The Effects of Color Enhanced Information Presentations

Lee L. Gremillion
Indiana University

A. Milton Jenkins
Indiana University

Follow this and additional works at: http://aisel.aisnet.org/icis1981

Recommended Citation
http://aisel.aisnet.org/icis1981/8
The Effects of Color Enhanced Information Presentations

Lee L. Gremillion
School of Business
Indiana University

A. Milton Jenkins
School of Business
Indiana University

ABSTRACT

Color coding is becoming an important dimension in the presentation mode capability of many information systems. Although color coded output, either alphanumeric or graphic, is of considerable intuitive appeal, there is little experimental data on its use or effects. An experiment is reported which focuses on the effects of color enhanced information presentations. Information is presented to two groups of subjects differing the presentation mode only along the dimension of color coding. The subjects are later tested on the material presented. Analysis of the test results indicates a significant difference in the subjects' performance which is apparently attributable to the color coding of the presentation materials. Directions for further investigations are suggested.

INTRODUCTION

One of the more interesting developments in computer technology in the last few years has been that of color in computer outputs. Devices to produce output in various colors, particularly cathode ray tube terminals, are now available at relatively low prices (Durett, 1980). Programming to drive this equipment, while lagging and hardware itself, is likewise becoming more readily available. For example, the basic video screen formatting software currently used in conjunction with IBM's main teleprocessing monitors gives any user with the appropriate hardware the capability to highlight data items in any of seven colors (IBM, 1980). Color has become another variable to manipulate in establishing a "presentation mode" for information (Mason and Mitroff, 1971).

A number of writers, particularly in the trade literature, have made dramatic claims for the benefits to be achieved from color. For example, one writer (Friend, 1980) states that using color graphics:

Transmission of information from the computer to the user increases so dramatically that the tedium of wading through written reports is replaced with a higher level of intellectual activity—depth of understanding, creative interaction with the data and conceptual problem solving.

Intuitively, there is considerable appeal to such a claim. A picture is worth a thousand words (and worth even more numbers), and a color picture is even better. Our general preference for colored
pictures, for example, has largely made black and white television sets obsolete except for special purposes.

But does color really enhance the assimilation of information as much as the popular and trade literature suggests? When it is not a matter of entertainment, but rather of purposeful information, particularly management information, what exactly does color coding, in and of itself, add? The statement quoted above—written by the president of a consulting firm which specializes in color graphics, and based apparently on his personal experience—is typical of the claims often made for color-coded computer output. Unfortunately, the research literature on the effects of color coding on human performance lends only weak and qualified support to such claims.

**RESEARCH ON COLOR CODING EFFECTS**

A great deal of research has been conducted over the past three decades into the effects on learner performance of employing color coding or cuing in instructional materials. Lamberski (1980) and Chute (1979), in independent reviews of this body of research, point out that the results reported have been contradictory and largely inconclusive. Chute concludes that "color's importance as a cuing device evidently is considerably less than what might naturally be expected." Lamberski points out that literally hundreds of experiments have failed to establish that color cuing and coding will consistently or significantly enhance learning when used in media presentations. Instead, the research indicates that a large number of intervening variables—such as information content, learner aptitude, presentation complexity, and presentation pacing—strongly influence the effectiveness of color as a learning aid. Although learners do generally prefer color instructional materials over monochrome materials, making that color effective is apparently much easier said than done.

Barker and Krebs (1977) surveyed the research literature dealing with the effects of color coding on human subjects' abilities to detect, locate, identify, compare, count, and/or track objects in a display. They also concluded that the effectiveness of color coding was highly dependent on the task being performed. Color was found most useful in aiding subjects detect a target whose position on the display was unknown. In many cases, they point out, other cuing devices (shapes, light intensity, blinking, etc.) were as effective as, or more effective, than color coding.

The reported research thus suggests that determining how to use color effectively in visual displays is not a trivial task, and that success in such use is far from guaranteed. Unsupported assertions which appear in articles on color computer output, such as "Color is an engaging medium..." "Color line graphs are much easier and more natural for the eye to follow," and "After the novelty wears off color displays may prove distracting, but I doubt it" (Eddy, 1980), must be strongly questioned. At any rate, they provide little or no guidance to the manager or system designer pondering the payoffs and pitfalls of using color computer output. Nevertheless, it is a fact that color display units are being purchased and employed in management information systems with increasing frequency. Clearly, experimentation is called for which focuses on the effects of color coding in the types of applications to which color is likely to be applied in information systems.

**BACKGROUND: SOME IDEAS FROM COGNITIVE PSYCHOLOGY**

The intuitive appeal of using color in presenting information is easy to understand. Most of us, if asked, would readily respond
that color would probably make the information more interesting, make it easier to understand, perhaps even make it easier to remember. A review of the cognitive psychology literature shows that there is empirical evidence to support these views. Three items are particularly pertinent.

A number of experiments have shown that people are more likely to remember material when they are able to meaningfully interpret that material. For example, Bower, Karlin, and Dueck (1975) had subjects study semi-abstract pictures which they were later asked to redraw. Some subjects were given no explanation of the meaning of the pictures, others were given labels for each picture which helped make them meaningful. Recall was much better for this latter group. Extending this finding to the use of color, we might conjecture that effective color coding can make presented information more meaningful, and therefore help people remember it. This is particularly true where color is used to communicate relationships and dependencies in the information (such as the often cited use of red to indicate losses or deficits). Salomon (1972) refers to this use of color coding as a "supplanting function"--in effect, the color cue supplants a covert mental operation that the learner or observer would otherwise have had to activate on his own. This is a use for color which previous research has suggested may be effective (Chute, 1979).

Experiments have also shown that when information is committed to memory it is often "elaborate" with additional redundant information which later facilitates recall by providing additional retrieval paths. A subject's memory for a piece of information, therefore, can be improved by manipulations that increase the amount of elaboration performed. For example, Hyde and Jenkins (1973) tested subjects on their ability to recall words which had been presented to them. Some of the subjects were asked to rate, to themselves, the "pleasantness" of each word as it was presented. This group showed much better ability to later recall the words than did the other group which performed no such elaboration. Extending this idea to the use of color, one might conjecture that color-coding in information provides another dimension along which elaboration might be performed, and thereby enhance its recall from memory.

Finally, there is evidence that a person's memory capacity for visual information is significantly greater than for verbal information. For example, Shepard (1967) found that subjects could recognize pictures that they had seen before with much greater accuracy than they could recognize sentences they had seen or heard. Since the use of color would be a visual enhancement to information presentation one might conjecture that it would provide reinforcement to this inherent visual memory strength.

While none of this directly addresses the use of color, it certainly suggests reasons why, in fact, the use of color in presenting information might be expected to be helpful, at least in storing and retrieving from memory the material presented, particularly if the color coding was used to reinforce the presentation of relationships. This paper reports an experiment designed to directly address this hypothesis. In the experiment, meaningful information is presented to two groups of subjects, differing the presentation mode only along the dimension of color coding. The subjects are then tested to determine whether any effects of color coding can be discerned.

**THE EXPERIMENT**

The experiment was conducted using students in an undergraduate course entitled "The Computer in Business" during the Spring Semester, 1981. Approximately 900
students were enrolled in twenty sections of the course. Most of the students were sophomore and junior business majors. The course employed a programmed instruction text supplemented by two and one-half hours of lecture per week.

The module entitled "Overview of Software" was selected as the subject matter for the experiment. This section covers system software and how it interacts with application software and the computer hardware. Types of system software and operating system functions are discussed, as well as sources of various types of software. This module was selected because the subject matter lent itself to the use of color coding to highlight and dramatize distinctions—a role often suggested for color in computer output. The material was aimed at enabling the students to recognize and distinguish between hardware and the two types of software and their functions. According to the reasoning developed above, one would expect color coded visuals to help the student understand and remember these distinctions.

For the treatments, a set of nine overhead transparencies were designed by the researchers to be used as a presentation aid. These were executed by the graphics designers in the university's Instructional Resources Department in two versions. The color version employed a color scheme in which hardware was keyed to red, system software to green, and application software to blue. Table I shows two of these and explains the coloring scheme. In general, on the color transparencies both labels and drawings were done in the red, green, or blue, on a clear background, depending on what topic (hardware, system software, or application software) was being referred to or depicted. The black and white version of the transparencies was identical to the color version except that all drawings and lettering consisted of black lines on a clear background.

Ten of the twenty sections of the course were randomly selected to use the color transparencies, ten the black and white. To help control for instructor bias, each of these ten section groups were further divided into two five section groups. One of these groups had the presentation made by their regular instructor, the other by one of the researchers. In all cases, the instructor presenting the material followed a detailed presentation outline, almost a script, in giving the lecture. The lecture itself took from thirty-five to fifty minutes to present.

For all classes, whether receiving the black and white or color treatment, the use of professionally prepared transparencies as a teaching aid was new. Thus any "novelty effect" would have been felt by both groups. Substitute instructors, whether using the black and white or color transparencies, were introduced with the comment "we're trying out some new instructional aids, so X is going to give the lecture today." The color and black and white transparency sets were used in exactly the same manner—no reference was made in the lecture to the coloring itself.

Figure 1 shows the number of students receiving each treatment. Attendance was rather low during the two days on which the lecture was given because of a pair of external factors. First, these were the first two class days following the spring break, traditionally a time of low attendance. Second, the university's basketball team played for the national championship the evening of the first class day, greatly distracting the students from all academic matters. This did, however, have the serendipitous effect of providing a substantial "control" group of students who would take the tests on the material having received neither the black and white nor the color treatment. (This was not a control group in the usual sense of the term. Members of this group were self-
This was the seventh of the nine transparencies. In it, the word "processor" and the outlines of the hardware (processor, main memory, and files) were in red. The main memory areas occupied by system software were shaded in green, and the labels (O/S, SORT, etc.) in green lettering. The main memory areas occupied by application software were shaded in blue, and the lettering in these program names done in blue.
Table 1 (continued)

<table>
<thead>
<tr>
<th>Type</th>
<th>Source</th>
<th>Written By</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>User Organization</td>
<td>Application</td>
</tr>
<tr>
<td>Software</td>
<td>Software Vendors</td>
<td>Programmers</td>
</tr>
<tr>
<td></td>
<td>(&quot;Canned Software&quot;)</td>
<td></td>
</tr>
<tr>
<td>Support Software</td>
<td>Computer Vendor</td>
<td>Systems</td>
</tr>
<tr>
<td></td>
<td>Independent Software Vendor</td>
<td>Programmers</td>
</tr>
<tr>
<td></td>
<td>User Organization</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(rarely)</td>
<td></td>
</tr>
</tbody>
</table>

This was the ninth of the nine transparencies. In the color version, the terms "Application Software" and "Application Programmers" were in blue, while the terms "Support Software" and "Systems Programmers" were in green.
### Figure 1. Distribution of Subjects According to Treatments

<table>
<thead>
<tr>
<th>Treatment</th>
<th>None</th>
<th>B&amp;W</th>
<th>Color</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular</td>
<td>147</td>
<td>105</td>
<td>95</td>
<td>347</td>
</tr>
<tr>
<td>Substitute</td>
<td>74</td>
<td>91</td>
<td>108</td>
<td>273</td>
</tr>
<tr>
<td>Total</td>
<td>221</td>
<td>196</td>
<td>203</td>
<td>620</td>
</tr>
</tbody>
</table>

(Treatment = none refers to students who took the quiz after having missed the lecture.)
selected rather than assigned by the experimenters. However, upon examination, this group appeared to be representative of the experimental subject population—their overall course grades were neither better nor worse than those of the other student subjects. Consequently, this group was utilized as a control group in this study.

To determine the effects of the treatments, students were subsequently tested on the material. The primary test was a short unannounced quiz administered in the class meeting following that one in which the presentation was given (i.e., 48 hours later). Table 2 shows this quiz. Questions on the material were also included in a departmental exam given to all sections two weeks later.

The specific hypothesis to be tested in this study is:

$$H_0:$$ There is no difference in test results between the students receiving the color presentation and those receiving the black and white presentation other than those attributable to random sampling error.

The alternative hypothesis, therefore, would be that there is a difference which might be explained by the treatments.

Note that the dependent variable is the subject's ability to recall the presented information upon testing rather than the ability to use it, as in a decision making situation. Recall is used because it is readily testable and because experimentation using recall as the dependent variable can be related to a large body of previous research. And, although the exact relationship between recall and use of information remains to be specified, it is clear that an individual must be able to recall information previously presented to be able to use it.

**RESULTS**

The first data analyses were done using the overall quiz grade, which was calculated as the number of correct answers on the quiz. Figure 2 shows the sample size, mean quiz score, and standard deviation for each treatment group, along with the t-statistic value between group means and its significance. Significance may be interpreted as the probability that samples with greater differences between their means would be drawn from a common population.

The same relationships between group means are seen if the groups are further subdivided by substitute versus regular instructor, as shown in Figure 3. Since the general relationship (i.e., students who saw the color transparencies had significantly higher scores than students who saw the black and white ones, who, in turn, had significantly higher scores than those who saw neither) holds for both regular and substitute instructor groups, the remaining analysis is based on compound data. (An analysis of variance on grade by treatment, none versus black and white versus color) and instructor (substitute versus regular) showed no interaction effect between instructor and treatment.

Clearly the color treatment had an effect, and we must reject the null hypothesis. To explore this effect further, we can look at the differences between the treatment groups on individual questions. These are summarized in Figure 4. Some interesting patterns may be seen. The questions in which the differences between black and white and color groups were the most significant were questions 1, 3, 8, and 10. All these deal with support software, and three of them deal with the operating system and its functions. Several of the transparencies (such as the one shown in Table 1) graphically depict the operating system in low main memory with other programs loaded above it. The color coding in the colored visuals helped distinguish it from
Table 2. Quiz Used as a Posttest

Fill in the blanks:

1. The practice of having the processor switch back and forth between execution of many programs, all of which are in main memory at the same time is known as ________________.

2. We give instructions or information to the operating system via ________________.

3. Basic data management functions are provided by ________________.

Multiple Choice:

4. All the following are examples of hardware except
   a. Printer
   b. Tape Drive
   c. Terminal
   d. Monitor

5. Which of the following are examples of support software?
   a. the COBOL compiler
   b. IU's payroll and personnel system
   c. RCA's inventory control programs
   d. a communications program

Matching:

6. _____ Keep(s) track of system usage statistics

7. _____ Often written by the user organization's programmer

8. _____ perform(s) often needed functions such as sorting and copying files

9. _____ the disk and tape drives

10. _____ handle(s) job management functions

Not for Grade: The classroom presentation on this material was (select one)

_____ Very helpful to me in understanding the material
_____ Somewhat helpful to me in understanding the material
_____ A little helpful to me in understanding the material
_____ Not at all helpful to me in understanding the material
<table>
<thead>
<tr>
<th>Treatment</th>
<th>n</th>
<th>Mean</th>
<th>S.D.</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>203</td>
<td>5.68</td>
<td>2.3</td>
<td>3.74*</td>
</tr>
<tr>
<td>B &amp; W</td>
<td>196</td>
<td>4.87</td>
<td>2.1</td>
<td>-3.04*</td>
</tr>
<tr>
<td>Control</td>
<td>221</td>
<td>4.25</td>
<td>2.1</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. *t* Test of Quiz Scores for Color, Black and White, and Control Groups

<table>
<thead>
<tr>
<th>Regular Instructor</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>n</td>
</tr>
<tr>
<td>Color</td>
<td>108</td>
</tr>
<tr>
<td>B &amp; W</td>
<td>91</td>
</tr>
<tr>
<td>None</td>
<td>74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Substitute Instructor</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>n</td>
</tr>
<tr>
<td>Color</td>
<td>95</td>
</tr>
<tr>
<td>B &amp; W</td>
<td>105</td>
</tr>
<tr>
<td>None</td>
<td>147</td>
</tr>
</tbody>
</table>

*p .001  
**p .01  
***p .05

Figure 3. *t* Test of Quiz Score for Color, Black and White, and Control Groups by Regular and Substitute Instructor
### Question

<table>
<thead>
<tr>
<th>Question</th>
<th>No Treatment Mean</th>
<th>Black &amp; White Transparencies Mean</th>
<th>Color Transparencies Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The practice of having the processor switch back and forth between execution of many programs, all of which are in main memory at one time is known as (multiprogramming)</td>
<td>.27 .42 .58</td>
<td>.27 .42 .58</td>
<td>.27 .42 .58</td>
</tr>
<tr>
<td>2. We give instructions or information to the operating system via (JCL)</td>
<td>.12 .18 .26</td>
<td>.12 .18 .26</td>
<td>.12 .18 .26</td>
</tr>
<tr>
<td>3. Basic data management functions are provided by (operating system)</td>
<td>.30 .38 .55</td>
<td>.30 .38 .55</td>
<td>.30 .38 .55</td>
</tr>
<tr>
<td>4. All the following are examples of hardware except: (a) printer, (b) tape drive, (c) terminal, (d) monitor.</td>
<td>.52 .58 .52</td>
<td>.52 .58 .52</td>
<td>.52 .58 .52</td>
</tr>
<tr>
<td>5. Which of the following are examples of support software? (a) the Cobol compiler, (b) IU's payroll and personnel system, (c) RCA's inventory control programs, (d) a communications program.</td>
<td>.28 .33 .42</td>
<td>.28 .33 .42</td>
<td>.28 .33 .42</td>
</tr>
<tr>
<td>6. Matching: Keep(s) track of system usage statistics: (the operating system).</td>
<td>.57 .64 .63</td>
<td>.57 .64 .63</td>
<td>.57 .64 .63</td>
</tr>
<tr>
<td>7. Matching: Often written by the user organization's programmers: (application software).</td>
<td>.51 .56 .66</td>
<td>.51 .56 .66</td>
<td>.51 .56 .66</td>
</tr>
<tr>
<td>8. Matching: Perform(s) often needed functions such as sorting and copying files: (service programs).</td>
<td>.39 .39 .55</td>
<td>.39 .39 .55</td>
<td>.39 .39 .55</td>
</tr>
<tr>
<td>9. Matching: The disk and tape drives: (hardware).</td>
<td>.86 .89 .89</td>
<td>.86 .89 .89</td>
<td>.86 .89 .89</td>
</tr>
<tr>
<td>10. Matching: Handle(s) job management functions: (the operating system).</td>
<td>.40 .52 .65</td>
<td>.40 .52 .65</td>
<td>.40 .52 .65</td>
</tr>
</tbody>
</table>

*Since all questions were coded 1/0 (1 = correct; 0 = wrong) the mean score indicates the fraction of the group which answered the question correctly.

Figure 4.
the other entities (hardware, other support programs, application programs) present in the diagram. Apparently the visual representation of relationships, highlighted with color coding, was helpful.

The results on question 4 reinforce this interpretation. This question also refers to the operating system (synonym: monitor), but here the difference was not very significant, and actually in the wrong direction. In the presentation, the word "monitor" never appeared on any of the visuals. The lecturer merely said that monitor is another term for operating system. Since the students did not get a visual image of the word "monitor" in either treatment, no significant effect would be expected, and none is evident.

Three of the questions (1, 2, 3) are fill-in-the-blanks, or memory recall questions. In general, recall is a harder task than mere recognition, such as is required to answer the multiple choice or matching questions (e.g., Wolford, 1971). And, as we would expect if color coding was an effective aid to learning the material, the color seems to have its greatest effect with the most difficult questions.

Three multiple choice questions on the material were placed on the departmental exam given to the students two weeks after the presentations. No significant differences were detected between treatment groups on these questions. The most likely explanation for this is that any residual differentiation after the students deliberately studied for the test was either non-existent or too small to be captured with these questions.

CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

This experiment offers evidence that the use of color in the presentation materials significantly affected the subjects' ability to recall the information presented. It is not clear, however, precisely why this is so. It could be that the color coding made the material itself more comprehensible as it was being presented to the subject—Salomon's supplanting function. Perhaps the color presentation aids were more easily read—this explanation is supported by LeCourier's Legibility Table (Daniels and Yates, 1971), where green, red, and blue on a white background all have a higher order of legibility than black on a white background. It could be that associating pieces of the material with color codes enhanced the subject's ability to store it into or retrieve it from memory, or both. Finally, it may be a combination of these effects, and perhaps the direct and combined effects of variables not yet identified.

This experiment should suggest to the managers or system developers considering the use of color computer output that they proceed with caution, especially in attempting to cost justify color equipment. In the experiment, color was shown to be useful for emphasizing distinctions and relationships in the information, but the effect, while statistically significant, did not have strong operational significance. This suggests that the use of color equipment in an MIS is likely to result in limited benefits (and without attention to the design and application, perhaps none) while having a definite cost.

Unfortunately, determining the cost/benefit ratio in such cases is very difficult. While the cost of obtaining color capability is usually clear, the benefits it may provide are not. Despite all the extravagant claims in the trade press, we usually just do not know ahead of time how effective color will be in a given information presentation. In large part this is because we do not have a very good theoretical basis for manipulating presentation mode variables—something that would suggest what media are appropriate for what tasks, for
example. Chute (1979) suggests that research should be aimed at the development of taxonomies of unique media effects on learners to allow instructional designers to predict with some assurance that the use of a particular medium will lead to specific learning outcomes. Similar taxonomies of presentation mode effects (including those of color) are needed to enable MIS designers to predict the outcome of using different presentation modes in information systems.

One line of research to be pursued, therefore, should be to refine our understanding of exactly how color is effective—under what circumstances, in what sort of task situations, in combination with what other variables. Will it make a difference, for example, if used with material which is readily understood by the subjects however it is presented? Questions such as this might be pursued by replication of experiments from cognitive psychology, introducing color as a new independent variable in the experiments. The better we understand the exact role of color, the more effectively we should be able to employ it.

A second area is perhaps of more interest to the MIS community. The experiment reported here used recall of information as its dependent variable. But what effect would color have on a person's ability to use the information, particularly to aid in making a decision. This is the justification for most color graphics equipment—to make information more readily usable by a decision maker. Yet we know of no empirical evidence in the literature that supports this contention. To pursue this, the experimentation must be moved into a decision making setting, most likely a management simulation. Again, this might be initiated by replicating experiments already done on presentation mode of information (e.g., Dickson, Senn, and Chernavny, 1977) with the introduction of color as a variable.

The use of color, with both alphanumeric and graphic representations of information, is inevitable in the near future. Color's intuitive attractiveness alone will prompt many organizations to at least experiment with color computer output. If MIS professionals are to make the best use of this new tool, however, they must possess more than an intuitive understanding of its effects. Based on this preliminary study, it appears that color can be effective, but that much research is needed to develop frameworks to guide its use.

REFERENCES

Eddy, D. "Application of Color Displays to Problems in the Undergraduate Laboratory," Behavior Research Methods and Instrumentation, Volume 11, Number 2, pp. 131-134.


