510-TM-2

DRD-64-131

APPARENT MOVEMENT PHENOMENA ON CRT DISPLAYS

FACILITY FORM 602

N65-36429	
(ACCESSION NUMBER)	(THRU)
76	1
(PAGES)	(CODE)
CR 67527	04
(NASA CR OR TMX OR AD NUMBER)	(CATEGORY)

GPO PRICE $ ______

CSFTI PRICE(S) $ ______

Hard copy (HC) 3.00

Microfiche (MF) .75

ff 653 July 65

PREPARED FOR:

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

WASHINGTON, D. C.

CONTRACT NO. NASw-954

DUNLAP AND ASSOCIATES, INC.

DARIEN, CONNECTICUT

Rpt 09344

510-TM-2
DRD-64-131

APPARENT MOVEMENT PHENOMENA ON CRT DISPLAYS

by:

H. M. Bowen
G. V. Guinness

J. G. Wohl, Program Director

October, 1964

Prepared for:

National Aeronautics and Space Administration
Washington, D. C.

Contract NASW-954

Dunlap and Associates, Inc.
Darien, Connecticut

ABSTRACT

36429

The appearance of movement can occur on CRT's which use a short persistence phosphor (e.g., the P-31 phosphor decays to 10% in 40 µs) and PRF rates between 20 and 50 cps. Analysis indicates that the apparent movement phenomena are due to movements of the eye; under suitable conditions, a stroboscopic interaction may be established between the succession of photic pulses and their impact at different locations on the retina. Symbols in an alphanumeric display can appear to "jump" or "dance," depending upon the extent and type of eye movements ("jumps" are analyzed to be associated with saccadic movements, and "dance" with eye tremor). The magnitude of the apparent movements vary from a threshold value, where they can hardly be seen, to an extreme state, where they are a predominant property of the display; variations in the magnitude of the effects depend primarily upon the space/time characteristics of writing and symbol brightness.

Informal observation of the phenomena suggest that, at their worst, they can render displays extremely difficult and wearing to observe. However, in less extreme form, the phenomena may be judged to have nuisance rather than detrimental value.

An experiment was conducted to estimate whether certain contemporary equipment which uses a P-31 fast phosphor and which evidences the phenomena in muted form is associated with lowered proficiency. Under controlled conditions, observers undertook data processing tasks on a fast-phosphor (P-31) and a slow-phosphor display. No definitive conclusions were made; however, a trend towards a higher error rate with the fast phosphor when the task was prolonged (and when the subject observers reported a feeling of fatigue) was noted. Also, a majority of observers preferred working with the P-12 phosphor.

We concluded that there is a high probability that when the apparent movement phenomena are present in sufficiently potent form, performance will tend to degrade, especially when the observer's task is long and arduous.

Author

TABLE OF CONTENTS

APPARENT MOVEMENT PHENOMENA ON CRT DISPLAYS

1. Discovery

During our study of display requirements for prelaunch checkout (conducted under subcontract to The RAND Corporation as part of NASA Contract NASr-21(08)), we observed the existence of certain apparent movement phenomena on CRT display units planned for use in Apollo prelaunch checkout. These phenomena were first observed by J. Wohl during installation and experimental use of a dd51 unit at the John F. Kennedy Space Center and later during factory testing of specific display units.

Our observations were not unique; as noted later in this report, the existence of the phenomena and a small-scale study of them were reported in 1960. Personnel at Data Display, Inc. (the designers and manufacturers of the units utilizing the CRT's) and operating personnel at the John F. Kennedy Space Center also observed the phenomena and at first attributed them to instabilities or noise in the equipment. However, it quickly became apparent that the major aspects of the phenomena were not due to any such faults; the locus of the phenomena is, rather, in the human eye and its interaction with certain temporal characteristics of the display.

The phenomena and their probable causal locus were first reported to NASA in February of 1964; in June, the Manned Spacecraft Center and the Office of Manned Space Flight, NASA, requested Dunlap and Associates, Inc., to study these phenomena and to conduct an experiment to investigate their effects upon operator performance. During the first quarter of the current contract, we have analyzed these visual effects in an effort to understand their underlying mechanisms, and we have conducted an experiment to evaluate their impact on performance.

2. The Nature of the Phenomena

a. Introduction

We believe it important to emphasize at the start of this report that we are not dealing with a single phenomenon. As will become apparent, there are a number of phenomena; we are not even certain of this number or of the most appropriate way of classifying them. Furthermore, each phenomenon occurs with varying degrees of magnitude or potency; it can vary from being barely noticeable to being the dominant property in a display in carefully arranged instances. In the latter case, all observers found viewing the phenomena very "trying," to the extent that it was unpleasant to continue looking at them.

We emphasize the plurality and varying potencies of the phenomena in order that the limitations of the controlled observations and experiments that are reported herein will not be misleading. In the future, various CRT displays will be used, and it will be desirable to design the total electronic package to avoid the appearance of any of the phenomena. Hence, we emphasize in this report the analytic understanding of the phenomena and have attempted to place the empirical observations within such a context.

The reader is cautioned that any evaluation of the phenomena raises difficult questions concerning criteria. The nature of the phenomena is to make the display appear spatially unstable; elements in it appear to jump or move.

The simplest criterion is the experiential one of, "Is there movement?" A variation of this is, "Will you (the operator) accept this degree of movement?" We would be asking for a report on the observer's perceptions or his reaction to his perceptions. For such reports to serve as a basis of policy, we, the scientists, must believe that the phenomena are noxious and will cause deterioration against some overt criterion related to the proficiency of the individual or the system. Such deterioration in proficiency need not necessarily come about by a direct causal chain, e.g., eye strain leading to reduced visibility of the stimuli. Indirect causal chains are possible, indeed probable. The fact of an unclear display may cause a degradation in attitude and motivation on the part of some observers which could lead to lowered standards of work and general antipathy towards the task. Hence we may believe that a display whose quality falls beneath the expectations of the user in terms of stability, clarity and legibility will engender, in the long run, a lowering of proficiency. On this basis, the task becomes one of ridding the display of all mal-qualities, at least to the point where the user is satisfied that he will suffer no ill effects. We should notice that the use of experiential criteria is very common, and that there is widespread belief in their validity. This fact in itself warrants paying attention to user reports and predisposes us to a solution which eradicates all display properties which the user may feel are either opposed to good performance or offend his sense of well-being.

The scientist within us may well rebel against placing trust in introspections and speculative causal chains. The behavioral scientist, especially, is committed to data which do not go beyond demonstrating some behavioral consequence of a situation. There will be no dispute, presumably, that controlled experimentation is the only sure way of determining the facts of a situation. However, in practice, experiments often are confined to fairly narrow aspects of a problem and, while illuminating one aspect, may not provide a basis for establishing a solution with respect to the whole problem. In theory, therefore, we would wish to rely on the criterion of behavioral consequence; in practice, however, we must acknowledge that our knowledge of behavioral consequence is often not sufficient to justify a general solution or the establishment of a general operational policy.

These considerations are most important, we believe, for placing the content of this report in perspective and for the wise use of the analyses, observations and experimental results.

b. The Particulars of the Phenomena

1) Flicker and the Perception of Change

It appears useful to separate general visual phenomena caused by intermittent light sources into two classes--those occurring above and those occurring below CFF (Critical Fusion Frequency).

An intermittent light occurring below CFF is seen to flicker. The human visual apparatus is able to distinguish light and dark periods, and it interprets them (completely or mainly) in terms of temporal change. The apparent simplicity of the situation is belied, however, by the fact that the observer perceives a light to flicker at or about 20 cps., regardless of the objective flicker rate. Some variation is perceptible, as demonstrated by the finding that most observers can discriminate three subjective flicker rates: fast, medium, and slow. In general there is a concensus that, where the attempt is being made to provide the observer with continuous information, a flickering display is to be avoided. Under the conditions which obtain on a typical fast-phosphor CRT display, very few flicker effects remain over 40 cps.; however, one of us (H. M. Bowen) has noted some lingering flicker effects on a display unit having a refresh rate of 50 cps. These effects occurred in the periphery of vision, which has higher CFF rates than has the center of vision.

The CFF point is determined by a multiplicity of factors which combine in complex ways. Primary factors are brightness, area, and duty cycle; the total range of CFF values extends from about 3 cps. to about 55 cps. In general, the CFF point is driven higher with increases in brightness and in area illuminated and with smaller light/dark ratios. It happens that the characteristics which are desired on alphanumeric CRT displays for prelaunch checkout use are exactly those conditions which are conducive to high CFF values; hence it becomes difficult to eradicate all traces of flicker.

Turning now to perceptual phenomena which occur at supra-CFF points; virtually all past observations and experiments have been concerned with the situation in which the temporal succession is associated with a physical change of location. The "moving" neon or electric bulb sign capitalizes on the fact that a space/time discontinuity gives rise to the appearance of movement. The literature establishes this phenomenon of apparent movement (the so-called "phi" phenomenon) and relates the various degrees and ways in which the phenomenon can be seen to the parameters of temporal succession, proximity of successive stimulations, size and brightness of the stimuli, attitudes of the observer, etc. A reading of the literature impresses one with the variety and complexity of the apparent movements that may be seen and serves to demonstrate that the conditions and expectations of the observer are almost as important as the parameters of the physical stimulus in determining exactly what is perceived.

In summary, some kind of movement may be perceived when two different parts of the retina are stimulated in succession--in the limiting case, by instantaneous succession. The human visual system must be capable, therefore, of differentiating space/time brightness discontinuities, even, as we shall see, very small ones.

The case with which we are concerned, however, is dissimilar to the classic case because there is no change in the physical position of the image (or only changes so small that their effect is considered to be negligible). Nevertheless, observers see various forms of apparent movement. The phenomena themselves must have been noticed at least by 1960, because in that year a report* appeared which stated that some CRT displays seemed spatially unstable; a small-scale experiment was conducted using a PRF rate of 60 cps. on a CRT, but no evidence of a degradation in reading speed or accuracy was found.

2) Description of the Phenomena

Observation of a variety of display formats on a particular equipment using a P-31 phosphor (the dd60) led one of us (H. M. Bowen) to categorize the phenomena into three main classes.

a) Jump

The term "jump" was introduced in an earlier report** for The RAND Corporation to describe the very rapid apparent displacement of display content when the viewer's eyes are moved. The symbols appear to jump and then return to their normal positions. The appearance of this jump is

*Crook, M., and Wade, E. Effect of periodic luminance reintensification on the reading of visual displays. Medford, Mass.: Tufts University, 1960.

**Pepler, R.D. and Wohl, J.G. Display requirements for prelaunch checkout of advanced space vehicles. The RAND Corporation, Santa Monica, Calif. RM-4200-NASA, June 1964, Appendix II, page II-4 and II-9. (Also available from Dunlap and Associates, Inc., Darien, Conn., Report No. 409-1.)

simple in the case where only one or very few symbols are present; the single symbol seems to jump from its normal position and then to return after a very short time interval. Sometimes more than one displaced symbol is seen so that one symbol is transiently converted into a string of replications of itself. When the display is composed of a number of lines of symbols, the phenomenon becomes perceptually more complex. The structure of the format exerts a restraining influence on the jump so that parts of the display seem to move while other parts do not. This phenomenon is difficult to describe, partly because of its unfamiliar nature, and partly because its particulars are different according to the direction, length and end-points of the (saccadic) movements of the viewer's eyes.

b) Half-Page Shift

This refers to the apparent relative motion that can occur between portions (in the specific instances observed, the top and bottom halves) of a display format. The phenomenon is dependent upon there being a pause between the writing of the top and bottom halves. In contrast to the jump phenomenon, the half-page shift occurs when the observer gazes steadily at the display without moving his eyes. The most pronounced effect occurs between the bottom line of the top half and the top line of the bottom half. The symbols in these lines appear to undertake a random "dance" with respect to one another. The movement is not confined to one line moving with respect to the other, although this is a predominant motion, and most of the movement is lateral rather than horizontal. When the observer's eyes are moved, an apparent movement of all (or of some large area) of one half-page with respect to the other can also occur. This movement is similar to the jump movement described above, but it takes place with respect to large areas and is seen as a definite break between the top half and the bottom half.

c) Veiling

This refers to the perception of shifting patches of filmy light moving over the tube face. The phenomenon is associated with a fairly dense format of symbols. We believe it is due to variations in total light flux in different areas of the tube face. It is most pronounced when one views the tube face steadily and is decreased when the eyes are moved around the display. The effect can merge into a more systematic sweep of light associated with the refreshing of the symbols. It can also merge with any flicker that may be perceived, most noticeably in the periphery of vision.

Added to these effects, a "visual beat" can occur due to the interaction of a single, bare, 60-cycle fluorescent light used for ambient illumination with the refresh rate (e.g., 50 cps.) on the CRT. This effect can easily be eradicated by the provision of a number of fluorescents of differing phase and/or by using fluorescents whose light emission is relatively continuous.

3) Analysis of the Phenomena

a) The Jump Phenomenon

The jump phenomenon occurs only when the observer's eyes are in motion. When scanning freely (as opposed to tracking a moving target), the eyes move in a series of sudden motions, called saccades, which are interspersed with fixations.

Observation of the jump phenomenon shows that the direction of the jump seems to be opposite to the direction of motion of the eye. Hence, the basis of the phenomenon must be that successive light pulses strike the rotating retina at different locations. For example, if the eyes should move from right to left, a stimulus would be seen first in the left visual field and then in the right visual field, and its apparent movement would be from left to right (i.e., in the direction opposite to the eye movement). While there can be little doubt that this is the basic mechanism of the phenomenon, our confidence in its correctness is strengthened by the observation that the jump appears more pronounced when the saccade is made over large angles rather than small angles. The relevance of this observation is that peak angular velocities increase as saccade angle increases. Sample figures are:

Saccade Extent	Peak Velocity*
15°	320° per second
30°	540° per second
60°	680° per second
90°	720° per second

Hence it is to be expected that the amplitude of the apparent jump should be correlated with the angular distance the eye moves in the time elapsing between successive pulses.

It also seemed (at least during our observations) that when the proper conditions were in effect, the jump phenomenon always occurred during long saccades but sometimes did not occur during short saccades. This is

*Hyde, J. Some characteristics of voluntary human ocular movements in the horizontal plane. J. Amer. Opthal., 1959, 48, 85-94.

consistent with the finding that the length (in angular distance) of a saccade is correlated with the time duration of the saccade. Average figures are:*

Saccade Extent	Time Duration
2°	15 ms
5°	29 ms
10°	39 ms
15°	48 ms
20°	55 ms
30°	80 ms
40°	100 ms

Thus, when the refresh interval is, say, 25 ms, there will be a fair number of saccades (e.g., those used in reading test, which may average out at about 3° or 4° in extent) which will be completed in less time than the refresh interval.

To summarize, the jump phenomenon is thought to be due to successive pulses falling on different points of the retina; the magnitude of the phenomenon is related to the velocity of the angular rotation of the eye, which is highest for long movements; and the probability of the phenomenon's occurring is related to the duration of the eye movement, which is longest for long movements.

For the jump phenomenon to be seen, it is necessary for the eye to see while in motion. In spite of a mistaken notion, still widespread, that the eye goes blind during saccades (the apparent reason for our not seeing a blur between fixations while scanning), the fact is that the eye does see during saccades, but not very well. In terms of brightness threshold, to be seen during a saccade, a stimulus may have to be brightened by a factor of 10 (as compared to brightness required when the eye is still). Hence, in general, we can expect the jump phenomenon to be perceived most strongly when the stimulus is bright and when the stimulus/background contrast is high. Observation of different brightness conditions on the dd60 equipment seemed to bear out this expectation.

The format of the visual stimulus also seems to exert some influence on the appearance of the phenomenon. In general, our observations led us to the conclusion that a simple element appearing in isolation has the

*Woodworth, R.S. and Schlosberg, H. Experimental Psychology (rev. ed.). Henry Holt and Company, New York, 1954.

greatest tendency to jump and that large blocks of densely packed symbols have less tendency to jump. There seemed to be a competition in the latter case for the symbols to jump on the one hand and to remain still and cohere with the rest of the material on the other hand. However, on the whole, we felt that the factors tending to cause some appearance of jump could easily overwhelm the factors predisposing to stability, so that the format variables took on major significance only when the jump conditions were weak.

Of all the variables which enter the phenomenon, the strongest is undoubtedly the distribution of light over time. For the conditions of display on the dd60 equipment, jump could not be seen when refresh intervals were at or about 15 ms; with dense format, 20 ms was the minimum interval. Jump became very obvious at 25-ms intervals and continued to grow worse up to and through the CFF point. It is not clear at this time what factors are responsible for setting the threshold points. The P-31 phosphor decays to a 10% value in 40 μs and thus is short enough, one might think, for jump to be seen at intervals shorter than 14 or 20 ms. Perhaps the explanation lies in the fact that the angular extent of movement is small (2° to 4°) in these time intervals and that the eye has sufficient spatial integrating capability to avoid seeing the extents as jumps. This explanation is not very plausible when it is recalled that the visual acuity of the still eye is a few seconds of arc and the threshold for movement (in central regard) is one to two minutes of arc per second. Further analysis and experimentation is required before an adequate understanding of the phenomenon can be achieved.

b) The Half-Page Shift Phenomenon

The nature of the half-page shift phenomenon has already been described. Viewing the characters through a microscope (on the dd60 equipment) satisfied us that the apparent movement was not due to the objective movement of the symbols. The amplitude of such movement was not greater than about 1 or 2 mils; and in any event, the objective movement within a given half-page was constant, and no similar movement phenomena were seen elsewhere in that half page.

Of the various observations we made, the following are the most indicative of the conditions of the phenomenon:

- The phenomenon persisted when the display was reduced to one character in the top half and one character in the bottom half; time interval between painting was about 10 ms.

- The phenomenon grew less noticeable as the top and bottom lines were separated and was absent when they were an inch apart.

- The apparent movements did not follow any set pattern and were characterized by a large degree of randomness in time and space.

While, in general, the phenomenon seems to be due to the fact of succession, which the eye interprets as movement, the details of the phenomenon do not conform to any published accounts which we have read. For one thing, the time interval is much shorter than the time interval previously reported to be conducive to the perception of apparent movement. Bartley* reports the optimum interval to be about 60 ms. However, neither he nor other workers has reported the threshold conditions for apparent movement. Some reports would lead one to believe that movement can be seen when stimulus A is turned off and at the same point in time stimulus B is turned on and when A and B are at separate but adjacent points. If one accepts this possibility, then apparent movement will typically be seen on a P-31 phosphor when the writing interval is greater than about 40 μs (the decay time to 10% of brightness) and when the brightness and juxtapositioning of the adjacent stimuli are conducive to the perception of movement.

The whole story, however, must be more complex than this, for it takes more than 100 μs to write each line; and, it is recalled, no movement phenomena comparable to the half-page shift occur within a half page. In other words, there must be some time interval value between about 0.1 ms and 7.0 ms (the shortest time interval at which we have observed the half-page shift) which brings about the half-page shift phenomenon.

At this time we believe the basis of the effect is due to an interaction between time/energy characteristics of the illuminated characters and normal eye tremor. In a laboratory demonstration it was possible to observe two small neon lights appear to jump apart and then return at certain combinations of flash rates and phase relationships, but only during saccadic eye movements.

---

* Bartley, S. H. The Principles of Perception. Harper and Row, New York, 1958.

3. Discussion

The crux of the problem can be thought of in terms of the relative amounts of photic integration that are performed in the phosphor and in the eye. When the phosphor decay persists long enough so that only a small decrement in brightness occurs before the next refreshing signal, no movement phenomena occur. When the brightness decay is large, and when time intervals are suitable, the apparent movement phenomena appear.

Selecting a longer persistence phosphor will eradicate the movement phenomena at the cost of some decrease in display brightness and some "smear" which occurs when the display is updated. Other possible solutions, in terms of extending the character writing time and/or of increasing the PRF rate, carry the penalties of reducing the amount of information that may be written.

Some discussion of these trade-offs is contained in a letter report from General Electric (Apollo Support Dept.) to NASA (Mr. M.E. Dell, PP-8), of 10 August 1964. This report also notes that on the dd74H production model the refresh rate has been increased to 47-1/2 cps. which has had the effect of minimizing the spatial instabilities. However, the desirability of using a larger-persistence phosphor was not discounted. A more complete discussion is contained in the RAND report previously mentioned.*

Thus, the particular problem which initiated this study has been partly resolved by the fortunate circumstance that later versions of the display/computer units have used higher PRF rates. Apparently a refresh interval of 21 ms is short enough, combined with the persistence of the phosphor, to avoid a dark period between successive writings. However this current situation does not resolve the more general problem of specifying the boundaries within which the phenomena can occur; this topic will be investigated in a laboratory experiment to be conducted at Dunlap and Associates, Inc.

During the time that these observations and analyses were being made, it was decided that an experiment should be conducted to determine whether any behavioral deterioration occurred associated with the visual phenomena. Specifically, the experiment was planned to determine whether displays having properties as close as possible to those appearing on the dd74G prototype would cause a slowing down or an increased error rate in reading data off the displays. This experiment is reported briefly in the following section, with a full account provided in Appendix I. A final discussion is postponed until after the description of the experiment.

*op. cit., Appendix II, page II-9.

4. Experiment

The following is a summary account of the experiment; a detailed description is presented in Appendix I.

a. Purpose

The purpose of the experiment was to determine whether a close approximation to the display conditions existing on the dd74G operational prototype equipment would cause performance impairment compared to a control condition.

b. Conditions

The experiment was conducted at Data Display, Inc., Minneapolis (the display manufacturer), and used their dd60 equipment. The equipment was modified so that two different CRT's were driven by the same symbol generation program. One CRT used the P-31 "fast" phosphor, the same as that used on the dd74G; the other CRT used a P-12 phosphor. The P-12 is a relatively slow phosphor; for the electronic conditions which obtained, its persistence was long enough to give the display an appearance of complete stability. Both CRT's were driven by identical electronic parameters which were made as close as possible to the electronic parameters existing on the dd74G (see Appendix I).

The task which the subjects had to perform consisted of answering questions concerning data which was written on the CRT's. This data consisted of 26 lines of alphanumeric material, actually nonsensical but having a face resemblance to technological data printouts.

In order to do the task, the subject had to read a question from a sheet mounted on the left of the CRT, scan the data on the CRT, and, on some question, refer to a data card above the CRT. This arrangement ensured a large amount of eye movement. Speed and error scores were derived from performance on this test.

While the subject was performing this primary task, he had to attend also to a simulated message stream over a headphone. The message stream consisted of numbers interspersed with names. Whenever the subject heard a name, he had to write it down. Error scores were derived from this task, and it was hoped that if the P-31 CRT absorbed more of the attention capacity of the subject than the P-12 CRT, more errors would be committed on the secondary task when the P-31 CRT was used.

c. Experiment I

In Experiment I, eight subjects were used. The subjects were male and represented approximately in age and general experience the type of man who uses the operational equipment. Each subject did four 25-minute tests on each CRT display; the sequence was counterbalanced to spread learning and sequence effects as evenly as possible.

The results of this experiment showed that performance on the two CRT's was essentially the same, thus corroborating the results of Crook and Wade.* The data is summarized in the following table:

Primary Task				Secondary Task	
Items attempted per minute		Errors expressed as % of items attempted		Errors per 100 signals	
P-31	P-12	P-31	P-12	P-31	P-12
** 2.96	2.87	6.63	6.76	1.80	1.98

**Average (of 8 subjects)

On being asked which CRT they preferred, five subjects chose the P-12, two the P-31, and one had no preference.

d. Experiment II

In an effort to make the task more onerous on the subjects and to make any negative properties of the displays have a pronounced effect, a second experiment was conducted in which four subjects (drawn from the original eight) undertook two 1-1/2-hour tests, one on the P-12 CRT, and one on the P-31 CRT. The results are summarized below.

*op. cit.

Primary Task				Secondary Task	
Items attempted per minute		Errors expressed as % of items attempted		Errors per 100 signals	
P-31	P-12	P-31	P-12	P-31	P-12
* 3.74	3.80	4.58	3.43	1.42	1.00

*Average (of 4 subjects)

While speed of performance remains about equal for the two displays, there is some suggestion that errors, although reduced over-all due to practice, tended to be greater on the P-31 CRT.

e. Statistical Confidence

None of the pairs of scores relating to the P-31 and P-12 CRT's are statistically different from one another, taken singly or in groups. Statistical delving (see Appendix I) indicates that there is some difference between the displays but that this difference is tied to interaction terms between the subjects, the sequences of viewing the two CRT's, and the displays. These second-order effects are not large and do not allow any confident assertion that one display is superior to the other.

f. Discussion

The interpretation of this experiment must be related to the physical conditions of the displays. The PRF was 51 cps., which was fast enough to obviate on the P-31 CRT virtually all traces of flicker and to reduce the jump phenomenon to minor proportions. The transfer time from the top half to the bottom half of the display was 7.4 ms (compared to 15.0 ms on the dd74G), and this was short enough to diminish the "shift" phenomenon, although it was clearly present. Also, brightness was reduced on the experimental display compared to the operational one, and increased brightness will tend to accentuate the phenomenon. It happened, then, that in simulating approximately the conditions on the dd74G, the "jump" and "shift" phenomena were considerably reduced compared to their maximum condition. Furthermore, the reduction in display brightness and in transfer time from upper to lower half in the experimental condition compared to the operational condition on the dd74G served to attenuate the phenomena even further.

These considerations, combined with the findings of the experiment (which, in summary, suggest that the conditions obtaining on the P-31 CRT may engender deteriorating performance in some subjects when the task is onerous), lead to the conclusion that:

- The P-31 phosphor will probably be satisfactory, provided that electronic conditions are chosen which minimize or eradicate the anomalous phenomena.

- When the P-31 phosphor is driven by electronic parameters which cause the anomalous phenomenon to be present, there is likely to be some performance deterioration, especially when the task period is long.

5. Conclusions

The anomalous phenomena giving rise to the perception of apparent movement on fast-phosphor CRT's are attributed to stroboscopic effects. In the course of moving the eye, successive impulses strike the retina at different places. The differentiating power of the eye in the general case is attested to by previous work on the perception of apparent movement, and we may regard the present phenomena as special variations of the general case.

In assessing the effect of these phenomena on proficiency of performance, we assert that, in general, a display which has components which are irrelevant to the task to be accomplished tends to produce worse performance than a display without such distractions. This assertion is substantiated by a large literature in display design. In particular, a display which has rapid transient terms or appears to move tends to be peculiarly distracting because of the potency of motion to capture attention. Hence we conclude that there is a strong *a priori* case to be made for ridding displays of the phenomena and for believing that if the phenomena are present in sufficiently potent forms, proficiency will suffer.

The results of the experiment are disappointing in that they do not confirm unmistakably this position. However, if we assume that the phenomena were present at a reduced potency level, the experimental results do not deny our interpretation. On a short test (Experiment I), there was no difference between a display with the phenomena and one without. However, when the task burden was increased (Experiment II), the results began to form a pattern in favor of the display without the phenomena. This interpretation is bolstered, in our view, by the observations that we and others have made of the phenomena at much increased potency levels. These displays were deplored by all and gave one the sense of being physically difficult to regard.

Therefore, we conclude that:

- These phenomena, when present in sufficiently potent form, will degrade performance on long and difficult tasks.
- The boundary conditions for the appearance of the phenomena should be investigated and established.
- Specifications should be derived from the boundary conditions, which establish the electronic parameters which obviate the appearance of the phenomena.

APPENDIX I

DETAILED DESCRIPTION OF THE EXPERIMENT

APPENDIX I. DETAILED DESCRIPTION OF THE EXPERIMENT

1. General

This study was an experimental investigation of operator performance using P-31 and P-12 phosphors with the dd60 display unit simulating a dd64G display unit. The experiment was conducted at the plant of Data Display, Inc., Minneapolis, Minnesota.

The task and the environmental conditions were set up to be as similar as possible to those encountered in the operational situation.

The purpose of the experiment was to compare the speed and accuracy of reading data on a P-31 phosphor CRT and a P-12 phosphor CRT. It was hypothesized that because the P-31 CRT was associated with the apparent movement phenomena (discussed in the text) while the P-12 CRT produced a stable display, performance would be worse on the P-31 CRT than on the P-12 CRT. The experimental conditions were made to resemble as closely as possible the conditions obtaining on the operational dd64G equipment so that the results would be directly relevant to contemporary field conditions.

2. Method

a. Task Description

The total task was composed of a primary and a secondary task. The primary task required the subject to read questions from a test card placed adjacent to the CRT display, scan information presented in the CRT scope format and on a permanent placard directly above the scope, locate and interpret the information required to answer each question, record each answer on an answer sheet located on the display console shelf.

Each test card contained twenty typewritten questions which the subject answered in order. Four test cards (eighty questions) were prepared for each scope format. A scope format contained twenty-six lines of alphanumeric information displayed in the center 6" x 6" area of the scope. The number and organization of the characters were selected to replicate the essential properties of a scope format during operational checkout. Six formats, all of similar design, were used in the experiments. The permanent placard contained seven lines of alphanumeric information, with content and organization similar to that of the scope formats. The same placard was used with all formats and test cards. Figure 1 shows the placard and representative sections of a test card and a scope format.

	Address	Sign	Code	Sender	Message
A	9P8320	-	19	JAV	RR STRBD. ROCK. EMPT.
B	53LZC	-	8		KL JAW EXPRESSWAY 953
C	629N10		025	TIS	ESC. VELOC 25 KM
D	32532Q	L	5	KO	ALEN FIN. LITE OK
E	JU8185	+	21	BIM	EEG BETA WGHT 24Z WGHT 21
F	59LP219	=	0	ARW	SHARD FREQ 90 KC
G	65210	-	AGC		SUPER HET. SINK SIGNAL -

PLACARD

12. Write Sender of Signal 6
13. Is the sign appearing in Placard Signal C the same as the sign appearing in Signal 5?
14. Which signal refers to BOATED?
15. Write out the part of the signal which comes before... ... B > 9I
16. Add the digits (single numbers) appearing last in the addresses of Signals 20, 9, and 15

TEST CARD

93A661	+	3	GO	WIRE BLOOD RATE	66LB	
91320B		9	ING	BOATED EFFECT	633 FT	
22321B	+	3	TUB	CREWDOW	CHGE +	63G
23322B	+	6	TH	FIL. FEEDTHRU	9 KV AT 1	
33333B		4	EBE	CONTROLLER	GRIP	6 LB
42C865	+	0	ST	PRINK ALARM	02.00	HR
89D625	-	19	WU	MERCURY FALLOUT	1%	K

FORMAT

Figure 1. Placard and Sample Sections of a Test Card and Scope Format

As shown in Figure 1, task material varied from meaningless alphanumeric groups to characters to words and statements having a contextual face validity. The questions were designed to require large saccadic eye movements between test card, placard, and format and both large and small movements within the format. Also, questions ranged in difficulty from those requiring simple transcription, e.g., Question #12, to those requiring elemental quantitative skills, e.g., Question #16. A complete format and one associated test card is shown in Appendix II.

During testing the test cards were placed adjacent to the scope in use and positioned to mask the unused scope. Figure 2 shows the arrangement of test card and placard for testing with the right (P-31) scope. The scope displays a typical format.

The secondary task, which was performed concurrently with the first task, required the subject to monitor audio communications continuously. The taped audio program presented an input to the subject once every ten seconds. Inputs consisted of one digit or one of twenty common male first names, each having a unique first letter. The subject was instructed to record on his answer sheet the first letter of any name he detected. On the average, two names were presented per minute.

Considered in combination, these tasks presented the subject with a total task comparable in critical behavioral characteristics to the actual task as described by operational personnel. The primary task required the appropriate scanning, reading, cognitive and transcribing behavior, while the secondary task demanded concurrent attentiveness to critical audio inputs.

Both experiments used this combination of tasks, with the major differences being task duration and the work-rest cycle. These specific differences are discussed in Section e, Procedures.

b. Equipment Description and Layout

The study was conducted in a 13' x 10' area of the equipment checkout room at DDI. The area was curtained off from the remainder of the room, where normal activity was going on. Subjects could not see this activity, but could hear the noise it produced (76 db $\pm$ 4 db). The room was air-conditioned, with temperature held at 68-72^{o}F. Ambient illumination, provided by overhead fluorescent lighting, produced five foot-candles incident to the plane of the scope face.

A digital computer (Control Data 160) located adjacent to the experimental area was used to program a dd60 display unit. Seven scope formats were

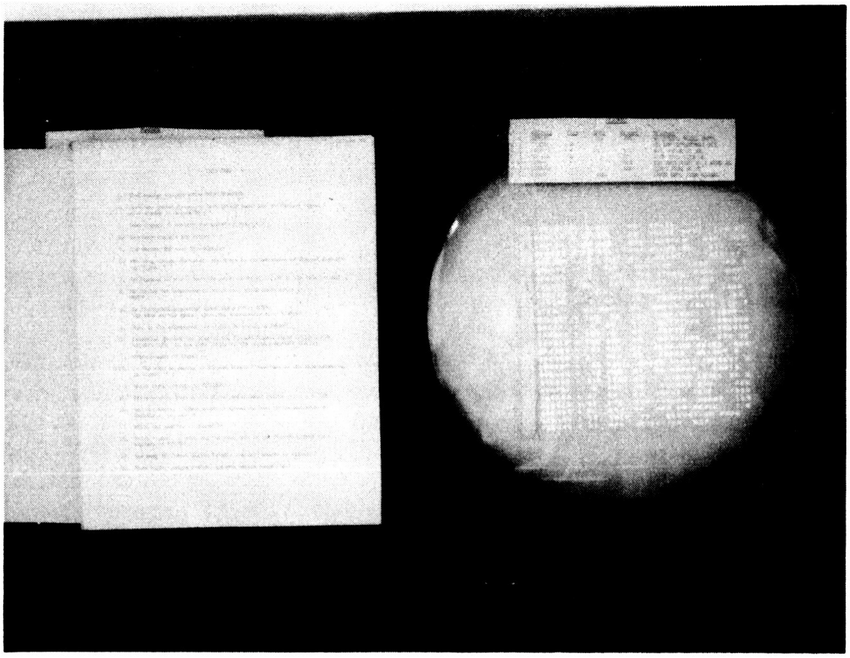


Figure 2. Arrangement of Test Materials

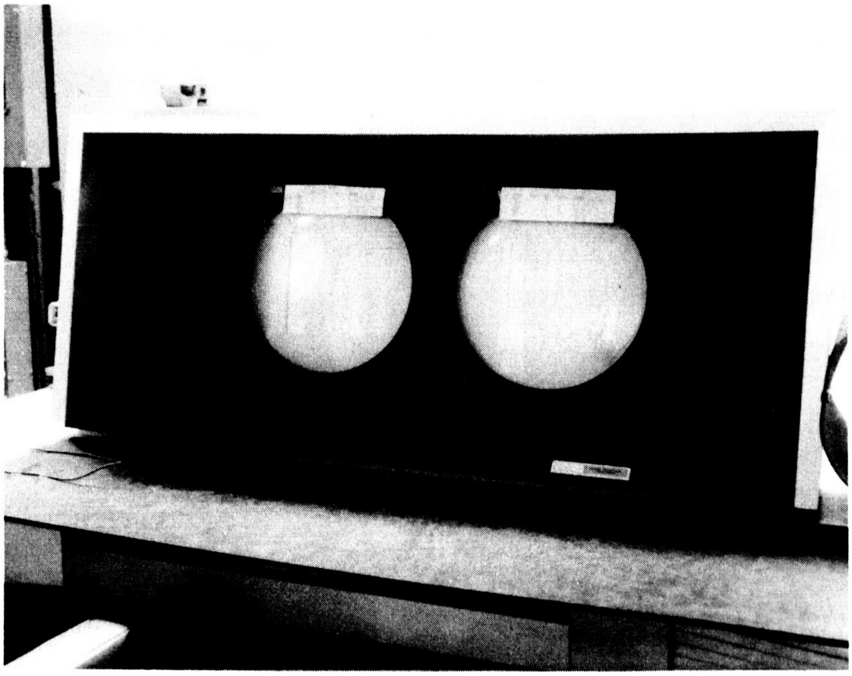


Figure 3. dd60 Display Unit with Placard and Typical Scope Format

programmed on the computer -- six formats for testing, one format for practice. Format selection was controlled by the experimenter using selection switches under the writing shelf of the display unit. The unit consisted of two 12-inch CRT's, as shown in Figure 3. The permanent placard is shown above each scope, and the scopes display a typical format. The left and right CRT's were coated with the P-12 and P-31 phosphors, respectively. Critical dimensions of the display unit were:

Scope height (floor to center)	39"
Scope tilt (back from vertical)	10^{o}
Console shelf depth	12"
Scope depth (shelf front to scope center)	18"

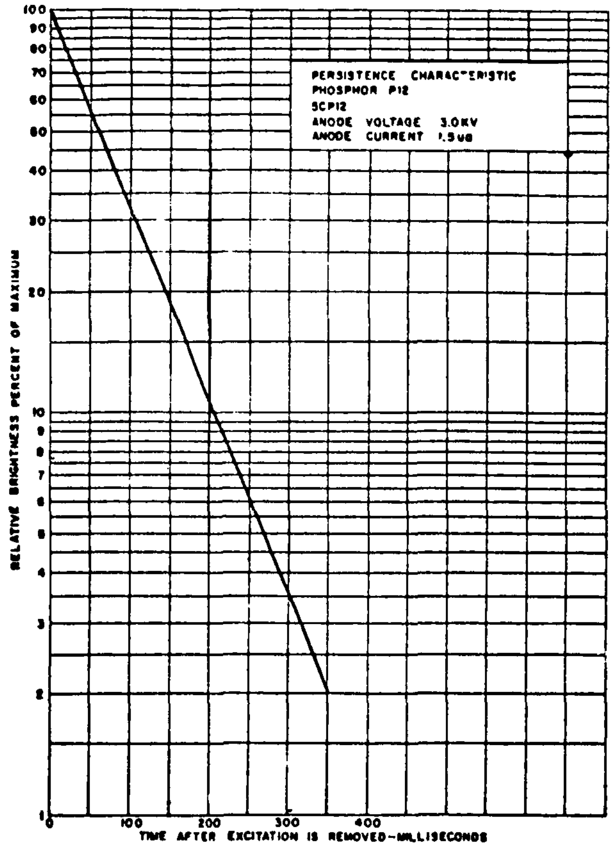
The spectral energy and persistence characteristics of the phosphors studied are shown in Figure 4. As indicated, the persistence is much shorter for the P-31 phosphor -- 10% of maximum after .04 ms, vs. 10% of maximum after 200 ms for the P-12. Also, the spectral energy distributions of each phosphor show considerable difference. The P-12 (orange) phosphor has maximum energy at 5250 angstroms, while the P-31 (green) phosphor has maximum energy at 5875 angstroms. This inherent difference in color did not affect the brightness match of the phosphors, since the energy required to produce equal apparent brightness is essentially the same for both spectral distributions.

Electronic parameters of the scope displays were selected to reproduce the essential characteristics of the dd74G presently in operational use. Because of limitations inherent in the computer simulation and the unavailability of a dd51 or dd74 unit, complete replication was not possible.

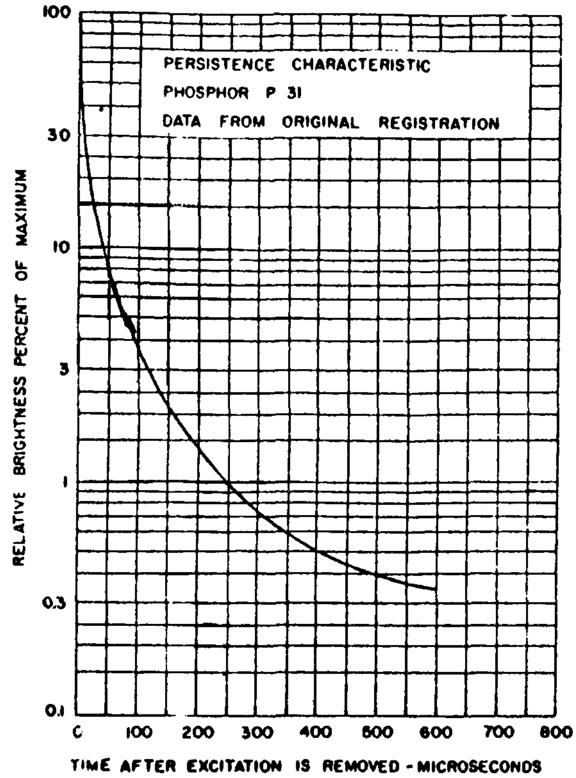
The major electronic characteristics of the dd60 and dd74G are compared below:

	dd74G	dd60
Pulse repetition rate	19.7 ms	19.7 ms
Character time	4.8 μs	7.2 - 7.4 μs
Flyback time	.8 μs	.8 μs
Half page print time (lines 1-13 or 14-26)	2.5 ms	6.1 ms
Transfer time - upper half to lower half	15.0 ms	7.4 ms

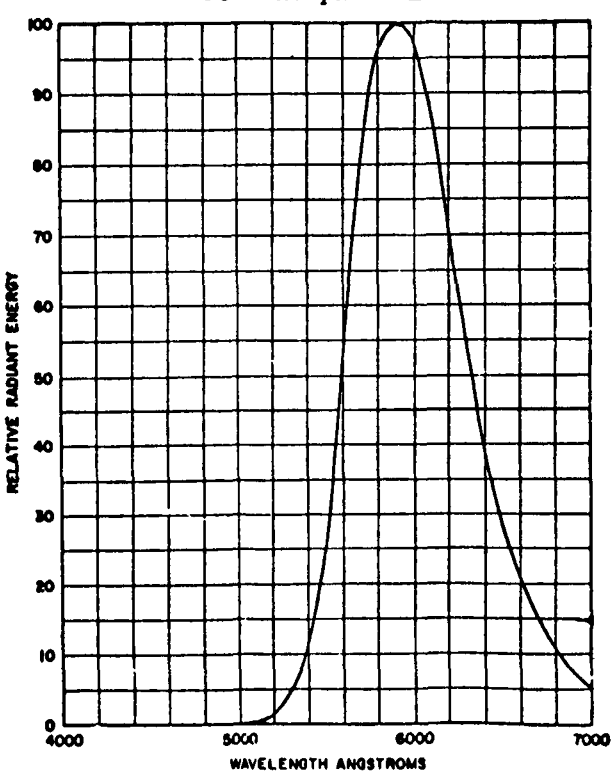
Persistence Characteristic - P12



Persistence Characteristic - P31



Spectral Energy Distribution For Phosphor P12



Spectral Energy Distribution for Phosphor P31

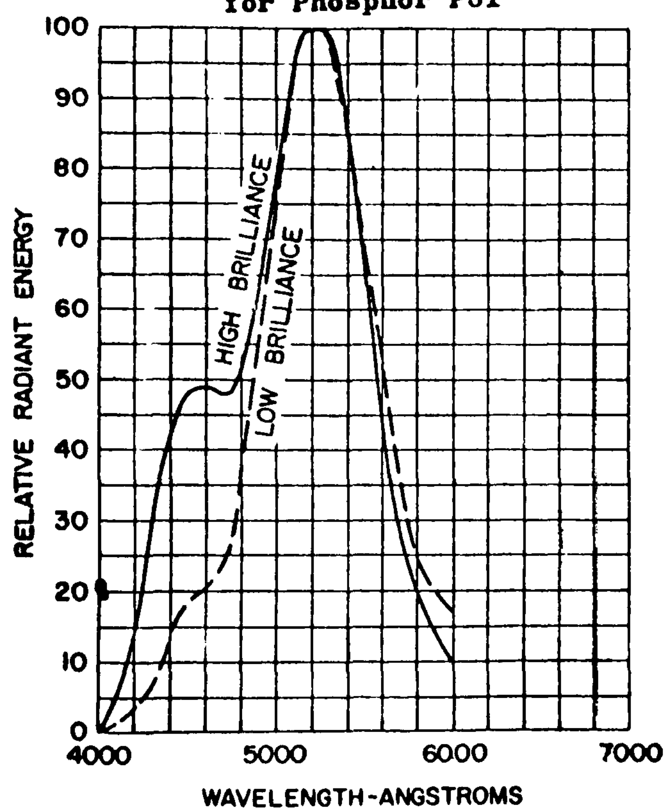


Figure 4. Persistence Characteristics and Spectral Energy Distributions for P-31 and P-12 Phosphors

In terms of reproduction of the apparent movement phenomena, the experimental display was thrğought to be somewhat deficient. In particular, the jump phenomenon occurred very weakly; and the shift phenomenon, while being quite unmistakeable, was not very pronounced. Because the purpose of the experiment was to evaluate the P-31 CRT display as used in the field in the dd74G equipment and not to evaluate the apparent movement phenomena per se, it was felt that no changes should be made in the direction of amplifying the apparent movement phenomena.

As shown, the major differences between each display unit are the half page print time and transfer time. If our analysis of the apparent movement phenomena is valid (see main text), then the shorter transfer time on the dd60 would result in a relative attenuation of the half page shift phenomenon when using this display. Our opinion was that the shift does not occur as frequently on the dd60 display unit. The practical consequences of this unavoidable disparity between operational and experimental conditions was to increase the probability that no performance differences would occur as a function of scope phosphor.

Illumination variables were precisely calibrated and monitored throughout the study. As indicated above, ambient illumination was 5 foot-candles at the scope; background brightness at both scopes was 3 foot-lamberts; spot brightness of a specific area was maintained at 10 foot-lamberts with daily maximum adjustment of 8%. Between-scope differences were minimal -- less than 5%. Some small change in brightness for other areas of the scopes could be expected. However, area by area, the scopes had identical brightness $\pm$ 5%.

It should be noted that ambient illumination under operational conditions (using a P-31 phosphor on a dd 54 display unit) was measured at 40 foot-lamberts (Complex 37 blockhouse). Spot brightness of the tube was not measured. However, if we assume contrast ratio between ambient illumination and spot brightness was approximately the same for the operational and experimental situations, then our best estimate of spot brightness for the operational conditions under which the apparent movement phenomena were first observed is 80 foot-lamberts. Thus, it is almost certain that spot brightness was about one order of magnitude greater under operational than under experimental conditions. This difference was a necessary consequence of the need to match brightnesses of the P-12 (which was seen at maximum brightness) and P-31 phosphors for valid comparative study. In another sense, however, the difference between P-31 operational and experimental spot brightness may have been unfortunate. If the analysis of the apparent movement phenomena (see main text) is valid, then reduction in spot brightness served to attenuate the phenomena under study; and this, as with the differences in the timing parameter, serves to increase the probability that no performance differences would occur as a function of scope phosphor.

In addition to the display unit, equipment within the enclosure included a console-type adjustable chair for the subject, a work table and chair for the experimenter, and a Revere table-model tape recorder used as the auditory stimulus generator for the secondary communications task. A plan view sketch of the equipment layout is shown in Figure 5.

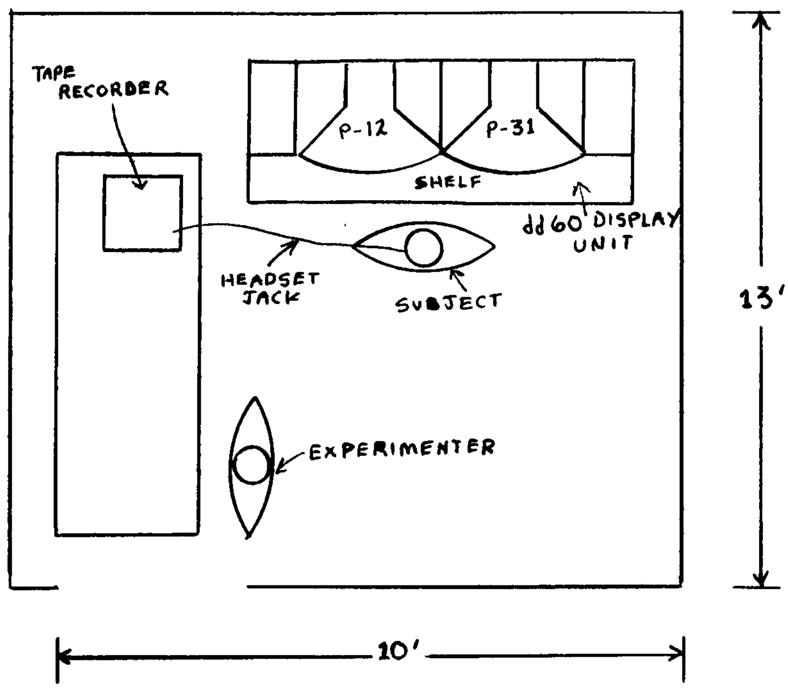


Figure 5. Equipment Arrangement

c. <u>Subjects</u>

Eight subjects were used in the first experiment. Four of these subjects were also used in the second experiment. All subjects were male college graduates recruited through contact with local public schools and graduate schools of colleges in the area. Several subjects had advanced degrees or were matriculants for advanced degrees. Six subjects were public school teachers; two subjects were ministers. The average age of the eight was 32 years, with a range from 27 to 43 years. Five subjects had normal vision; three subjects had corrected normal vision.

Subjects were paid $2.25 per hour plus a bonus of $4.00 for successful completion of each experiment. Average payout per subject was $60.00 for

Experiment I and $25.00 for Experiment II. All subjects were punctual and cooperative, exhibited high motivation, and successfully completed the program. In general, their demonstrated ability in performing the task, in combination with their educational and professional backgrounds, indicated their output after training was comparable to the output and general work style of operational personnel.

d. Experimental Design

In both experiments, subjects were run as their own controls in a counterbalanced design. Repeated measurements were taken in Experiment I across the scope conditions. In Experiment II each subject was given one extended trial on each scope condition with the order of conditions counterbalanced across subjects. Other variables counterbalanced in both experiments were: format order, test card order, tape order (the communications task), and time of testing. The design, as applied to two subjects in Experiment I, is shown below.

Table 1. Experiment I Design (Two Representative Subjects)

Subject	Trial	Test	Time of Test	Scope	Format	Test Card Order	Tape No.
	1	1	9:30a	P-31	2	ABCD	1
	1	2	10:00a	P-31	3	ABCD	2
	2	1	1:00p	P-12	4	ABCD	2
1	2	2	1:30p	P-12	5	ABCD	1
	3	1	9:30a	P-12	6	ABCD	1
	3	2	10:00a	P-12	7	ABCD	2
	4	1	1:00p	P-31	3	DCBA	2
	4	2	1:30p	P-31	2	DCBA	1
	1	1	9:30a	P-12	5	DCBA	2
	1	2	10:00a	P-12	6	DCBA	1
	2	1	1:00p	P-31	7	DCBA	1
3	2	2	1:30p	P-31	2	DCBA	2
	3	1	9:30a	P-31	3	DCBA	2
	3	2	10:00a	P-31	4	DCBA	1
	4	1	1:00p	P-12	6	ABCD	1
	4	2	1:30p	P-12	5	ABCD	2

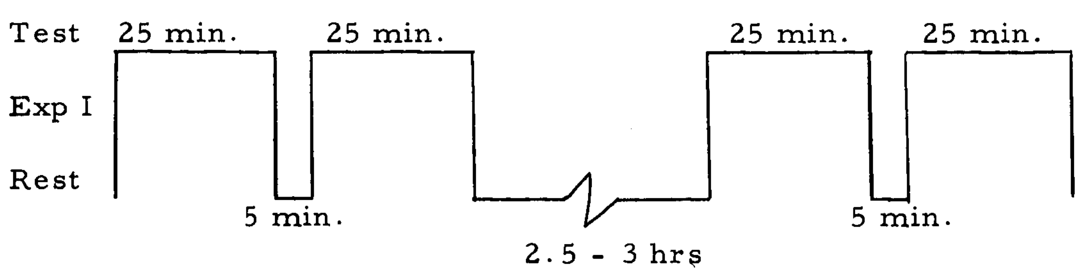
As indicated in Table 1, testing was completed for each subject in two consecutive days. With four trials of two tests per trial, it was necessary to repeat two of the six scope formats. The two formats repeated were varied systematically across the eight subjects -- they were always the first two formats used and were presented the second time in reverse order.

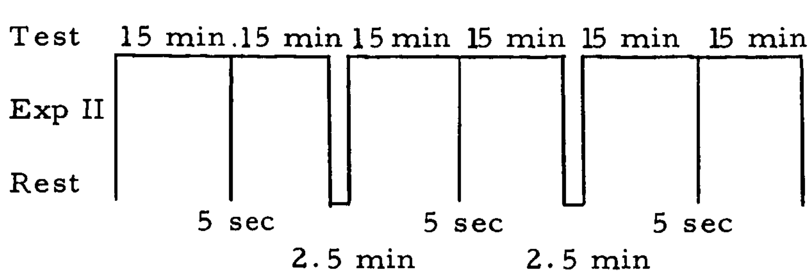
The same basic design was used in Experiment II with one exception -- one extended trial per scope was used in place of repeated measurements. Thus, order of scope presentation was counterbalanced across subjects and not across and within subjects, as permitted in Experiment I.

e. Procedure

Prior to experimentation, all subjects were given a three-hour group orientation to the study. In addition to discussion of administrative, scheduling, and procedural problems, the subjects were told a purpose for the study and were each given a 15-minute familiarization trial on the task. The stated purpose of the study was intended to be deceptive. Subjects were told that the scopes varied in their engineering design such that one (not specified) was cheaper to produce while the other was easier to maintain and that their performance, in combination with cost considerations, would determine which scope would be manufactured. Post-test interviews indicated all subjects operated under this assumed purpose.

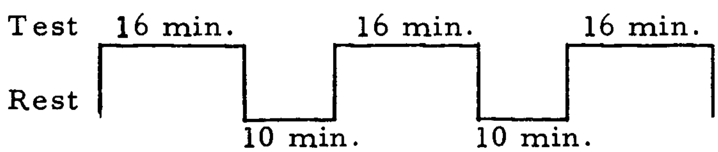
In Experiment I, each subject was given four trials (two per day), with each trial consisting of two 25-minute tests. The rest period between trials was several hours; between tests it was 5 minutes. In Experiment II, each subject was given two trials (one per day), with each trial consisting of six 15-minute tests. Rest periods between tests averaged less than 2 minutes. A time-line plot of the work-rest cycle for one day in each experiment is shown below.



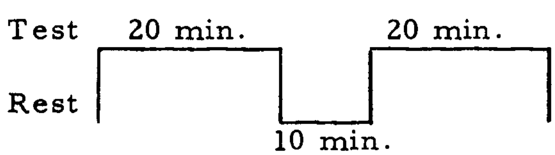


As indicated, the intent of Experiment II was to load the task by minimizing duration of rest in the work-rest cycle. Post-test questionnaires indicated that the procedure was effective in that all subjects reported the task in Experiment II to be more demanding.

It is of interest to compare the work-rest cycles in the above and in the Crook and Wade experiments, which found no differences between behavior on a steady and on a "refreshed" display. Their experiment required subjects to read scrambled letters from steady and "refreshed" displays, and the work-rest cycle was as follows:



Their second experiment, which required subjects to identify specific forms in a matrix of nonsense forms, used the following cycle:



It is clear that both experiments in the present study had higher task loading as measured by the duty cycle. Further, the present study required subjects to perform a concurrent audio monitoring task; no such secondary task was employed in the Crook and Wade experiments.

Therefore, while the supporting data are not abundant, we could reasonably conclude that the present study used more difficult tasks, i.e., our subjects experienced greater burden and probably greater fatigue.

f. Measures

In each experiment, four performance measures were obtained on each test. One speed and one accuracy measure were obtained on the primary task (the 80-item test). The speed measure was expressed as a rate -- average number of items attempted per minute. The accuracy measure was expressed as percent errors -- total items incorrect over total items attempted. Two accuracy measures were obtained on the secondary task (the communications task). The measures were percent errors of commission and percent errors of omission, as based on the total number and type of critical audio inputs programmed for each test. Speed measures on the communications task were inappropriate, since the rate of critical inputs was not under the subject's control. In addition, a post-test questionnaire was administered at the completion of each experiment. The questionnaire was identical for both experiments, and the analysis included study of responses as an indication of differential subjective experiences in each experiment.

3. Results and Discussion

This section presents and discusses the results of the study in the following sequence:

- Experiment I - Major analysis of all data, including post-test questionnaire
- Experiment II - Major analysis of all data, including post-test questionnaire
- Summary Discussion - Discussion of the major findings from each experiment and inferences, trends, etc., which may be apparent in comparing across both experiments

Results are presented according to the above headings and, where appropriate, in supporting appendices (Appendices III and IV). In addition to these results, there is a considerable quantity of descriptive statistics which were a by-product of the major inferential analyses. This material was judged of small value to the issues in question, but it is available in computer output form as a single-copy supplement to this report.

a. Experiment I - Analysis of Performance Measures

As discussed in Section 2.f., four performance measures were obtained on each test. These measures were submitted to an analysis of variance to determine if performance was significantly different for each scope phosphor. The results of the analysis are presented and discussed in the following order: Primary Task - Speed; Primary Task - Accuracy; Secondary Task - Accuracy (errors of commission); and Secondary Task - Accuracy (errors of omission).

1) Primary Task - Speed

All tests in Experiment I consisted of an 80-item test and a concurrent audio monitoring task. The subject was allowed 25 minutes for completion of the 80 items. The resulting speed score was computed as a rate -- the average number of items attempted per minute. These scores were submitted to an analysis of variance, with results as shown in Table 2.

Table 2. Analysis of Variance Summary Table - Experiment I
Dependent Variable, Primary Task - Speed

Sources of Variation	Sums of Squares	d.f.	Mean Squares	F. Ratio	Significance
Subjects (1)	15.612	7	2.230	22.14	$p < .01$
Scopes (2)	0.130	1	0.130	1.28	----
Trials (3)	9.410	1	9.410	93.42	$p < .01$
Tests (4)	0.797	1	0.797	7.91	$p < .05$
(1) x (2)	0.772	7	0.110	1.09	----
(1) x (3)	0.819	7	0.117	1.16	----
(1) x (4)	0.418	7	0.060	0.59	----
(2) x (3)	0.004	1	0.004	0.04	----
(2) x (4)	0.183	1	0.183	1.81	----
(3) x (4)	0.168	1	0.168	1.67	----
(1) x (2) x (3)	4.817	7	0.688	6.83	$p < .05$
(1) x (2) x (4)	0.241	7	0.034	0.34	----
(1) x (3) x (4)	0.444	7	0.063	0.63	----
(2) x (3) x (4)	0.026	1	0.026	0.25	----
Residual	0.705	7	0.101		----
Total	34.544	63			

As shown in Table 2, it was possible to partial the total variance into four main effects and their resulting interactions. With no within- or between-conditions error variance, the fourth order interaction term becomes the residual variance for testing the significance of variance attributable to other sources.

Of major interest in this and all succeeding analyses is the effect of scopes on the performance measure. This analysis indicates that scopes are not significantly different in terms of speed of performance on the primary task. Thus, the observed difference (2.96 vs 2.87 items per minute for P-31 and P-12, respectively) cannot be attributed to other than chance variability.

All other main effects showed significant differences. Differences between subjects are not uncommon in tasks of this type, and they were treated as a main effect, primarily to partial out the variance attributable to them and thus provide a more sensitive test of other effects. Their average scores are shown below.

Subject	Average Speed (Items per Minute)	
	P-31	P-12
1	1.93	2.06
2	2.55	2.52
3	2.98	2.50
4	2.87	2.83
5	3.14	2.89
6	3.17	3.20
7	3.56	3.36
8	3.47	3.59
Average	2.96	2.87

These differences simply indicate that while all subjects were motivated and performed well, they varied in their relevant skills and abilities.

Performance also varied as a function of trials. The average performances (2.53 vs. 3.30 items per minute for Trial 1 and Trial 2, respectively) indicate that learning improved performance.

The significant difference between tests (2.80 vs. 3.02 items per minute for Test 1 and Test 2, respectively) indicate that within a trial, learning was again affecting performance. Thus, fatigue effects, if present, were more than offset by increased learning.

One interaction, which involved subjects, scopes, and trials, was significant. Such higher order interactions are often difficult to interpret, but if we examine the relative percentage changes in performance as a function of the three effects, a meaningful interpretation is possible. The data on these percentage changes is shown below.

	% Improvement from Trial 1 to Trial 2	
	4 Best Subjects	4 Worst Subjects
P-12 Scope	32.2	31.1
P-31 Scope	34.7	21.6

The interpretation of this data is as follows: While all subjects improve with experience on both scopes, the average improvement is greater on the P-12 scope, because the worst subjects improve less on the P-31 scope. This finding argues for the P-12 scope, but because the interaction is complex (third order) and not supported by the main effect, it is not very compelling.

2) Primary Task - Accuracy

The accuracy measure of performance was computed for each 80-item test and expressed as percent errors -- number of items incorrect over total items attempted. As above, these scores were submitted to an analysis of variance. The results of that analysis are shown in Table 3.

Table 3. Analysis of Variance Summary Table - Experiment I
Dependent Variable, Primary Task - Accuracy

Sources of Variation	Sums of Squares	d.f.	Mean Squares	F. Ratio	Significance
Subjects (1)	317.587	7	45.370	7.87	$p < .01$
Scopes (2)	0.282	1	0.282	0.05	----
Trials (3)	70.917	1	70.917	12.30	$p < .01$
Tests (4)	44.240	1	44.240	7.67	$p < .05$
(1) x (2)	22.606	7	3.230	0.56	----
(1) x (3)	76.678	7	10.954	1.90	----
(1) x (4)	70.151	7	10.022	1.73	----
(2) x (3)	0.963	1	0.963	0.17	----
(2) x (4)	0.007	1	0.007	0.001	----
(3) x (4)	0.546	1	0.546	0.09	----
(1) x (2) x (3)	28.081	7	4.012	0.70	----
(1) x (2) x (4)	116.600	7	16.657	2.89	----
(1) x (3) x (4)	61.050	7	8.721	1.51	----
(2) x (3) x (4)	0.610	1	0.610	0.11	----
Residual	40.366	7	5.767		----
Total	850.683	63			

The variance was partialled as in the analysis of speed on the primary task. The results are also similar to the results in that analysis. Subjects showed differences in their ability to perform the task accurately. The average performance of each subject is shown below.

Subject	Average Accuracy (% Errors)	
	P-31	P-12
1	8.83	7.57
2	9.62	7.29
3	7.56	10.00
4	4.24	6.89
5	7.63	7.79
6	5.63	2.90
7	5.79	5.06
8	3.75	6.60
Average	6.63	6.76

As in the analysis of speed, subjects became more accurate with trials and with tests within each trial. The average performance on trials and tests is shown below.

Trial	Average Accuracy (% Errors)	Test	Average Accuracy (% Errors)
1	7.75	1	7.53
2	5.64	2	5.86

On the major independent variable (scopes) performance was not significantly different. Thus, the small difference in accuracy (6.63 vs. 6.76 % errors for P-31 and P-12, respectively) cannot be attributed to other than chance variability.

3) Secondary Task - Accuracy (Errors of Commission)

The audio monitoring task was scored in terms of errors of commission and omission. An error of commission was noted whenever the subject responded to a critical input but his response was in error, i.e., the wrong alphabetic character was recorded. Commission errors were expressed in percent terms -- the number of commissions over the total number of critical inputs. These scores were submitted to variance analysis, with results as shown in Table 4.

Table 4. Analysis of Variance Summary Table - Experiment I
Dependent Variable, Secondary Task - Accuracy
(Commission Errors)

Sources of Variation	Sums of Squares	d.f.	Mean Squares	F. Ratio	Significance
Subjects (1)	24.216	7	3.460	1.56	----
Scopes (2)	3.231	1	3.231	1.46	----
Trials (3)	17.914	1	17.914	8.07	$p < .05$
Tests (4)	3.222	1	3.222	1.45	----
(1) x (2)	22.973	7	3.282	1.48	----
(1) x (3)	17.249	7	2.464	1.11	----
(1) x (4)	16.556	7	2.365	1.06	----
(2) x (3)	9.257	1	9.257	4.17	----
(2) x (4)	0.423	1	0.423	0.19	----
(3) x (4)	4.431	1	4.431	2.00	----
(1) x (2) x (3)	31.201	7	4.457	2.01	----
(1) x (2) x (4)	28.829	7	4.118	1.85	----
(1) x (3) x (4)	30.593	7	4.370	1.97	----
(2) x (3) x (4)	1.626	1	1.626	0.73	----
Residual	15.541	7	2.220		----
Total	227.260	63			

The results indicate only one source of variation is significant. Performance over trials (1.63 vs. 0.57% errors for Trials 1 and 2, respectively) indicate improvement in accuracy with practice. The major variable (scope) shows no systematic difference in performance. The difference in average performance on each scope (0.87 vs. 1.32% errors for P-31 and P-12, respectively) must again be treated as a chance finding.

4) <u>Secondary Task - Accuracy (Errors of Omission)</u>

An error of omission was noted whenever the subject failed to make any response to a critical audio input. The errors were expressed in percent terms -- the number of omissions over the total number of critical inputs. The analysis of variance results for these scores is shown in Table 5.

Table 5. Analysis of Variance Summary Table - Experiment I Dependent Variable - Secondary Task - Accuracy (Omission Errors)

Sources of Variation	Sums of Squares	d.f.	Mean Squares	F. Ratio	Significance
Subjects (1)	19.198	7	2.743	1.01	----
Scopes (2)	1.180	1	1.180	0.44	----
Trials (3)	2.239	1	2.239	0.83	----
Tests (4)	5.935	1	5.935	2.20	----
(1) x (2)	9.086	7	1.298	0.48	----
(1) x (3)	42.530	7	6.076	2.25	----
(1) x (4)	12.343	7	1.763	0.65	----
(2) x (3)	0.173	1	0.173	0.06	----
(2) x (4)	2.187	1	2.187	0.81	----
(3) x (4)	0.000	1	0.000	0.00	----
(1) x (2) x (3)	10.093	7	1.442	0.53	----
(1) x (2) x (4)	16.734	7	2.391	0.88	----
(1) x (3) x (4)	13.931	7	1.990	0.74	----
(2) x (3) x (4)	0.002	1	0.002	0.00	----
Residual	18.918	7	2.703		----
Total	154.549	63			

As indicated, the analysis yields no significant sources of systematic variation. Most importantly, the difference in average omission-type errors between scopes (0.92 vs. 0.65% errors for P-31 and P-12, respectively) is again due to chance variation.

b. Experiment I - Analysis of Questionnaire Data

Immediately upon completion of experimentation, each subject was given a post-test questionnaire. The complete results of the questionnaire as recorded by the subjects are given in Appendix III. A summary of the major findings is presented below.

1) Scope Preference

- Five subjects preferred the P-12 scope; two subjects preferred the P-31 scope; one subject had no preference.

- The five subjects preferring the P-12 scope gave the following reasons: "better color, easier on the eyes, easier to read, no jumping of characters (as on P-31), and lettering clearer and steadier."

- The two subjects preferring the P-31 scope gave the following reasons: "easier to read, and easier to find the data."

- The one subject registering no preference stated that the characters on the P-31 wavered but this did not affect his performance.

2) Fatigue

- Three subjects reported the task fatiguing; five subjects reported no fatigue.

- Of the three subjects reporting fatigue, one subject reported more fatigue on the P-31, one subject reported fatigue only on the P-31, and one subject reported no scope differences.

c. Experiment I - Discussion

The major conclusions from Experiment I can be summarized quite briefly. There is no significant systematic difference in performance, under the task conditions described in the experiment, between the P-12 and P-31 scope phosphors. Other independent variables are frequently significant. However, these differences have no practical import to the main issue, i.e., they are not design-relevant. Subjective data indicate a distinct preference for the P-12 phosphor. This preference is based, in part, on some increase in reported fatigue when using the P-31 phosphor.

d. Experiment II - Analysis of Performance Measures

Four performance measures were taken for each test in Experiment II. The measures were identical to those used in Experiment I. Similar analyses were performed on these measures, and the results are reported in the same manner. It will be seen that the analyses of variance for this experiment involve only three main effects and their resulting interactions. Trials were not an effect here, since only one extended trial comprising six tests was administered for each subject for each scope condition.

1) Primary Task - Speed

The data obtained on speed of performance for the primary task were analyzed using a three-way analysis of variance. Results of this analysis are shown in Table 6.

Table 6. Analysis of Variance Summary Table - Experiment II Dependent Variable, Primary Task - Speed

Sources of Variation	Sums of Squares	d.f.	Mean Squares	F. Ratio	Significance
Subjects (1)	11.375	3	3.792	36.29	$p < .01$
Scopes (2)	0.041	1	0.041	0.39	----
Tests (3)	1.333	5	0.266	2.55	----
(1) x (2)	2.466	3	0.822	7.87	$p < .01$
(1) x (3)	1.762	15	0.118	1.12	----
(2) x (3)	0.214	5	0.643	0.41	----
Residual	1.567	15	0.104		----
Total	18.758	47			

The analysis indicates the only main effect of statistical significance is subjects. The average performance of each subject is shown below.

Subject	Average Speed (Items per Minute)
1	4.28
2	3.14
3	4.21
4	3.45

The major variable (scopes) was not a significant effect. Therefore, no interpretation other than random variability can be attached to the observed difference in average performance (3.74 vs. 3.80 items per minute for P-31 and P-12, respectively).

One interaction which involved subjects and scopes was significant. The interaction can be interpreted by examining percentage changes in subjects' performance as a function of scope condition. The necessary data for interpretation is shown below.

Subject	% Improvement Using P-12 vs. P-31
1	- 6.1
2	+ 8.6
3	+18.1
4	-12.5

Based on the above, the interaction indicates that subjects whose performance is better using a P-12 phosphor exhibit a greater relative difference between scopes than subjects whose performance is better using the P-31 phosphor. While not as conclusive as a significant difference in main effects, this finding does argue for a P-12 phosphor. Stated more directly, the finding says: Subjects who do better on a P-12 do much better, whereas subjects who do better on a P-31 do better but not by as much.

2) Primary Task - Accuracy

An analysis of variance was performed on the scores measuring accuracy of performance on the primary task. The results of the analysis are shown in Table 7.

Table 7. Analysis of Variance Summary Table - Experiment II
Dependent Variable, Primary Task - Accuracy

Source of Variation	Sums of Squares	d.f.	Mean Squares	F. Ratio	Significance
Subjects (1)	274.915	3	91.638	10.85	$p < .01$
Scopes (2)	15.916	1	15.916	1.88	----
Tests (3)	28.960	5	5.792	0.68	----
(1) x (2)	21.053	3	7.018	0.83	----
(1) x (3)	98.243	15	6.550	0.77	----
(2) x (3)	22.155	5	4.431	0.52	----
Residual	126.682	15	8.445		----
Total	587.926	47			

As indicated, the only significant effect was subjects. Their average performances are shown below.

Subject	Average Accuracy (% Errors)
1	2.78
2	2.36
3	8.14
4	2.74

Scope conditions, the critical independent variable, was not significant. Therefore, the average performance for each scope (4.58 vs. 3.43% errors for P-31 and P-12, respectively) must be attributed to chance variability.

3) Secondary Task - Accuracy (Errors of Commission)

Errors of commission on the secondary task were submitted to analysis of variance. The results of the analysis are shown in Table 8.

Table 8. Analysis of Variance Summary Table - Experiment II Dependent Variable, Secondary Task - Accuracy (Commission Errors)

Sources of Variation	Sums of Squares	d.f.	Mean Squares	F. Ratio	Significance
Subjects (1)	51.546	3	17.182	5.99	$p < .01$
Scopes (2)	3.848	1	3.848	1.34	----
Tests (3)	2.796	5	0.560	0.19	----
(1) x (2)	6.111	3	2.037	0.71	----
(1) x (3)	13.946	15	0.930	0.32	----
(2) x (3)	4.692	5	0.938	0.33	----
Residual	43.009	15	2.867		----
Total	125.948	47			

The results are identical to those found in analysis of errors on the primary task. Subject effects are the only significant effects. Their average scores are shown below.

Subjects	Average Accuracy (% Errors of Commission)
1	0.0
2	0.0
3	0.85
4	2.54

Average performance on each scope indicates 1.31 vs. 0.57% errors for P-31 and P-12, respectively. However, we must ascribe this difference to random variability.

4) Secondary Task - Accuracy (Errors of Omission)

The variance analysis of omission errors is shown in Table 9.

Table 9. Analysis of Variance Summary Table - Experiment II Dependent Variable, Secondary Task-Accuracy (Omission Errors)

Sources of Variation	Sums of Squares	d.f.	Mean Squares	F. Ratio	Significance
Subjects (1)	2.780	3	0.927	1.16	----
Scopes (2)	0.265	1	0.266	0.33	----
Tests (3)	13.265	5	2.653	3.33	$p < .05$
(1) x (2)	4.833	3	1.611	2.02	----
(1) x (3)	19.750	15	1.317	1.65	----
(2) x (3)	1.224	5	0.245	0.31	----
Residual	11.952	15	0.797		----
Total	54.070	47			

The results show tests to be the only significant source of variation. Average performance on each test is shown below.

Test	Average Accuracy (% Errors of Omission)
1	0.88
2	0.00
3	1.29
4	0.00
5	0.00
6	0.00

The data clearly indicate that subjects become most accurate with increased practice.

The difference in performance on each scope (0.29 vs. 0.44% errors for P-31 and P-12, respectively) is again a result of chance variability.

e. Experiment II - Analysis of Questionnaire Data

A questionnaire identical to that used in Experiment I was administered to the subjects at the completion of testing. Complete results of the questionnaire are presented in Appendix IV. The major findings are summarized below.

1) Scope Preference

- Two subjects preferred the P-12 scope; two subjects preferred the P-31 scope.

- The two subjects preferring the P-12 scope gave the following reasons: "steadiness - on P-31 there was flicker, unsteadiness, and upper and lower halves slipped across one another, lettering clearer and steadier."

- The two subjects preferring the P-31 scope gave the following reasons: "less fatigue and performed better, easier to find information."

2) Fatigue

- All four subjects reported the task to be fatiguing.

- One subject reported more fatigue with the P-31; one subject reported more fatigue with the P-12; two subjects reported no difference in fatigue using either phosphor.

f. Experiment II - Discussion

As with Experiment I, we can summarize the major conclusions quite briefly. The objective measures of performance indicate that scope conditions, considered as a main effect, do not systematically affect performance. Some of the analyses show other minor effects to be significant. However, these differences have no direct bearing on phosphor design. One interaction effect involving scopes was significant and argues for the P-12 phosphor. Subjective data indicate no differential preference for either phosphor and no difference in reported fatigue as a function of type of phosphor.

g. Subsidiary Analyses

It was deemed worthwhile to delve more deeply into the error scores associated with the primary task in both experiments. It was thought that the analysis of variance technique could be relatively insensitive to the small differences observed due to (1) the relatively few observations entered into the analysis, and (2) the possibility that the errors were distributed in a non-normal way, possibly according to the Poisson distribution. However, it was found that the distribution of error scores did not conform to the Poisson distribution. While there were some differences between the two displays in the ways that errors were distributed with respect to successive trials and different subjects, these differences were not systematic and did not, in sum, amount to more than a hint that if more observations were available more systematic differences might appear.

4. Summary Discussion of Experiments

As apparent from the above presentation, the results do not indicate a marked superiority for either phosphor. However, on balance, the results favor a phosphor of the P-12 type. The reasons for this statement are as follows:

- The only significant source of variation involving scopes (phosphors) argue for the P-12.

- As task loading increases (Experiment II vs. Experiment I), there is a trend in the descriptive statistics which favors the

P-12, i.e., in Experiment I three of the four performance measures favor the P-31, while in Experiment II three of the four performance measures favor the P-12.

- To the degree that experimental conditions failed to replicate operational conditions (half page print time, transfer time, and spot brightness), the scope differences were reduced (the apparent movement phenomena associated with the P-31 were attenuated).

We must emphasize that the conclusion is an on-balance judgment, and we must acknowledge the following facts which, in themselves, argue for no difference between phosphors:

- There were no significant main effect differences between scopes in any of the analyses.

- For the four subjects used in both experiments, the scope preferences were equally divided.

APPENDIX II

SAMPLE SCOPE FORMAT AND TEST CARD

SCOPE FORMAT NO. 4

1)	56A732	+ 5	THIS	SWITCHES OK 1
2)	95C786	+ 5	ON	GYRO DIAL - 272 DEGREE
3)	832BJK	- 4	EIS	ACCELEROMETER 325 FPS
4)	93A661	- 3	GO	WIRE BLOOD RATE 66 LB
5)	21320B	> 9	ING	BOATED EFFECT 633 FT.
6)	22321B	+ 3	TUB	CREWDOW CHGE. +63 G
7)	23322B	+ 6	TH	FIL. FEEDTHRU 9 KV AT 1
8)	33333B	= 4	EBE	CONTROLLER GRIP 6 LB.
9)	42C865	+ 0	ST	DRINK ALARM 02.00 HR.
10)	89D625	- 19	WU	MERCURY FALLOUT 15 K
11)	95JW62	- 12	NYE	PILOT FATIGUE 2 MK
12)	8631	≤ 21	TUW	PARTIAL CRPE 10N
13)	899CZ53	< 32	AYT	FLT NO. 6
14)	41K623	= 45	AN	DEPART 5.30 P
15)	4K1263	+ 10	DSE	ARRIVE 5.35 A
16)	842HL6	+ 21	EI	PASSENGERS 123
17)	956A93	- 22	FTH	AIR. TYPE 707F
18)	0621CW	- 29	ISIS	INTEGR. ASCE 51 K
19)	9ZA62	- 92	NT	EXH. TEMP. 30 K
20)	85JJ615	+ 65	SEW	FUEL MIX 16 - 15.3
21)	54K321	+ 4	RU	ATTITUDE - 16D
22)	892MO1	< 131	SA	REMAIN FAX 3 J
23)	629N10	> 025	TIS	ESC. VELOC. 25 KM
24)	51E432	- 1	FYE	BALANCE 060 MG
25)	60F219	- 36	DNO	DECREASE 5 P
26)	00501	- 265	WBUD	ESC PEMECH 1

TEST CARD 4 C

1. What number appears after DEPART?

2. Does the symbol appearing last (right-hand end) in Placard Signal D also appear in Signal 1?

3. Does Signal 8 contain the symbols L, E, 6, 3, B?

4. How many signals are signed < ?

5. Did Sender TIS send two signals?

6. How many common characters are there in the addresses of Placard Signal C and Signal 7?

7. Do Signals 20 and 12 have the same number of digits in the code column?

8. Write out the remainder of the signal which comes after . . .
CHGE . . .

9. Is the symbol appearing immediately after THIS the same as the symbol appearing immediately after 325FP?

10. What is the address of the signal referring to FLT NO ?

11. Determine whether the code of the Placard Signal which contains JAW appears ONCE in the code column of the signals on main display.

12. Write Sender of Signal 6.

13. Is the sign appearing in Placard Signal C the same as the sign appearing in Signal 5?

14. Which signal refers to BOATED?

15. Write out the part of the signal which comes before . . . . B > 9 I

16. Add the digits (single numbers) appearing last in the addresses of Signals 20, 9, and 15.

17. Write the address of Signal 4.

18. Which signal, if any, has the same code as the Placard Signal whose address is 53LZC?

19. How many different digits (single numbers) appear in Signal 5?

20. What digit (single number) appears before Sender EIS?

APPENDIX III

POST-TEST QUESTIONNAIRE RESULTS
EXPERIMENT I

Subject 1
Experiment I

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   P-31 scope caused more eye strain.
   P-12 scope (color) seemed to make the precise data needed "jump out" more readily.
   More orderly presentation of data (in columns) made reading easier.
   P-31 scope color (green) more pleasant to look at but not to extract data from. Scanning type of viewing gave biggest difference between two scopes (P-12 better).

2. Did you experience any fatigue in performing the task?

   Yes - P-31 scope.

   If Yes, then:

   a. What kind(s)?

      Fuzziness - letters appeared to blend (greater concentration needed to overcome difficulty).

   b. When experienced?

      Near 3rd set of test cards (whenever using P-31 scope - seemed to recover by the 4th set of cards - got 2nd wind)

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      Yes - on P-31 scope only.

   d. Was the rest period between experimental trials sufficient to overcome any fatigue?

      Yes.

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Yes, improved after first trial after task was learned.

   b. Format:

      More orderly presentation of data in columns made reading easier and faster.

   c. Test cards?

      No.

   d. Scopes?

      Yes - improved on P-12 scope.

   e. Time of day?

      No.

   f. Other?

4. Do you have a scope preference?

   Yes.

   If Yes, why?

   P-12 scope (orange color) seemed to make the precise data needed "jump out" more readily.

Subject 3
Experiment I

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   P-31 seems easier to read but has more distracting flicker. Non-glare glass would facilitate reading.

2. Did you experience any fatigue in performing the task?

   No, but eye fatigue would soon develop if the task was sustained for longer periods of time. Large print would have helped.

   If Yes, then:

   a. What kind(s)?

      Became nervous toward end of trials, began to speed up, hurry responses, etc., when reached 4th card in attempt to complete in the time period.

   b. When experienced?

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      No.

   d. Was the rest period between experimental sessions sufficient to overcome any fatigue?

      Yes.

Post-Test Questionnaire (Continued)

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Yes - learning the task.

   b. Format:

      Yes (the ones that lined up better and had more sensible material were easier).

   c. Test cards?

      No.

   d. Scopes?

      No - after task was learned (by end of second trial).
      Yes - on first two trials judged performance superior on 2nd trial compared to first because of learning (experience) and because of shift to P-31 scope.

   e. Time of day?

      No.

   f. Other?

      Physical fatigue - general, not specific to any muscles.

4. Do you have a scope preference?

   Yes - P-31.

   If Yes, why?

   Seemed easier to read.

Subject 5
Experiment I

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   The scope to the left (P-12) seemed much easier to read. The one to the right (P-31) seemed to have a shaky break in the middle and also appeared to have a slight flicker.

2. Did you experience any fatigue in performing the task?

   Yes.

   If Yes, then:

   a. What kind(s)?

      Eyes got a little tired.
      Felt a little frustrated at times when I couldn't find something.

   b. When experienced?

      In the afternoon during the second trial.

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      Yes - the scope to the right (P-31) was more fatiguing. The scopes with full messages were more difficult to read - to find the correct answer. Light reflection on the scopes hindered reading also.

   d. Was the rest period between experimental trials sufficient to overcome any fatigue?

      Yes, but 15 minutes between tests would be better with a chance to move around.

Post-Test Questionnaire (Continued)

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Yes - more experience helped me find the answers faster and I felt I became more efficient.

   b. Format:

      Yes, the formats with spaces in the messages were easier to read.

   c. Test cards?

      After I learned the basic form of the questions I didn't need to read the entire question.

   d. Scopes?

      I felt the scope to the left (P-12) was easier on the eyes and easier to read.

   e. Time of day?

      I felt more alert and efficient in the morning.

   f. Other?

      I enjoyed the experiments.

4. Do you have a scope preference?

   Yes.

   If Yes, why?

   The one to the left (P-12) was steady - easier on the eyes; more pleasing color (orange) for this type of work.

Subject 7
Experiment I

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   Personally, the gold lettering (P-12) was easier to read than was the green. When I shifted my eyes from left to right and vice versa, the green lettering appeared to shift a line at a time also.

2. Did you experience any fatigue in performing the task?

   No.

   If Yes, then:

   a. What kind(s)?

   b. When experienced?

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

   d. Was the rest period between experimental sessions sufficient to overcome any fatigue?

      Yes.

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Definitely! After two trials I set up a system so as to facilitate the search for the information. I could also remember approximate locations of bits of information from one time to another.

Post-Test Questionnaire (Continued)

b. Format:

The displays in which codes, senders, etc., were not vertically (more or less) arranged presented more of a problem.

c. Test cards?

No difference here.

d. Scopes?

No difference.

e. Time of day?

Mornings preferred, especially when a reasonable length of time has elapsed following breakfast.

f. Other?

4. Do you have a scope preference?

P-12 scope.

If Yes, why?

Lettering appears clearer and more steady. By this I mean that the green lettering (P-31) seems to reflect unequal activation of the sensitive material on the scope screen.

Subject 8
Experiment I

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   I did experience some jumping of the letters on the green screen (P-31), during my first testing period. Due to this I believe I prefer the gold lettering (P-12).

2. Did you experience any fatigue in performing the task?

   My eyes seem to be a little bit tired at the end of the first testing period.

   If Yes, then:

   a. What kind(s)?

      Letters seem to be jumping a little on the P-31 scope.

   b. When experienced?

      During first testing period.

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      No.

   d. Was the rest period between experimental sessions sufficient to overcome any fatigue?

      Yes.

Post-Test Questionnaire (Continued)

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Yes - primarily getting to know requirements of task.

   b. Format:

      Yes - less information the better.

   c. Test cards?

      No.

   d. Scopes?

      Yes - some prediction (not a strong one) that performance would be better on left scope (P-12).

   e. Time of day?

      No.

   f. Other?

4. Do you have a scope preference?

   Yes. P-31.

   If Yes, why?

   Because of my experience during first testing period.

Subject 9
Experiment I

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   The right-hand scope (P-31) seemed to waver more than the other. Nevertheless, I don't feel that this had a great deal of effect on my performance.

2. Did you experience any fatigue in performing the task?

   No.

   If Yes, then:

   a. What kind(s)?

   b. When experienced?

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

   d. Was the rest period between experimental sessions sufficient to overcome any fatigue?

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Yes - once the questions were understood and I became familiar with the format of the scope material it was much easier to perform the task.

Post-Test Questionnaire (Continued)

b. Format:

c. Test cards?

d. Scopes?

e. Time of day?

f. Other?

4. Do you have a scope preference?

No.

If Yes, why?

Subject 10
Experiment I

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   I found that finding data with the scope having green printing (P-31) was easier. In scanning the scope it was easier to get an impression of needed information and to remember where the impression was last sensed. (i.e., when looking for one word notice others and remember them if a later question refers to them.)

2. Did you experience any fatigue in performing the task?

   No.

   If Yes, then:

   a. What kind(s)?

   b. When experienced?

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

   d. Was the rest period between experimental sessions sufficient to overcome any fatigue?

      I believe that any fatigue that was being built was lost during the rest period.

Post-Test Questionnaire (Continued)

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Yes. The more experience I gained, the less time it took to look for things on the placard and I felt more certain that the answer put down was the one asked for.

   b. Format:

      No.

   c. Test cards?

      No.

   d. Scopes?

      Yes. As previously indicated, I felt the color of printing facilitated my finding much of the information.

   e. Time of day?

      No.

   f. Other?

      None that come to mind.

4. Do you have a scope preference?

   Yes - P-31.

   If Yes, why?

   This has already been indicated.

Subject 11
Experiment I

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   The P-12 scope seems a little easier to read - possibly a little easier on the eyes.

   Bottom half of P-31 scope jumps.

2. Did you experience any fatigue in performing the task?

   Noticed fatigue after completed - not during.

   If Yes, then:

   a. What kind(s)?

      Eyes tired.

   b. When experienced?

      When relaxing after a trial.

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      Not noticed.

   d. Was the rest period between experimental sessions sufficient to overcome any fatigue?

      Yes.

Post-Test Questionnaire (Continued)

3. Do you think your performance varied as a function of any of the following:

a. Experience (learning the task, the cards, etc.)?

Yes - at first thought had to be given to the meaning of the question and the learning of the terms and their location. Later this became second nature and all you looked for was the specifics in the question.

b. Format:

Yes - formats that were crammed or did not line up in columns were more difficult.

c. Test cards?

No.

d. Scopes?

P-12 scope - a little easier to read.

e. Time of day?

Not quite as sharp later in the day.

f. Other?

4. Do you have a scope preference?

Yes - P-12.

If Yes, why?

Easier to read.

APPENDIX IV

POST-TEST QUESTIONNAIRE RESULTS
EXPERIMENT II

Subject 3
Exp II

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   P-31 flicker becomes distracting over time. P-12 stability was welcome.

2. Did you experience any fatigue in performing the task?

   Yes.

   If yes, then:

   a. What kind(s)?

      Coordination of the two tasks. It became harder to hear the signals. Also harder to pick up exact letter and number combination. Found myself rechecking answers and referring back to questions more often. More pronounced toward end of each trial (esp. 2nd trial).

   b. When experienced?

      Increasingly difficult over time.

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      No.

   d. Was the rest period between experimental trials sufficient to overcome any fatigue?

      Not toward the end.

Post-Test Questionnaire (Continued)

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Essentially, task was learned in first exp.

   b. Format?

   c. Test cards?

   d. Scopes?

   e. Time of day?

      Yes -- more physically fatigued during the evening trials compared to trials in Exp. I, which were run earlier.

   f. Other?

4. Do you have a scope preference?

   Yes.

   If yes, why?

   P-12 picture was steadier. The flicker, unsteadiness of P-31 scope gets to you after awhile; upper and lower halves slip across one another.

Subject 7
Exp II

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   Same -- except the display (P-12) did jump several times at the beginning of the experiment.

2. Did you experience any fatigue in performing the task?

   None today. Some fatigue on 1st trial.

   If yes, then:

   a. What kind(s)?

      Eye fatigue on 1st trial.

   b. When experienced?

      Middle of trial.

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      No.

   d. Was the rest period between experimental trials sufficient to overcome any fatigue?

      No.

Subject 7
Exp II

Post-Test Questionnaire (Continued)

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Some learning of information on cards.

   b. Formats?

      Some formats easier than others because of organization.

   c. Test cards?

      No.

   d. Scopes?

      No.

   e. Time of day?

      Am convinced a full stomach is bad for testing demands, i.e., first trial.

   f. Other?

4. Do you have a scope preference?

   P-12.

   If yes, why?

Subject 8
Exp II

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

2. Did you experience any fatigue in performing the task?

   My eyes seem to be more fatigued after taking the P-12 scope.

   If yes, then:

   a. What kind(s)?

   b. When experienced?

      After taking the test on the P-12 scope.

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      Yes, different scopes.

   d. Was the rest period between experimental trials sufficient to overcome any fatigue?

      No.

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Some learning of task and cards.

Post-Test Questionnaire (Continued)

b. Format?

Formats with less information were easier.

c. Test cards?

No.

d. Scopes?

Seem to perform faster on P-31 scope.

e. Time of day?

f. Other?

4. Do you have a scope preference?

Preference for P-31, but not a real strong preference.

If yes, why?

I experienced less eye fatigue and I also thought I performed faster on this scope.

Subject 10
Exp II

POST-TEST QUESTIONNAIRE

1. Based on your experience as a subject in this study, tell me anything you think might be relevant to the selection of one of these scopes.

   Same as Exp. I.

2. Did you experience any fatigue in performing the task?

   Yes.

   If yes, then:

   a. What kind(s)?

      1) Eyes became tired.
      2) Neck became tired.

   b. When experienced?

      1) While using P-12 scope (eyes).
      2) While using P-31 scope (neck).

   c. Was fatigue in any way different for different experimental conditions (i.e., different test cards, different scopes, different formats on the scopes, etc.)?

      1) When using format with considerable information on it (greater fatigue -- eye).

      2) Using P-31 scope (greater fatigue -- neck).

   d. Was the rest period between experimental trials sufficient to overcome any fatigue?

      Did not notice fatigue until after rest period. Five-minute rest instead of 2.5-minute rest would be better.

Post-Test Questionnaire (Continued)

3. Do you think your performance varied as a function of any of the following:

   a. Experience (learning the task, the cards, etc.)?

      Same as Exp. I.

   b. Format?

      When there was considerable information on format, I think it resulted in slower performance.

   c. Test cards?

      No.

   d. Scopes?

      Same as Exp. I.

   e. Time of day?

      Same as Exp. I.

   f. Other?

      Yes. I feel that I worked slower on trial 2, probably because of lack of practice over the weekend.

4. Do you have a scope preference?

   Yes.

   If yes, why?

   Same as Exp. I. (P-31)