UNCLASSIFIED

AD NUMBER
AD857871
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies only; Administrative/Operational Use; NOV 1968. Other requests shall be referred to Army Electronics Command Night Vision Lab., Fort Belvoir, VA 22060.**
AUTHORITY
USAECOM ltr 8 Nov 1974

THIS PAGE IS UNCLASSIFIED

7919-32-T

SOME TESTS OF NIGHT AND DAY VISION

By

C. A. Omarzu

Infrared Physics Laboratory
Willow Run Laboratories
Institute of Science and Technology
The University of Michigan
Ann Arbor, Michigan

November 1968

Prepared for the Advanced Research Projects Agency, Washington, D. C., ARPA Order No. 774, and administered by the Night Vision Laboratory, U. S. Army Mobility Equipment Research and Development Center, Fort Belvoir, Virginia, under Contract DA-44-009-AMC-1494(T).

STATEMENT #3 UNCLASSIFIED

Each transmittal of this document outside the ~~agencies of the~~ U.S. Government must have prior approval of Army Electronics Command, Night Vision Lab. Fort Belvoir Va 22060

NOTICES

Sponsorship. The work reported herein was conducted by the Willow Run Laboratories of the Institute of Science and Technology for the Advanced Research Projects Agency, Washington, D. C., under ARPA Order No. 774. This project is being administered by the U. S. Army Night Vision Laboratory, U. S. Army Mobility Equipment Research and Development Center, Fort Belvoir, Virginia, under Contract DA-44-009-AMC-1494(T). Contracts and grants to The University of Michigan for the support of sponsored research are administered through the Office of the Vice-President for Research.

Final Disposition. After this document has served its purpose, it may be destroyed. Please do not return it to the Willow Run Laboratories.

7919-32-T

SOME TESTS OF NIGHT AND DAY VISION

By

C. A. Omarzu

Infrared Physics Laboratory
Willow Run Laboratories
Institute of Science and Technology
The University of Michigan
Ann Arbor, Michigan

November 1968

Prepared for the Advanced Research Projects Agency, Washington, D. C., ARPA Order No. 774, and administered by the Night Vision Laboratory, U. S. Army Mobility Equipment Research and Development Center, Fort Belvoir, Virginia, under Contract DA-44-009-AMC-1494(T).

ACKNOWLEDGMENTS

The author received helpful suggestions and assistance from several members of the Infrared Physics Laboratory of Willow Run Laboratories regarding the implementation of the tests and the procedures used therefor, most notably from Dr. G. J. Zissis, Dr. W. T. Pollock, Mr. R. E. Miller, Mr. M. Bender, and Mr. R. Stancliff. The author also profited from several conversations with Dr. H. Leibowitz, Professor of Psychology, the Pennsylvania State University, especially in connection with the form-recognition tests. The author acknowledges the helpful advice given by Dr. E. Jebe of the Radio Science Laboratory of Willow Run Laboratories in regard to the sample-size estimates presented in section 5.

The slides used in the tests were prepared by Mr. J. Cooper, Head of the Photography Laboratory at Willow Run Laboratories. The American University of Beirut test results reported in section 4 were obtained by Dr. N. A. Haddad, co-principal investigator with Dr. B. Faris, and their assistant, Mr. M. El Geddawi. Mr. Stancliff obtained most of the University of Michigan test results reported in section 4.

Finally, though the author received helpful suggestions and assistance from the sources just named, the decisions involved in the design and implementation of the tests were essentially his own and the author accepts full responsibility for these as well as for the opinions expressed herein.

FOREWORD

This report was prepared by the Infrared Physics Laboratory of Willow Run Laboratories, a unit of The University of Michigan's Institute of Science and Technology. It is an unscheduled technical report submitted under Contract DA-44-009-AMC-1494(T), "Night Vision Aids for Counterinsurgency." The contract is funded by the Advanced Research Projects Agency, Washington, D. C., as part of Project AGILE, and is being administered by the Night Vision Laboratory, U. S. Army Mobility Equipment Research and Development Center, Fort Belvoir, Virginia. While the tests described herein were being developed (during the period January to October 1967), the ARPA-AGILE project monitor was Lt. Col. E. I. Golding, USAF. The present project monitor is Lt. Col. W. Y. Cole, USAF.

The contract supports research and evaluation that will help in the design and selection of night-vision aids for use by counterinsurgent military forces, particularly those of governments other than that of the United States. The program is especially concerned with improving the ability to predict the performance of a night-viewing system by measuring various contributing factors such as the viewing device, operator performance, and environment (target, background, and illumination). The purpose of the tests described herein is to measure the visual performance of individuals and to establish from such measurements the average visual performance of a potential user population. To this end, the tests are now being given as part of a pilot vision survey which involves a sample of a Middle Eastern population. This work has been subcontracted by The University of Michigan to the American University of Beirut, Lebanon, where N. A. Haddad, M. D., and B. Faris, M. D., are the Principal Investigators. The work began on 1 January 1968 and is expected to continue through at least 31 December 1968.

BLANK PAGE

ABSTRACT

This report introduces some tests of night and day vision which, though devised for a special purpose, are potentially useful in a much wider variety of applications. The tests are mostly of binocular visual acuity for far vision, measured either during the course of dark adaption or after adaption to various light levels has been completed. In addition, a special form recognition test has been developed to investigate the relationship between visual acuity data and the visibility of a "real-life" shape, namely, the silhouette of a man.

The tests are implemented via specially developed slides to be used in a commerical vision test instrument. The latter has been adapted to provide a range of discrete nighttime and daytime light levels (through the use of neutral density filters). Unlike the slides normally provided with this instrument, those specially developed cover a wide range of contrast values. Thus, acuity may be tested over a wide contrast range at each of the nighttime and daytime light levels.

While the tests must still be conducted in a darkroom, having the tests implemented as they are enables the darkroom to be considerably smaller and more primitive than would otherwise be necessary, and the tests and darkroom can even be made "portable." Basically, this results from the fact that the instrument used is a closed stereoscope type device which produces the far vision test distance (20 feet) by otpical means.

The report covers the design and development of the tests, presents some results obtained with them, and gives some estimates of the sample sizes required to make comparative population surveys, the purpose for which the tests were originally intended. Detailed equipment operating instructions and test procedures are not included, however, since they are not germane to the discussion. Instead, these are covered in another document designed to accompany the test equipment.

BLANK PAGE

CONTENTS

FIGURES

SOME TESTS OF NIGHT AND DAY VISION

1
INTRODUCTION

The purpose of this report is to introduce some tests of night and day vision which, though devised by The University of Michigan to meet a special need, stand on their own as being potentially useful in a variety of applications, both military and civilian. The tests are mostly of visual acuity for far vision, measured either during dark adaptation or after adaptation to various light levels has been completed. These employ a conventional type of acuity target, the Landolt C, as the test stimulus. In addition, a special form recognition test has also been developed to investigate the relationship between (Landolt C) acuity data and the visibility of one "real-life" shape, namely, the silhouette of a man. All the tests are implemented via "slides" to be used in a commercial vision test instrument, the Bausch and Lomb Modified Ortho-rater, which has been adapted to enable testing at various nighttime and daytime light levels. (The latter are achieved through the use of neutral density filters.) Unlike the acuity test slides normally provided with this instrument, the specially prepared Landolt C slides cover a number of (negative) contrasts which not only include the conventional "black-on-white" case but also range almost to zero contrast (i.e., "white-on-white"). Besides providing a more comprehensive means of assessing the effect of those physiological mechanisms which limit "contrast perception" (as distinguished from those which limit "size perception"), the use of a variety of contrasts more nearly accords with what is encountered in everyday seeing. Also, as will be explained in more detail later, by making the human form slides of the same contrast as the Landolt C acuity slides, an unusual test of a form recognition principle has been enabled.

The primary reason for implementing the tests via a Modified Ortho-rater was to take advantage of the main feature this instrument is designed to provide, a reduction in the distance actually needed for testing far vision (which entails a commensurate reduction in target sizes also.)* By using stereoscope optics and slides, this instrument makes the test objects appear to be at a "far" distance of 20 feet though in fact they are only about 1 foot from the observer. Although the tests must still be conducted in a darkroom, the latter can be considerably smaller

---

* For reasons which will be made evident later, only the ability to see objects at far vision distances was of interest; accordingly the tests described were so restricted. This incidentally permitted cannibalizing parts from the near vision portion of the Modified Ortho-rater for use in altering its light source to achieve the range of light levels desired. While the near vision portion of the machine was thus rendered inoperative, this need not have been the case. Indeed, the near vision portion could also have been adapted to achieve a range of light levels, had such been desired.

and more primitive than would otherwise be necessary, and the tests and darkroom can even be made "portable."

Besides the special test slides mentioned above, all the far vision test slides normally provided with the Modified Ortho-rater can still be used. These enable tests of lateral and vertical phoria, acuity (with checkerboard type targets), depth perception, and color discrimination. In addition, all those slides used in the "Armed Forces Vision Tester," a more elaborate type of "Ortho-rater," can also be adapted for use in the Modified Ortho-rater. (The Armed Forces Vision Tester was not used as the basic means of implementing the UM tests because it could not be as readily adapted to obtain the various light levels as could the Modified Ortho-rater. The same is true for other commercial instruments similar to the "Ortho-rater.") Finally, slides used for examining drivers license applicants and even school children are also available for use in the Modified Ortho-rater.

The immediate reason for developing the tests was to supply the means for conducting a vision survey and comparison of the U. S. and at least one other world population, namely, that of the Middle East. This in turn was to obtain information in support of a much broader UM study on night vision aids. More specifically, it was thought that significant differences might exist in the vision of various world populations which would at least be reflected in their performance with night vision aids and which might even necessitate special designs for each. The idea that various populations have different visual characteristics was based in part on anecdotal evidence, but more importantly on awareness of such things as the varying degrees to which populations experience dietary deficiencies and diseases which can affect their vision.

To explore the concept just outlined with regard to the Middle East population, a pilot "Vision Study Program" has been established jointly between The University of Michigan and the American University of Beirut, Lebanon. Though some reference will be made to this program in the following discussion, it is not the purpose of this document to describe it in detail. Rather, this document is more to explain how the purpose of such a program, the night vision aids application, led to development of the UM tests, and to indicate how the tests might be used in other studies involving other applications.

The basic principles underlying the tests are discussed next, in section 2. Following this, section 3 tells how these principles were translated into specific tests and reviews the major considerations involved in their implementation. Section 4 presents some results which have been obtained at UM and compares these with previous work which guided the development of the tests. Some preliminary AUB results are also presented in section 4 and these are compared with the UM results. Pursuant to using the tests in other studies, section 5 presents estimates of the sample sizes needed to determine whether various "population" differences exist. Finally, in section 6, some applications of the tests other than that connected with night vision aids are discussed.

2
PRINCIPLES UNDERLYING THE TESTS

2.1. VISUAL ACUITY

The basic principle on which visual acuity tests are founded is, the smaller the objects one can see, the better is one's vision. Such tests measure minimum visual angle,* which depends on the contrast between the test object(s) and the background, the general light level (usually specified in terms of the background light level), and many other factors. A typical wall-chart vision test displays only black silhouette targets against a uniform white background and involves only one light level, but this usually reflects only the convenience and simplicity of making up such a test: the test objects need not be silhouettes, the background and test objects could just as well be in shades of grey, and the background could be darker than the objects instead of lighter. Moreover, any two of the above parameters—visual angle, contrast, and light level—may be fixed and the third used as the dependent test variable to generate a threshold visibility surface in the three-dimensional "seeing space" defined by these parameters; this surface would divide the space into "see" and "no see" regions. Such tests are properly called visual acuity tests only if the visual angle is the dependent variable. If, instead, contrast were varied, then the terms "liminal brightness increment" or "contrast threshold" might more properly be used to indicate the nature of the test. (There seems to be no special terminology for the use of light levels as the dependent test variable.) No matter which parameter is the dependent variable, however, the same threshold surface would presumably result. Thus, to speak of minimum visual angle would logically be no different than to speak of threshold contrast or threshold brightness. (This statement must be qualified where the visual angle is such that only one retinal receptor is excited or where the large-area contrast approaches a certain low value established by the physiology of the eye; nevertheless, it remains a useful generalization and is appropriate for the discussion which follows.) It is in this extended sense that the tests measure visual acuity; i.e., they employ visual angle as the dependent variable to generate a threshold visibility surface over a wide range of contrasts and light levels.

The fact that the angle measured is a minimum angle is basically the result of deliberately designing or selecting each test object used so that it has some critical feature which must be "just seen" before the object itself can be "just seen." If all acuity tests made use of objects

---

*Visual angle is defined as the angle (usually in minutes of arc) subtended at the eye bv some linear object dimension. For objects displayed at some fixed distance (as in a typical acuity test), the minimum visual angle is directly proportional to the minimum size one can see and thus serves as an equivalent measure of size. Actually, the technical definition of "visual acuity" is the reciprocal of the minimum visual angle. In any case, the distance involved is important. Disregarding the effect of the intervening medium on object/background contrast, minimum visual angle is essentially constant for distances of 20 feet or more (far vision) but not necessarily so for lesser distances because of differences in the degree of accommodation and convergence.

containing the same critical feature, then the same results would obtain regardless of which test was used; however, this is generally not the case, and it does not appear to be a simple matter to make it so. The common wall-chart test in which a variety of letters must be read is a case in point. Though all letters of the same size presumably are constructed to be equally visible, some are slightly more difficult to read than others. Moreover, there is evidence that such discrepancies are not constant but depend at least on light level [1]. The fact that these discrepancies exist, however, does not invalidate the idea that just seeing objects of this type is equivalent to just seeing some critical feature they possess; it simply emphasizes the improbability of being able to tell on an a priori basis just how critical any one feature is, even though it is known (by design) that the feature must in some sense be critical. As far as acuity tests are concerned, self-contained discrepancies can obviously be removed by designing a test which uses a single type of test object, and many such tests have been developed which do just this. Indeed, the UM tests were deliberately chosen to be of this type and for just this reason. Even so, the likelihood of establishing precisely how critical an object feature is has important ramifications not only in comparing the results of acuity tests but also in the more realistic visibility problems to be considered next.

2.2. REAL-LIFE VISIBILITY

If one were to apply classical silhouette-target visual acuity data to predicting the visibility of objects in natural scenes, then at least four additional discrepancies would be encountered:

(1) At light levels high enough for color vision, differences in color between objects and backgrounds would obviously come into play and might be the sole determinant of object visibility. This is not accounted for in the classical version of visual acuity which deals only with achromatic targets and backgrounds.

(2) Real-life objects would not usually appear "uniform" (i.e., would not be silhouettes).

(3) Real-life backgrounds would not usually appear uniform.

(4) Even if the objects and backgrounds were uniform, it still might be virtually impossible to know, a priori, which features of the objects were likely to be critical.

Because of the above, it is true that, in general, there is no simple relation between the classical version of visual acuity and the "ordinary seeing ability" it is often assumed to describe; however, it is also true that, were it not for items 2 and 3, at least an ad hoc relation could be established which would at most require a straightforward extension of the classical approach to include color.* For the achromatic case, this would result from the ability simply to equate the threshold visibility of some acuity target and of a real-life object's silhouette for the same background luminance and contrast conditions. For the chromatic case, the acuity

---

*If, contrary to 2 and 3, the objects were silhouettes and the backgrounds uniform, then only two colors could be involved, one for the target and one for the background.

target and silhouette would have to be of the same color as well, as would their respective backgrounds. In either case, from the known distance at which the real-life object's silhouette was presented as well as from the known visual angle subtended by the critical dimension of the acuity target, a critical dimension could be computed for the ordinary object. This dimension would then be an ad hoc property of the object seeing task for the particular contrast, background luminance, and color conditions involved even though there might well be no obvious feature of the object itself to which the dimension could be sensibly referred.

If the critical dimension so determined were then proved to be independent of contrast and background luminance, an invariant property of the object seeing task would have been established. Further, if this principle of invariance were truly valid for all such tasks, then any ad hoc critical dimension would be an invariant, and it would only be necessary to determine such a dimension for one set of contrast and luminance conditions. Once this was achieved for various object seeing tasks, the acuity data could be used as a standard to predict visibility under different parametric conditions, or to assess the practical effect of individual differences in seeing ability, or both. While objects other than acuity targets could be used, the fact that acuity targets have critical features that can easily be identified makes them most attractive for this purpose. Put another way, to be able to determine an unknown critical dimension by using one which is known is especially satisfying. Because the known critical dimension of an acuity target is subject to the vagaries described above, however, the critical dimension of the object would not be independent of the kind of acuity target used in its definition.

2.3. FORM RECOGNITION

Although invariance with respect to contrast and background luminance seems never to have been explored for the chromatic case, previous achromatic-case experiments indicate that ad hoc critical dimensions may indeed be independent of these parameters. Because only the achromatic case is of particular importance in the night-vision-aids context, the UM form-recognition tests were restricted to being of this type and were developed in part to see if a sounder basis for the invariance could be established for at least one silhouette of interest, that of a man.

Another reason for developing the form-recognition tests was the following: Since seeing even silhouettes of ordinary forms involves perceptual and cognitive processes of a higher order than those used in seeing the more primitive acuity targets, it seemed possible that differences in individual ability or capacity for such higher-order processes might be revealed from such tests; i.e., individuals with the same acuity might have to see different critical dimensions for the same object seeing task. As will be explained later, the attempt to achieve a relatively pure measure of such effects partly accounts for the forms used in the UM tests. The extent to which higher-order processes are involved may be better appreciated by considering how the nature of the seeing task determines the critical dimension for a given object. Consider, for example, the case of a silhouette of a man. If an observer had only to sense the presence of this form as something in one of a number of space-time locations, with nothing in the other locations, then it is easy to imagine that a comparatively large critical dimension would result. (Though the observer might be able to see more than "something," this would be irrelevant.) On the other hand, if an observer had to recognize a silhouette of a man amidst a variety of somewhat similar forms (as in the UM test), then a smaller critical dimension would be expected. The more similar the other forms were to the man, the more discriminating the observer would have to be, and the smaller the expected critical dimension. For objects at a fixed distance, the visual angle required would, of course, become smaller as the discriminations grew more difficult and would thus serve as a measure of the degree of difficulty of the task. However, different tasks could be of equivalent difficulty provided the objects involved were placed at different ranges, since it is the critical dimension and not the degree of difficulty that is defined by the task.

2.4. NIGHT-VISION AIDS AND THE SINE-WAVE-RESPONSE APPROACH

Thus far, the discussion has been concerned only with the classical concept of visual acuity and what can be done with it. Despite some lingering uncertainties, it has been suggested that at least a primitive connection can be established between classical acuity data and "reality," in terms of the critical dimensions of real-life silhouettes. Assuming this to be true, what remains to be established is the appropriateness of the approach to the basic purpose for which the UM tests were intended: obtaining data useful for predicting operator performance with night-vision aids. For this purpose, it is necessary to consider the appearance of scenes when viewed with night-vision aids and the way such devices operate on the scenes. (This subject will lead to a discussion of a modern version of visual acuity, the sine-wave-response approach.)

The night-vision aids of primary concern are those which provide a real-time viewing capability. Besides ordinary optical aids such as binoculars or telescopes, these include electro-optical aids such as television systems (which allow remote viewing) and direct-view image-intensifier devices. Whereas optical aids improve object visibility essentially in only one way, by optical magnification, electro-optical aids can also raise the apparent scene luminance by electronic amplification and may amplify the contrast as well.* Color imagery is also possible with electro-optical aids, but the added complexity and cost of providing this feature generally

---

*For image-intensifier devices, this would result from deliberately selecting an input spectral sensitivity such that for given target and background spectral characteristics, the apparent contrast of the target would be increased. Besides the possibility of employing this approach, television systems can amplify contrast electronically. In either case, however, the input spectral sensitivity coupled with the spectral characteristics of the scenes would determine the input contrast to the device, and this, by accident or design, might not be the same as the direct visual contrast.

has been regarded as unjustified for most military needs; moreover, at low nocturnal light levels (e.g., starlight), the optical filters used at the input of electro-optical aids to sample colors reduce the input signal to such an extent that color imagery is marginal if not impossible for state-of-the-art detectors. This subject has been discussed by Paprocki and Miller [2] for the case of image intensifiers aids. They point out that such devices can now be made to produce colors at intermediate moonlight levels. For ordinary optical aids, natural nocturnal light levels are not in themselves high enough to enable a viewer to see colors, but the active use of supplementary artificial illumination is often undesirable, at least in the military context. Due to the kinds of target-seeing tasks of greatest military importance, most targets would be at a far-vision distance when seen and would naturally appear to be so to an observer using an optical aid. While the latter is usually not the case for television systems and need not be the case for image-intensifier devices, both may be equipped with viewing optics to make it so, and, indeed, such is usually the case for image-intensifier devices, which are of greatest interest to this study. (If television systems do not make objects appear at their true far distances, neither do they make them appear at what is conventionally used in acuity testing as a near-vision distance, i.e., a normal reading distance; in fact, the acuity tests most appropriate to viewing television under conditions similar to those in the home would be of the far-vision type.)

Although these features do not cover every conceivable type of night-vision aid or every viewing situation, they do apply to the vast majority of cases. Further, they readily translate into gross requirements for acuity testing: that the scenes be achromatic; that they require the use of far vision; and that a range of light levels, contrasts, and visual angles be covered. With regard to light levels, it should be emphasized again that whereas optical aids essentially do not affect the apparent scene luminance, electro-optical aids are capable of raising the apparent luminance to daytime values; thus, to account for electro-optical as well as optical aids, acuity at daytime as well as at nighttime light levels must be considered.

We now turn to the areas where the use of night-vision aids conflicts with the conventional acuity-test approach. When a silhouette target is viewed through a night-vision aid, it appears no longer to be a silhouette, because all night-vision devices characteristically introduce a certain amount of blurring; the edges of the form appear less sharp, or separable parts of the form appear to be merged, or both. This is in addition to any geometric distortion or lack of uniform response. The visibility of conventional acuity test silhouettes, therefore, cannot be exactly matched to the visibility of the blurred and possibly nonuniform or distorted images of such targets, and the use of the conventional data to predict operator performance with a night-vision aid would be similarly inexact. Although geometric distortions can be eliminated by design, and lack of uniformity might have a negligible effect, there would still be no solution to the problem of blurring unless one additional condition were satisfied: If the device were linear, i.e., if the overall output luminance were linearly related to input radiance (or luminance), then the sine-wave-response approach could be employed. This approach is based on the fact that a spatial luminance distribution, modulated along some direction in a sinusoidal manner, would be imaged as such; i.e., its contrast might be diminished and some spatial phase-shifting encountered, but its image would still be a sine-wave pattern. Thus, acuity data derived from spatial sine-wave patterns could be exactly matched to the output images of such patterns. (If the sine-wave pattern were finite in extent and its boundaries were within the field of view, then each 180^0 phase shift would also result in the loss of one fringe. This could be important to the comparison with acuity data if only a few fringes remained.) Further, since many optical and electro-optical aids are linear, this solution could, at least in principle, have widespread applicability.

Given that the acuity problem can thus be solved, what of the visibility of real-life shapes? For unaided* (i.e., well-corrected) vision, sine-wave-pattern acuity might be substituted for any of the classical types of acuity data in deriving critical dimensions for silhouette forms, but what of aided vision where such shapes are distorted? In principle, these questions could be answered as follows. A device may be completely characterized (over its linear operating range) by its ability to pass spatial sine-wave patterns. This may be done in terms of an optical transfer function, which indicates how both the contrast and the phase of spatial sine waves of any frequency are affected by the device. (That part of the optical transfer function which gives contrast information alone is called the modulation transfer function.) If, as is often the case, there were no phase shifting, then the transfer function could be compounded with sine-wave acuity data in such a way as to unambiguously indicate the object-space contrast required to see sine waves of various frequencies, i.e., to produce a kind of aided visual acuity data. This, together with the fact that the luminance distribution representing an object may be decomposed by Fourier analysis into an entirely equivalent spatial sine-wave-pattern content, means that the results would be the same as if unaided acuity were changed. Thus, aided visual acuity data could be coupled with the critical dimension of a silhouetted form (determined from sine-wave acuity) to predict the visibility of that form under changing parametric conditions.

The spatial sine-wave-pattern approach was strongly considered as the means for performing the tests. Though it was rejected principally because of the practical difficulties involved in making such patterns, there were a number of other reasons as well. First, the fact remains that some devices are not linear. Moreover, those that are often exhibit nonuniform response and geometric distortion as well as phase shifts. In addition to such loss of generality due to device "perturbations," serious questions can be raised about the way sine-wave acuity tests are conducted. Specifically, while the conventional "method-of-limits" approach to obtaining sine-wave acuity may be satisfactory for some laboratory work, it does not seem ap-

---

*The term "aided" does not refer to the wearing of eyeglasses but rather to the use of optical or electro-optical aids.

propriate for survey tests.* While this problem could in principle be solved by techniques analogous to those used in the classical acuity tests, there was no time to develop a specific alternative. More important than these limitations, however, was the question of how exact the tests should be. As discussed in more detail in section 5, the statistical spread in any type of acuity data requires that larger and larger samples be taken to detect smaller and smaller differences in acuity; in fact, the samples can become impractically large. It seemed evident that only large differences would be of interest and that therefore the conventional acuity test employing silhouetted forms would be adequate.

2.5. DYNAMIC VISUAL ACUITY AND OBSERVATION TIME

Though the preceding discussion essentially covers the principles underlying the UM tests, a few more points deserve mention. Whereas the classical silhouette approach can only marginally be extended to the case of nonuniform (but achromatic) targets and backgrounds (by using averages), the sine-wave-response approach is capable of more completely describing such situations, at least in principle, for linear devices. Even so, the nonuniform case is itself not especially relevant to the problem at hand simply because no particular nonuniform scene is likely to be common to different geographic areas and populations; moreover, even if such a scene were identified, analyzing the visibility of the objects within it by means of the sine-wave approach would be an unnecessary complication. Neither the visibility of moving targets nor the amount of time allowed for observing targets has been discussed. With regard to the former, the tests involve only static targets, primarily because there was no time to devise or give tests with dynamic targets, but also because it was not at all clear what kinds of motion should be given these targets. As for the observation time, whereas acuity is affected when the time is less than approximately 0.2 seconds [3], there was no compelling practical reason to base the tests on this fact and no simple method of implementation if there <u>were</u> reason. Hence, subject to completing each observation in a "reasonable" amount of time, the tests allow as much time as is needed. (Based on the AUB program, the maximum time allocated for giving the tests is four hours.)

3
IMPLEMENTATION OF THE TESTS

3.1. SUMMARY

Having decided that a conventional acuity test should be used, there remained the problems of selecting the type of target and the procedures and equipment for implementing the

---

*The method-of-limits approach requires the observer to tell if he can see something he knows, a priori, to be present. A better approach would be to require the observer to pick out a sine-wave pattern from several alternatives, much as an observer has to pick out the checkerboard from several other greyish squares in the standard Ortho-rater acuity test.

tests. Consistent with the goal of obtaining data for predicting operator performance with night-vision aids, the range of light levels which might profitably be covered had first to be established. By referring to published data on natural nocturnal illumination levels, the reflectivity of natural backgrounds, and the luminance gains obtainable for various types of electro-optical aids, representative values of apparent scene luminance were deduced for electro-optical as well as optical aids. These values ranged over 7 orders of magnitude from a low of $\sim 10^{-5}$ mL, corresponding to the optically aided (or unaided) viewing of green grass in starlight, to a high of $\sim 10^{2}$ mL, corresponding to the electro-optically aided viewing of snow in the light of a full moon. The latter assumes a representative maximum luminance gain of $\sim 10^{4}$ for electro-optical aids based on what can be done with certain image-intensifier devices. With this luminance range as a guide, the next step was to determine how much of this range had been covered by conventional acuity tests and of what these tests consisted. While it was found that tests had been performed over portions of the range using a variety of conventional targets, it was also found that these tests usually involved only one (high) contrast, which may be designated black-on-white. Two well-known exceptions, however, are the tests by Blackwell [4, 5] and by Conner and Ganoung [6] and Cobb and Moss [7, 8]. Whereas Blackwell made use of disk-shaped test targets, Conner and Ganoung used Landolt C targets at low light-levels to "extend" the acuity measurements which Cobb and Moss had made at high light-levels with two-bar patterns. Both of these tests include the black-on-white case, but they also cover lower contrast values, ranging almost to zero. In addition, Blackwell [4] shows that for the same background luminance and disk size, disks in negative contrast (dark-on-light background) and disks in positive contrast (light-on-dark background) are about equally visible.* A similar conclusion for Landolt C targets was reported by Vos, Lazet, and Bouman [9]. (Their tests also covered a broad luminance and contrast range and would also have provided a useful precedent except for ambiguities in their report in connection with specifying the luminances used.)

Since it had always been intended that more than one contrast be included in the tests (this was first suggested by Dr. Haddad in his July 1966 Proposal), the next step was to decide which of the above precedents to use, i.e., whether to employ disk or Landolt C targets. This is not to say that some other conventional approach could not have been considered, but rather that there was no compelling reason to do so. Both approaches seemed quite compatible with one implicit requirement of the AUB tests—that they be free of literacy and linguistic barriers (e.g., that they not use letters of the alphabet). Moreover, it was evident that an equally valid test could be given with either approach; both use an identical procedure in which less than an even-chance opportunity for success can be associated with every correct test response. For

---

*The equivalence of positive and negative contrast seems likely to be subject to the same qualifications as mentioned in connection with visibility threshold surfaces (i.e., more than one retinal receptor must be involved, etc.); however, for most practical situations, this remains a useful generalization.

the Landolt C target, this means that the subject would have to tell the actual location of its gap when the gap has been randomly placed in one of more than two fixed locations (e.g., top, bottom, right, or left). Since the ring portion of the C is, by design, more visible than its gap, the Landolt C essentially provides its own self-contained reference frame for this purpose. On the other hand, a separate (more visible) reference frame must be provided for the disk to create the same kind of test situation. Because of practical implementation problems and some uncertainties associated with this separate reference frame, the disk approach was ultimately eliminated.

To know that a test score represents the best an observer can do requires that the test procedure provide some means of detecting "deceptions," whether they are accidental or deliberate, and whether they lead to better or worse results than those representing an observer's true abilities. Deceptively good results can be guarded against by adopting a procedure in which there is less than an even-chance opportunity for making a correct response. There is no procedure based on the design of the target for guarding against deceptively bad results. Whereas the detection of deceptively good results is often all that is considered important (e.g., in drivers license or job applicant tests), this need not be the case, and the provision of both kinds of safeguards for the tests was initially a matter of some concern. In deciding on such matters, however, consideration had to be given to the fact that, even without any safeguards (as, for example, with the method of limits), it has frequently been demonstrated that repeatable results can be obtained [7, 8]. Though such results do not necessarily represent the best that can be done, neither do they represent the worst, and, in any event, they provide one means of making population comparisons. To be sure, there would be no way of proving that those results obtained with [4, 6] and those obtained without safeguards represent the same kinds of response for each of the populations being compared. Nevertheless, as long as there is nothing about the test situation to motivate deliberate deception, it would seem unlikely that deception would be prevalent. So-called objective techniques, by means of which deceptively bad results (as well as deceptively good results) may presumably be detected, have apparently been restricted to the use of involuntary physiological responses (e.g., optokinetic nystagmus and electrical signals from the cortex) as indicators of subjective perception. (The use of hypnosis or "truth serum" might also be contemplated but would seem manifestly unsuitable for survey testing.) Though consideration was given to using involuntary physiological responses, this was rejected as being an unnecessary complication. The procedure finally selected was the type which, by the nature of the stimulus used, guards against deceptively good results; it was chosen over the method-of-limits approach because it extended the potential applicability of the tests to those situations where such a safeguard would be important.

3.2. SELECTION OF THE MODIFIED ORTHO-RATER

It was immediately apparent that to use an actual far-vision-test distance (e.g., 20 ft) and to conduct the tests over the range of light levels being considered would require a specially designed, carefully controlled, and relatively large darkroom facility. It did not seem likely that such a facility could be easily acquired for the AUB case or for any other. Fortunately, rather than having to face this problem, an ideal alternative seemed at hand. There are a number of instruments available for testing daytime acuity which reduce the far-vision-test distance by employing the optical stereoscope principle—stereoscope optics and slides make the test objects appear to be at an appropriate far distance (e.g., 20 ft) while a smaller observer-target separation (e.g., 1 ft) is maintained—and while none of these instruments come ready-made with provisions for testing acuity over various contrasts and light levels, it was clear that any one could probably be so adapted. In deciding which of the instruments to select, consideration was given to the best method of changing the light levels. While the light output of the bulb used to back-light the slides could have been varied directly by changing the voltage applied to the bulb, this would also have changed the color of the light and would have been hard to control because of the highly nonlinear dependence of bulb luminance on applied voltage. A better and more natural solution was to use neutral-density filters to attenuate the light reaching the slides from the bulb (or, more correctly, from the intervening diffusing plate). The instrument which most readily lent itself to this approach was the Bausch and Lomb Modified Ortho-rater. Whereas other instruments effectively limit the number of slides which could be used (by putting them on a rotating drum around the bulb diffusing plate), the Modified Ortho-rater features manual insertion of each slide and thus permits showing any number of slides without disassembling the instrument. Since it was evident that the same feature could easily be provided for the neutral-density filters (because of the instrument's box-like structure), and since the number of new slides and filters needed was initially an unknown factor, the Modified Ortho-rater seemed a clear choice.* It was clear also that new slides would be needed. Even if the tests had been arbitrarily restricted to the single black-on-white contrast level of the acuity slides provided with the instrument, these would have covered too few visual angles over too small a range to have been of much use. Moreover, it was evident that the best solution to the problem of how to obtain a range of contrasts was simply to make slides with targets at various contrasts. While consideration was also given to using the well-known veiling luminance technique for obtaining reduced contrasts, implementation of this technique would have complicated both the equipment and the administration

---

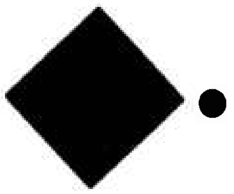
*Had it been decided to test over considerably fewer light levels (and contrasts), one of the other instruments might have been chosen and the reduced light levels achieved by physically combining neutral-density filters with the test slides. Indeed, this is essentially what was done to one drum-type instrument, the Armed Forces Vision Tester [10] (itself a military adaptation of a more elaborate version of the Ortho-rater), to produce the Army Night Seeing Tester, Model I [11]. The latter is understood to have been a device which provided limited tests of contrast discrimination and resolution, but at only one reduced light level (in the mesopic vision range).

of the tests; further, assuming the test answers were to be changed from contrast to contrast, it would not have reduced the number of slides required.

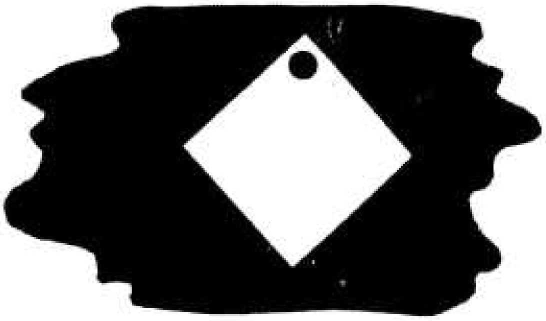
By choosing a compact, box-like, stereoscopic instrument, the size and baffling problems associated with a darkroom were avoided (though baffles were still required for the interior of the instrument), and the light integrity of the instrument could be readily ensured. This, together with the fact that, when the observer is looking into the instrument, his peripheral vision is largely obstructed, meant that although the tests would still have to be conducted in a darkroom, the light integrity of the darkroom could now be somewhat less than perfect. Thus, since the tests could be made portable simply by adding a self-contained, portable power supply, it was apparent that they could even be conducted outdoors where the "darkroom" would be the sky on a dark night. Though choosing this method meant that only one subject at a time could be tested, this seemed a small price to pay for the versatility obtained.

3.3. SELECTION OF THE TARGET AND QUANTIFICATION OF THE PARAMETERS

Since the selection of a disk target would require a separate reference frame, it was evident that this frame would have to be provided on the slides and be different from the type of frame that Blackwell [4] had used, and this would introduce uncertainty about the applicability of Blackwell's data. (Blackwell used a small red light at the center of an imaginary circular "orbit" along which the disks were placed). Figure 1 illustrates the problem: An n-sided



(a) Interior Reference Frame



(b) Exterior Reference Frame

FIGURE 1. EXAMPLE OF THE TYPE OF REFERENCE FRAME NEEDED FOR DISK-SHAPED TARGETS

regular geometric polygon would have had to be placed in either an interior or exterior relation to the disks, with the corners used to mark the proximate disk locations and with the value of n determined by reliability considerations. Even though a square exterior frame (fig. 1b) is used in one commercial instrument, it was decided that, to increase reliability,* an octagonal frame should be used for which no real precedent existed. More importantly, since there was not sufficient time to experimentally settle questions on the relative sizes and proximity of the reference frames and disks, it was thought better to use the Landolt C target, for which no real problems of this type existed and for which an n = 8 precedent did exist [6].

The next step was to decide what background luminance values, contrasts, and Landolt C sizes should be included in the tests. In this connection, it is important to recognize that selection of the method of implementation described above meant that each of the threshold parameters would be a discrete variable. To review, the light levels would be set by neutral-density filters and the contrasts and sizes by what would be on the slides themselves. The significance of discreteness is that it leads to built-in systematic errors in the tests. Therefore, selecting values for the parameters reduced to deciding how much systematic error would be tolerable; this in turn had to be weighed against practical problems such as limiting the number of filters and slides for purposes of handling. In order to size the built-in systematic errors, first two of the three threshold parameters had to be selected and then the values for these two parameters for various values of the third parameter had to be chosen. This amounted to choosing the most important two-dimensional orthogonal projection of the visibility threshold surface; visual angle versus background luminance (for various contrasts) was selected because these parameters vary when a natural scene is viewed, whereas contrast is essentially an invariant of the scene. Figure 2 shows the data in references 6-8 plotted this way. (For the data in refs. 7 and 8, the longest exposure time, 0.3 sec, was used since it was assumed that this gave the same results as if an indefinitely long time had been used, as was the case for the data in ref. 6.) Actually, the curves in figure 2 are interpolations from figure 3, where the data are more accurately represented by the projection contrast-versus-visual angle (for various background luminances). For both figures, contrast is expressed in terms of modulation:** this

---

*Test reliability may be understood as follows: If an observer were to surpass his true threshold by guessing, the chance of his exceeding his threshold by some amount would be 1/n for the first try, $1/n^2$ for the second try, and so on.

**Modulation, M, may be defined in terms of test objective luminance, B_T, and background luminance, B_B, as $M = (B_T - B_B)/(B_T + B_B)$. So defined, M is a qunatity which varies between -1 and +1 and does not change in kind with a sign reversal. Because of the assumption that Landolt C's in positive and negative contrast are equally visible, only absolute (i.e., always positive) values of M are considered in the discussion. In converting the Cobb and Moss [7, 8] version of contrast to modulation, their original data were also corrected to remove an apparent discrepancy which resulted from using a set of reflectivity values for computing contrast which differed from those stated to have been available.

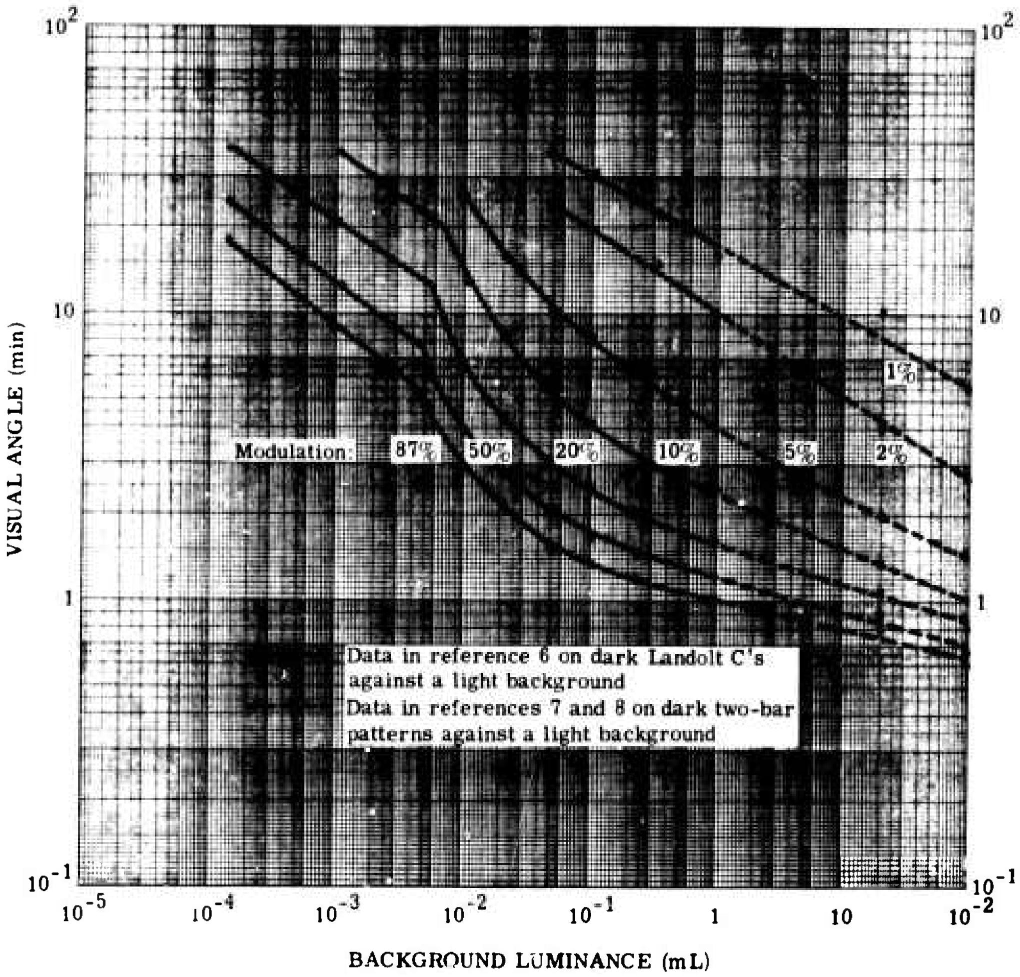


FIGURE 2. VISUAL ANGLE VERSUS BACKGROUND LUMINANCE FOR VARIOUS MODULATIONS [6-8]. These curves are based on interpolations from the data in figure 3.

differs from the expression used in references 6-8, but the conversion from the tabulations of the data presented there was easily and accurately made. (Aside from the conceptual theoretical advantages of using modulation, this version of "contrast" was chosen because it seems to be emerging as a standard.) The modulations given in figure 2 (100, 50, 20, 10, 5, 2, 1%) are the nominal values of those actually selected for the tests; though these values were chosen in part to be suitable for reproducing figure 3, the selection was otherwise arbitrary except for 1% and 2%, which are sometimes used in describing practical resolution limits for certain types of electro-optical aids.

The luminance values and visual angles were conjointly selected to enable reproduction of figure 2 and to realistically minimize the visual-angle error; i.e., the burden for systematic error was placed on visual angle. In achieving these goals, the following factors were considered: the light output of available bulbs, whether the slides should be in positive or negative contrast, and the density of available filters. When it was learned that the maximum output of available bulbs would give a luminance for the intervening diffusing plate no higher than the

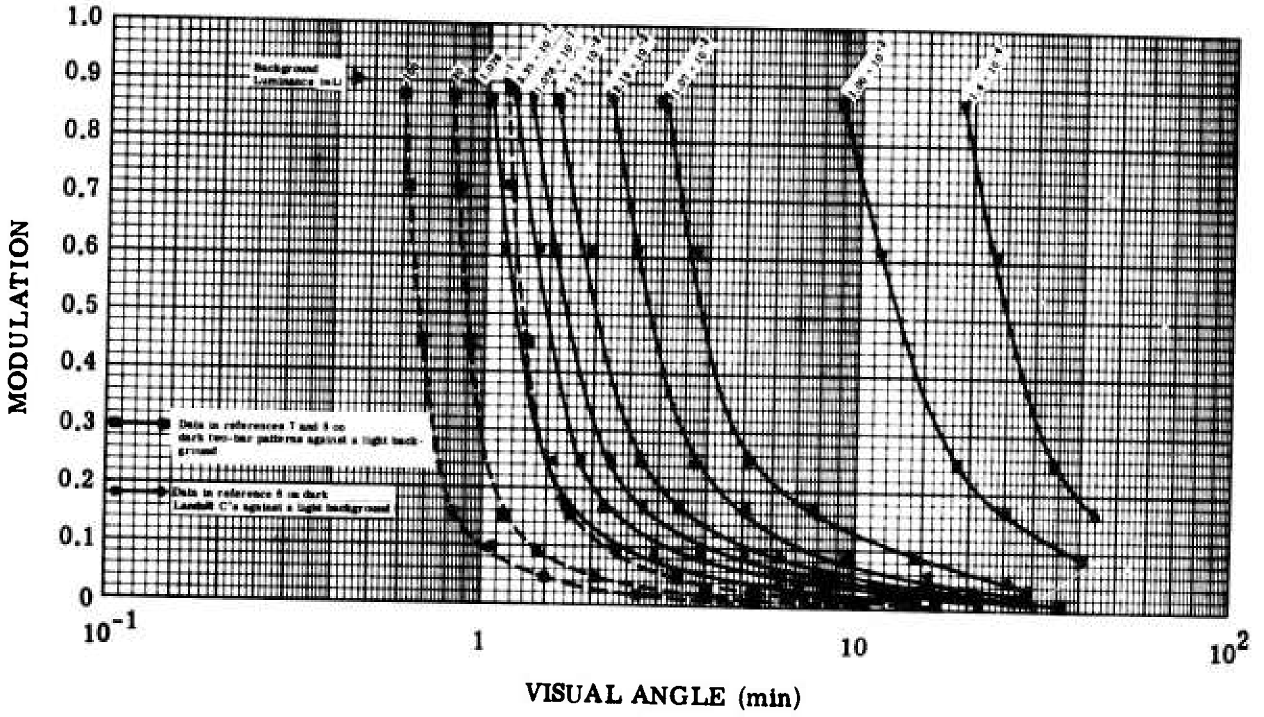


FIGURE 3. MODULATION VERSUS VISUAL ANGLE FOR VARIOUS BACKGROUND LUMINANCES [6-8]

highest value being considered for the tests (~ 10^2 mL), it became evident that the slides would have to be in negative contrast, i.e., the targets darker than the background and the density of the lighter, more transparent background the same low value (fog density) for each slide. Since neutral-density filters were most readily available in (diffuse) optical densities which varied in unit steps (transmission in decade steps), and since sampling in decade steps seemed to give reasonable coverage, this was the approach selected. To make room for the filters (which when inserted into the machine were placed in close contact with that side of the diffusing plate closest to the bulb) and to prevent their becoming overheated by the bulb, the bulb was moved away from the filters along the optical axis to a distance of approximately 3 1/4 in. To counterbalance the decrease in luminance which resulted from this change, the wattage of the bulb enclosed in the Ortho-rater was increased from 25 to 40. These things taken together resulted in the following luminances: 55 to 60 mL luminance maximum × 10^{-D}, where D is the (diffuse) optical density of the filters used and D = 0, 1, 2, 3, 4, 5, 6, 7. The range 55 to 60 mL accounts partly for variations in luminance due to innate manufacturing differences in the bulbs and partly for some aging of each bulb before it is replaced; i.e., 55 mL represents an arbitrarily selected lower acceptable limit for the tests. (A photometer is supplied with the equipment to check this.) The color temperature of the light from the bulb is approximately 2800°K. An internal microswitch in the Ortho-rater which operates only when a filter is completely inserted protects the observer's state of adaptation to darkness and accounts for the necessity of a zero-density filter.

The next problem considered was how close the visual-angle steps should be. The following scheme was adopted. Starting with 10 minutes of arc, the smaller angles differ by $\log_{10}$ unit steps of 0.05 (a factor of 1.122), whereas the larger angles differ by $\log_{10}$ unit steps of 0.10 (a factor of 1.259). Though use of the smaller (0.05 log unit) steps above 10 minutes of arc would have produced more accurate results in this range, the larger (0.10 log unit) steps were selected because (1) the additional accuracy did not seem worth the effort of making the extra slides that would have been necessary, and (2) the larger steps still fit well with the expected data. (Predictions for the data were based in large part on linear extrapolations from the data in ref. 6.) The systematic error thus established as tolerable can be understood by noting that the actual threshold could be within a step and, in a limiting sense, as large as a step. Further, since the angle subtended by an object varies inversely as the range of that object, selecting steps as indicated was equivalent to saying that predictions of the range at which an object could be seen might differ by the factors given and still be acceptable (see fig. 4). Besides displaying the parametric relationship between object size, angular subtense, and object range, figure 4 shows the actual visual-angle steps selected as well as some representative values for the critical dimensions of selected types of objects. The decision to use the larger steps was in part based on an assessment of the importance of the range differences for these selected objects. (The critical dimensions shown in figure 4 are based on some unpublished work by the author.)

Thirty-five Landolt C acuity slides were prepared consisting of five slides for each of the seven modulations selected (100, 50, 20, 10, 5, 2 and 1%). Each such set of five slides displays Landolt C's in the same way; i.e., C's of the same size (though of generally differing orientations) are displayed in the same order with the same number on each slide. A total of 35 C sizes were used which yield visual angles ranging from 0.794 through 79.4 minutes of arc in steps determined by the scheme described above. The exact visual angles, their Snellen rating equivalents, and the orientations used for the C's are given in table I. The paired modulations in the table result from the original intention to test each subject over only one of the following sets of modulations: 100, 50, and 5%; 100, 20, and 2%; or 100, 10, and 1%. This in turn was the result of overestimating the time required to test a subject over one of these sets. Consequently, to test subjects over all the modulations and to avoid carry-over from one modulation to the next, slides having C's of the same orientation are best given in different sessions.

As may be inferred from the table, when there were more than two C's on a slide, they were arranged in horizontal rows with their sizes diminishing from left to right in the top row, then from right to left in the next row, and so on. Thus, to observe the C's in consecutive order, they must be viewed in zig-zag fashion. This arrangement was adopted to permit close packing of the C's on the slides and to help minimize the total number of slides to be

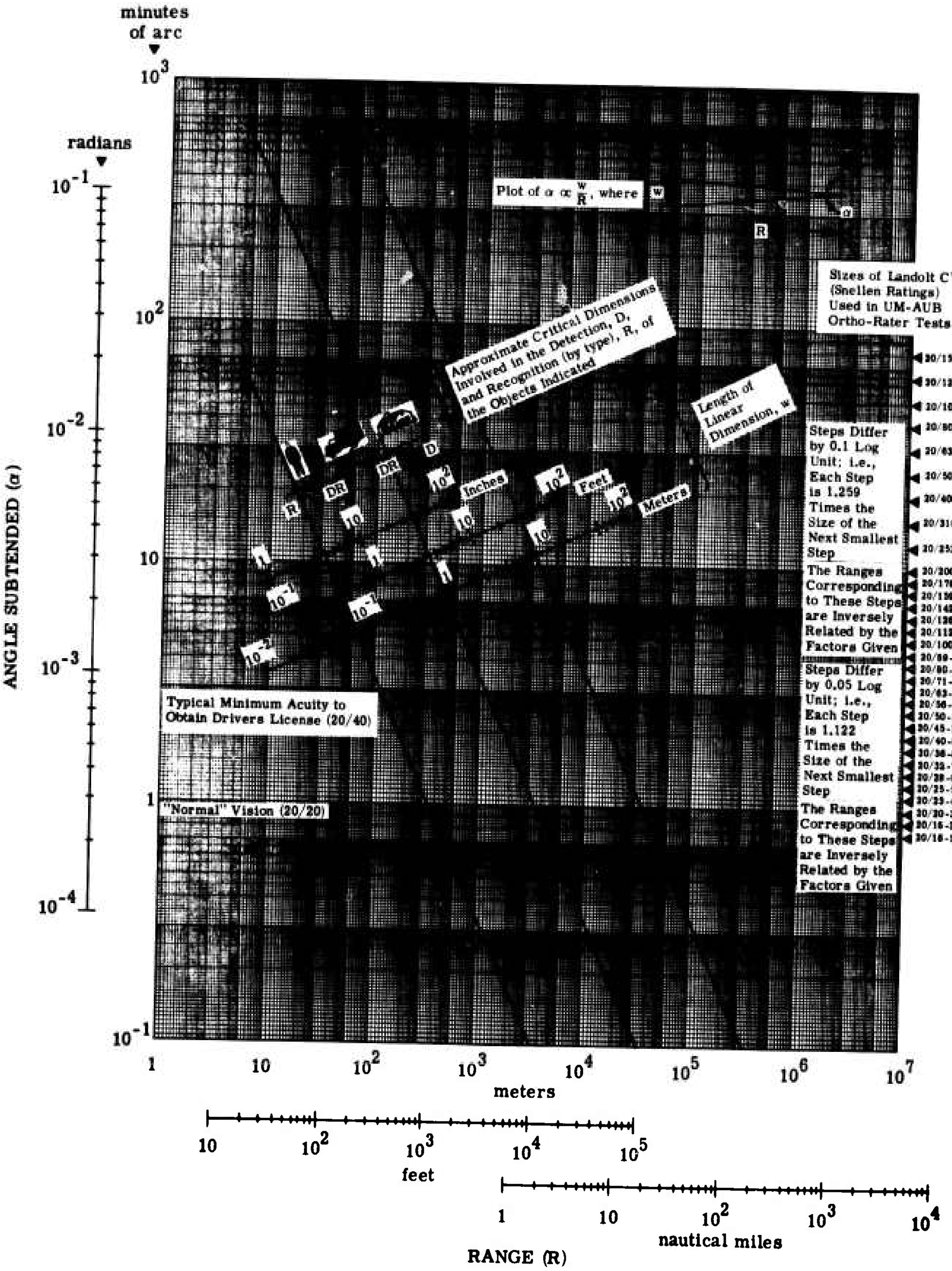


FIGURE 4. ANGLE α SUBTENDED BY A LINEAR DIMENSION w AT A RANGE R. Also illustrated here are practical visibility requirements and the coverage of visual angle provided by the UM-AUB acuity tests.

TABLE I. BASIC CHARACTERISTICS OF THE LANDOLT C ACUITY SLIDES

Slide Number	Step Number	Nominal Snellen Rating	Nominal Visual Angle (min)	Layout*		Orientations for Modulations Indicated			
						100%	50%, 2%	20%, 1%	10%, 5%
1	1	20/16	0.794	Bottom row	R	SE	E	W	S
	2	20/18	0.891		↑	SW	S	NE	SE
	3	20/20	1.00		↑	W	NE	N	S
	4	20/23	1.12		L	E	SW	NW	W
	5	20/25	1.26	Middle row	L	N	W	SE	SW
	6	20/28	1.41		↑	SE	NE	SW	NW
	7	20/32	1.58		↑	N	SE	S	E
	8	20/36	1.78		R	S	SW	SW	SE
	9	20/40	2.00	Top row	R	W	E	SW	N
	10	20/45	2.24		↑	SW	N	S	S
	11	20/50	2.51		L	W	E	NE	S
2	12	20/56	2.82	Bottom row	L ↑	SW	S	N	SW
	13	20/63	3.16		R	NE	W	N	NW
	14	20/71	3.55	Lower-middle row	R	S	N	SW	W
	15	20/80	3.98		↑	SE	SE	NE	E
	16	20/89	4.47		↑	SW	SW	SW	W
	17	20/100	5.01		↑	NW	E	SE	SW
	18	20/113	5.62		↑	SW	E	S	NE
	19	20/126	6.31		↑	W	S	NE	N
	20	20/142	7.08		L	N	NW	E	SE
	21	20/159	7.94	Upper-middle row	L	E	NE	W	SW
	22	20/178	8.91		↑	N	S	E	SE
	23	20/200	10.0		↑	NE	NW	NE	W
	24	20/252	12.6		↑	E	SE	SW	NW
	25	20/316	15.8		R	NW	SW	N	W
	26	20/400	20.0	Top row	R	SE	SW	S	S
	27	20/504	25.1		↑	NW	N	NE	SW
	28	20/632	31.6		L	NE	NW	NW	S
3	29	20/800	39.8	Lower right		NW	NE	NW	SE
	30	20/1000	50.1	Upper left		W	SW	SW	SW
4	31	20/1260	63.1	Center		E	E	NW	NW
5	32	20/1590	79.4	Center		N	NW	W	W

---

*R → L means from right to left, and L → R means from left to right.

handled. Figure 5 shows the layout of the slides. The fact that most of the slides display more than one C means that, to prevent the compromise of answers, the tests must be conducted beginning with the lowest light level (~ 10^{-5} mL) and then working successively upward, from one light level to the next. Indeed, recognition of the fact that the tests could be conducted in this way led to placing more than two C's on each slide, thus facilitating a considerable reduction in the number of slides required. Another advantage of this approach was that it avoided uncertainties about adaptation-time requirements that would have existed had the tests been conducted in exactly the opposite manner, i.e., by working downward in light level. (Whereas adaptation to a higher light level takes place nearly instantaneously, adaptation to a lower level does not; the time required depends on the initial and final light levels and generally increases as both of these decrease, and for the light levels under consideration, the precise times were not known.) While no true precedent could be found which described Landolt C acuity during adaptation to the lowest light level, investigations by Hattwick [12], Miles [13], and Brown et al. [14] indicate that 30 min would probably be more than enough time to achieve complete adaptation; accordingly, this was the time allocated.

To verify that complete adaptation had occurred, consideration was next given to measuring Landolt C acuity during the adaptation process by including a dark-adaptation test. Of the slides already prepared for the postadaptation tests, those with 100% modulation were chosen because only they seemed likely to cover a range of visual angles sufficient for a dark-adaptation test. In deciding upon a known starting point for this test, it seemed likely that the highest light level used in the tests would be high enough to overcome any previous state of adaptation (e.g., to room light); accordingly, it was decided that observers should preadapt to this light level (for 1 min) before taking the dark-adaptation test.

3.4. TEST PROCEDURE AND SCORING

The sequence of the tests is as follows. First, the dark-adaptation test is given. This test requires (by definition) a period of 1 min for preadaptation to the highest light level (~ 100 mL) and then 30 min for adaptation to the lowest light level (~ 10^{-5} mL). During dark adaptation, a record is made of the times required by an observer to correctly state the gap orientations of successively smaller 100% modulation Landolt C's. Upon completion of this test (i.e., at the end of 30 min), the postadaptation tests are given. These begin at the lowest light level and then work successively upward through the light levels to the highest, i.e., from filter 7 through filters 6, 5, 4, 3, 2, 1, and 0. At each light level, three contrasts are presented: first 100% contrast, then a medium contrast, and then a low contrast (specifically, 100, 50, and 5%; 100, 20, and 2%; or 100, 10, and 1%). Though a modification of this sequence was subsequently thought more desirable and tried, the original has proved the least problematic and remains the recommended approach. The total time required to complete this sequence is generally

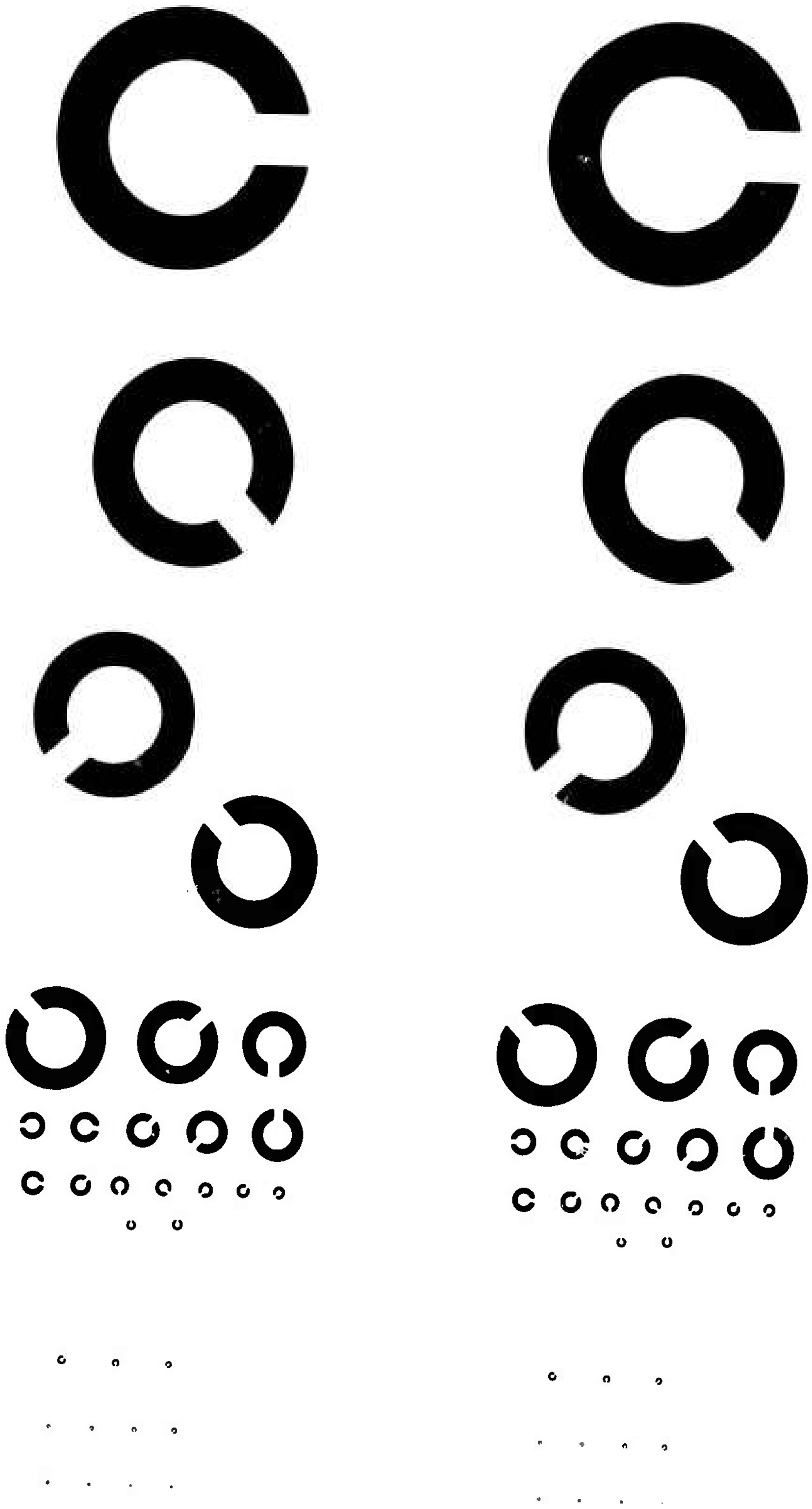


FIGURE 5. LAYOUT OF LANDOLT C TEST SLIDES

somewhat less than 2 hr; this includes instruction time and time for a rest pause generally taken after the scotopic vision postadaptation tests (filters 7, 6, 5) and before the photopic vision postadaptation tests (filters 4, 3, 2, 1, 0).

The scoring system adopted for the postadaptation tests is the same as that employed in the standard Ortho-rater tests. At each light level, tests with a given contrast are terminated when (1) the observer says he can go no further, or (2) the observer misses the orientations twice consecutively. The observer's score is the last correct answer. In the event the observer misses one C, views the next one correctly, and then misses again, the experimenter stops him and, without telling him that he has missed, retests beginning with the C two steps larger than the first one missed. Usually the observer corrects his first miss; however, if the original sequence persists, the first miss is ignored in scoring. (In this case, the experimenter later re-examines the slide for defects or dirtiness which might have caused the mistakes.) The observer is not told the scoring rules; his instructions are, "Keep indicating the orientations until you think you can go no further, or until you are told to stop." Allowing the observer to stop the test himself is a departure from the more proper psychophysical forced-choice test procedure, but this is mitigated by the size of steps from one target to another, making it entirely possible that the jump to the next smallest C would bring the observer to a chance success level anyway. It has been found that when some observers say they cannot go further, they are actually capable of doing better; therefore, each time an observer says he would stop, he is asked, "Are you sure you cannot go further?" It has been found that the improvement which may result usually does not exceed one or two steps.

For the dark-adaptation tests, the time it takes the observer to make correct responses is measured. If the observer makes an incorrect call, the experimenter first attempts to retest by removing and reinserting the same slide, or, if the slide has a single C which is not pointed east or west, by removing the slide, flipping it over so that the top becomes the bottom, and reinserting it. In the latter case, all northerly directions will become southerly and vice versa. If the observer again makes a mistake, this size C is skipped, and the experimenter proceeds to the next smallest one. The experimenter does not tell the observer whether he is right or wrong during adaptation, since afterwards the observer is tested over the same 100% contrast C's in accordance with the postadaptation rules. The observer may indicate which of the eight possible gap orientations is correct by using any one of the following methods:

(1) Verbal response
  - (a) Major compass points: N, NE, E, SE, etc.
  - (b) Directions: right, left, top, bottom; top right, top left; etc.
  - (c) Clock system: 12 o'clock, 1 o'clock, 6 o'clock, etc.

(2) Finger pointing (equivalent to verbal response): the subject's finger is wrapped with phosphorescent tape which gives off enough light to be seen by the experimenter.

To enable the recording of test scores in the dark, an auxiliary console unit was developed which incorporates a red illuminator light for use as a scoring table, a strip-chart recorder for use in recording adaptation times, and a constant voltage transformer for powering the Ortho-rater light bulb. For aid in locating the slides, the filters, and certain portions of the instrumentation in the dark, labels made of phosphorescent tape are used.* Figure 6 shows a typical test setup. The experimenter sits facing the console unit with the Ortho-rater to his

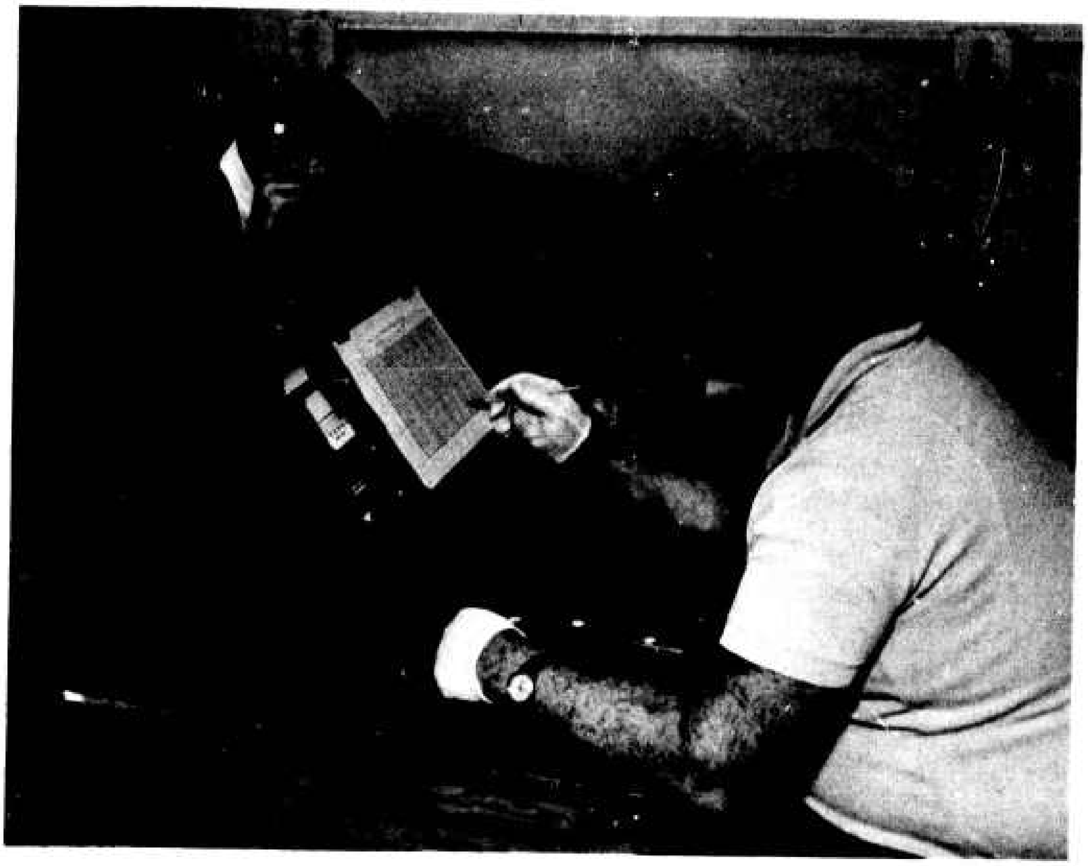


FIGURE 6. SETUP FOR ADMINISTERING THE TESTS

---

*The observer's state of adaptation to darkness is essentially unaffected by light from the phosphorescent tape or from the illuminator. Light from the tape rapidly decays in the dark to a very low level; moreover, the tape usually cannot be seen by the observer. Since the illuminator light is red (wavelength > 600 mμ) it has essentially no effect on the rod vision used below 3×10^{-3} mL (filters 7, 6, and 5). While this light could possibly affect the cone vision used at higher levels (filters 4, 3, 2, 1 and 0), it is deliberately kept at a low level, and, at any rate, the observer generally cannot see it.

right. One neutral-density filter has been inserted into the rear slot of the Ortho-rater, and the rest are readily accessible in the standard wooden base of the machine. One test slide is in place in the front slot, and the rest of the slides being used are on the table in front of the experimenter.

3.5. THE FORM-RECOGNITION TEST

The decision to develop a real-life form test was not reached until some time after the acuity tests and instrumentation were in the process of being implemented. Hence, although the latter influenced the design of the form-recognition test, the converse was not true. One way the form-recognition test was influenced was in the time that could be allocated for it; since obtaining Landolt C acuity data seemed likely to take most of the 4-hr period assumed as the limit for all testing, it was clear that the form-recognition test could not be allowed to take up much time. Further, the Landolt C slides were much more difficult to prepare than had been anticipated, so that considering time, manpower, and cost limitations (as well as the desire to minimize the total number of slides), it seemed best to plan on making as few real-life-form slides as possible. For these reasons, it seemed evident that no more than one real-life form and no more than one seeing task could be considered. Partly because of its obvious military importance and partly because it would be familiar to members of any population, the form selected was that of a man. The seeing task chosen was more difficult to decide upon, and still more difficult to implement. Though military seeing tasks run the gamut from "detection" of a man (i.e., sensing his presence) to a range of "recognitions" involving various critical dimensions, it seemed best to choose a task difficult enough to involve cognitive processes of a higher order than those involved in detection, and yet not too complicated.

The task chosen—recognition of the man as a man—seemed a good compromise, but there remained the problem of how to build validity and reliability into the test. With regard to reliability, it may be noted that the problem for real-life forms is essentially no different in kind from the problem of deception in acuity tests (see sec. 3.1). The method-of-limits approach was rejected, as it was for the acuity test, because here it seemed even more likely to lead to variation in interpreting the nature of the task. It was decided that the procedure should be similar to that adopted for the acuity test; i.e., the test should (1) give the observer the opportunity to make mistakes when he is not able to perform the required recognition task, and (2) give him less than an even-chance opportunity for guessing the correct answer. It should be emphasized that it would not have been valid to use a test which simply presented the form of a man in one of a number of space/time locations with nothing in the other locations. This would have been no different from the method of limits and probably would have amounted to a detection rather than a recognition test. (That some type of recognition might occur simultaneously with detection would be irrelevant; there would be nothing about the test itself to lend credence to the claim that the man had been recognized as a man.) To provide the kind of test desired, it was necessary to use false targets in the other locations, the visibility of which would be equal to but not greater than that of the target of interest and the appearance similar enough to compel the kind of recognition required. In other words, the false targets had to be such that when the contrast and background luminance were not high enough to enable the observer to see the critical dimension of the real target, then the false targets would appear to him to be forms of a man. If, on the other hand, the false targets were visible when the man was not, then the man could be identified by a process of elimination (since the observer would already know that the silhouette of a man was present by virtue of the fact that he had been shown an example of the slide), test validity would be jeopardized, and reliability would be reduced. (Reliability would, of course, depend upon the number of false targets used.)

A part of the implementation problem, then, was to decide what form these false targets should take. Though real-life forms other than that of a human being were considered, they were rejected for two reasons. First, there was no guarantee that observers would not be familiar enough with such forms that they could readily distinguish them from the form of a man (especially if they were allowed to see them once during the course of the test). Second, there was no assurance that the visibility of these forms could be made equal to or less than that of the target of interest. Indeed, as indicated in the discussion on acuity tests, controlling the visibility of any target by the design of its parts can at best be done only approximately. The decision taken, therefore, was to use forms that would be unfamiliar to any observer, not easily remembered, but designed to have roughly the same overall appearance and detail as the form of a man. Various nonsense forms were devised to meet these criteria; those which bore either too much or too little resemblance to a man were eliminated (subjectively, because of time limitations) and the remainder adopted for the test. Figure 7 shows the nonsense forms finally selected and the silhouette of the man. (Actually, this figure shows one side of a stereo slide layout used in the tests.) There are seven nonsense forms, so that the chance of identifying the man is one in eight, the same as for guessing the Landolt C gap orientation. The method used in designing the nonsense forms is an adaptation of Attneave and Arnoult's Method 1 [15], which involves connecting randomly chosen grid coordinates in such a way as to produce the closed boundary which outlines the silhouettes. In adapting this method, it was decided that not all the coordinates should be randomly chosen; instead, to achieve the desired degree of overall similarity to the silhouette of the man, five points (shown on the man in fig. 7) should be common to all the forms. (For one of the forms, the second from the left on the bottom line, points 2 and 4 were inadvertently shifted downward.) The other coordinate pairs were then randomly located within a rectangular grid no higher than point 3 and essentially no wider than points 2 and 4.*

(Footnote on following page.)

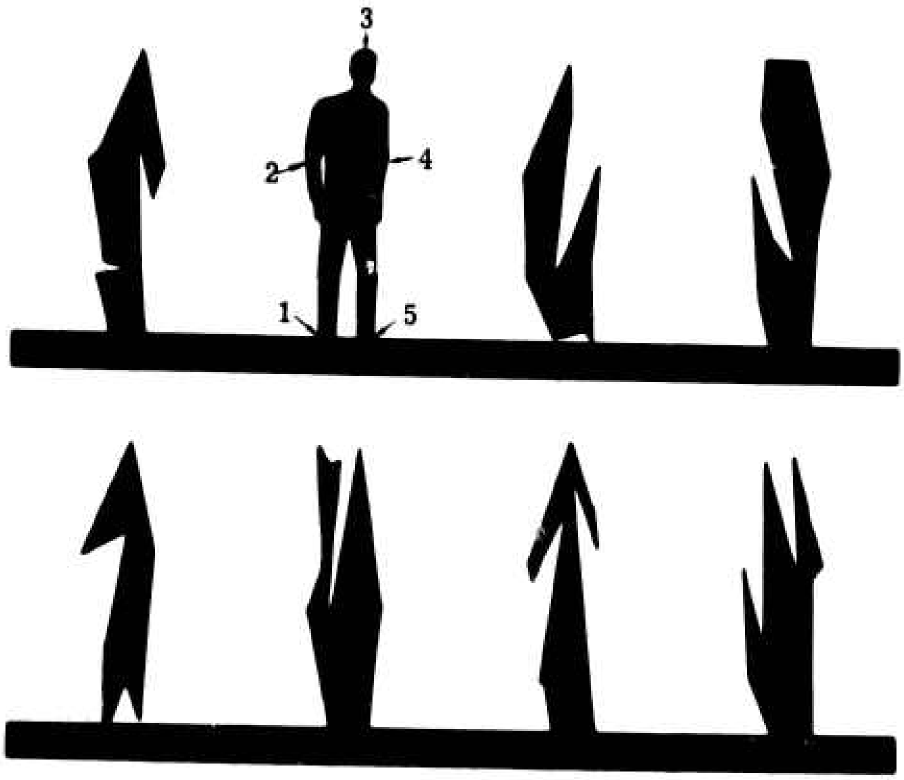


FIGURE 7. LAYOUT OF THE HUMAN-FORM-RECOGNITION TEST SLIDE

Consideration was next given to the procedure for administering the test. Given any one contrast value and any one of the light levels, an ad hoc critical dimension could be determined in the following manner: (1) make up slides with forms varying in size in some predetermined manner (say in the same way as for the acuity tests) and varying randomly in location from slide to slide; (2) have an observer pick out the form of the man; (3) knowing the actual size of the man (appropriately scaled down for the 20-ft Ortho-rater viewing distance), compute the critical dimension by matching the degree of difficulty of seeing the smallest man visible at the given light-level and contrast values with the minimum Landolt C visual angle for those values. Several slides (probably five to seven) would be required for this procedure, and to use it as the basis for establishing invariance, i.e., to repeat it for more than one contrast and light level, would necessitate an impractically high number of slides. As an alternative, a less

---

*Seven other coordinate pairs originally were plotted, based on the notion that these plus the five constant points would give as many major inflections on the boundaries of the nonsense forms as seemed present on the silhouette of the man. This proved not to be the case, however, because the method used required that "interior points" be connected to the "outer boundary."

precise test of invariance could be made using slides with forms of the same size but covering all seven (or less) contrasts used in the acuity tests (and thus requiring seven or less slides). Since five to ten slides seemed a practical limit to the number of additional slides to be made, it was necessary to decide between precisely establishing one ad hoc critical dimension or making a somewhat imprecise test of invariance. Since the former would result in precise knowledge about only one particular target-seeing task, and the latter would shed light on a far-reaching principle, the decision was in favor of the latter; moreover, it was noted that if invariance did hold, then this latter procedure would provide a desirable kind of redundancy not obtainable with the first procedure and possibly an even better measure of the critical dimension.

In keeping with the above decision, seven slides were prepared covering each of the seven contrasts used in the acuity tests. These slides are used as follows. Concurrently with the postadaptation-acuity tests, they are shown at each light level, starting with the lowest contrast and working upwards until the observer can correctly identify the location of the man. This gives a minimum contrast required to see the man, which, together with the light level, can be used to specify a task equivalent in difficulty to seeing a Landolt C target of the same contrast at the same light level; i.e., the visual angle for the Landolt C would indicate the difficulty of the task. Since the objects seen are all the same size, if they were to have the same critical dimension, then the same visual angle would result no matter what the contrast and light level. As shown in section 4, preliminary results suggest that this is the case and that invariance is thus confirmed.

The actual size selected for the targets was based on an assumption about the magnitude of the invariant critical dimension likely to be involved, a desire to cover as broad a contrast range as possible (commensurate with the contrasts available), and a desire to be "realistic." The critical dimension was assumed to be 3 in., or rather, 3 in. scaled down from a realistic target size to that actually presented at the apparent 20-ft viewing distance of the Ortho-rater. The visual angle chosen corresponding to this dimension was 4 minutes of arc (20/80 Snellen); this choice was arbitrary except that, referring to figure 2, it seemed likely to cut across the acuity curves for most of the contrasts, and it also was representative of viewing through electro-optical aids. To reduce the data from this test, then, requires that, for each minimum contrast given by an observer, the Landolt C visual angle for the same contrast and light level be tabulated. Thus, although the raw data are in terms of contrast, the final data are in terms of visual angle, and averages are computed in terms of the latter. The critical dimension, x, can then be determined from the ratio

$$\frac{x\ (\text{in.})}{\text{Experimental Landolt C Visual Angle (min)}} = \frac{3\ (\text{in.})}{4\ (\text{min})}$$

4
PRELIMINARY RESULTS

The results presented in this section are given for the purpose of illustration and do not necessarily represent normative data.

Figure 8 compares the averages for postadaptation acuity for 11 young male subjects at UM with a composite of the data in references 6-8. (After obtaining these results, it was discovered that the actual contrasts of the slides used differed from the nominal values given in section 3.4; the data in figure 3 were then used to generate the composite curves for the appropriate contrast values.) Since the "populations" involved were different, there was no a priori reason to expect the UM curves to match the composite curves. Nevertheless, the agreement seems remarkably good, except at 5.5×10^{-3} mL and for the lower contrasts. With regard to the former, the discrepancies might be due to the lack of adequate data in this range in reference 6. With regard to the latter, it can be demonstrated that the low contrast values for the composite curves [6-8] are suspect by plotting these data in terms of $\log_{10}$ modulation versus visual angle (in minutes) for various background luminances and observing the scatter at low contrasts; a similar plot for the UM data shows much more regular behavior for the curves at low contrasts. As can be seen in figure 8, when fewer than 11 subjects were able to respond, an average was still computed for those who were able in order to indicate possible trends. Such data points are denoted in figures 8-10 by fractions indicating the number of subjects out the total who were able to respond. Figure 9 shows logarithmic averages* for postadaptation acuity and form recognition for 10 male subjects at AUB; the form-recognition curve may be compared with the expected result, a constant 4 minutes of visual angle required to see the man. Figure 10 shows logarithmic averages for postadaptation acuity for the UM and the AUB tests. The same slides were used at both locations, with the exception of new 10% slides used at AUB, for which no comparable UM results have yet been obtained. Neither the human-form nor the 1% slides were available at UM in time to be included in its testing. Figure 11 shows the results of a dark-adaptation test administered to one male subject at UM; it is believed that these are representative of the results which can and should be obtained, although this test has proved to be very problematic and a final evaluation of its potential has yet to be made.

---

*Logarithmic averages are essentially based on the numbers of the visual-angle steps used in the tests. They are much easier to compute than true averages and usually do not differ enough from true averages to make them necessary in preliminary evaluations of the data.

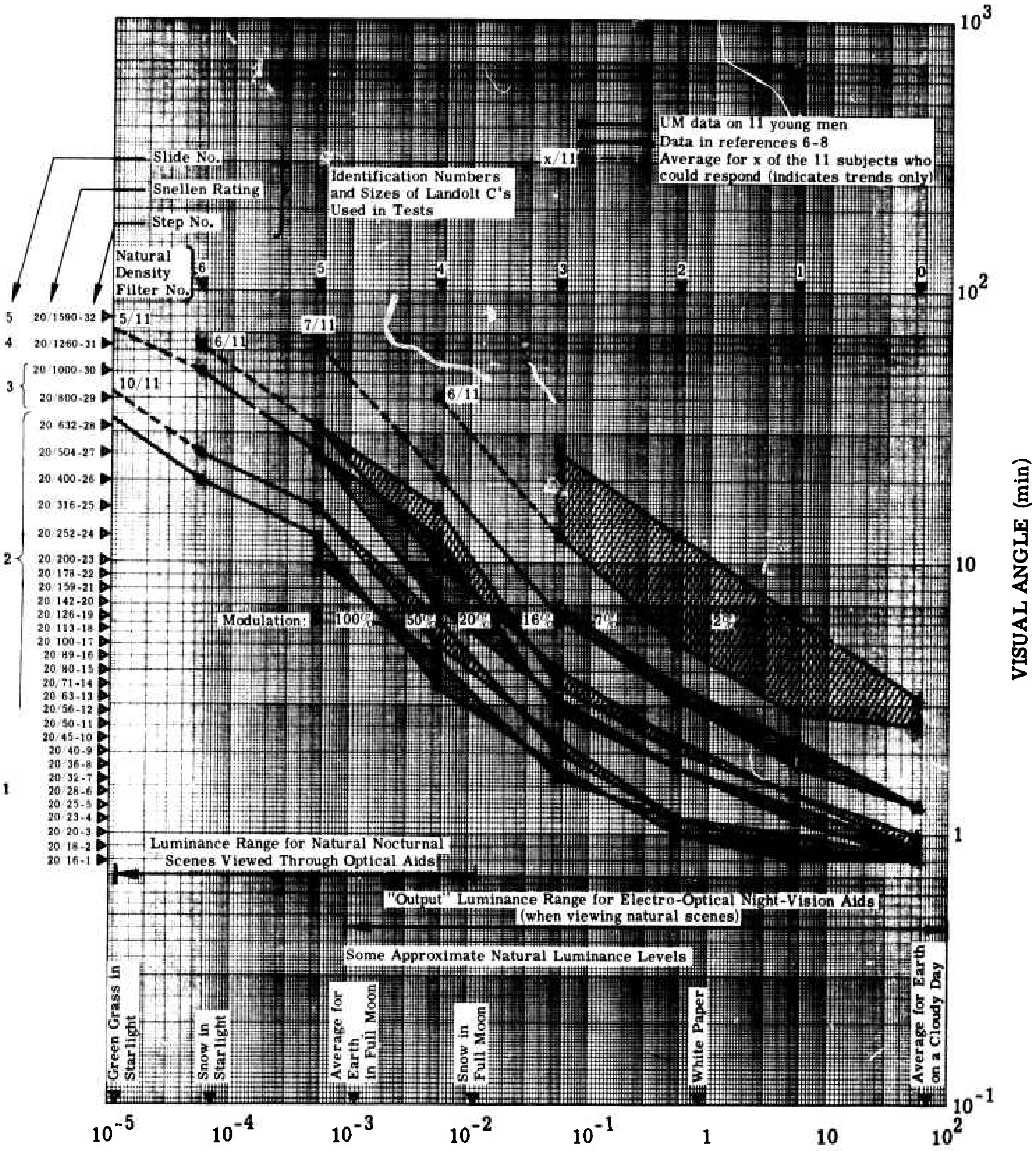


FIGURE 8. COMPARISON OF THE PRELIMINARY UM POSTADAPTATION-ACUITY TEST RESULTS WITH THE DATA IN REFERENCES 6-8. The UM data are true arithmetic averages rounded to the nearest visual-angle step. The data in references 6-8 are interpolated from the curves in figure 3 and are rounded to the nearest step.

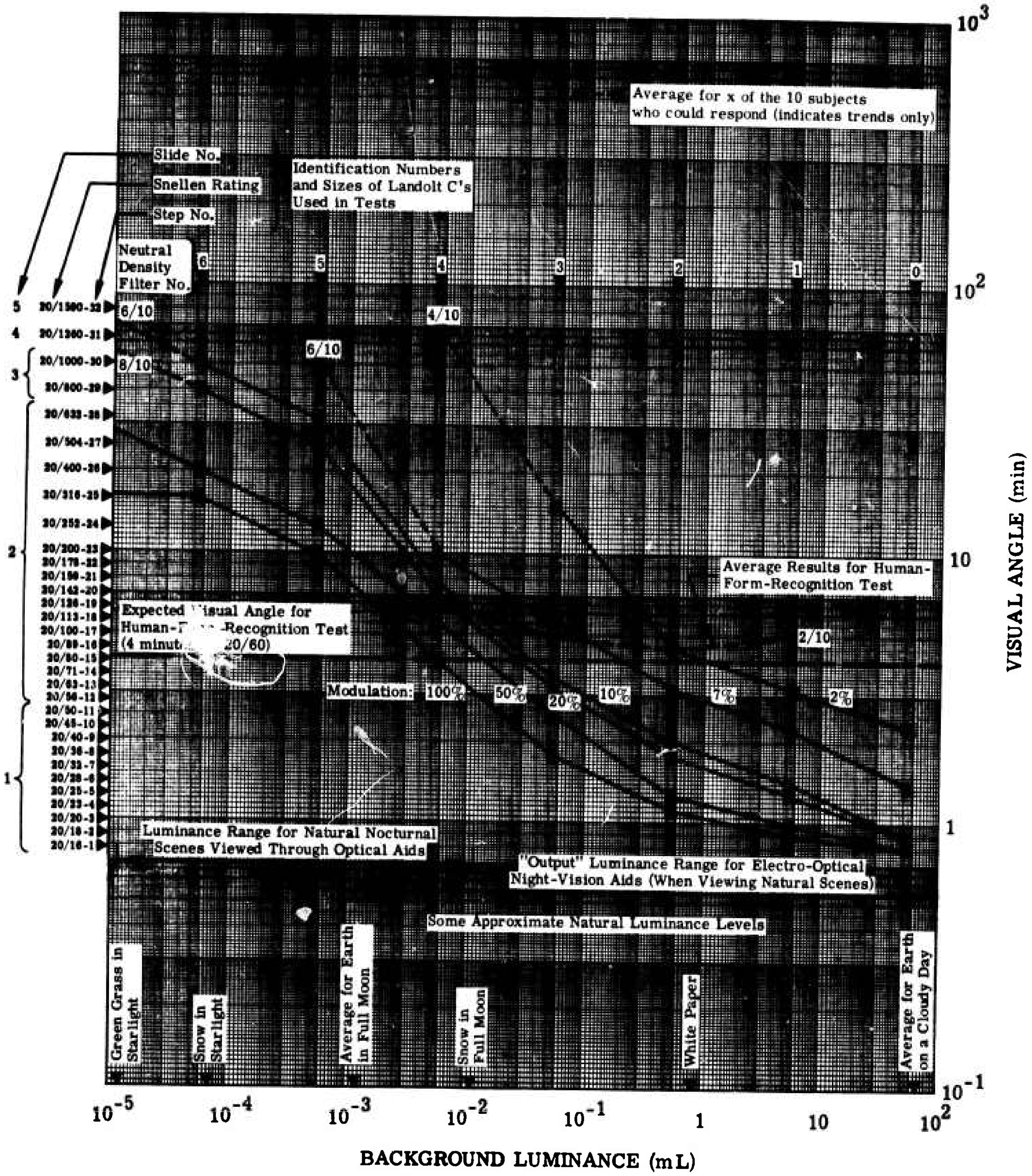


FIGURE 9. PRELIMINARY AUB POSTADAPTATION AND HUMAN-FORM-RECOGNITION TEST RESULTS. These data are logarithmic averages rounded to the nearest visual-angle step.

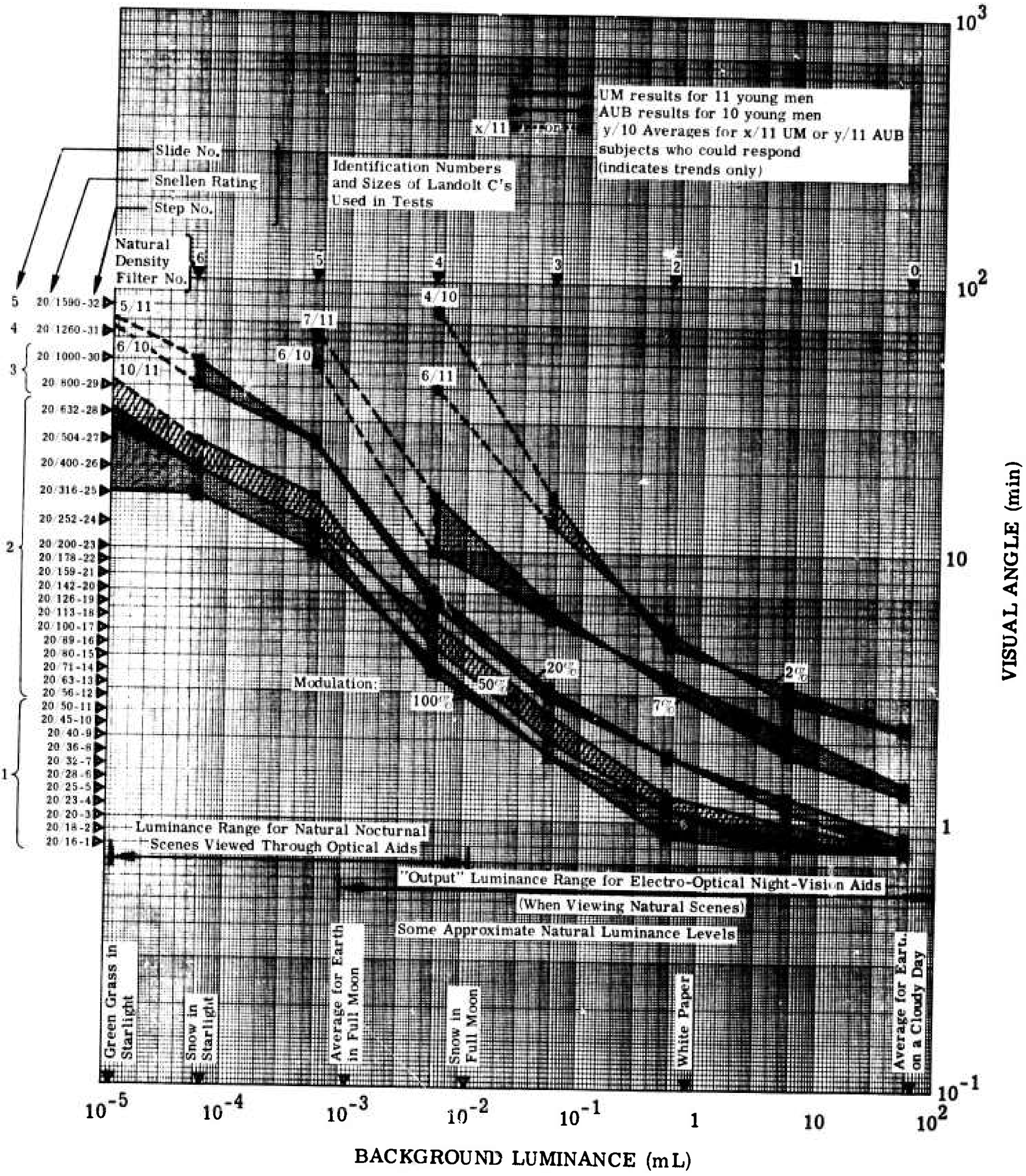


FIGURE 10. COMPARISON OF PRELIMINARY UM AND AUB POSTADAPTATION-ACUITY TEST RESULTS. Both sets of data are logarithmic averages rounded to the nearest visual-angle step.

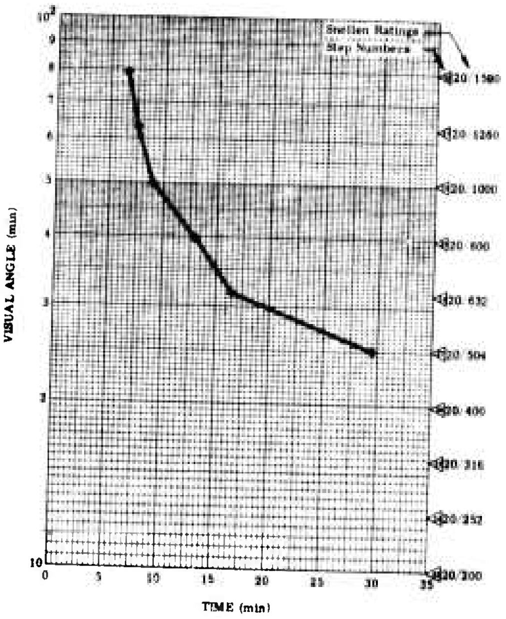


FIGURE 11. REPRESENTATIVE UM DARK-ADAPTATION TEST RESULTS (For one male subject)

5
SAMPLE-SIZE ESTIMATES

This section presents some estimates of the sample sizes needed to make statistically meaningful population comparisons with the UM tests. Strictly speaking, these estimates were prepared only for the postadaptation acuity tests, but they are also expected to apply to the form-recognition test and possibly to the dark-adaptation test. The average minimum visual angle for each population would be the basis for comparison in all the tests except the dark-adaptation test, for which the average times required to see various visual angles would be compared. The basic factors that determine sample size requirements are the minimum difference to be detected and the desired statistical quality of the comparison. For some fixed quality, the smaller the difference to be detected, the greater the total number of samples required, and conversely. Specifically, if the samples were unequal in size, then the total number of samples (i.e., the sum for the two populations) would increase more rapidly than if the samples were equal in size (though both totals would increase). On the other hand, the statistical quality of a comparison could be improved by increasing the minimum difference as well as the sample sizes, and conversely. Since these factors would not necessarily be the same for each application, the estimates cover a wide range of possibilities for each. The estimates also cover a wide range of relative sample sizes, since there are many instances when unequal sizes may be necessary and/or desirable.

The method employed is given by Natrella [16, p. T-16] to estimate "sample sizes required to detect prescribed differences between averages when the sign of the difference is not important." This method is conservative in that it leads to somewhat higher estimates than can be obtained when the sign of the difference is important, as may well be the case for some applications. It assumes that the distributions for the population parameters involved are normal and that the true standard deviations and true means of these distributions are known. An attractive feature of the method is that it provides an analytic expression for sample size which makes evident which of the parameters are important.

If the estimated number of samples needed for populations x and y is denoted by n_x and n_y, then n_x and n_y may be expressed as

$$n_x = \frac{(z_{1-\alpha/2} + z_{1-\beta})^2}{d^2} \tag{1}$$

and

$$n_y = \frac{n_x}{C} \tag{2}$$

where $C \geqq 1$, so that $n_y \leqq n_x$. The numerator of (1) effectively defines the statistical quality of the comparison. The left-hand term depends on α, the level of significance (i.e., the probability of concluding from a random sampling that the absolute difference between the true population means, $|\mu_x - \mu_y|$, exceeds some value when in fact it does not—this is also referred to as the probability of making an "error of the first kind," or the "false alarm" probability). The right-hand term depends on $1 - \beta$, the "power" of the estimate (i.e., the probability of concluding that $|\mu_x - \mu_y|$ exceeds some value when it really does—β represents the probability of making an "error of the second kind," or of concluding that $|\mu_x - \mu_y|$ does not exceed some value when it really does). Both the left-hand and the right-hand terms are of the form z_p, the values of which are tabulated by Natrella [16, p. T-3] for various values of p.

The denominator of (1) is a normalized measure of $|\mu_x - \mu_y|$; i.e.,

$$d = \frac{|\mu_x - \mu_y|}{\sqrt{\sigma_x^2 + C\sigma_y^2}} \tag{3}$$

where σ_i is the true standard deviation of population i. Combining (1) and (2) yields

$$n_x = \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 \left(\sigma_x^2 + C\sigma_y^2\right)}{|\mu_x - \mu_y|^2} \tag{4}$$

Equation 4 focuses attention on six parameters of importance: α, $1-\beta$, the σ's, C, and $|\mu_x - \mu_y|$, the difference regarded as of practical significance. The lower the value of α (i.e., the lower the false-alarm probability), the higher the value of $z_{1-\alpha/2}$ and thus the higher the values of n_x and n_y. The higher the value of $1 - \beta$ (i.e., the lower the probability of an error of the second kind), the higher the value of $z_{1-\beta}$ and thus the higher the values of n_x and n_y. Similarly, the higher the value of C or of the σ's, the higher the values of n_x and n_y. Finally, the smaller the value of $|\mu_x - \mu_y|$ one is willing to ignore, the greater the values of n_x and n_y.

From certain theoretical considerations and an analysis of some UM postadaptation acuity data, it may conservatively be assumed that

$$\sigma_i \leqq \frac{\mu_i}{2} \tag{5}$$

Taking the worst-case condition expressed by (5), namely equality, permits making a substitution for the σ's in (4). Inspecting the result suggests still another substitution to eliminate the μ's:

$$\mu_y = M\mu_x \tag{6}$$

where $M > 0$, but $M \not= 1$, so that $\mu_y > \mu_x$. The result of making both substitutions in (4) is

$$n_x = (z_{1-\alpha/2} + z_{1-\beta})^2 \frac{1 + CM^2}{4|1 - M|^2} \tag{7}$$

The closer M is to unity (i.e., the smaller the difference between the true means), the larger are the values of n_x and n_y.

Figure 12 is a graph of the estimated sample sizes, n_x and n_y, versus their ratio, $C = n_x/n_y$, for various values of $M = \mu_y/\mu_x$* and for $\alpha = 0.05$ and $1 - \beta = 0.90$. Ideally, one would choose values of α and $1 - \beta$ which were derived from an analysis of the rewards and penalties involved in the situation. Lacking such information, the values selected for figure 12 were chosen simply in a traditional ad hoc manner. Nevertheless, other values may be more appropriate; figure 13 shows multiplication factors which may be used to correct the n values in figure 12 for certain other common values of α and a range of $1 - \beta$. From figure 12, it may be seen that in the case of unequal sample sizes, a slight increase in the size of the smaller sample, n_y, may result in a very large decrease in the size of the larger sample, n_x, and in the total number of subjects, $n_y + n_x$. This fact may be quite useful in planning test strategies and should be kept in mind when dealing with unequal sample sizes.

---

*The values of n needed for the case $M = 1/Q$ are actually slightly smaller than for $M = Q$, but the difference is not important for the purposes at hand.

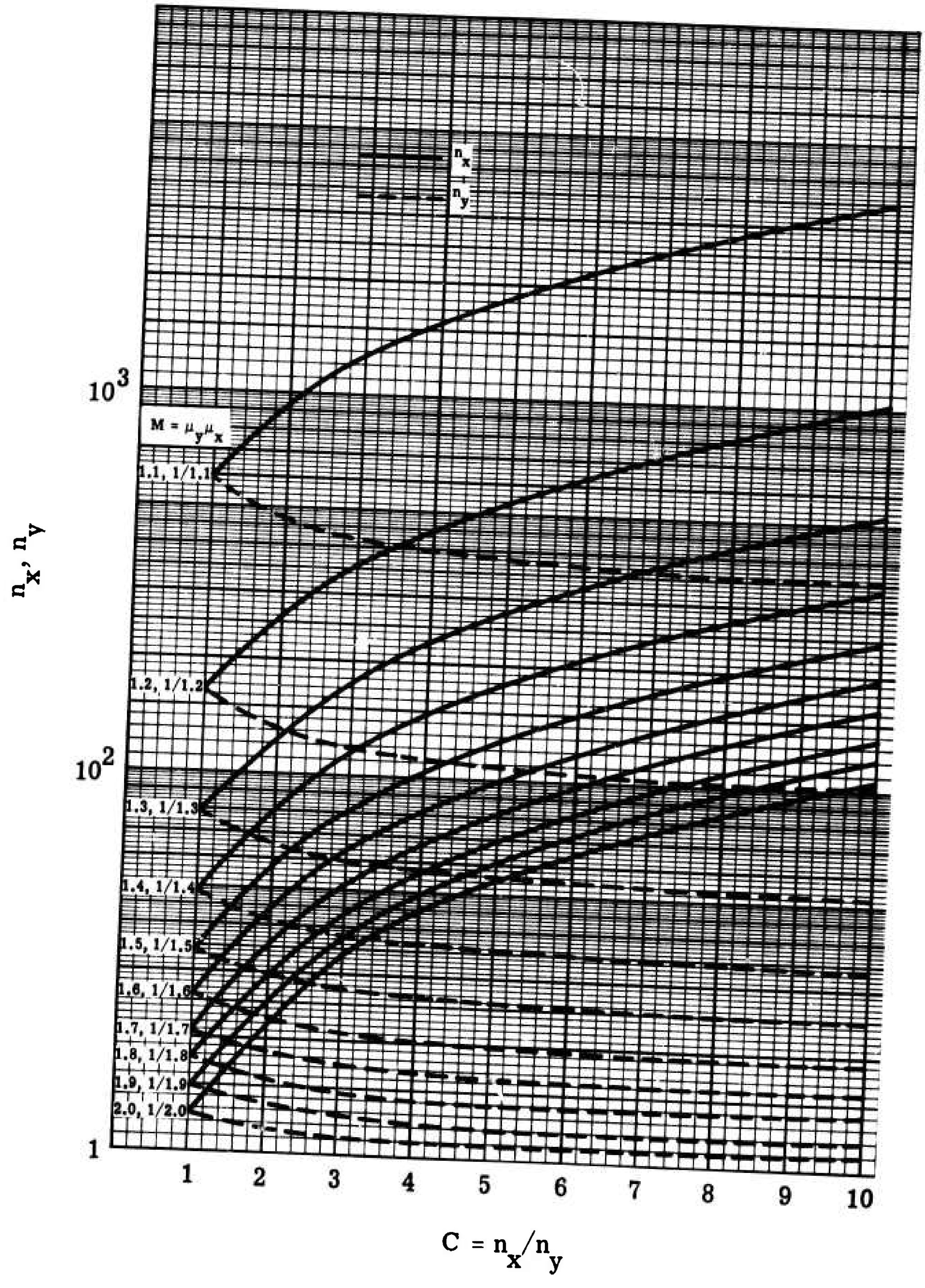


FIGURE 12. EXAMPLES OF SAMPLE-SIZE ESTIMATES

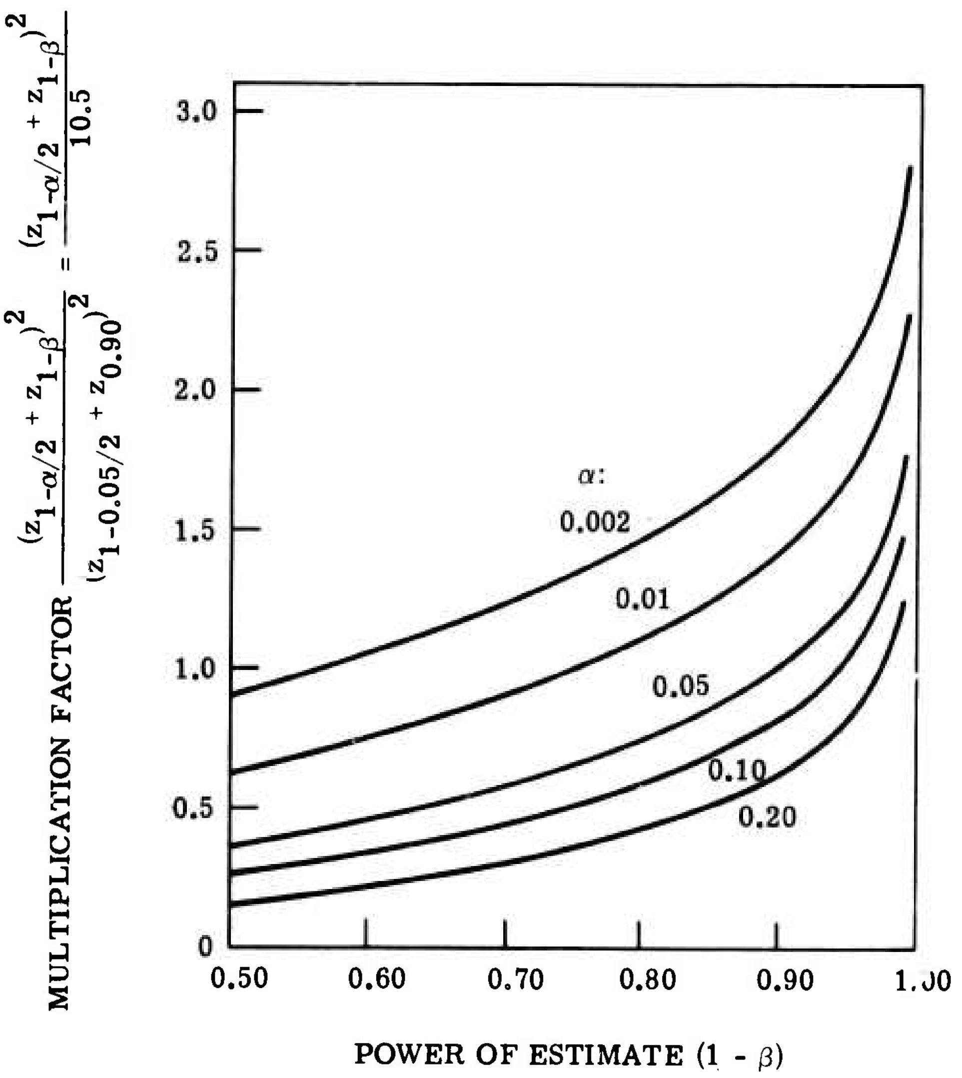


FIGURE 13. MULTIPLICATION FACTORS FOR CHANGING THE SAMPLE-SIZE ESTIMATES IN FIGURE 12. α = significance level, or false-alarm probability.

6

OTHER APPLICATIONS OF THE TESTS

Whereas the preceding sections have been mostly devoted to explaining the interrelationships between the tests and the night vision aids application, it now seems appropriate to conclude the discussion with a few brief remarks on some other possible applications of the tests and of the ideas involved therein.

In this vein, it must first be noted that, with regard to the UM-AUB study, a native U. S.-vs.-native Middle East population comparison is expected to be only one (small) part of its results. Besides establishing an apparently unprecedented amount of vision test data on a native Middle Eastern population (from which intrinsically valuable norms will be established), comparative data will also be acquired at AUB on nonnative but resident Middle Easterners in an effort to pin down any influence of the Middle East "environment." As a somewhat related effort, appropriate "characteristics" will be noted for each subject to enable studying the possible influence of such factors as smoking, diet, the wearing of sun glasses (on night vision and dark adaptation), etc. Finally, partly to obtain an increased understanding of the tests and partly to explore their potential as a diagnostic tool, a small number of subjects with certain known visual deficiencies will also be examined at AUB and their results compared with the norms obtained there.

In addition to studies of the above type, the tests could also be used to screen individuals, i.e., to select individuals who were visually well suited for some task from those who were not. Indeed, it is the latter which would seem to account for the greatest usage of instruments like the Ortho-rater; though again, the tests involved are typically nowhere near as comprehensive as those which have been described above. Nevertheless, with regard to the "in between" approach represented by the previously mentioned ANST-I*, it would be possible to reduce the number of contrasts and light levels discussed above so that some particular range of interest could be covered. (Of course, different targets might also be used).

Finally, regardless of the instrumentation used, it is believed that the nonsense-form approach to testing "recognition" may be valuable in establishing the critical dimensions of real-life targets involved in other applications. Moreover, if the invariance of critical dimensions to changes in contrast and light level holds as well as it now appears to do (for the man), then such critical dimensions could be established using only one contrast and light level, a considerable savings in effort.

---

*Army Night Seeing Tester

REFERENCES

1. S. Shlaer, "The Relation Between Visual Acuity and Illumination," J. Gen. Physiol., Vol. 21, 1937, pp. 165-188.
2. C. Paprocki and R. E. Miller, Use of Image Intensifiers for Real-Time Multi-Spectral Viewing, Report No. 7919-26-T, Willow Run Laboratories of the Institute of Science and Technology, The University of Michigan, Ann Arbor, July 1968.
3. L. A. Riggs, "Visual Acuity," Vision and Visual Perception, ed. by C. H. Graham, Wiley, 1965, p. 339.
4. H. R. Blackwell, "Contrast Thresholds of the Human Eye," J. Opt. Soc. Am., Vol. 36, 1946, pp. 624-643.
5. H. R. Blackwell and D. W. McCready, Jr., Foveal Contrast Thresholds for Various Durations of Single Pulses, Report No. 2455-13-F, Engineering Research Institute, The University of Michigan, Ann Arbor, June 1958.
6. J. P. Conner and R. E. Ganoung, "An Experimental Determination of the Visual Thresholds at Low Values of Illumination," J. Opt. Soc. Am., Vol. 25, 1935, pp. 287-294.
7. P. W. Cobb and F. K. Moss, "Four Fundamental Factors in Vision," Illum. Eng., Vol. 23, 1928, pp. 496-506.
8. P. W. Cobb and F. K. Moss, "The Four Variables of the Visual Threshold," J. Franklin Inst., Vol. 205, 1928, pp. 831-847.
9. J. J. Vos, A. Lazet, and M. A. Bouman, "Visual Contrast Thresholds in Practical Problems," J. Opt. Soc. Am., Vol. 46, 1956, pp. 1065-1068.
10. Armed Forces-National Research Council Vision Committee Secretariat, Manual of Instructions: Armed Forces Vision Tester, FSN 7610-721-9490, Department of Physics, University of Florida, Gainesville, July 1962 (revised).
11. J. E. Uhlaner and J. Zeidner, The Army Night Seeing Tester—Development and Use, Human Factors Research Branch Technical Research Report 1120, Technical Advisory Group, Research and Development Command, U. S. Army, May 1961, AD Div. 23/1,28/4.
12. R. G. Hattwick, "Dark Adaptation to Intermediate Levels and to Complete Darkness," J. Opt. Soc. Am., Vol. 44, 1954, pp. 223-228.
13. W. R. Miles, "Light Sensitivity and Form Perception in Dark Adaptation," J. Opt. Soc. Am., Vol. 43, 1953, pp. 560-566.
14. J. L. Brown, C. H. Graham, H. Leibowitz, and H. B. Ranken, "Luminance Thresholds for the Resolution of Visual Detail During Dark Adaptation," J. Opt. Soc. Am., Vol. 43, 1953, pp. 197-202.
15. F. Attneave and M. D. Arnoult, "The Quantitative Study of Shape and Pattern Recognition," Psychol. Bull., Vol. 53, 1956, pp. 452-471.
16. M. G. Natrella, Experimental Statistics, NBS Handbook 91, 1963.

Appendix

PROCEDURES USED FOR PRODUCING THE TEST SLIDES

by J. Cooper

The first step involved in making the Landolt C slides was to produce master negatives from which positive slide transparencies could later be made by a contact printing process. These negatives had to be of the same size as the positives and had to exhibit minimal distortion due to lens aberrations. They were obtained as follows. First, a matte white vertical surface was placed at one end of the photographic studio to define an object plane and to serve as a background against which black cardboard Landolt C's would ultimately be placed. Next, a carefully measured rectangle of black thread was formed on this surface. A 4 × 5 view camera with a process lens was then placed on a heavy tripod which in turn was set at an appropriate predetermined distance from the object plane.[*] The camera position was carefully measured to ensure that it was on a line perpendicular to the centerline of the object plane, and the camera was carefully leveled. A final check to ensure parallelism between the object and image plane was made by photographing the rectangle on the object plane. After processing, the negative was measured on a comparator. Adjustments were made to the camera back until parallelism was obtained. Also, a matched set of film holders was selected for use by employing a micrometer to sort out those with film planes at an identical "depth."

Two banks of lights were used to illuminate the object plane; these were carefully adjusted to that the illumination on the object plane was even within the reading accuracy of a sensitive CdS light meter. Then, from the set of black cardboard C's which had been provided for the layouts, the largest C was selected, taped to the object plane, and photographed. Cronar base film (COS 4) was used for this purpose because of its high dimensional stability. Also a standardized film processing procedure was adopted to avoid size changes due to processing variation. After processing and stabilization to room conditions, the negative depicting the large C was measured with a comparator to ascertain the diameter of the C. Camera positions were changed and realignment procedures repeated until this diameter was within tolerance. At this point, the camera, tripod, lights, and object plane were locked down to prevent any accidental shift in position. Also, register marks were fastened to the object plane for use in final assembly of the stereo slides.

---

[*] The distance was set so as to produce negatives (and thus slide positives) that would be of a proper size for use in the Ortho-rater. Consistent with this, the focal length of the lens used was selected such that the angle of the image formed by the lens was less than 1/6 the normal angle of the lens. This was done in order to produce images essentially without geometric distortion.

As needed for each slide, the C's were fastened to the object plane in the appropriate patterns, and, to make stereo slides, two negatives were made of each pattern. These were processed together, following the standardized procedure, using fresh developer for each pair. Then, by means of an alignment "jig" specially designed for the purpose, they were spliced together with the proper separation to form stereo masters. Since it seemed desirable to handle these masters as little as possible, duplicate negatives were then printed for use as working negatives. The latter were also checked to be sure there was no deviation from correct image size. (The masters and working negatives were all black-on-white, regardless of the contrasts of the slides for which they were intended.)

To make the positive transparencies needed for the slides, Dupont Aerial Duplicating Film was used. The positives were then made from the working negatives after preliminary tests were made to determine approximate exposure times for the required contrasts. A procedure was instituted giving a film gamma close to that which would result if the development were indefinitely prolonged; this involved development in fresh Dupont 53-D developer for 3 1/2 min at 68°F with carefully controlled agitation. During preliminary testing on the positives, it was found that stray light was falling on the printing area. To eliminate this light, the point light-source used in the contact printing operation was hooded with a 2-ft-long black cylinder to prevent its light from reflecting off the ceiling, floor, and walls of the darkroom. In addition, this area itself had to be "tented" in black nonreflecting material. Even the photographer had to wear dark clothes and stand a good distance away from the printing area during exposure. Failure to exercise these precautions caused variations in density of as much as 0.02 from one side of the positives to the other.

For the low-contrast slides (i.e., those of 1% and 2% modulation), a method of outlining the overall slides was needed, since the C's were barely visible and therefore could not be used for proper alignment. For this purpose, a combination mask and film holder was made to slip in the vacuum frame used in making the contact prints. This permitted the print area around the slides to be heavily exposed, thus ultimately making the overall outline of the slides visible. By removing the mask, the slide itself could then be exposed to the required density. Two slides (of the same contrast) could be printed together this way.

In addition to the precautions against stray light, a constant voltage transformer was used on both the point light-source and the exposure timer to guard against deviations in exposure (and thus contrast) due to line voltage fluctuations. The line voltage was monitored with a meter and no exposure was made if the meter showed any deviation from normal.

With these controls, preliminary exposures were made and the positives given standard processing and dryed. After stabilization, the densities were read with a Welch Densichron. (This was used throughout and was checked against a standard before and after each density reading.) Based on the density readings, the exposure was varied to make new test positives until the contrasts of these came within tolerance. After the proper exposure was established for any one contrast, all positives of that contrast were so exposed and processed, each in fresh developer, and the contrasts checked. Since two complete sets of slides were required, each positive was duplicated on the same sheet of material to ensure that the sets would be identical. Whenever the C's were too small to be read in the densitometer, a patch of negative material of the same density as the clear area of the working negatives was mounted between the slide areas, and the positive densities formed through it were read and assumed to be identical to that of the C's. To ensure that the density of the patch was exactly the same as that of the clear area of the working negatives, the film used to make the patch was given an exposure identical to that given a card covered with the same black paint used for the C's and was processed in the same way as the working negatives. Further, the films used for both the working negatives and the patch were from the same emulsion batch.

It was important to be precise in all stages of manufacture of these slides. Even the washing time for the positives had an effect on density. A set of slides with too high a density was rewashed for 20 min; this reduced the density by 0.04 and brought them to the required reading. Final preparation of the slides simply involved sandwiching the appropriately trimmed positives between two carefully cleaned sheets of cover glass and binding the edges with tape. Though a different camera setup was used in producing the human-form-recognition slides, the procedures were otherwise the same.

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Willow Run Laboratories of the Institute of Science and Technology, The University of Michigan, Ann Arbor	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

SOME TESTS OF NIGHT AND DAY VISION

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*

Charles A. Omarzu

6. REPORT DATE November 1968	7a. TOTAL NO. OF PAGES viii + 42	7b. NO. OF REFS 16

8a. CONTRACT OR GRANT NO. DA-44-009-AMC-1494(T)	9a. ORIGINATOR'S REPORT NUMBER(S) 7919-32-T
b. PROJECT NO. ARPA Order No. 774	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency, Department of Defense, Washington, D. C.

13. ABSTRACT

This report introduces some tests of night and day vision which, though devised for a special purpose, are potentially useful in a much wider variety of applications. The tests are mostly of binocular visual acuity for far vision, measured either during the course of dark adaption or after adaption to various light levels has been completed. In addition, a special form recognition test has been developed to investigate the relationship between visual acuity data and the visibility of a "real-life" shape, namely, the silhouette of a man.

The tests are implemented via specially developed slides to be used in a commercial vision test instrument. The latter has been adapted to provide a range of discrete nighttime and daytime light levels (through the use of neutral density filters). Unlike the slides normally provided with this instrument, those specially developed cover a wide range of contrast values. Thus, acuity may be tested over a wide range of contrasts at each of the nighttime and daytime light levels.

While the tests must still be conducted in a darkroom, having the tests implemented as they are enables the darkroom to be considerably smaller and more primitive than would otherwise be necessary, and the tests and darkroom can even be made "portable." Basically, this results from the fact that the instrument used is a closed stereoscope type device which produces the far vision test distance (20 feet) by optical means.

The report covers the design and development of the tests, presents some results obtained with them, and gives some estimates of the sample sizes required to make comparative population surveys, the purpose for which the tests were originally intended. Detailed equipment operating instructions and test procedures are not included, however, since they are not germane to the discussion. Instead, these are covered in another document designed to accompany the test equipment.

DD FORM 1 NOV 65 1473

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Visual acuity						
Landolt C						
Form recognition						
Critical dimension						
Light level						
Contrast						
Stereoscope instrument						