal colors pink, brown, orange, purple;  $t(22) = 2.86, p < .01$  (for correlated measures)' (1972b:15).<sup>13</sup>

<sup>13</sup> The absolute order of reaction times was not in exact agreement with the hypothesis. Whereas orange, purple, pink, and brown should have had the four longest response latencies, the actual rank order of latencies was (from shortest to longest) black, yellow, white, PURPLE, blue, red, pink, brown, GREEN, orange (Heider 1972b:15). But there are at least two difficulties with Heider's methodology which suggest caution in interpreting these results. First, she determined 'focal' colors by taking the 'geometric centers' of the areas enclosing all the focal choices in the B & K data (p. 9; Heider 1972b:12). Unfortunately, focal greens in B & K were over-extended toward blue because Vietnamese *xanh* 'grue', which has a near-unique blue focus, was misclassified as green. Thus the 'focal' green which Heider used, Munsell 7.5G 5/10, is rather distant in both hue and saturation from the 10GY and 2.5G, Munsell colors typically chosen as focal greens. This particular difficulty might help explain the especially low position of green in the response latency rankings.

Second, and more important, B & K obtained their focal color judgments, which Heider used to select her stimulus materials, under a standard illuminant A, while Heider conducted her tests with daylight fluorescent illumination. The difficulty here is that object colors can

Apart from the details of their formal (or cognitive) representation, another question raised by the emergence of derived categories in post-Stage V systems is the possible effect of the emergence of these categories on the primary categories from which they are derived. It was noted above that the encoding of a primary category is accompanied by decomposition of some previously existing composite category. What might seem a comparable effect of the encoding of a derived category would be contraction of the membership functions of the primary categories from which it is derived. This possibility is illustrated for orange, yellow, and red in Figure 12B.

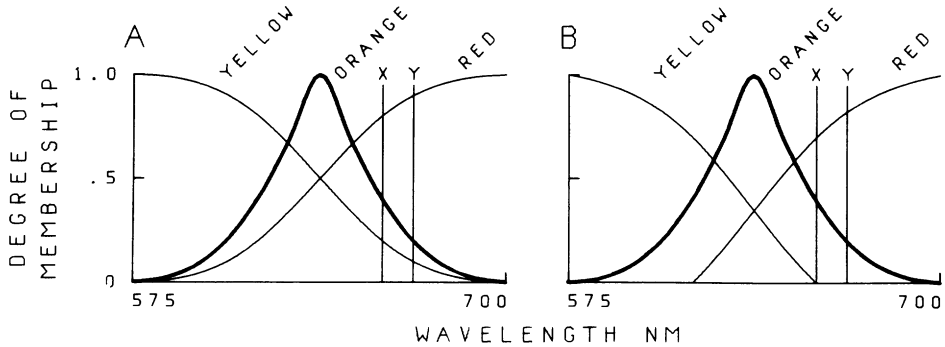


FIGURE 12. Note: Formulation A, in which there is no change in the primary categories, is preferred for reasons given in the text.

There are data, however, which indicate that the non-contracted membership functions for yellow and red shown in Figure 12A are retained by yellow and red even after orange is encoded. In an experiment conducted by Sternheim & Boynton, subjects were asked to name ten long-wavelength (530–620 nm) stimuli under a variety of naming conditions. One condition allowed subjects to use the names *yellow*, *red*, and *orange*. In this condition, the yellow and red naming-functions were like the retracted functions shown in Fig. 12B. A second condition required that subjects use only the terms *yellow* and *red*, although a null response was accepted if a subject felt that neither *yellow* nor *red* was appropriate. In this condition, the naming functions for yellow and red expanded to approximate the non-contracted functions shown in Fig. 12A. These results suggest that, as derived categories are encoded, the underlying membership functions of the primary categories remain unchanged, while competition between the memberships which colors have in the old primary categories and the newly encoded derived categories alters surface naming behavior in many (but not all) contexts.

Our intuitions about the use of primary basic color terms in comparative constructions also support the model of orange, yellow, and red depicted in Fig. 12A. An English speaker discussing the two oranges denoted by x and y in Figures 12A and 12B might distinguish them by saying that y is redder than x, since in

---

shift radically with changes in illuminant conditions. Thus Heider's relatively well-defined illuminant A focals may have been poor approximations to the focal colors when viewed under her daylight fluorescent illuminant conditions. A less than ideally constituted set of stimuli for the given experimental conditions may therefore be partly responsible for the somewhat equivocal nature of these results.

both figures  $f_{\text{red}}(y) > f_{\text{red}}(x)$ . By the model represented in Fig. 12A, the speaker could as well say that  $x$  is yellower than  $y$ , since  $f_{\text{yellow}}(x) > f_{\text{yellow}}(y)$ . However, if  $f_{\text{yellow}}$  is contracted, as in Fig. 12B, then  $f_{\text{yellow}}(x) = f_{\text{yellow}}(y) = 0$ , and our speaker won't be able to use relative yellowness to distinguish  $x$  and  $y$ . Thus, by the model in Fig. 12B, there is a region of the spectrum where a speaker can describe a first orange as redder than a second, but not the second as yellower than the first. An experiment might show this to be the case; but our intuition is that any orange that is redder than another is also less yellow, and therefore the second will always be describable as yellower than the first. This intuition is taken as further argument in support of the analysis of derived categories shown in Figures 11B and 12A.

While a simple intersection model of orange and the other derived basic categories has been rejected, it may be that simple intersection is the appropriate model for secondary color categories corresponding to non-basic terms such as *chartreuse*. Like derived categories, these categories are psychophysically mixtures of the fundamental response categories. Treating *chartreuse*, for example, as  $f_{\text{yellow AND green}}$  would locate  $f_{\text{chartreuse}}$  in the appropriate region of the color space with an intuitively acceptable membership function. In addition, this formal move would entail the empirical claim that, for speakers with *chartreuse* in their active vocabulary, any stimulus which they would judge as *chartreuse* would be judged as an equally good or better representative of yellow or of green. This speculation is not immediately rejected by introspection; however, we know of no systematic data that support it. If some subjects were found to react to appropriate stimuli with the labels *chartreuse*, *yellow*, and *green* in the manner just suggested, while others were found to use *chartreuse* in the way they generally use labels for basic (derived) categories such as orange, this could be taken as evidence for the assertion that *chartreuse* has achieved basic color-term status for the latter group, but not the former. Fuzzy set-theoretic representations of color categories might thus prove useful in yet another way, this time in their ability to aid in the sometimes difficult task of distinguishing basic and non-basic categories. (Several controlled field studies have shown orderly variation among speakers in a given community with regard to the number of basic color categories they have; see Kay 1975 for a summary.)

The use of characteristics of color-category membership functions to distinguish between basic and non-basic categories suggests a natural way to enrich the notion of basicness itself. The basicness of a color category could be taken as the degree to which its membership function approximates a characteristic (prototypic) membership function of one of the three types of basic category described above. We are not aware of any extant empirical evidence bearing on this speculation.

**4.5. SUMMARY OF THE FUZZY SET FORMULATION FOR BASIC COLOR CATEGORIES.** Basic color terms are universally associated with a small set of non-discrete though well-defined semantic categories. The formalism of fuzzy set theory provides a natural framework for the description of these universal and non-discrete categories. It also allows a formal set-theoretic characterization of the relations which these fuzzy categories bear to neural categories inherent in the perception of color.

All basic color categories are formed from the human visual system's six fundamental response categories by one of three fuzzy-logical operations: identity,



fuzzy union, or fuzzy intersection, sometimes along with one or more non-fuzzy operations. Identity with the six fundamental response functions is the basis of the primary basic color categories black, white, red, yellow, green, and blue. Fuzzy unions of fundamental response categories are the basis of the four composite basic-color categories light-warm, dark-cool, warm, and cool (grue). Fuzzy intersections of fundamental response categories are the basis of at least five derived basic color categories—brown, pink, purple, orange, and grey.<sup>14</sup> Thus where B & K described eleven universal basic color categories of a single logical type, there are in fact at least fifteen basic color categories of three types (McDaniel 1974, MS), distinguished by the relations which their semantic structures bear to the visual system's fundamental neural response categories for color. Table 2 is a summary listing of these categories, showing the three types of fuzzy set operations that relate them to the fundamental neural response categories.

NEURAL RESPONSE CATEGORIES	SEMANTIC CATEGORIES BASED ON IDENTITY
$f_{\text{black}}$	$f_{\text{black}}$ = black
$f_{\text{white}}$	$f_{\text{white}}$ = white
$f_{\text{red}}$	$f_{\text{red}}$ = red
$f_{\text{yellow}}$	$f_{\text{yellow}}$ = yellow
$f_{\text{green}}$	$f_{\text{green}}$ = green
$f_{\text{blue}}$	$f_{\text{blue}}$ = blue
	SEMANTIC CATEGORIES BASED ON FUZZY UNION
$f_{\text{black OR green OR blue}}$	$f_{\text{black OR green OR blue}}$ = dark-cool
$f_{\text{white OR red OR yellow}}$	$f_{\text{white OR red OR yellow}}$ = light-warm
$f_{\text{red OR yellow}}$	$f_{\text{red OR yellow}}$ = warm
$f_{\text{green OR blue}}$	$f_{\text{green OR blue}}$ = cool; grue
	SEMANTIC CATEGORIES BASED ON FUZZY INTERSECTION
$f_{\text{black + yellow}}$	$f_{\text{black + yellow}}$ = brown
$f_{\text{red + blue}}$	$f_{\text{red + blue}}$ = purple
$f_{\text{red + white}}$	$f_{\text{red + white}}$ = pink
$f_{\text{red + yellow}}$	$f_{\text{red + yellow}}$ = orange
$f_{\text{white + black}}$	$f_{\text{white + black}}$ = grey

TABLE 2.

The basic color categories formed by each fuzzy logical operation share a number of semantic properties. The fuzzy membership functions of the six primary basic categories, being based on identity with the fundamental response categories, share all the properties common to the fundamental categories. Each category has a single membership maximum, i.e. a single focus, at a physiologically defined unique hue point. Membership in these categories declines continuously as colors

<sup>14</sup> The formulation we have given for orange generalizes directly to purple, whose two constituent primary categories are both hue categories. The same may be true for grey, both of whose constituent primaries are achromatic. The algebraic formulation given in Equation 3 may perhaps generalize also to the categories pink and brown, each of which is based on one hue constituent and one achromatic constituent; but the special relations obtaining between the primary hues that permit the derivation of 4 from 3 do not apply in the case of pink and brown. Henceforth, we will use the symbol '+' to denote whatever function, known or unknown, relates two primary categories to produce a derived category. Thus we will write ' $f_{\text{orange}} = f_{\text{red + yellow}}$ ', ' $f_{\text{pink}} = f_{\text{red + white}}$ ', etc. The symbol '+' so used does not denote a single, well-defined fuzzy operation.

perceptually more distant from these maxima are considered, reaching zero at the unique hue points adjacent to the focal unique hue points. These characteristics of primary basic color-category membership functions are illustrated in Figs. 6–7.

In contrast, composite categories, based on fuzzy union, have multiple membership maxima (foci); and colors increasingly distant from a focus do not necessarily have lower degrees of membership. The unique hue points of the two or three fundamental response categories encompassed by a composite category are the (multiple) foci of the category. Membership in the category does decline to zero for colors between one of the focal unique hue points and a non-included unique hue point. For the colors between unique hue points that define foci, membership values decline away from each focus, but do not reach zero. Instead, a positive minimal membership value is reached somewhere between the two foci. Fig. 9 shows a membership function with these characteristics.

Derived category formation from modified fuzzy intersections produced membership functions with structural characteristics analogous to those of the primary basic color categories. As with the primaries, derived category membership functions reach maxima at single points in the color space, declining continuously from these foci to zero at the unique hue points adjacent to them. But while the primary and derived categories share these structural characteristics, they differ significantly in their patterns of neural association. In particular, while the foci of the primary (and composite) categories are associated with physiological unique hue points, the foci of the derived categories are not. Derived category foci are associated with points in the color space perceptually equidistant between the two unique hue points that define the category's boundaries. Simple fuzzy intersections yield functions that assign the correct colors positive degrees of membership in the derived categories, but a scalar multiplication of these functions is necessary to produce the actual membership values appropriate for these categories. Figs. 11B and 12A illustrate a membership function of this type.

In general, formulating basic color categories as fuzzy sets, rather than in terms either of discrete features or of foci and boundaries, allows us to construct models of these categories more in accord with our empirical knowledge of their semantic structure. It also provides descriptions of basic color categories that can be derived directly from the neural response patterns that underlie the perception of color. As discussed below, the fuzzy set formalism also allows a succinct formal restatement of B & K's model of basic color-term evolution.

#### THE EVOLUTION OF BASIC COLOR CATEGORIES

**5.1. COMPOSITE, PRIMARY, AND DERIVED BASIC COLOR CATEGORIES AND THEIR RELATIONS TO THE PROCESSES AND PATTERN OF BASIC COLOR-TERM EVOLUTION.** AS noted in the discussions above, primary, composite, and derived basic color categories differ not only in their characteristic semantic structures, and in the relations they bear to the fundamental neural response categories, but also in the stages of basic color-term development with which they are associated. Composite categories are found only prior to Stage V. No primary categories exist at Stage I, and only at Stage V and beyond are all the primaries encoded. The derived categories begin appearing only after Stage V. Figure 13 illustrates these relations,

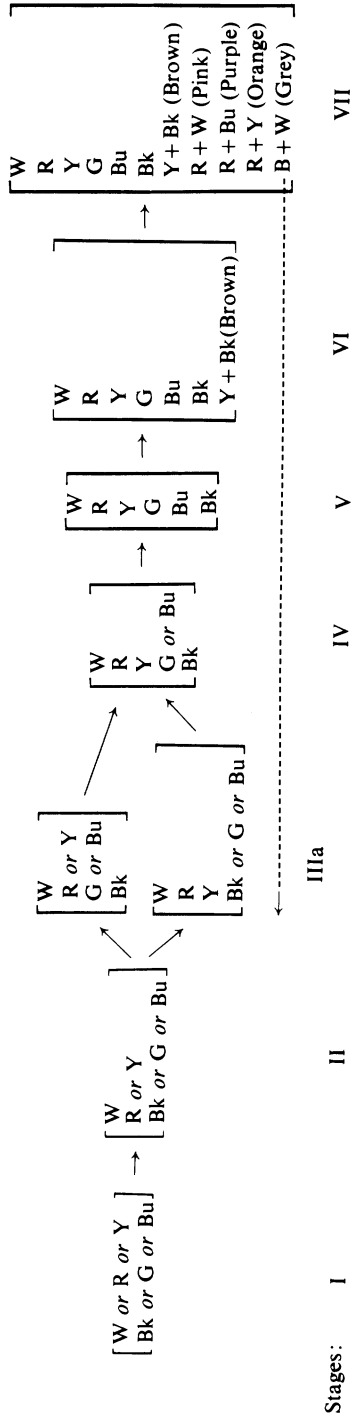


FIGURE 13.

recasting the original B & K evolutionary sequence in terms of the fuzzy set-theoretic basic color categories described above.<sup>15</sup>

Fig. 13 embodies a re-interpretation of the evolutionary sequence that views the development of basic color-term lexicon not as the successive encoding of foci, but as the successive differentiation of previously existing basic color categories (McDaniel 1974, ms). Beginning at Stage I with the composite categories light-warm and dark-cool (as illustrated by Dani *mola* and *mili*; Heider 1972a), Stages II–V in the evolutionary sequence involve the decomposition of composites into their constituent primaries. First the white component of the light-warm category is lexically distinguished, producing the white, warm, dark-cool configuration of Stage II systems. The Bellonese color system described by Kuschel & Monberg is of this type. Next, either dark-cool is decomposed into black and cool, or warm is split into its red and yellow constituents. The first alternative results in a Stage IIIa configuration, as described for Aguaruna by Berlin & Berlin and for West Futunese by Dougherty 1974, 1975. The second leads to a stage IIIb system, as Hage & Harkes describe for Binumarien. Whichever of these decompositions is not accomplished at Stage III is achieved in the transition to Stage IV, producing a system that lexicalizes the primaries black, white, red, and yellow, as well as the most enduring of the composites, grue (cool). Such Stage IV systems have been experimentally observed among more acculturated speakers of Aguaruna (Berlin & Berlin) and Futunese (Dougherty 1975), and among the Mam (Harkness); from the investigation of general ethnographic and linguistic sources, Stage IV systems appear to include the majority of New World languages (Bornstein 1973a,b; Hays et al. 1972). Stage V is achieved with the differentiation of cool into its green and blue primary constituents, completing the decomposition of the composites. Mam Spanish (Harkness) is an example of such a system.

At Stage VI and beyond, differentiation proceeds through the encoding of intersections of the primaries. At Stage VI, brown encodes the intersection of yellow and black. Stage VII extends the privilege of name to white+red (pink), red+yellow (orange), red+blue (purple), and usually to black+white (grey).<sup>16</sup> This view of the later-stage, derived categories suggests that if no language investigated so far has more than eleven basic color terms, with the possible exception of Russian (see B & K, 35–36), this is more an accident of the present moment in world history than a theoretical inevitability. Russian *goluboy* ‘light blue’ (white+blue) is a potential instance of a twelfth basic color term; it is surely a basic term for some Russian speakers, though probably not for all.<sup>17</sup> There is no certainty that *goluboy* will not, at some point in the future, achieve basic term status for all

<sup>15</sup> The empirical facts relevant to the restatement of the encoding sequence represented in Fig. 13 are presented in the works cited in fn. 4, above.

<sup>16</sup> The dotted arrow in Fig. 13 indicates that grey may occur (albeit infrequently) ‘as a wild card at various points in the sequence’ (B & K, 45). Whereas B & K originally guessed that grey might occur, ‘say at any point after Stage IV’, more recent information shows that it may occur at any stage from IIIa onward, or possibly even earlier (Barry Alpher, p.c.; MacLauray ms). For further discussion, see Kay (1975a:261).

<sup>17</sup> Cf. Daly ms. Other Slavic languages have monolexic terms for ‘light blue’, but these appear to be basic terms for very few speakers, if any.

speakers of Russian. Similarly, it is possible that several now non-basic color terms in English, used to name intersections of fundamental response categories, will become basic in the future, e.g. *aqua/turquoise* (green + blue), *maroon/burgundy* (black + red), and *chartreuse/lime* (yellow + green). Some of these may already be basic terms for some speakers. The process that characterizes derived category formation has not been logically exhausted by any known language; so there is no apparent reason to believe that the process will not continue, extending basic color-term lexicons beyond their present eleven terms.

**5.2. FUZZY PARTITIONS: A FORMAL CHARACTERIZATION OF THE ENCODING SEQUENCE.** In formalizing the discussion of the evolutionary sequence, it is desirable to extend fuzzy set theory in a minor way by developing the concept of fuzzy partition. In standard set theory, a collection of sets partitions a set  $S$  just if (a) each set in the collection is a subset of  $S$ , (b) everything in  $S$  is a member of one of these subsets, and (c) nothing in  $S$  belongs to more than one of these subsets.<sup>18</sup> Such subsets are often called the cells or blocks of the partition.

There is an intuitive use of the word 'partition' in which it is obvious that every color terminology 'partitions' the universe of color percepts, and in which the transition from stage to stage in the evolutionary sequence (until Stage V is reached) involves moving from a coarser to a finer partition. This intuitive usage can be explicated as follows. We note first that each color category is a fuzzy subset of (is fuzzily contained in) the set of all color percepts. This follows from the fact that, for each individual in the domain, its degree of membership in the fuzzy set 'color percept' (which is unity) is greater than or equal to its degree of membership in any given color term.<sup>19</sup> Hence any color category is formally a fuzzy subset of the set of color percepts. Thus, in speaking of fuzzy partitions, we can directly take over condition (a) of standard set theory that the cells of the partition be subsets of the set being partitioned. Condition (b), the exhaustiveness condition, can also be directly taken over into the definition of fuzzy partition, if we translate 'belongs to' as 'having a non-zero degree of membership in'.<sup>20</sup>

<sup>18</sup> For example, the cells of a jail normally provide a partition of the set of prisoners, since every prisoner is assigned to a cell, and no prisoner is assigned to more than one cell.

<sup>19</sup> The set of color percepts is, of course, a standard set and is therefore a fuzzy set, since every standard set is also a fuzzy set—in particular, the special case of a fuzzy set in which all values of the characteristic function are either zero or unity.

A fuzzy set  $A$  is contained in (is a subset of) a fuzzy set  $B$  if, for all  $x$  in the domain of discourse,  $f_B(x) \geq f_A(x)$ .

<sup>20</sup> Zadeh cautions: 'the notion of "belonging" [membership], which plays a fundamental role in the case of ordinary sets, does not have the same role in the case of fuzzy sets. Thus, it is not meaningful to speak of a point  $x$  "belonging" to a fuzzy set  $A$  except in the trivial sense of  $f_A(x)$  being positive' (1965:342). We will henceforth use 'belong' in just this sense, since with respect to color the notion is not trivial. Zadeh appears to have in mind applications in which few, if any, individuals in the relevant domain will have zero membership in any of the fuzzy sets under discussion. Such is not the case in color. For example, no color percept belongs both to red and to green. The same holds for blue and yellow. In discussing color percepts and categories, it is often of interest to know if a given percept belongs (to ANY POSITIVE degree) to a certain category. Similarly, it is often of interest to know whether two categories have members in common, like green and yellow, or are disjoint, like green and red.

While conditions (a) and (b) can be taken over from standard set theory, condition (c) must be modified, since distinct but adjacent basic categories have considerable overlap. As shown above (see Fig. 6), every green above 505 nm (unique green) is also to some positive degree yellow, and every green below 505 nm is to some degree blue. Thus all greens except the unique green at 505 nm are to some degree also yellow or blue. These are not suppositions of the formalism, but facts for which the formalism must account.

These facts can be dealt with by replacing the mutual exclusion condition (c) of standard set theory by a statement to the effect that each cell of a fuzzy partition has at least one member that belongs to no other cell. Thus, given a fuzzy set  $S$  and a collection of fuzzy sets  $F = \{F_1, F_2, \dots, F_n\}$ ,

- (5)  $F$  is a fuzzy partition of  $S$  just if
- (i) each  $F_i$  in  $F$  is a subset of  $S$ ;
  - (ii) each  $x$  in  $S$  has a degree of membership greater than zero in at least one  $F_i$  in  $F$ ; and
  - (iii) Each  $F_i$  in  $F$  contains at least one member  $x_i$  such that the degree of membership of  $x_i$  in every other member  $F_j$  of  $F$  ( $F_j \neq F_i$ ) is zero.

From this definition of fuzzy partition, it follows that the set of fundamental response categories (**black**, **white**, **red**, **yellow**, **green**, and **blue**) is a partition of the universe of color percepts. Since every color category is a subset of the set of color percepts, condition (i) is satisfied. As evidence for the satisfaction of (ii), we note the fact that subjects given labels for black, white, red, yellow, green, and blue can successfully classify any color stimulus (Hering; Sternheim & Boynton). The evidence that condition (iii) is met is that, for each of these categories, color stimuli exist that are classifiable in that category and not in any of the other five.

Having shown that the fundamental response categories are a partition of the universe of color percepts, we can show that every basic color-term lexicon is a partition, or contains a partition, of the universe of color percepts. To see how this is so, we consider in turn Stage V systems, systems later than Stage V, and finally systems prior to Stage V.

Stage V systems consist of just the six primary color categories whose identity with the six fundamental response categories makes them ipso facto a partition of the set of color percepts.

Systems later than Stage V contain all the Stage V primary categories. If we assume, as argued in section §4.4, that the membership functions of the primary categories do not contract with the encoding of derived categories, these systems therefore contain a partition of the color space.

Systems earlier than Stage V are made up of primary and composite categories that together exhaust the list of fundamental response categories. It is easy to show that, if we start with a set  $F$  of subsets of a set  $S$ , where  $F$  is a fuzzy partition of  $S$ , and create a new set of subsets of  $S$  by taking fuzzy unions of the members of  $F$  in such a way that every member of  $F$  occurs in exactly one of the unions, then the resulting collection of subsets of  $S$  is also a fuzzy partition of  $S$ .<sup>21</sup> Since the early-

<sup>21</sup> We want to prove that if (a)  $S$  is a fuzzy set, (b)  $F$  is a fuzzy partition of  $S$ , and (c)  $F'$  is a set whose members are unions of the cells of  $F$ , then  $F'$  is a fuzzy partition of  $S$ . Obviously the

stage systems can be viewed as formed from the fundamental response category partition in just this manner, basic color category Stages I–IV also constitute fuzzy partitions of the domain of color percepts. Thus, at any stage, the basic color categories comprise or contain a set of categories that partitions the entire color space. From this it follows that, at every stage, basic color-term vocabularies provide terms for all colors.

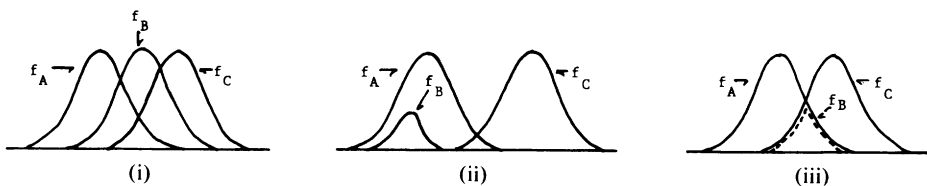
Given the fuzzy partition perspective, Stages I–V can be seen as a continuous refinement of partitions of the color domain—where ‘refinement of a partition’ refers to the creation of a new partition which classifies separately everything classified separately in the old partition, and in addition classifies separately at least two individuals classified the same in the old partition. The addition of derived basic color categories subsequent to Stage V does not in this sense refine the partition, since the derived categories do not satisfy condition (iii) of fuzzy partition. This follows from the argument made above that primary basic categories do not contract with the emergence of derived categories, and the resulting fact that no color sensation belongs to a derived category if it does not also belong to a primary category. In fact, every stimulus in a derived category belongs to both of the categories from which it is derived. That derived categories do not refine the partition of the color space, in the sense given above, supports the observation that derived categories are fundamentally less important than primary categories. Derived categories are thus gratuitous in two senses: (a) any color sensation can be referred to without using any of them, and (b) no color sensation can be referred to exclusively by any one of them.<sup>22</sup>

---

members of  $F'$  are subsets of  $S$ ; e.g., *grue* is a subset of the set of color percepts. So condition (i) is satisfied. With respect to condition (ii), for each  $x$  in  $S$ ,  $x$  belongs to whatever member of  $F'$  contains  $F_1$  as a subset; e.g., if some  $x$  in  $S$  belongs to *green*, then it belongs to *grue*. So condition (ii) is satisfied. With respect to condition (iii), each member of  $F'$  contains a subset  $F_1$  which is guaranteed to have a member with the desired property, by virtue of the fact that  $F$  satisfies condition (iii); e.g., *grue* contains *green* as a subset, and the unique hue point of *green* has the desired property. Thus condition (iii) is satisfied, and the proof is complete.

<sup>22</sup> As noted, the way we have defined fuzzy partition excludes derived categories from being possible cells of a partition, because they do not meet the mutual exclusion condition (iii). Fuzzy partition may be alternatively defined with a weakened mutual exclusion condition which is met by derived color categories as well as by the primaries.

Recall that, in the original statement of fuzzy mutual exclusion (iii), a collection of fuzzy sets meets this condition just if, for each fuzzy set, there is an individual that has positive membership in this set and zero degree of membership in each other set in the collection. This may be weakened by defining a collection of fuzzy sets as mutually exclusive just if, for each fuzzy set, there is an individual that has a higher degree of membership in this set than in any other set in the collection. In Figure (i) below, fuzzy sets  $A$ ,  $B$ ,  $C$  partition the domain of individuals, represented by the abscissa; but in Figures (ii) and (iii) there is no partition.



An alternate definition of fuzzy mutual exclusion and fuzzy partition.

In this context, it may be asked whether derived categories should be considered basic color categories at all. In defining 'basic color term', B & K had tacitly in mind a standard set-theory model when they wrote: 'A basic color term's significance is not included in that of any other color term' (p. 6). We can ask whether a basic but non-primary category like orange satisfies the translation of this criterion into fuzzy set theory. The answer is that it does. Orange, e.g., is not fuzzily contained in either red or yellow (or any other color category). A fuzzy set, say that defined by  $f_{\text{red}}(x)$ , contains another, say that defined by  $f_{\text{orange}}(x)$ , just in case  $f_{\text{red}}(x)$  exceeds  $f_{\text{orange}}(x)$  for any  $x$ . But, as derived orange has been defined (Equations 3-4), this holds neither for red and orange nor for yellow and orange (see Figs. 11B and 12A.) Thus the notion that non-primary, derived categories are nonetheless basic color categories survives the translation into the fuzzy set model.

#### CONCLUSIONS

6. We have reviewed evidence that the semantics of color display substantial linguistic universals; and that these semantic universals, which explain a considerable range of both synchronic and diachronic linguistic fact, are based on pan-human neurophysiological processes in the perception of color. We interpret these findings as placing strict limits on the applicability of the Sapir-Whorf hypothesis and related hypotheses of extreme linguistic/cultural relativity.

We have found further that the facts of color semantics are modeled felicitously in fuzzy set theory, and are not readily modeled in the traditional theory of discretely contrasting semantic features. This finding casts doubt on the general usefulness of the feature model, and suggests that more powerful formalisms, employing a range of structures much broader than the restricted Boolean algebra implicit in the discrete semantic-feature approach, are probably necessary to provide realistic accounts of the meanings of words (Fillmore 1975, Kay 1975b, Lakoff 1972).

#### REFERENCES

- BERLIN, B., and E. A. BERLIN. 1975. Aguaruna color categories. *American Ethnologist* 2.61-87.
- , and P. KAY. 1969. *Basic color terms: their universality and evolution*. Berkeley & Los Angeles: University of California Press.
- BORNSTEIN, M. H. 1973a. Color vision and color naming: a psychophysiological hypothesis of cultural difference. *Psychological Bulletin* 80.257-85.
- . 1973b. The psychophysiological component of cultural difference in color naming and illusion susceptibility. *Behavior Science Notes* 8.41-101.

---

Note that none of the situations depicted in these figures conforms to the definition of fuzzy mutual exclusion (and hence fuzzy partition) given in condition 5iii, since no set has a member with zero membership in each of the other sets. In particular, B has no such member. If this weakened version of fuzzy mutual exclusion and fuzzy partition is adopted, then addition of derived categories after Stage V does further refine the partition of the color domain. This formulation would appear to characterize the systems of those individuals for whom derived categories such as brown, pink, orange, and grey are on a perceptual/conceptual par with the primaries. Note that, in the alternate definition of fuzzy partition, categories such as crimson (cf. Fig. B) and chartreuse (cf. Fig. C) still do not participate in a refinement of the partition.



- . 1975. The influence of visual perception on culture. *American Anthropologist* 77.774-98.
- CAIRO, J. E., II. MS. Spectral color terms: a physiological theory. State University of New York, Binghamton.
- COLLIER, G. A. 1973. Review of Berlin & Kay 1969. *Lg.* 49.245-8.
- , et al. 1976. Further evidence for universal color categories. *Lg.* 52.884-90.
- DALY, T. MS. Color terms in the Slavic languages. University of California, Berkeley.
- DE VALOIS, R. L.; I. ABRAMOV; and G. H. JACOBS. 1966. Analysis of response patterns of LGN cells. *Journal of the Optical Society of America* 56.966-77.
- , and G. H. JACOBS. 1968. Primate color vision. *Science* 162.533-40.
- DIMMICK, F. L., and M. R. HUBBARD. 1939. The spectral location of psychologically unique yellow, green and blue. *American Journal of Psychology* 52.242-54.
- DOUGHERTY, J. W. D. 1974. Color categorization in West Futuna: variation and change. Paper presented to American Anthropological Association, Mexico, D.F.
- . 1975. A universalist analysis of variation and change in color semantics. Berkeley: University of California dissertation.
- FARIS, J. 1974. A theory of color conception. Paper presented to American Anthropological Association, Mexico, D.F.
- FILLMORE, C. J. 1975. An alternative to checklist theories of meaning. *Proceedings, 1st Annual Meeting, Berkeley Linguistics Society*, 123-31.
- FISHMAN, J. A. 1960. A systematization of the Whorfian hypothesis. *Behavioral Science* 5.323-39.
- GATSCHET, A. S. 1879. Adjectives of color in Indian languages. *American Naturalist* 13.475-85.
- GLEASON, H. A. 1961. An introduction to descriptive linguistics. New York: Holt, Rinehart & Winston.
- HAGE, P., and K. HAWKES. MS. Binumarien color terms. Salt Lake City: University of Utah.
- HARKNESS, SARA. 1973. Universal aspects of learning color codes: a study in two cultures. *Ethos* 1.175-200.
- HAYS, D. G., et al. 1972. Color term salience. *American Anthropologist* 74.1107-21.
- HEIDER, E. R. [= E. H. Rosch.] 1972a. Probabilities, sampling, and ethnographic method: The case of Dani colour names. *Man* (n.s.) 7.448-66.
- . 1972b. Universals in color naming and memory. *Journal of Experimental Psychology* 93.10-20.
- HEINRICH, A. C. 1972. A non-European system of color classification. *Anthropological Linguistics* 14.220-27.
- HERING, EWALD. 1920. *Grundzüge der Lehre vom Lichtsinn*. Berlin: Springer. [English version: *Outlines of a theory of the light sense*. Translated by L. M. Hurvich & D. Jameson. Cambridge, MA: Harvard University Press, 1964.]
- JACOBS, G. H. 1964. Single cells in squirrel monkey lateral geniculate nucleus with broad sensitivity. *Vision Research* 4.221-33.
- JAMESON, D., and L. M. HURVICH. 1968. Opponent-response functions related to measured cone photo-pigments. *Journal of the Optical Society of America* 58.429-30.
- KATZ, J. J. 1964. Analyticity and contradiction in natural language. The structure of languages, ed. by J. A. Fodor & J. J. Katz, 519-43. Englewood Cliffs, N.J.: Prentice-Hall.
- . 1966. *The philosophy of language*. New York: Harper & Row.
- KAY, P. 1975a. Synchronic variability and diachronic change in basic color terms. *Language in Society* 4.257-70.
- . 1975b. Tahitian words for race and class. (Language Behavior Research Laboratory, Working paper 40.) Berkeley: University of California.
- KUSCHEL, R., and T. MONBERG. 1974. 'We don't talk much about colour here': a study of colour semantics on Bellona Island. *Man* (n.s.) 9.213-42.

- LAKOFF, G. 1972. Hedges: a study in meaning criteria and the logic of fuzzy concepts. Papers from the 8th Regional Meeting, Chicago Linguistics Society, 183-228.
- MCDANIEL, C. K. 1972. Hue perception and hue naming. A.B. honors thesis, Harvard College.
- . 1974. Basic color terms: their neurophysiological bases. Paper presented to the American Anthropological Association, Mexico, D.F.
- . ms. Universals in color terms semantics and their neuropsychological sources.
- MACLAURY, R. ms. Reconstructions of the evolution of some basic color term lexicons. University of California, Berkeley.
- MILLER, G. A., and P. JOHNSON-LAIRD. 1976. Language and perception. Cambridge, MA: Harvard University Press.
- NEWHALL, S. M.; D. NICKERSON; and D. B. JUDD. 1943. Final report of the Optical Society of America Subcommittee on the Spacing of the Munsell Colors. *Journal of the Optical Society of America* 33.385-418.
- RATLIFF, F. 1976. On the psychophysiological bases of universal color terms. *Proceedings of the American Philosophical Society* 120.311-30.
- ROSCH, E. H. [= E. R. Heider.] 1973. On the internal structures of perceptual and semantic categories. Cognitive development and the acquisition of language, ed. by T. E. Moore, 111-44. New York: Academic Press.
- . ms. Human categorization. To appear in *Advances in cross-cultural psychology*, I, ed. by N. Warren. London: Academic Press.
- SAHLINS, M. 1976. Colors and cultures. *Semiotica* 16.1-22.
- STEPHENSON, P. H. 1973a. Color: its apprehension and symbolic use in language and culture. M.A. thesis, University of Calgary.
- . 1973b. The evolution of color vision in the primates. *Journal of Human Evolution* 3.379-86.
- . 1976. A strophe on structural analysis and neurology: color salience and the organization of the tricolor traffic signal. Paper presented to the Canadian Ethnological Society Meetings, Victoria, B.C.
- STERNHEIM, C. D., and R. M. BOYNTON. 1966. Uniqueness of perceived hues investigated with a continuous judgemental technique. *Journal of Experimental Psychology* 72.770-76.
- WITKOWSKI, S. R., and C. H. BROWN. 1977. An explanation of color nomenclature universals. *American Anthropologist* 79.50-57.
- WOOTEN, B. R. 1970. The effects of simultaneous and successive chromatic constraint on spectral hue. Doctoral dissertation, Brown University.
- ZADEH, L. A. 1965. Fuzzy sets. *Information and Control* 8.338-53.
- . 1971. Quantitative fuzzy semantics. *Information Sciences* 3.159-76.
- ZOLLINGER, H. 1972. Human color vision as an interdisciplinary research problem. *Platte* 40.1-7.
- . 1973. Zusammenhänge zwischen Farbennennung und Biologie des Farbensehens beim Menschen. *Vierteljahrschrift der Naturforschenden Gesellschaft* 118.227-55.
- . 1976. A linguistic approach to the cognition of color vision in man. *Folia Linguistica* 9.265-93.

[Received 31 October 1975.]