The World Color Survey Database: History and Use

Richard S. Cook
University of California, Berkeley

Paul Kay[1]
University of California, Berkeley

Terry Regier
University of Chicago

[1] Corresponding author: kay@cogsci.berkeley.edu. International Computer Science Institute, 1947 Center Street, Berkeley, CA 94704, USA.

The World Color Survey [WCS] is a research project that was undertaken to validate, invalidate or – most likely – modify the main findings of Berlin and Kay (1969) [B&K]: (1) that there exist universal cross-linguistic constraints on color naming, and (2) that basic color terminology systems tend to develop in a partially fixed order. To this end, the WCS collected color naming data from speakers of 110 unwritten languages. The data have recently been compiled into a unified data archive, available online at http://www.icsi.berkeley.edu/wcs/data.html. In this chapter, we review the history of the WCS, including the creation of the online data archive, and describe our recent use of the archive to test the universality of color naming across languages.

The WCS: History and Methodology

The WCS was begun in 1976 to check and expand the findings of B&K in a full-scale field study. B&K had investigated the color terminology systems of twenty languages in the following way. The stimulus array used by Lenneberg and Roberts (1956), consisting of 320 Munsell chips of 40 equally spaced hues and eight levels of lightness (Value) at maximum saturation (Chroma) for each (Hue, Value) pair, was supplemented by nine Munsell achromatic chips (black through gray to white) – the resulting stimulus array is shown in Figure 1a[2]. First, without the stimulus array present, the major color terms of the collaborator's native language were elicited by questioning that was designed to find the smallest number of simple words with which the speaker could name any color (*basic color terms*)[3]. Once this set of basic color terms was established, the collaborator was asked to perform two tasks. In the *naming* task the stimulus array was placed before the speaker and for each color term *t*, a piece of clear acetate was placed over the stimulus board and the collaborator was asked to indicate, with a grease pencil on the acetate sheet, all the chips that he or she could call *t*. In the *focus* task the stimulus array was shown as before and the collaborator was asked to indicate the best example(s) of *t* for each basic color term *t*. The boundaries of categories showed great variability, perhaps because of the vagueness of the instruction of the naming task: probably some subjects took the instruction to call for all the chips that were more *t* than anything else, while others appear to have taken it to call for all chips in which any trace of *t* was visible.[4] The focal choices of the B&K subjects were much more clustered and led to the conclusion that

> ... [1] the referents for the basic color terms of all languages appear to be drawn from a set of eleven universal perceptual categories, and [2] these categories become encoded in the history of a given language in a partially fixed order (Berlin and Kay 1969: 4f).[5]

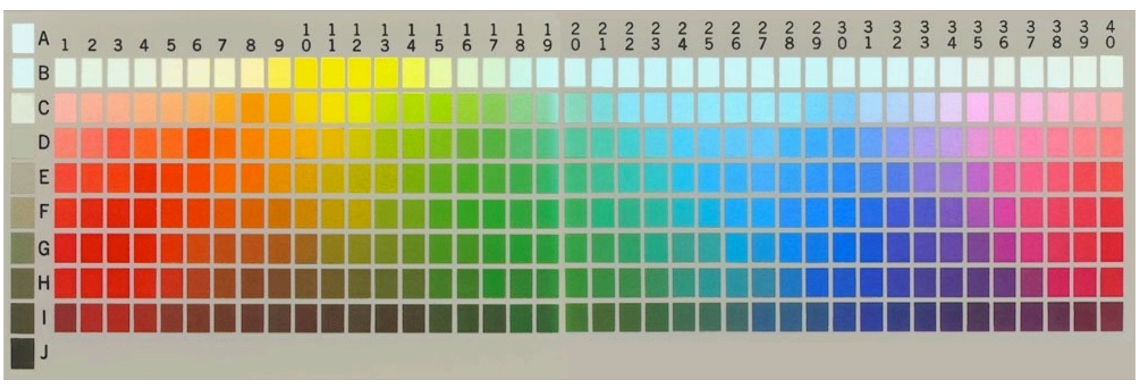


Figure 1a. The WCS stimulus array.

In retrospect, the B&K study – only twenty languages directly assessed with calibrated color stimuli and all the work done in the San Francisco Bay Area – can be viewed as a pilot project for the WCS.[6] The B&K results were immediately

---

[2] Actually, Figure 1a shows the very slightly modified stimulus palette used in the WCS. The B&K stimulus array lacked achromatic chip A0.

[3] For a review of the B&K basicness criteria, as well as of the notion of basic color terms in B&K and other literature, see Maffi (1990).

[4] MacLaury later demonstrated that speakers can often be induced to increase the number of chips they will indicate as belonging to a given term simply by asking them if there are "any more," and frequently speakers increase the size of a named category several times in response to this "mapping" task (MacLaury 1997: 77-84 *et passim*).

[5] B&K extended their findings on the twenty languages assessed experimentally to another seventy-eight reports of color terminology systems they found in the literature.

challenged, mainly by anthropologists, on the grounds that the sample of experimental languages was too small, too few collaborators per language were questioned (sometimes only one), all native collaborators also spoke English, the data were collected in the San Francisco Bay area rather than in the homelands of the target languages, certain regions of the world and language families were under represented or over represented in the sample of twenty, and that the sample of twenty had too few unwritten languages of low technology cultures (Hickerson 1971, Durbin 1972, Collier 1973, Conklin 1973). The results were nevertheless supported by various ethnographic and experimental studies conducted after 1969[7] and were largely accepted by psychologists and vision researchers (e.g., Brown 1976, Miller and Johnson-Laird 1976, Ratliff 1976. See also Kaiser and Boynton 1996: 498ff, Boynton 1997: 133 ff).

In the late 1970s, through the cooperation of SIL International (then the Summer Institute of Linguistics), which maintains a network of linguist-missionaries around the world, data on the basic color term systems of speakers of 110 unwritten languages representing forty-five different families and several major linguistic stocks were gathered *in situ*. Field workers were provided with a kit containing the stimulus materials (330 individual chips in glass 35 mm slide sleeves for the naming task and the full stimulus board for the focus task) as well as coding sheets on which to record collaborators' responses. The included instructions requested that fieldworkers collect data from at least twenty-five speakers, both males and females, and urged them to seek out monolingual speakers insofar as possible. The modal number of speakers actually assessed per language was twenty-five and the mean number was twenty-four. (A facsimile of the WCS instructions to field workers and of the original coding sheets is included as an Appendix.) The aim was to obtain names, category extent and best examples of basic color terms in each language – basic color terms being described in the instructions as *"the smallest set of simple words with which the speaker can name any color."*

The WCS methodology coincided with that of the B&K study in the use of a standardized set of Munsell color chips, consisting of 320 chromatic chips representing forty equally spaced hues at eight levels of lightness (Munsell Value), each at maximum available saturation (Munsell Chroma). One white chip was added in the WCS study that was whiter than any chip available at the time of the B&K study, making for a total of ten achromatic chips and an overall total of 330 chips, as shown in Figure 1a. The Munsell notations of the chips employed and the simplified notation used for precisely this pallete by the WCS project are shown in Figure 1b.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	18	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	
A	0																																									9.5
B	0	2	2	2	2	2	2	2	2	4	6	6	6	6	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	9.0
C	0	6	6	6	6	6	6	8	14	16	14	12	12	12	10	10	8	8	6	6	6	6	4	4	4	4	4	4	4	6	6	4	4	4	4	6	6	6	6	6	6	8.0
D	0	8	8	10	10	10	14	14	14	12	12	12	12	12	12	10	10	10	8	8	8	8	8	6	6	6	6	6	8	8	8	6	6	6	6	8	8	10	10	8	8	7.0
E	0	12	12	12	14	16	12	12	12	10	10	10	10	10	10	12	12	10	10	10	10	8	8	8	8	8	8	8	10	10	10	8	8	8	8	10	10	10	10	12	12	6.0
F	0	14	14	14	16	14	12	10	10	8	8	8	8	8	8	10	12	12	10	10	10	10	8	8	8	8	8	8	10	12	12	10	10	10	10	10	12	12	12	14	14	5.0
G	0	14	14	14	14	10	8	8	6	6	6	6	6	6	6	8	8	10	10	10	10	8	8	8	8	6	6	8	8	10	10	12	10	10	10	10	10	10	10	10	10	4.0
H	0	10	10	12	10	8	6	6	6	4	4	4	4	4	4	6	6	8	8	10	8	6	6	6	6	6	6	6	8	10	10	12	10	10	10	10	10	10	10	10	10	3.0
I	0	8	8	8	6	4	4	4	2	2	2	2	2	2	2	4	4	4	6	6	6	4	4	4	4	4	4	6	6	6	8	10	8	8	8	6	6	8	8	8	8	2.0
J	0																																									1.5
		2,5	5	7,5	10		5		10		5		10		5		10		5		10		5		10		5		10		5		10		5		10		5		10	
			R				YR				Y				GY				G				BG				B				BP				P				RP			

Figure 1b. Munsell and WCS coordinates for stimulus palette. The leftmost column and the top row give the WCS coordinates for lightness and hue respectively. The rightmost column and the bottom two rows give the Munsell coordinates for Value and Hue, respectively. Entries in the body of the table show the corresponding Munsell Chroma numbers. (With regard to the A and J rows, there are no Munsell hues at the extremes of Value (lightness): 9.5 (white) and 1.5 (black).)

The WCS differed from B&K in the technique for eliciting naming responses. In the WCS procedure, no preliminary interview was administered to establish a set of basic color terms, and in the naming task the 330 individual color stimuli were shown to each cooperating speaker, one by one, according to a fixed random order, and a name elicited for each (in contrast with the B&K procedure of presenting the entire stimulus array at once to elicit naming responses). Field workers were instructed to urge observers to respond with short names (although, depending on the morphology of the language, particular field

[6]Initial support for the WCS was in the form of NSF grant BNS 76-14153. Subsequent NSF support was furnished by grants BNS 78-18303, BNS 80-06808, SBR 94-19702, BCS 01-30420, and BCS 04-18283. NSF support of the WCS project is gratefully acknowledged, as is additional support by the University of California, Berkeley, the Summer Institute of Linguistics (now SIL International) and the International Computer Science Institute. We would also like to express our most sincere gratitude to the many field linguists of the SIL who unselfishly devoted long hours to what for many must have often been an unwelcome task.

[7] For example, Berlin and Berlin (1975), Dougherty (1975, 1977), Hage and Hawkes (1975), Harkness (1973), Heider (1972a, 1972b), Heider and Olivier (1972), Heinrich (1972), Kuschel and Monberg (1974), MacLaury (1986, 1987, 1997), Maffi (1990b), Monberg (1971), Senft (1987), Snow (1971), Turton (1978, 1980).

circumstances and local culture, there was considerable variation in the degree to which the field investigators were able to satisfy these desiderata). Identification of basic color terms, therefore, was done by the field worker as a result of the naming task itself, rather than through prior elicitation. The best example (focus) responses were elicited in the same way in both studies: once a set of basic color terms was isolated, the native observer was presented with the full palette (in WCS, a physically improved version of the original Munsell chip board, devised by Collier *et al.* 1976) and asked to indicate the chip (or chips) that represented the best example of each term, one by one.

Data processing and analysis

Once data gathering was completed (circa 1980), data processing, quality control, and analysis were undertaken at University of California, Berkeley, and at SIL in Dallas.

Computer programs were developed for both data entry and data analysis. The original processing yielded, for each language, a preliminary data summary that included the following information:

- Language name and location.
- Name, age, sex and other vital statistics of each speaker interviewed.
- List of terms used, each with a tentative gloss and a typographical symbol representing it in the naming and focus arrays. (See Figure 2a for an example. All the examples in Figure 2 are for the Niger-Congo language Wobé of the Côte D'Ivoire. The information shown in Figure 2 is not that of the initial data entry and preliminary processing but of the final results of checking, following corrections to the original data entry and preliminary analyses, as described below)
- Individual naming arrays, structured by the form of the full stimulus array shown in Figure 1 and presenting, for each speaker, the full picture of his or her use of color terms from the naming task. (See Figure 2b for examples.)
- Individual focus arrays, presenting, for each speaker, the full picture of his/her focal (best example) choices from the focus task. (See Figure 2c for examples.)
- Aggregate naming arrays, also in the form of the stimulus array, presenting the aggregated results of the naming task across all speakers, at various levels of inter-speaker agreement. For example, for a language with twenty-five native observers the 40% naming aggregate shows for each stimulus chip *c* (1) the symbol for the most popular name given *c*, if at least ten speakers gave *c* that name or (2) a blank if no single name was given to *c* by ten or more speakers. (See Figure 2d for Examples.)

Symbol	Term	Gloss
o	*kpe'*	black/green/blue
+	*"pluu-*	white
#	*-sain'*	red/yellow

(Diacritics stand for tones: " = very high; ' = high; - = low.)

Figure 2a. Wobé basic color terms, with glosses and key symbols

```
          ***  Speaker  15    ***
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B +++++++++#++#++++++++++++++++++++++++++++ B
C +#############++oo+++++++++++++++++++#### C
D +############ooooo+ooo+o+o++oooo+++###### D
E +###########*#ooooooooooooooooooooo###### E
F o###########ooooooooooooooooooooooo###### F
G o#########ooooooooooooooooooooooooo###### G
H o######o#oooooooooooooooooooooooooo###### H
I o####o#oooooooooooooooooooooooooooooo#### I
J ooooooooooooooooooooooooooooooooooooooooo J
```

```
          ***  Speaker  17    ***
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B +++#++#++###+##++++++++++++++++++++++++++ B
C +#################oooo+o+o++o+o++++###### C
D +###############ooooooooooo+ooooo#+###### D
E o##############oooooooooooooooooo######## E
F o#############oooooooooooooooooooo####### F
G o##########oo##o#ooooooooooooooooo####### G
H o########oooooooooooooooooooooooooo###### H
I o######oooooooooooooooooooooooo+o#o###### I
J ooooooooooooooooooooooooooooooooooooooooo J
```

```
          ***  Speaker  22    ***
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B +++++++++++++++++++++++++++++++++++++++++ B
C +##############oooooooo+++++++++++++##### C
D +##############oooooooooo#ooooo+o++###### D
E +#############ooooooooooooooooooo######## E
F o############oooooooooooooooooooooo###### F
G o############oooooooooooooooooooo######## G
H o###o######ooooooooooooooooooooooo####### H
I o######ooooooooooooooooooooooooooo####### I
J ooooooooooooooooooooooooooooooooooooooooo J
```

```
          ***  Speaker  23    ***
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B ++++++++++++++++++++++++++++++++++++#++++ B
C +#############ooooo+++++o++++++++++++#### C
D +####+#########ooooooooooooo+ooo++++##### D
E +#############oooooooooooooooooooooo##### E
F o#############oooooooo#ooooooooooo+###### F
G o######## #oooooooooooooooooooooooo###### G
H o#######o#*#oooooooooooooooooooooooo##### H
I o####ooooooooooooooooooooooooooooo#oo#### I
J ooooooooooooooooooooooooooooooooooooooooo J
```

Figure 2b. Naming responses for four Wobé speakers (See previous figure for terms denoted by typographical symbols.)

```
           ***  Speaker  15    ***
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B                                           B
C                                           C
D                                           D
E                                           E
F                                           F
G #                                       # G
H                                           H
I                                           I
J ooooooooooooooooooooooooooooooooooooooooo J
```

```
           ***  Speaker  17    ***
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B                                           B
C                                           C
D                                           D
E                                           E
F                                           F
G #                                       # G
H                                           H
I                                           I
J ooooooooooooooooooooooooooooooooooooooooo J
```

```
           ***  Speaker  22    ***
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B                                           B
C                                           C
D                                           D
E                                           E
F                                           F
G #                                       # G
H                                           H
I                                           I
J ooooooooooooooooooooooooooooooooooooooooo J
```

```
           ***  Speaker  23    ***
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B                                           B
C                                           C
D                                           D
E                                           E
F                                         # F
G #                                         G
H                                           H
I                                           I
J ooooooooooooooooooooooooooooooooooooooooo J
```

Figure 2c. Focus (best example) responses for four Wobé speakers

```
   40% Agreement Level, 10 of 25 speakers
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B +++++++++++++++++++++++++++++++++++++++++ B
C +##############oooo++++++++++++++++++#### C
D +##############oooooooooooo+ooo++++###### D
E +#############ooooooooooooooooooo######## E
F o#############ooooooooooooooooooooo###### F
G o##########oooooooooooooooooooooo######## G
H o########oooooooooooooooooooooooooo###### H
I o####ooooooooooooooooooooooooooooo####### I
J ooooooooooooooooooooooooooooooooooooooooo J
      Coverage is 100% (330 of 330 chips)
```

```
   100% Agreement Level, 25 of 25 speakers
            1         2         3         4
  01234567890123456789012345678901234567890
A +++++++++++++++++++++++++++++++++++++++++ A
B ++  ++ ++      +++++++++++ +++++ +++ ++   B
C + #     ###                    + +      # C
D +# ## #####                         ##### D
E  ########                           ##### E
F  ###### #                          ###### F
G o#####                    o        ###### G
H  ###          o                      #### H
I o         o   o      o     o              I
J                                           J
      Coverage is 31% (102 of 330 chips)
```

Figure 2d. Aggregate naming arrays for 25 Wobé speakers. (Note that at the 40% level of agreement all 330 chips were named. That is, at least ten speakers gave the modal response for each of the 330 chips. Wobé was a high consensus language.)

Subsequently, an additional kind of array was produced, called a *term map*. A term map for a given term furnishes a visual picture of the relative frequency of that term's usage over the stimulus space in the form of a kind of contour map. A term map is thus a display of the denotation of a color term. Conceptually, one can imagine a 3D histogram for a term t in which the stimulus surface constitutes the floor plane and the height of the column over each stimulus chip c represents the proportion of speakers using t to name any chip that used t to name c. We represent 2-dimensionally a contour map of such a 3-D histogram, viewed from above: this is what we call a term map. Specifically, a term map for a term t is a display of the stimulus surface where the symbol appearing on chip c is

- #, if 81% or more speakers who used *t* named *c* with *t*,
- +, if 61-80% of the speakers who used *t* named *c* with *t*,
- –, if 41-60% of the speakers who used *t* named *c* with *t*,
- ¨, if 21-40% of the speakers who used *t* named *c* with *t*,
- nothing (blank), if 20% or fewer of the speakers who used *t* named *c* with *t*,
- @, if the % of *t*-users who used *t* to name *c* equaled or exceeded the percentage of *t*-use by *t*-users for any other chip. The numerical value of @, the consensus level for *t*, is given at the bottom of each term map, as well as the number of collaborating speakers of the language and the number using the term mapped.

The density of the symbols as visual objects increases as the proportion of respondents they represent increases; thus a term map gives a somewhat iconic representation of a term as a gradient category, where proportion of speakers using *t* to name *c* is taken as a proxy for the degree of membership of color *c* in the gradient category named by term *t*.[8]

Figure 3 shows the term maps for Wobé.

```
o: kpe' 'black/green/blue'

            1         2         3         4
  01234567890123456789012345678901234567890
A                                           A
B                                           B
C                ---+......                 C
D               .++####++#-+-.-++           D
E              .+###########+####-+.        E
F -           ..-##################--       F
G @         .-+++###########@######..       G
H #      ..-+###@##################+-.      H
I @    -+###@###@######@#####@#####+---     I
J ######################################### J

  25 of 25 speakers searched; 25 used term
 Consensus level is 100% (25 of 25 speakers)
```

```
+: "pluu- 'white'

            1         2         3         4
  01234567890123456789012345678901234567890
A @@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@ A
B @@##@@#@@++-++#@@@@@@@@@@@#@@@@@#@@@#@@## B
C @   .       . ...-.+-+++++####+@#@#++     C
D @                    .  ...--..+##+       D
E +                              ..         E
F .                                         F
G                                           G
H                                           H
I                                           I
J                                           J

  25 of 25 speakers searched; 25 used term
 Consensus level is 100% (25 of 25 speakers)
```

```
#: -sain' 'red/yellow'

            1         2         3         4
  01234567890123456789012345678901234567890
A                                           A
B           .-..                            B
C  #@#+###@@@#+#.                    ..###@ C
D  @#@@#@@@@@@##+-                   +@@@@@ D
E  @@@@@@@@@###++.                 -+#@@@@@ E
F  @@@@@@#@#++++.                  -.@@@@@@ F
G  @@@@@##++-..                    -+@@@@@@ G
H  @@@###+-.                       .-+#@@@@ H
I  ####-.                          .---+### I
J                                           J

  25 of 25 speakers searched; 25 used term
 Consensus level is 100% (25 of 25 speakers)
```

(# = 81-100% agreement, + = 61-80 % agreement, - = 41-60% agreement, . = 21-40% agreement)

Figure 3, Term maps for the three terms of Wobé. (The large number of chips receiving perfect consensus for the white/light and red/yellow/warm terms, indicated by @, is unusual.)

Cleaning the data

Before the initial data entry had been checked, a set of microfiche summaries, containing preliminary versions of the data arrays described above was accidentally made available to the public in 1991. This release of unchecked data was unfortunate because we subsequently discovered that these summaries contained many errors, especially in the assigning of similar spellings to same or different terms. Coders had frequently made snap judgments about probable morphological variations in unfamiliar languages. Such errors of primary interpretation, along with simple input errors – using the same abbreviation for two different terms, using two different abbreviations for a single term, and the like, introduced significant inaccuracy into the data of 1991 microfiche summaries. The current archived data are based on a number of data-cleaning procedures subsequently adopted: (1) checking the electronic record against the original paper coding booklets, (2) carefully surveying fieldworkers' notes regarding spelling variations and morphological structure, (3) rerunning various summary programs to see if, for example, two distinct 'terms' with similar spellings had virtually identical term maps (indicating mere spelling variation), (5) corresponding with the original investigators in some cases and with other specialists in the same or related languages in other cases regarding morphological analysis of the recorded color terms, and (6) in general checking all

[8] The conventions for representing the various displays in Figures 2 and 3 were developed in an age of typewriter technology.

possible information from whatever source to make as certain as possible that what our data listed as the roster of color terms of a language was an accurate rendition of the color terms of that language as recorded by the field investigators.

Original format of the data

The early work of converting the handwritten coding sheets (prepared in the field) to electronic format was split between a team of researchers at SIL in Dallas, Texas, and a team at UC Berkeley. The two halves were eventually joined together when the SIL data came to Berkeley in the mid-eighties. At that point, the WCS data were not stored in a single database but rather in 110 separate directories, one for each language. Within a language directory, the data were stored in four main files (with some subsidiary files): (1) an Informants file, containing the name and vital statistics of each native collaborator for that language, along with an identification number and some ancillary information regarding other languages spoken, etc. (2) a Dictionary file, containing the color terms of the language and one or more abbreviations with which they could be referred to in other files, (3) a Naming data file, containing the naming data for each collaborating speaker, and (4) a Focus file, containing the best example(s) data for each speaker. The WCS data remained in these 110 separate and unlinked directories until December of 2000.

Creation of the WCS Online Data Archive

The primary copies of the World Color Survey electronic computer files, described above, were housed at UC Berkeley on an aging hard disk connected to an old computer, with accompanying fragile removable media back-up disks. In December of 2000 those files were all compressed into a single archive, copied to a campus server and then burnt to CD-ROM. In early 2002, work began to extract the data from that archive, to create an operating system-independent, public online archive. We describe the archive creation process here, so that users of the archive may have a clear picture of the nature of the data, and their origin.

Initial study of the electronic source files revealed that a number of components in different formats would have to be processed separately for integration into a coherent whole. A major reason for this was that the SIL and Berkeley teams had applied different conventions to digitization of their respective data.

The results of the naming task (presented in the current online archive in the file “term.txt”) constituted by far the bulk of the data, totaling nearly one million lines of text. These naming data, in 110 separate files, had received the greatest attention of all the existing files, and by the year 2000 had attained a fairly stable state. Other components of the data included focus data (from the best-example task), speaker data (personal information on the native observers), language data (geographical information on the languages), dictionary data, and analyses of the color terms present for each language. Compared to the naming data, each of these latter components had received considerably less attention over the years. Unlike the naming data, the separate Berkeley and Dallas focus data had never been consolidated into a coherent whole. Instead, portions of each had been partially processed, and in 2002 we began to reassemble the pieces and complete the processing.

The source files fell into two major groups, labeled “new” and “old”. The “new” files represented the effort to combine the Berkeley and Dallas data. The old files themselves were still split into “Berkeley” and “Dallas” groups. There were omissions in the “new” files of focus data only available in the “old” files, and there was a certain amount of overlap and variability in formatting among all three. The task then became one of identifying omissions in the “new” focus data, filling in these gaps, and verifying the new focus data where it was present. Gaps in the new data were filled in from the primary electronic source of the missing data (Berkeley or Dallas). In the event that no electronic data was available, it was necessary revert to the coding sheets and input the data afresh, interpreting the fieldworkers’ conventions as best these could be determined. Fortunately, recourse to the original coding sheets was necessary only for a relatively small number of languages.

The data from the coding sheets had originally been entered into the computer systems via a process that involved creating an electronic “dictionary”, that is, an inventory of the terms attested on all coding sheets for all speakers of the language. A unique abbreviation (WCS code) was then assigned to each of the terms, often bearing an obvious relation to those used by the original fieldworker. Different fieldworkers had used different transcriptional and notational conventions. Thus, the electronic Dictionary data contained both an ASCII (or other encoded) interpretation of the fieldworker’s transcription of the native color term, and a one or two-letter abbreviation to be used in the input of both naming and focus data. These three repositories of term abbreviations had to ultimately agree among themselves, and also bear transparent relations to the rendering of the terms on the coding sheets. Where there was disagreement, correction had to be applied, based on assessment of all of the available documentation.

The original fieldworkers themselves had employed various collection and organizational techniques, all of which contributed another level of variability to the data. In examination of the coding sheets, it became apparent that some fieldworkers were fastidious workers and had clearly put great effort into selecting their native collaborators, collecting the responses, and copying out the final coding sheets in pen or with a typewriter, using IPA for their transcriptions. At the other extreme, data on the coding sheets was sometimes barely legible and sometimes internally inconsistent, as when the same abbreviation was used for evidently different terms or two different abbreviations were used for the same term. Once the relations among the various data formats became clear, we compiled the data together into a single data archive composed of four files in tab delimited plain text format: one containing the naming data ("term.txt"), one containing best example data ("foci.txt"), one describing the languages ("lang.txt"), and one describing the individual speakers of the languages ("spkr.txt"). The various "dictionary" files also have been combined into a single file compatible with the other four, and integrated into the relational database system. This dictionary data appears in the online archives in UTF-8 format under the name "dict.txt". Online documentation concerning the formats of these files is included with the archive. The work to prepare the data consumed all of 2002, and it was not until January 2003 that the first portion of the online data was released to the public.

Uses of the WCS archive

The WCS data archive has been used in investigating two broad questions, one concerning *universals,* and other concerning *variation*, of color naming.

Universals of color naming

Since B&K found evidence for universals in color naming across languages, the existence of such constraints has generally been accepted in the scientific community. However, there have always been dissenters from this consensus (e.g. Hickerson, 1971; Durbin, 1972), and this dissenting view has recently gained some prominence (e.g. Lucy, 1992; 1996; 1997; Saunders & van Brakel, 1997; Davidoff, Davies, & Roberson, 1999; Roberson, Davies, & Davidoff, 2000). Criticisms of the universalist position have come in two major varieties. The first points out that B&K's findings were never objectively tested, as they relied on visual inspection of color naming data. Lucy (1997) challenges such a methodology as hopelessly subjective:

> "[Work in the B&K tradition] not only seeks universals, but sets up a procedure which guarantees both their discovery and their form. … when a category is identified ... it is really the investigator who decides which 'color' it will count as … What appears to be objective - in this case, a statement of statistical odds - is [not]." (p. 334)

On this view, B&K's subjective methodology allowed them to impose their own universalist assumptions on their data – so the universals are actually in the minds of the investigators, not in the languages of the world. The second strand of criticism points out that B&K's data were drawn primarily from written languages, and thus may not be representative. This point is coupled with analyses of particular unwritten languages, which are claimed to counterexemplify universal constraints (e.g. Berinmo: Davidoff et al., 1999; Roberson et al., 2000; Hanunóo and Zuni: Lucy, 1997). Kay (1999) has responded to this with counter-analyses of these languages, arguing that each fits neatly into the universal pattern. Disputes of this sort over conflicting interpretations of individual color naming systems could continue indefinitely without resolving the main issue of whether universal, cross-language constraints on color naming systems actually exist. We wished to resolve this issue in a manner that would respond to both varieties of criticism.

To that end, Kay and Regier (2003) used the WCS database and the B&K data to test objectively the hypothesis that color terms across languages cluster together more tightly in color space than would be expected by chance. This was done as follows. First, for each term in each WCS language, the centroid (i.e., center of mass) of each speaker's naming distribution was calculated, after translation of Munsell coordinates into CIE L*a*b* coordinates (Wyszecki & Stiles, 1967).[9] We refer to the resulting point as the "speaker centroid" for that speaker and that term. For each term, we then calculated the "term centroid:" the centroid of the speaker centroids for that term. This produced a point representation of the term in CIE L*a*b* space. Each centroid was coerced to the nearest point in the stimulus array, so that our point representation of the term resided within the set of points out of which it was constructed. The speaker centroids were plotted over the stimulus space, yielding the picture shown in Figure 4a. Intuitively, the speaker centroids are not distributed randomly or evenly over the stimulus space. Figure 4a shows sharp peaks and broad valleys in the distribution of speaker centroids. We showed this clustering of speaker centroids to be statistically significant by the results of a Monte Carlo simulation on the *term* centroids, depicted in Figure 4b – this time demonstrating universality across color terms without regard to their frequency of use, since each term is now represented by one centroid, whatever its frequency. In that simulation, a measure of dispersion (the opposite of clustering) of term centroids was defined as the sum, across languages, of the distances between each term (centroid) in that

[9] CIE L*a*b* is a 3-dimensional color space; its creators made a systematic effort to assure that Euclidean distance corresponds to perceptual dissimilarity.

language and the closest term (centroid) in another language. This measure was calculated in the WCS data and in 1,000 hypothetical randomized datasets, each created by rotating the actual WCS naming centroid distribution a random degree of hue angle in CIE L*a*b* space (to maintain the shape of the distribution while randomizing its location in perceptual space). In Figure 4b it can be seen that the dispersion measure for the actual WCS dataset falls well below the lower bound of the distribution of 1,000 hypothetical WCS datasets, indicating a probability less than .001 that the degree of clustering in the real WCS dataset is the result of chance. A similar Monte Carlo test also revealed that color terms in the unwritten languages of the WCS tend to cluster near the color terms of written languages (Berlin & Kay, 1969). (For further explanation, see Kay and Regier 2003.)

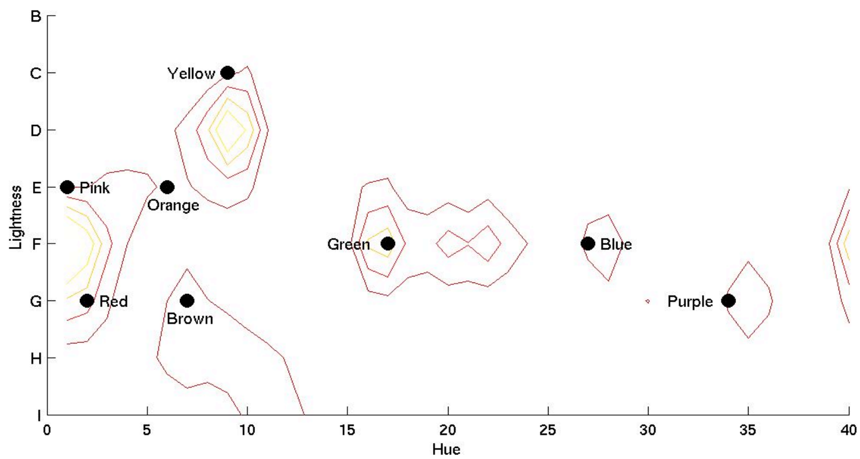


Figure 4a: Contour plot of WCS speakers' naming centroids, compared with English naming centroids (black dots; Source for English naming centroids: Sturges and Whitfield 1995). The outermost contour represents a height of 100 centroids, and each subsequent contour represents an increment in height of 100 centroids. (Source: Kay and Regier 2003)

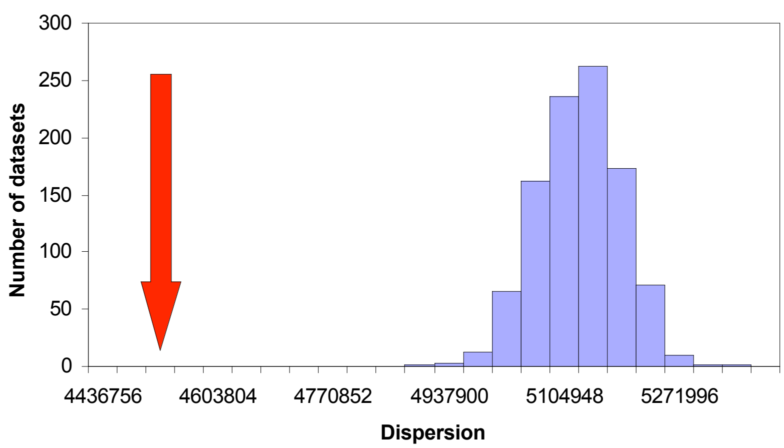


Figure 4b: Monte Carlo test for clustering within the WCS data. The distribution of dispersion values shown in blue was obtained from 1000 randomized datasets. The red arrow indicates the dispersion value obtained from the WCS data. (Source: Kay and Regier 2003)

The above results concern the naming data from the WCS. Statistical tests of the degree of clustering of WCS best example (focus) choices remain to be performed, but a preliminary plot of the focus data suggests strongly that such tests will turn out to be statistically significant. (Black dots represent the naming centroids for the English terms indicated.)

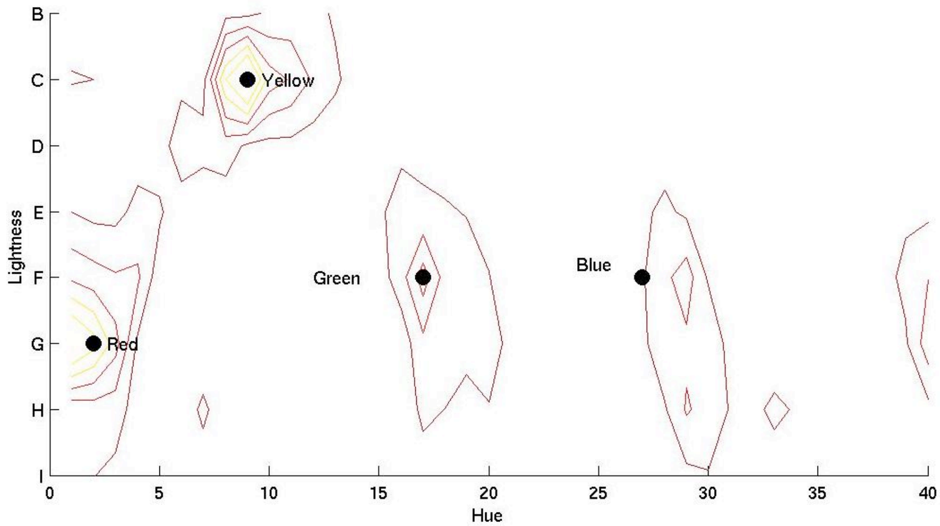


Figure 5. Contour plot of WCS chromatic focus peaks compared with English naming centroids. (Source for English naming centroids: Sturges and Whitfield 1995.)

In the WCS focus distribution, the chips receiving the highest numbers of focus choices were J0 (black) and A0 (white), not shown in Figure 5. In Figure 5, restricted to the WCS chromatic chips, two of the four major focus peaks fall, one each, on the English yellow and English green naming centroids. A third WCS focus peak falls on a chip adjacent to the English red naming centroid and the fourth major peak in the WCS best example distribution falls two chips away from the English blue naming centroid. These observations suggest strongly that objective tests will show a non-chance association between the highest peaks of the WCS focus distribution and points in color space favored by English color naming.

Variation in color naming

The above results demonstrate universal constraints in color naming. Yet there is also considerable cross-language variation, and it is still an open question *why* languages vary as they do in the naming of colors.

Lindsey and Brown (2002) provided for one aspect of this question a provocative answer, whose investigation has employed the WCS online data archive. Some languages have separate terms for blue and green, while others have compound green-or-blue ("grue") terms – they asked why this should be. They suggested that grue terms may derive from a sunlight-induced yellowing of the ocular lens: with a yellowed lens, short wavelengths are disproportionately filtered and blue stimuli would appear green, and would be named by the word for green. In other words, grue terms on this view are really words for green, and they extend to what normal eyes see as blue only because the yellowed lens distorts the perception of color. In support of this hypothesis, they noted that the proportion of languages with grue terms (rather than separate green and blue terms) is well-predicted by the amount of UV-B radiation from sunlight that strikes the earth's surface where those languages are spoken – as would be predicted if grue is ultimately traceable to sunlight-induced lens yellowing. They also showed that speakers of English who were shown stimuli that artificially simulated sunlight-induced yellowing of the lens extended the English word "green" to include stimuli presenting a spectral distribution comparable to that of a blue stimulus viewed through a yellow filter.

Intriguingly, Lindsey and Brown note that their hypothesis has the potential to "explain away" some recent findings suggesting a Whorfian influence of language on color cognition. Davidoff *et al.* (1999) examined color naming and memory in speakers of Berinmo, a language that has an enlarged yellow term, extending into the region that would be named "green" in English; this enlarged yellow category shares a border with a grue category. Davidoff *et al.* examined how well Berinmo speakers remembered colors straddling the boundary between these two Berinmo categories – and found that their performance was better for these colors than it was for colors straddling the boundary between English yellow and green. English speakers showed the opposite pattern. These findings suggest that a language's color terms may influence color cognition for speakers of that language (see also Kay & Kempton, 1984). However, Lindsey and Brown suggested a different interpretation of these data. They argued that since Berinmo has a grue term, its speakers may have yellowed lenses. This would explain why Berinmo's yellow term expands into green: because yellowish-greens are seen as more yellow through this lens. And it would also explain why color memory covaries with color naming across English and Berinmo: because both memory and naming are shaped by color perception, and that perception may be distorted by a yellowed lens in Berinmo speakers, relative to English speakers.

Regier and Kay (2004) tested this lens-yellowing hypothesis further, by probing a prediction it makes concerning the *best examples* of grue categories. If grue is really a green category that extends into blue because of a distorted perceptual color space, then there should be a single peak in the best example choices for grue (since there is a single peak for green), and it should fall somewhere between focal green and focal blue. However, if grue is instead a genuine abstraction over green and blue in an undistorted perceptual color space, the best examples for grue should peak either at green, or at blue, or at both. This prediction was easily tested using the focus data from the WCS data archive. We found that best examples for grue terms peak at English "green" and very near English "blue" (Figure 6; see also MacLaury 1997: 234-5; compare also with Figure 5, which shows focal choices from all WCS languages combined, including those that distinguish between green and blue). This suggests that the lens-yellowing hypothesis is incorrect. In doing so, it also indirectly supports the Whorfian hypothesis, by removing a competing explanation for the findings from Berinmo.

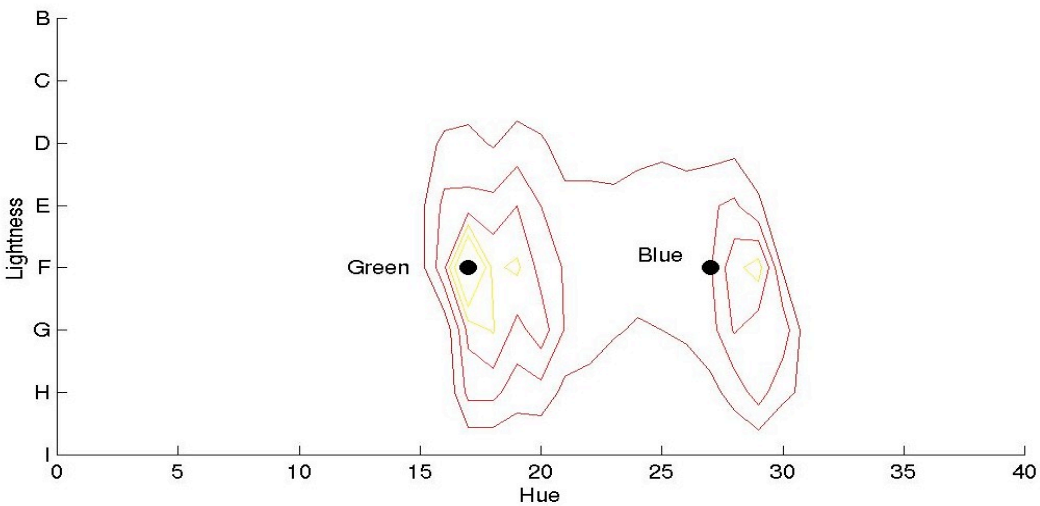


Figure 6. Contour plot showing the distribution, over chromatic stimuli, of best examples of grue terms in the WCS. Outermost contour represents a height of 10 hits; each subsequent inner contour represents a height increment of 10 hits. (Source: Regier and Kay 2004)

In response, Lindsey and Brown (2004) have presented further analyses of WCS focus data. Their findings confirm that grue best example choices tend to peak at focal green and focal blue – which argues against their hypothesis. But at the same time, their findings also leave open the possibility that there may be a subset of speakers whose best example choices are between universal focal green and blue, and thus consistent with their hypothesis. Further analysis will be needed to establish whether the observation of a small number focal choices between green and blue provides support for the Lindsey and Brown UV-B hypothesis. In any case, the hypothesis seems imperiled on other grounds as well. A recent finding has shown that color naming in English by individuals with *naturally* yellowed ocular lenses does not differ from that by individuals with non-yellowed lenses (Hardy, Frederick, Kay, & Werner, in press). The probable reason for the difference observed between Lindsey and Brown's simulation of yellowed optical media and naturally yellowed optical media is that in the latter case long-term processes of adaptation have time to operate (Schefrin and Werner 1990, Schefrin and Werner 1993, Neitz *et al.* 2002, Delahunt *et al.* in press), in order to perceptually compensate for the increased yellowing of the lens. This makes it quite unlikely that lens-yellowing could account for grue terms in the world's languages.

Conclusion

The WCS data archives are a publicly accessible resource, available to all who wish to pursue questions related to color categorization across languages. We have provided this background to orient potential users of the archive – to give them a sense for where the data came from, how the data were compiled into an archive, and what sorts of questions the data can be used to investigate. We hope the archive proves to be a useful and flexible research tool for the scientific community as a whole.

References

Berlin, Brent and Elois Ann Berlin (1975) Aguaruna color categories. *American Ethnologist* 2:61-87.

Boynton, Robert M.

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

Brown, Roger W.

1976 Reference. *Cognition* 4: 125-153.

Collier, George A.

1973 Review of Basic Color Terms. *Language* 49: 245-248

1976 Further evidence for universal color categories. *Language* 52: 884-890.

Conklin, Harold C.

1955 Hanunóo color categories. *Southwestern Journal of Anthropology* 11: 339-344.

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

Davidoff, Jules, Ian Davies, and Debi Roberson (1999). Colour categories in a stone-age tribe. *Nature* 398: 203-204.

Delahunt, P.B., Webster, M.A., Ma, M., & Werner, J.S., (in press) Long-term renormalization of chromatic mechanisms following cataract surgery. *Visual Neuroscience*.

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

Dougherty, Janet W.D. (1977) Color categorization in West Futunese: Variability and change. In *Sociocultural Dimensions of Language Change*. B.G. Blount

Durbin, Marshall

1972 Review of Basic Color Terms. *Semiotica* 6: 257-278.

Hage, Per and Kristen Hawkes (1975) Binumarin color categories. *Ethnology* 24: 287-300.

Hardy, Joseph L., Christina M. Frederick, Paul Kay, and John S. Werner (in press). Color naming, lens aging, and grue: What the optics of the aging eye can teach us about color language. *Psychological Science*.

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

Hickerson, Nancy

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

Kay, Paul, Brent Berlin, and William R. Merrifield

1979 *Human color categorization: A survey of the world's unwritten languages*. Research proposal submitted to the National Science Foundation (Renewal).

Kay, Paul and Willett Kempton (1984). What is the Sapir-Whorf Hypothesis? *American Anthropologist* 86, 65-79.

Kay, Paul (1999) The Emergence of Basic Color Lexicons Hypothesis. in *The Language of Color in the Mediterannea*n. ed. by Alexander Borg. Stockholm: Almqvist and Wiksell International.

Kay, Paul and Terry Regier (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences* 100: 9085-9089.

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

Lenneberg, Erik H. and John M. Roberts (1956) The Language of Experience: A study in Methodology. Memoir 13 of International Journal of American Linguistics.

Lindsey, Delwin and Angela Brown (2002). Color naming and the phototoxic effects of sunlight on the eye. *Psychological Science* 13: 506-512.

Lindsey, Delwin and Angela Brown (2004). Sunlight and "blue": The prevalence of poor lexical color discrimination within the "grue" range. *Psychological Science* 15, 291-294.

Lucy, John A. (1992) *Language Diversity and Thought: A Case Study of the Linguistic Relativity Hypothesis*. Cambridge, England: Cambridge University Press.

Lucy, John A. (1996) The scope of linguistic relativity: an analysis and review of empirical research. In *Rethinking Linguistic Relativity*. ed. by John J. Gumperz and Stephen C. Levinson. Cambridge, England: Cambridge University Press.

Lucy, John A. (1997) The linguistics of color. In *Color Categories in Thought and Language*. ed. by C.L. Hardin and Luisa Maffi. Cambridge, England: Cambridge University Press.

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

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

MacLaury, Robert E. (1997) *Color and Cognition in Mesoamerica*. Austin: University of Texas.
Maffi, Luisa (1990) Somali color term evolution: Grammatical and semantic evidence. *Anthropological Linguistics* 32: 316–334.
Miller, George A. and P. Johnson-Laird (1976) *Language and Perception*. Cambridge, Massachusetts: Harvard University Press.
Monberg, Torben (1971) Tikopia color classification. *Ethnology* 10: 349-358.
Neitz, J., Carroll, J., Yamauchi, Y., Neitz, M. & Williams, D.R. (2002) Color perception is mediated by a plastic neural mechanism that is adjustable in adults. *Neuron* 35, 783-792.
Ratliff, Floyd (1976) On the psychophysiological bases of universal color terms. *Proceedings of the American Philosophical Society* 120: 311-330.
Regier, Terry and Paul Kay (2004). Color naming and sunlight: Commentary on Lindsey and Brown (2002). *Psychological Science* 15: 289-290.
Roberson, Debi, Ian Davies and Jules Davidoff (2000). Colour categories are not universal: Replications and new evidence from a stone age culture. *Journal of Experimental Psychology: General* 129: 369-398.
Saunders, B.A.C. and J. van Brakel (1997). Are there non-trivial constraints on colour categorization? *Brain and Behavioral Science*s, 20: 167-228.
Schefrin, B.E. amd Werner J.S. (1990) Loci of spectral hues throughout the lifespan. *Journal of the Optical Society of America* 7, 305-311.
Schefrin, B.E. and Werner J.S. (1993) Age-related changes in the color appearance of broadband surfaces. *Color Research and Application* 18, 380-389.
Snow, D.L. (1971) Samoan color terminology: A note on the universality and evolutionary ordering of color terms. *Anthropological Linguistics* 13: 385-390.
Tornay, Serge (ed., 1978) *Voir et nommer les couleurs*. Nanterre: Publications du Laboratoire d'Ethnologie et de Sociologie Comparative, Université de Paris X.
Turton, David (1978) La catégorization de la couleur en Mursi. In Tornay, ed. 1978.
Turton, David (1980) There's no such beast: Cattle and colour naming among the Mursi. *Man* 15: 320-338.

Appendix: Facsimile of the WCS instructions to field workers and of the original coding sheets

World Color Term Survey

INSTRUCTIONS TO FIELD WORKERS

PLEASE READ THESE INSTRUCTIONS THROUGH. IF YOU HAVE ANY QUESTIONS, THEY CAN BE ANSWERED BY THE PERSON FROM WHOM YOU RECEIVED THESE MATERIALS. WE HOPE YOU FIND THIS MATTER INTERESTING. WE APPRECIATE YOUR COOPERATION VERY MUCH.

The object of the study which you are participating in is to obtain the meanings of the words for colors in a variety of languages around the world. The meanings from the different languages will then be compared in order to answer questions regarding such things as the extent to which color words are translatable across languages. It has probably been your experience that difficulties may be encountered in translating words for color between European languages such as English and the languages of non-European peoples. Both the differences between languages in color words which make translation difficult and the universals in color naming across languages which make translation possible are scientifically interesting and may help advance our understanding of the process of translation.

You are asked to gather data from twenty-five adult informants, as nearly as possible equally divided between males and females unless this imposes a great burden in your particular field circumstances. In so far as possible they should be monolingual.

You have been provided with two sets of color stimuli. Both are contained in the 14-1/2" x 7-1/2" metal box. Please open that box now. You will find inside the box a piece of cardboard on which a large number of circular color patches of 1/4" diameter have been pasted. Please put this aside for the moment. In the box you will now see that there are six trays containing color chips enclosed in glass slide cases. There are 330 of these chips. Each of the first five rows contains 56 chips and the sixth row contains 50 chips. You will note that each slide case has a white side and a gray side. The chips have been packed with the white sides facing the front of the box and the grey sides facing the hinges. The color patch shows through the grey side. On the back (white) side of each chip there has been written a number between 1 and 330. The chips in the first row have been numbered, from front to back, 1-56, in the second row 57-112, in the third row 113-168, in the fourth row 169-224, in the fifth row 225-280, and in the sixth and last row 281-330. You will always present the chips to informants in the numbered order. The set of 330 loose chips contained in the slide cases constitutes the first stimulus set and will be used in the naming task, to be described below.

The second set of stimulus materials consist of the piece of cardboard on which a large number (410) circular color patches have been glued. These are exactly the same colors as contained in the first stimulus set. The reason there are 410 of these as against 330 of the loose chips is, as you may have already noted, that the entire top and bottom rows of the array consist of 40 copies of the pure white and pure black chips, respectively.

You will be asked to perform one task with each set of stimulus materials. As indicated above, the first task is called the naming task.

The Naming Task

The object is to get the informant to name, preferably with a single word, each of the 330 chips. This task will take about 40 minutes. (The second task should take less than five minutes.) It should be performed on a sunny day if possible in the shade, not in direct sunlight.

Please place on a table or on the ground between you and the informant (a) the box containing the chips and (b) a copy of the CODING BOOKLET for the study of color categories in unwritten languages.

First, please fill in the information on the front page of the booklet regarding the informant's language, name, age, sex, etc. You may wish to precede or follow this with a certain amount of informal conversation to put the informant at ease. Depending on your personal relationship with the informant and his degree of sophistication in such things, you may wish to explain more or less about the nature and purpose of the study.

When the informant appears reasonably at ease, please open the booklet to page 2 and proceed with the naming task. As you may be aware, many languages do not contain a word meaning "color". Yours may well be one of these. Experience has shown, however, that it is always possible to find some verbal formula to elicit color responses. Sometimes these translate to, "How has it been dyed?" or "How does it strike the eye?" or "What is its appearance with respect to red, blue, etc.?" and so on. Probably none of these three is just the thing needed for your language, but with a little experimentation you should be able to find a question that elicits color words. You will be aided in this by the stimulus objects themselves, which differ from each other only with respect to color.

Please draw chip 1 from the box and hold it up so that the color patch faces the informant and the number faces you. Ask the informant to name the chip and record the name after the number 1 on the answer sheet. Then replace chip number 1 at the back of the first row of the metal box. Continue the same procedure with chips 2-56, until the informant has named all the chips in the first row and you have recorded all the names. (The chips in the first row are now back in their original order and ready for use on the next informant.) Now follow the same procedure with the chips in rows two through six. When recording the informant's responce to each chip, please do a visual-mental check that the number on the back of the chip corresponds to the number of the line of the coding sheet on which you record the informant's response.

Note that the procedure has been designed so that you never have to search for a chip and you never have to rearrange the chips. The next chip to be presented is always the chip at the front of the row you are working on, and after the full task is finished the chips are in their regular order and ready for presentation to the next informant. (Of course, if the chips should get spilled, they will have to be put back in the original order.)

This task has been designed for a single worker, but if you have someone available to help you, it can be done even more easily by two workers, one handling the chips and the other recording the informant's responses.

A Note on the Kind of Response We Are Looking For

In this study we are interested primarily in basic color terms, that is the simple, frequently used, most general one-word designations of color. In English these include black, white, red, blue, etc., and exclude blond, greenish blue, scarlet. Blond lacks generality because it can be used only of hair and furniture; greenish blue is linguistically complex; scarlet lacks generality in that it is a kind of red. Of course, any such descriptive expression as the color of sea water is not a basic color term. We have found that in every language we have investigated so far there is a short list of between two and eleven simple words with which every color can be named. These are the basic color terms of the language. Usually it is easy to tell which are the basic color terms for a language, but occasionally one or more terms show up which require some judgement. Furthermore, it sometimes occurs that a given term is basic for some speakers but not basic for others. (In general, you may be surprised to find more variation between informants than you might have expected.) So, the shorter the informant's responses, the better for our purposes. We are interested in responses like "red" and not in responses like "the color of the blood of a toucan that has been dead for a few hours." You will find that when the informant limits himself to short responses, the task is much easier, quicker, and less strained for both of you.

You may find it convenient to make one- or two-letter abbreviations for the basic color terms in the language you are recording, since a great many chips will receive the same name and this will save you a lot of time in recording the informant's responses. This is quite acceptable as long as you provide us with a clear key to the abbreviations. But please be consistent and always abbreviate the same way.

Defining the Concept Basic Color Term

Every language has an indefinitely large number of expressions that denote the sensation of color. Note, for example, the following English expressions: (a) crimson, (b) scarlet, (c) blond, (d) blue-green, (e) bluish, (f) lemon-colored, (g) salmon-colored, (h) the color of the rust on my aunt's old Chevrolet. But psychologists, linguists, and anthropologists have long operated with a concept of basic color term, or basic color word, which excludes forms such as (a)-(h) and includes forms like black, white, red, and green. Essentially, the basic color terms for a given speaker are the smallest set of simple words with which the speaker can name any color. However, the expression 'basic color term' does not have a unique operational definition. We use the following procedure for the determination of basic color terms. Ideally, each basic color term should exhibit the following four characteristics:

(i) It is monolexemic; that is, its meaning is not predictable from the meaning of its parts. This criterion eliminates examples (e)-(h) and perhaps also (d).

(ii) Its signification is not included in that of any other color term. This criterion eliminates examples (a) and (b), which are both kinds of red for most speakers of English.

(iii) Its application must not be restricted to a narrow class of objects. This criterion eliminates example (c) which may pe predicated only of hair, complexion, and furniture.

(iv) It must be psychologically salient for informants. Indices of psychological salience include, among others, (1) tendency to occur at the beginning of elicited lists of color terms, (2) stability of reference across informants and across occasions of use, and (3) occurrence in the ideolects of all informants. This criterion eliminates all the examples (a)-(h), most particularly (h).

These criteria (i-iv) suffice in nearly all cases to determine the basic color terms in a given language. The few doubtful cases that arise are handled by the following subsidiary criteria:

(v) The doubtful form should have the same distributional potential as the previously extablished basic terms. For example, in English, allowing the suffix -ish, for example, reddish, whittish, and greenish are English words, but *aquaish and *chartreus(e)ish are not.

(vi) Color terms that are also the name of an object characteristically having that color are suspect, for example, gold, silver, and ash. This subsidiary criterion would exclude orange, in English, if it were a doubtful case on the basic criteria (i-iv).

(vii) Recent foreign loan words may by suspect.

(viii) In cases where lexemic status is difficult to assess (see criterion (i)), morphological complexity is given some weight as a secondary criterion. The English term blue-green might be eliminated by this criterion. (The preceding discussion of the notion basic color term is taken, with some minor additions, from Brent Berlin and Paul Kay, Basic Color Terms. Berkeley: University of California Press. 1969. pp. 5-7.)

We mentioned above that different informants may have different sets of basic color terms. Some informants may lack entirely terms which are basic for others or they may have these terms but functioning as secondary terms. Situations like this occur frequently in languages that are undergoing rapid change, often in response to contact with European languages. For example, the traditional system may contain a single term covering both green and blue. But some informants, often younger ones, may have taken over either the term for green or the term for blue from a European language and restricted the meaning of the traditional term to that part of its original signification not included in the borrowed European term. This is just one example of many complex situations of this sort that may arise. Once again, we will welcome any comments you may wish voluntarily to make about situations such as this if they arise, but your much appreciated cooperation in the study certainly does not require you to so comment.

The Focus Mapping Task

As you may have surmised, the naming task is designed to elicit the informant's widest extension of his color terms. In contrast, the focus mapping task is intended to elicit the best example(s) of his basic color terms. A moment's reflection reveals that in addition to being able to say whether a given color is, say, red, we can also judge whether or not it is a good, typical, or ideal red. Please look now at the board with the ¼" circular patches and ask yourself, "Which is the best example of red?" If you do this you will find it a fairly easy question to answer. Our experience shows that this is true generally across languages. People have an easy time looking at the board and pointing to one or a few chips which best exemplify each of their basic color terms.

We would like you to elicit in the focus mapping task your informants' judgements of the best example or examples of their basic color words. If you have spent a long time working on your language, you will probably already know which terms fulfill the criteria which we have outlined above, and so deciding which terms are basic will be a straightforward matter. If you have just recently begun to carry out translation work on your particular language or if for any other reason you have doubts as to which terms are basic, please consider as basic terms for the purpose of this task any term elicited for five or more chips in the naming task. (You will recall that you have tried to discourage responses such as light blue, greenish blue, robin's egg blue, bluish, etc. These are all modifications of a basic term and may be indicated in various languages by reduplication, addition of particular affixes or modifying words, and so on. If the total number of occurrences of a simple term plus its modifications exceeds five, please count the simple term as a basic color term for the purpose of the focus mapping task.)

In performing the focus mapping task you will probably want to do something like the following. Place the fixed array of 410 ¼" circular color chips before the informant, hand him a pencil or a stick, and say something like, "Would you please indicate the best example of (native color term)?" Responses should be recorded on the back page of the scoring booklet labeled FOCUS MAPPING CODE SHEET. You will see that the coding sheet reproduces the array. Please indicate which chip or chips the informant gives as best examples for each color term by circling the dot(s) corresponding to the chosen chip(s) and connecting the native term to the circled dot(s) with a line. (You may consult the sample on the last page of the instructions.) This task should also be performed on a sunny day in the shade.

Try to restrict the informant to a single chip in so far as possible, but if the informant insists that several chips are equally good representatives of a color you may accept this as final. Generally, if you can get the informant to choose a single chip as a judgement of applicability of the color term, please do so, but don't force him beyond a point where he appears to be mentally "tossing a coin".

Summary of Instructions

1. Sample. Select twenty-five adult informants, in so far as possible monolingual.

2. Face Sheet. Interview each informant and fill out the name, age, and other questions.

3. Naming Task.

   a. Place on table or floor the full box of chips and the scoring booklet opened to the CODING SHEET FOR NAMING TASK.

   b. Starting with chip number 1: remove the chip; ask the informant to name it; record the response in the CODING BOOKLET; return the chip to the end of the row it came from.

   c. Complete row one in this way; then do in the same way rows two through six in order.

   d. Remember, two workers will make this task go faster.

4. Focus Mapping Task.

   a. Determine which of the expressions the informant gave as answers on the Naming Task are basic color terms. (When in doubt, consider as a basic color term any expression that -- together with its modifications -- occurred five or more times. A modification is an expression like light blue, bluish or greenish blue.)

   b. Have the informant indicate the best example(s) of each basic color term on the array of circular color chips, noting the chip(s) indicated, along with the native word, on the FOCUS MAPPING CODE SHEET on the back of the booklet. (Please see the sample on the last page of these instructions.) Try to restrict the informant's response to a single chip if possible.

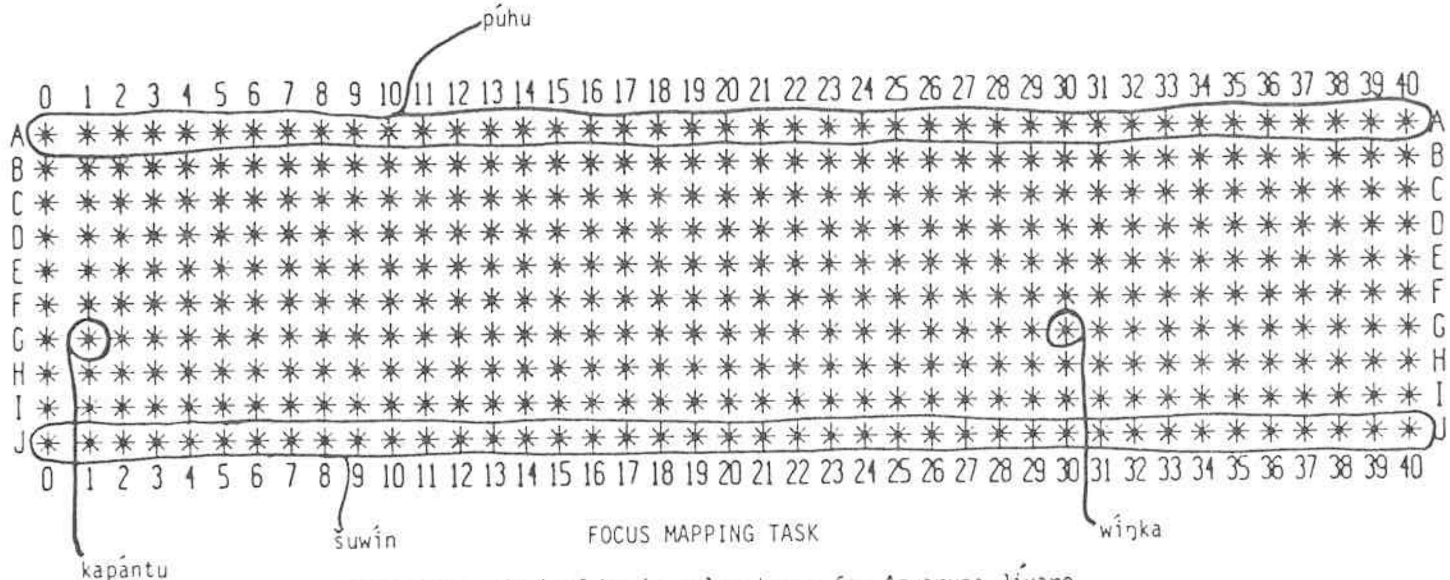
5. When you have finished you should have twenty-five completed coding booklets, one for each informant. Each coding booklet should contain (a) "Face-sheet" information on the name, age, sex, etc., of the informant, (b) a completed CODING SHEET FOR NAMING TASK, (c) a completed FOCUS MAPPING CODING SHEET.

6. We will be deeply grateful for any comments you may have on any aspect of your data. Also if you have any questions or comments regarding the study as a whole, we will appreciate receiving them and will answer them to the best of our ability.

Let us thank you again for your cooperation in this study. Without you, there would have been no way of doing it. We hope the results may be useful in many areas of linguistics, including the art and science of translation. If you have any comments or suggestions regarding this study we will be delighted to receive them.

Brent Berlin and Paul Kay
Department of Anthropology
University of California
Berkeley, California
94720

William Merrified
Summer Institute of Linguistics
7500 West Camp Wisdom Road
Dallas, Texas
75211



FOCUS MAPPING TASK

(Example of foci of basic color terms for Aguaruna Jívaro
Peru, South America)