PERCEIVED VERSUS ACTUAL VALUE OF COLOR-CODING

Carl Karner

University of Oklahoma

ABSTRACT

Color-coding has become a widely-used method of information input coding. Unfortunately, little is known regarding general statements concerning under what conditions color-coding will facilitate performance. For this reason, a decision is often made on the basis of judgment if empirical data is not available.

The present study is designed to examine performance in a particular short-term memory task and to assess the actual value of color-coding in the task. This actual value is then compared to the participants' judgments of the value of color-coding in that task. The task required the subject to keep track of the current state of each of several variables. One group used a color-coding system while a second group did not. The results showed performance to be significantly better without color-coding; however, all subjects from each group in a post-test interview that they felt that color-coding would be beneficial. The implication is that humans may have rather poor insight regarding the facilitating effects of color-coding.

INTRODUCTION

In recent decades engineers have created complex machines that require constant monitoring by human operators. Far too often chaos results from operators being unable to keep track of all the information revealed on display systems. Design engineers of the past relied on their common sense to determine how much information should be monitored by an operator and how the information should be presented. Such a method is no longer adequate. It is necessary to determine how much information an operator can handle as well as the most efficient method of presentation. One scheme used to present information is to color-code the data. Unfortunately, the decision to use color-coding is often made in an unscientific manner with the idea that "it couldn't hurt". The purpose of this paper is to compare the actual and perceived value of color-coding in a task involving short-term memory.

There are many examples of color-coding being used to aid human operators. The electronics industry color-codes many parts including wires, resistors, and lights. Libraries color code many of their reference material. Offices often color code the forms they deal with and pharmacies color code drugs.

An example of a system which would be more closely related to the experiment discussed in this paper would be pipe line monitoring and control. Modern pipe lines depend on remote control of pump stations. Quite often all needed information is automatically relayed from unmanned stations to one that is manned several miles away. Usually, there are several bits of information being relayed from each station at once to a single display board. In some cases the information from one station is distinguished from the information from another station through the use of color-codes.

LITERATURE BACKGROUND

Among the early work done concerning the effect of the number of variables on short-term memory used tape recordings with word pairs. This method showed that short-term retention is a function of average storage load (Reid, 1961). Later, D. B. Yntema used cards to show that the same was true using sight rather than sound (Yntema, 1963). He also showed that the method of presentation could make a difference in the amount of information a person could retain. He found it is better to have a few variables with few states. In order to keep the type of information from biasing his results, Yntema tried to present the same information when testing few variables with many states as when he tested many variables with few states.

The usefulness of color-coding in the presentation of information is not easy to determine. To quote E. J. McCormick, "Although color seems to be a very useful coding dimension in some contexts, it is apparent that it is not a universally preferred scheme, since in certain studies...other coding dimensions were found to be superior or at least equal" (McCormick, 1970). McCormick goes on to say that it is the attention-getting characteristic that makes color particularly useful.

Researchers have compared the effects of various colors on reaction time. It was determined that numerical codes evoked the speediest and most accurate responses (Alluisi, 1958). Experiments have indicated that color-codes do not appear to be suited for situations that demand rapid and precise identification, whereas they are valuable in decreasing search-time with locate-type tasks (Jones, 1962). In 1970 an experiment was performed that found that color did enhance short-term memory and that the use of certain colors with certain words was even more meaningful. It might be added, however, that this change was so slight that the results were not statistically significant. One other fact about

227

color should be brought out at this point. Colors seem to change hue due to changes in light, and people seem to see some colors differently (Helson, 1952). That fact was kept in mind when choosing the colors to be used in the following experiment so that colors would not be used which were easily confused.

DESIGN OF EXPERIMENTAL EQUIPMENT

The design of the experimental equipment was such that it tried to incorporate the ideas found in previous research. Two decks of cards were constructed for the experiment with each card divided into two parts. The top half of each card contained the color (variable) coding while the bottom half contained a symbol (state). The color-coding of one deck of cards was a patch of a color painted on the card. In order to avoid any bias due to the type of information being presented, both decks of cards contained the same information with only the method of color-coding being different. Cards were used to insure that the subjects would get a good enough view of each color, word, and symbol that he or she would be able to distinguish them clearly. In this way the difficulties of rapid and precise identification would be minimized. The cards were made using the colors (variables) red, blue, yellow, green, orange, and white. In order to keep contrast constant, each card was separated into two parts rather than print a symbol (state) on a colored card.

The following symbols were used to represent the states of the colors: O , ◻ , ✗ , ⊛, △ , ✱ . These symbols were chosen because they are simple and relatively free of color connotations. An example of a symbol that would be associated with color would be a ♡ which is associated with red. The use of numbers as symbols was avoided because previous research has shown that people will keep track of numbers rather than colors when presented with both (Kanarick, 1971). In other words, the purpose of this experiment was to determine the effects of color-coding, not the effect of the type of symbol used.

Two additional sets of cards were made. These cards were of the same size and had the same general appearance as each deck of cards described above. The only difference was that these cards did not have one of the usual symbols. These cards contained the symbol "?" and were used to ask questions of the subjects in a random manner.

The number of cards of each color-coding to be made was determined randomly by the use of a die. The number of each symbol to be assigned to each of the color-codes was also determined randomly through the use of a die. TABLE A shows the number of each type of card that was made. It should be pointed out that in order to keep the information consistant the same number of each card was made for both color-coding decks.

TABLE A

Symbol / Color	✱	O	◻	△	⊛	✗	?	Total
Orange	2	5	5	4	2	3	5	26
Blue	6	4	3	5	3	2	5	28
Green	2	3	4	5	4	5	4	27
Red	3	4	4	5	2	4	4	26
White	3	3	3	5	3	4	5	26
Yellow	5	2	3	4	3	3	7	27
Total	21	21	22	28	17	21	30	160

EXPERIMENTAL PROCEDURE

An outline of the experimental procedure can be found in TABLE B.

TABLE B

Outline of Experimental Procedure

I. Familiarize the subject with the card deck to be considered.

   A. Give the subject a pencil and paper and ask him/her to:

      1. List the colors (variables) being considered.
      2. List the symbols (states) to be considered the first time through.
      3. Number paper from one to twenty.

   B. As part "A" is being done show the subject examples of each variable and state from the deck.

   C. Show the subject a "?" card and ask them to write the last symbol they remember seeing with that color code during the experiment.

   D. Show the subject about ten cards to be sure he/she understood the instructions.

   E. Tell the subject to guess if he/she doesn't remember the last symbol for a given color.

   F. Begin the experiment by showing the subject one card at a time until 20 questions have been asked.

   G. Perform the second part of the experiment after:

      1. Increasing or decreasing the number of states.
      2. Shuffling cards.
      3. Not telling subject how he or she did.

   H. After the second portion is finished, grade the papers and tell the subject how he/she did.

II. Inform each subject about the other deck and ask which deck would be easier to use.

A total of sixteen subjects were used, eight were tested on each of the two card decks. Fortuitous sampling was used in selecting subjects, that is, the subjects received only a "thank you" for their trouble. All subjects were college students ranging from 20 to 25 years of age. Two females were used on the colored card deck while four of the eight subjects used on the printed deck were females. All subjects used with the colored cards were screened for colorblindness before being considered as a possible subject (Ishihara, 1966). The subjects were not told that another deck of cards existed until after the experiment. They also were not told how they were doing or how others did on the experiment until they were finished. Each subject was allowed to write at the top of his answer page the colors and symbols he was being asked to keep track of. They were also told they would have two sets of 20 questions and were asked to number their papers before the experiment began.

Each subject was tested twice, once with 4 states (symbols) and once with 6 states but always on the same color-code-type deck. The tests were counter-balanced to avoid systematic bias. Four subjects in each group were tested with 4 states and then with 6. The other four were tested with 6 states before being tested with 4 states. The experiment began after a key was prepared containing the correct answers of the initial conditions. This kept the time lost on each subject at a minimum because it was not necessary to check their answers before adding or subtracting cards for the second part of the experiment. The key for the second part of the experiment was made after it was performed and both tests graded at the same time.

Just prior to the first test the subject was shown an example of each color and symbol to be used. During the test, the subjects were shown one card at a time. Each subject was instructed to write down the last symbol (excluding question marks) he saw for a given color when he saw a question mark on a card in place of a symbol. The subjects were also instructed to guess when they did not know the answer. The question cards were randomly shuffled into the deck along with the other cards so that their occurrence would be random. On the four state test, the Δ's and $\square$'s were removed from the deck. Ten question cards were also removed to keep the proportion of questions to non-question cards constant for the trials of both states. It should be noted that before each test the deck was checked to make sure that cards defining each variable (color) occurred before a question on that variable appeared.

After both tests had been graded, the subject was told what his score was. The subject was then told that another deck of cards existed. The difference was explained. The subject was then asked which deck of cards he thought would be easiest to keep track of. In all cases the subject concluded that the cards with a physical color painted on them would be easier than cards with the names of colors printed on them. This decision was constant regardless of which deck of cards the subject had been tested on.

THE RESULTS

The results of the tests were the number of errors each subject made as shown in TABLE C:

TABLE C

Colored Cards

Subjects	4 States	6 States	Total
1	7	9	16
2	8	11	19
3	10	10	20
4	11	9	20
5	13	10	23
6	9	8	17
7	12	12	24
8	12	10	22
Total	82	79	161

Printed Cards

Subjects	4 States	6 States	Total
9	14	8	22
10	7	10	17
11	6	5	11
12	10	7	17
13	6	7	13
14	11	9	20
15	7	6	13
16	8	6	14
Total	69	58	127

STATISTICAL ANALYSIS

The type of design used in the statistical analysis of the results was a two-factor analysis of variance with repeated measures on the second factor (Chapanis, 1965). The question considered by this analysis was whether the use of painted colors significantly changes the ability of the subjects to keep track of several things at once as compared to the use of printed names of colors. This analysis also indicated whether it is easier to keep up with four states as compared to six states when the states are distributed in six variables.

An α level of .05 was chosen for this analysis. The value of the F ratio for 1 and 14 degrees of freedom is 4.60. The analysis of variance produces TABLE D:

TABLE D

| Source	SS	df	MS	F
Between Subjects	114	15		
Variables	36.125	1	36.125	6.49
Subjs. w/Categories	77.875	14	5.563	
Within Subjects	48	16		
States	6.125	1	6.125	2.53
Variables x States	8.0	1	8.0	3.31
States x Subjs. w/Variables	33.875	14	2.42	
Total	162			

229

DISCUSSION OF RESULTS

From the Analysis of Variance Table we note that 6.49 > 4.60. This means that there is a significant differnece between the use of printed cards and colored cards. Closer inspection reveals that the printed cards gave generally better results than the colored cards. This indicates that the printed words were easier for people to encode and remember than information presented by colors. A possible explanation for this finding is the fact that college students were used as subjects and that they are accustomed to trying to remember information presented as written words. Another possible reason could be that interpreting and encoding colors are more difficult than interpreting and encoding words.

It is interesting to note that these results are contrary to what the subjects thought would be the case. Recall that when told about the two decks of cards the subjects felt sure that the colored deck would be easier to keep track of than the printed deck. Of course, the results of the experiment were just the opposite. In this case the use of statistics was not needed because 100% of the subjects thought the use of colors would be superior to the use of the names of those colors. This result shows that one can't depend upon their common sense to determine the most efficient method of presenting data even in the most "obvious" cases. The use of colors were really "non-sense" catagories since there was no natural relationship between the variables (colors) and states (symbols). Even so a significant difference was found between the types of coding used. One would expect variables that suggest the states to be even more superior than the use of colors. In the case of the pipe-line example used earlier, one would expect fewer errors to occur using variables such as the location that each state is concerned with than a row of colors to distinguish a given location.

Referring again to the Analysis of Variance Table we find that 2.53 < 4.60. This means that the number of states used did not have a significant impact on the results. It can therefore be concluded that the level of difficulty of six states is not significantly different from four states.

Still another look at the Analysis of Variance Table shows that 3.31 < 4.60 which is not significant. It is therefore concluded that varying the type of card (printed or colored) did not significantly affect the number of errors when the number of symbols were changed. That is to say that the pattern of errors vs states was unchanged by the type of card used (see FIGURE A). Also, the large amount of variation and overlap of results probably had much to do with the insignificant finding concerning the difference between 4 and 6 states.

FIGURE A
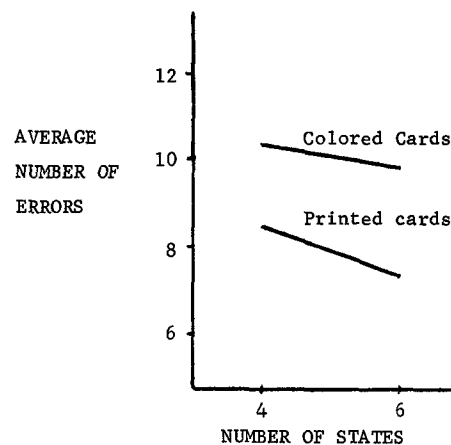
SUMMARY AND CONCLUSIONS

The results of this experiment can be put to good use by engineers who are designing information display systems. It is clear that one should not design such systems with colors as the method of distinguishing between variables. In the pipe line monitoring example discussed earlier, a method other than color-coding should be used to indicate the various remote stations on the control panel.

Perhaps even more important than this is the fact that this experiment shows that one should not depend on his common sense to determine the best way to present information. All the subjects felt certain that it would be easier to keep track of several things at once with color-codes than printed codes. All of the subjects were wrong. Clearly then, an engineer designing a display should run an experiment to determine how to best present information. He should not envoke color-coding just on an opinion that it would be helpful. He should run an experiment to determine whether color-coding or any other type of coding for that matter would be beneficial in presenting the type of data he is concerned with.

This experiment also shows that one is justified in dividing the information being presented by six variables into six states. This can be concluded by the fact that this experiment shows no significant difference in accuracy between six and four states. As a result a design engineer can break each variable of information into six rather than four states. In this way he will increase the preciseness of each reading without losing a significant amount of accuracy.

230

It can also be noted from this experiment that the negative effect of color-coding is not dependent upon the number of states involved for each variable. Therefore, one would conclude that this negative effect of color-coding can be expected to apply to any number of states not just to 4 and 6 states.

RECOMMENDATIONS

The purpose of this experiment was to investigate only one type of information task. There are many experiments that should be conducted to further investigate this topic. One such experiment would be to use colors as states and symbols as variables. Such an experiment might reveal that color-coding shouldn't be used in this kind of task either.

Another interesting experiment would be to vary the number of states and variables. Such a series of experiments could determine at what point the human mind becomes overloaded. Design engineers knowing such information would be able to design systems for the maximum amount of information without overloading the operator.

One could change the type of information being presented. This would enable an engineer to determine the best means to use in presenting information. This could be done in conjunction with varying the number of variables and states in order to determine what type of display systems should be used when many things must be kept track of.

As it was conducted, the experiment shows that one should not rely only on his common sense. Also, he should not routinely use color-coding unless there is clear evidence that it would be beneficial. However, all the subjects used were college students. Since college students are accustomed to memorizing printed words it would be of interest to repeat the experiment using people other than students. In this way it would be possible to generalize the results or discover that there may be a difference in the way college students keep track of things than do people in general.

REFERENCES

1. Alluisi, E. A. and Muller, P.F., Jr. "Verbal and Motor Responses to Seven Symbolic Visual Codes: A study in S-R Compatibility." Journal of Experimental Psychology. 1958, 55, pp. 247-254.

2. Chapanis, Alphonse. Research Techniques in Human Engineering, paper-back edition, 1965, p. 138.

3. Helson, H. "Vision", Annual Review of Psychology, Vol. 3 (1952), 55-84.

4. Ishihara, Dr. Shinobu. Ishihara's Tests for Colour-Blindness, Concise Edition, 1966.

5. Jones, M.R. Color Coding. Human Factors, 1962 4, pp. 355-365.

6. Kanarick, Arnold F. and Peterson, Ronald C. (Honeywell Inc., Systems Research Center, Minneapolis, Minn.) "Redundant Color Coding and Keeping Track of Performance." Human Factors, 1971 (April), Vol. 13 (2), pp. 183-188.

7. Lee, David M. (U. of Calif. L.S.). "Does Color Enhance Short-Term Memory?" Dissertation Abstracts International, 1970 (Feb.), Vol. 30 (8-B), pp. 3890-3891.

8. McCormick, Ernest J. Human Factors Engineering, Third Edition, McGraw-Hill, 1970, p. 106.

9. Reid, L.S., Lloyd, K.E., Brackett, H.R., and Hawkins, W.F. "Short-Term Retention as a F Function of Average Storage Load Reduction." Journal of Experimental Psychology, 1961, 62, 518-522.

10. Yntema, D.B. "Keeping Track of Several Things at Once." Human Factors, 1963, 5, 7-17.