AD A124505

OBJECTIVE COLOR MEASURING SYSTEM

BY
ALVIN O. RAMSLEY
THERESE R. COMMERFORD
LISA B. HEPFINGER

SEPTEMBER 1982

UNITED STATES ARMY NATICK
RESEARCH & DEVELOPMENT LABORATORIES
NATICK, MASSACHUSETTS 01760

INDIVIDUAL PROTECTION LABORATORY

DTIC
SELECTE
FEB 1 8 1983

E

DTIC FILE COPY

83 02 016 014

Approved for public release; distribution unlimited.

Citation of trade names in this report does not
constitute an official indorsement or approval of the
use of such items.

Destroy this report when no longer needed. Do not
return it to the originator.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NATICK/TR–83/005	2. GOVT ACCESSION NO. AD-A124505	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* OBJECTIVE COLOR MEASURING SYSTEM		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER NATICK/TR–83/005
7. AUTHOR(s) Alvin O. Ramsley Therese R. Commerford Lisa B. Hepfinger		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS US Army Natick Research and Development Laboratories Countersurveillance Section, IPL Natick, MA 01760		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Materials Testing Technology
11. CONTROLLING OFFICE NAME AND ADDRESS US Army Natick Research and Development Laboratories ATTN: DRDNA–ITC Natick, MA 01760		12. REPORT DATE September 1982
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

COLORS	FABRICS	MILITARY REQUIREMENTS
MEASUREMENT	DYES	SPECTROPHOTOMETERS
TEXTILES	CLOTHING	CALIBRATION

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

A project was established to design a practical instrumental method for judging color acceptability of textiles. The present visual method leads to subjective judgments that often are equivocal. The project was divided into four phases. Phase I determined that three commercial spectrophotometers are capable of meeting the optical requirements. In Phase II a prototype system was designed. This included development of a procedure for mathematical expression of acceptability; a rigorous system-wide calibration procedure; and computer software to control calibration and operation, to make necessary computations, to store required data,

DD FORM _{1 JAN 73} 1473   EDITION OF 1 NOV 65 IS OBSOLETE

20. ABSTRACT (cont'd)

and to communicate between the master unit and remote stations. An expandable two-unit system was procured; the master unit is located at NLABS and one remote unit is located at the Defense Personnel Support Center in Philadelphia. This report outlines all four phases of the program; summarizes in some detail the results of Phase I and II, which are complete; and briefly discusses the beginnings of work on Phase III. Phases I and II of the project were funded under the Material, Testing Technology Program. Manufacturing Methods Technology funds have been requested to support Phases III and IV.

EXECUTIVE SUMMARY

Progress on Phases I to III of the Army's program to convert procurement of military fabrics to an objective color measurement system is reported. As of May 1982, research has been conducted on assessing the various commercial color measuring instruments for their usefulness in the objective measurement of textiles (Phase I) and the requirements for such an instrument have been established, as well as the development of a method for establishing acceptability criteria for textiles and a fail-safe calibration procedure (Phase II).

The program has progressed to the point where Phases I and II are complete and Phase III, procurement of an instrument and validation studies, is underway. The instrument which has been procured is the Applied Color Systems Spectro Sensor, which is now undergoing repeatability testing. Work is also being done on establishing acceptability parameters for the four-color Woodland Camouflage pattern in order to run a trial procurement using the new system.

Accession For		
NTIS GRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By_		
Distribution/		
Availability Codes		
	Avail and/or	
Dist	Special	
A		

1

PREFACE

Design, acquisition, and evaluation of a prototype color measurement system has been supported under the Materials Testing Technology program. Procurement of the system begins the third of four phases in the program. It is deemed appropriate to provide a progress report at this time.

In particular the writers acknowledge the essential contributions in generating and administratively supporting the program by Mr. John V. E. Hansen, Director, Individual Protection Laboratory, Mr. Charles R. Williams, Chief, Clothing and Equipment Division, and Mr. Kenneth A. Reinhart, Chief, Textile Research and Engineering Division. They also acknowledge the considerable technical assistance provided by Mr. Williams throughout the first two phases of the program. Of particular help on the scientific aspects of the program was the National Research Council Committee on Color Measurement chaired by Dr. David L. MacAdam of the University of Rochester. Other members of the committee, each of whom made significant contributions, are:

Franc Grum, Eastman Kodak
Robert F. Hoban, Sandoz
Michael E. Breton, Visual Physiology Dept., Wills Eye Hospital,
                    Philadelphia, PA
John J. Hanlon, Retired (formerly Mohasco)
Dr. Ellen C. Carter, Consultant

TABLE OF CONTENTS

LIST OF FIGURES

4

LIST OF TABLES

OBJECTIVE COLOR MEASURING SYSTEM

1. Introduction

The Department of Defense annually spends several hundred million dollars for dyed textiles. These are subsequently used in the myriad textile items needed by the services; items as varied as combat clothing, field packs, parachutes, armored vests, formal attire, tents, tarpaulins, and vehicle covers. Depending on the environment, up to 30 different fabrics are needed just to equip a soldier for combat. To meet all military needs, there are over 500 standard fabrics in a variety of colors, composition and construction.

Before a production lot is accepted, the Government tests the fabric in many ways to assure that the quality defined by the specification is actually delivered. Most of the tests performed are either physical or chemical and lead to numerical data that can be verified by others. Problems arise when the results are borderline, but these problems are usually resolved by the objective, measureable nature of these tests.

Procuring agencies have stated that the most vexing problems they face are those that relate to color acceptability. Despite all the standardization built into the test procedure, the results produced by the present method are opinions based on visual observation. It is not surprising, therefore, that in borderline cases, disagreements arise between the producer and the procuring agency. Such disagreements rarely are reconciled simply by repetition of the subjective, non-quantitative procedure.
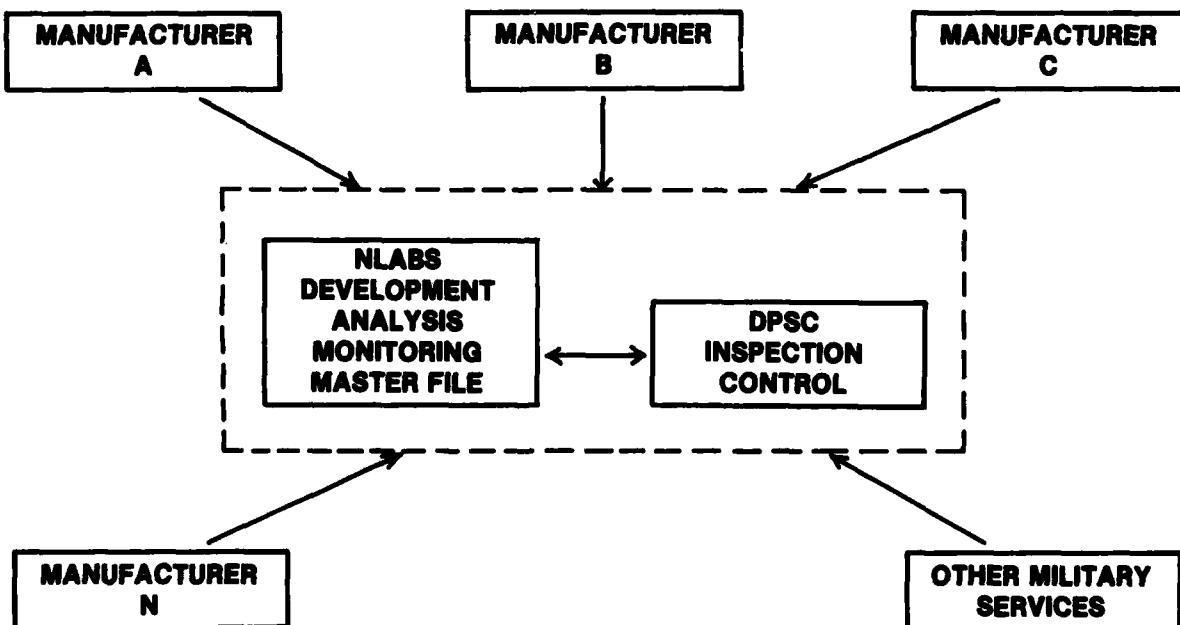

Figure 1. Configuration of a Color Measurement System

7

This report describes a program to design and validate an objective system to determine acceptability of textile colors. The system is based on modern optical units with associated computational means that lead to objective, reliable, numerical definition of color differences from standard shades. Figure 1 portrays one configuration of units that make up an integrated system. The center of the system is a spectrophotometer and extensive computational and data storage capability located at the US Army Natick Research and Development Laboratories (NLABS). A second unit with an identical spectrophotometer and similar associated equipment is located at the Defense Personnel Support Center (DPSC) in Philadelphia. Satellite units will also be located at contractor plants, where primary measurements and computations will be made and transmitted to both DPSC and NLABS. It is expected that this system will be used in a process that at first will supplement and ultimately will replace the method now used.

2. Present Method

In the method now being used, a Government representative samples the production at a contractor's mill and submits specimens to a central laboratory at DPSC in Philadelphia. There a technologist compares the specimens visually to the standard fabric, using carefully controlled illumination and viewing conditions.[1] A typical booth is illustrated in Figure 2.

Figure 2. Typical shade board simulating daylight and horizon/twilight illumination

[1] Federal Standard, CCC–T–191b, Test Method 9010.

Specimens are judged acceptable if they are as close or closer in color to the standard than a series of specimens previously established as the limits of acceptability. These limit samples or tolerances differ from the standard in hue, saturation and lightness. This is basically the same procedure most Government suppliers use to monitor their production. When production is under control and all the submitted specimens are acceptable, this process works well, although there is typically a lag of about two weeks between production and acceptance testing.

Serious problems arise when production does not satisfactorily match the standard and a production lot is rejected for shade by the visual inspection process. If the rejection occurs early in a contract period or the mismatch can be corrected, the producer can usually minimize his losses and deliver an acceptable product to the Government. Even then there are costs and time is lost.

If the deviation from standard cannot be corrected, major problems arise for both the Government and the producer. First, the Government does not get timely delivery of materials urgently needed for conversion into end items. Although a penalty may be assessed, the Government is often forced by circumstances to accept sub-standard materials. Secondly, the penalty may place a severe financial burden on the contractor in a highly competitive industry with narrow profit margins. If production is beyond the limits of acceptance even with penalties and a lot is totally rejected, the consequences for the producer can be disastrous. Moreover, in that case the Government must begin the procurement process all over and months can be lost.

One major cause of the difficulties lies in the visual process itself; different persons do not see color alike, even those whose color vision is considered normal. Moreover, different persons do not use the same illuminating and viewing conditions when they make their judgments. The result is that differences of opinion arise as to what is an acceptable color match. The procuring agency and its technologists are caught between an industry seeking to relax the requirements and a military user who prefers to see no "tolerance" at all. Another source of difficulty lies in the communication of data acquired completely by subjective methods. A pass-fail judgment without comment or guidance is of little help to the producer who needs to correct his process.

In the past, rejected materials were often converted to civilian use by redyeing to darker shades. For two reasons, this option is not as available as it once was. First, civilian fashions have not recently featured dark colors as much as they once did. Second, fabrics are often delivered with functional finishes; a water-repellent or a durable-press finished fabric cannot be redyed. The factors of time and cost require more rigorous control of production processes, more timely and reproducible inspection procedures, and a quantitative language of acceptability. These needs cannot be met by the subjective methods still being used.

Some of the many textile mills are already using optical instruments in various ways to improve their performance. Levels of sophistication vary widely. The dyer can sample production as it proceeds, measuring color of the production relative to the standard, taking corrective action if needed, and submitting only those samples to the Government that the data assure him are acceptable. Using this method, one mill has reduced its rejection rate

from 15 percent (the usual industry-wide rejection rate) to less than five percent. Higher levels of sophisticated color control include computer colorant formulation and on-line monitoring systems that result in even better control and quality goods.

3. The Program

The program was divided into four major phases, each with its own goals, as summarized below. To assist NLABS in the program, a National Research Council (NRC) Committee on Color Measurement was established. This committee has provided an independent, highly qualified review at each important step in the program from establishment of component criteria to the design of the prototype system and guidance on design and evaluation of three subsidiary studies.

The program has progressed to a point where Phases I and II are complete, and Phase III has begun. These two phases will be discussed in some detail in later sections. The following outline of the total program sets forth the objectives of each of the four phases.

a. **Phase I. Survey of Commercial Instruments.** It was known that good colorimetric results were being obtained in industry with three new, relatively inexpensive spectrophotometers. It was not known, however, whether the specific needs of the NLABS program could be met by any of these new instruments. The major goals of Phase I were to evaluate these instruments and to identify any specific changes that would be needed, either optical or computational. The instruments evaluated were:

> Macbeth MS—2000
> Diano Match-Scan
> Hunter D54P—5

To aid in meeting the objectives of Phase I, the following criteria were developed with help from the NRC committee. Although some of the criteria are severe, little benefit could result from the development if they could not be met. The criteria are:

(1) Must be as sensitive as the eye, to small color differences,

(2) Must be more consistent than the eye,

(3) Must be reproducibly adaptable to a variety of textures,

(4) Must be reproducibly adaptable to a variety of sample sizes,

(5) Must use CIE standard illuminant and viewer data:

(a) Illuminants D_{65} and A,

(b) Both 1931 Standard Observer (2-degrees) and the 1964 Supplementary Colorimetric Standard Observer (10-degrees),

(c) Polychromatic illumination, for valid results with possibly fluorescent materials,

(d) Direct illumination must not injure the specimen,

(6) Must include specular component of reflection in the measurement.

(7) Must provide automatic computation and system-wide calibration.

(8) Must compute color differences both in terms of acceptability and perceptibility.

(9) Must be durable, rugged and serviceable for use in an industrial environment.

(10) Must be rapid; not more than 15 seconds per determinations,

(11) Must be simple enough for a technician of modest training to operate,

(12) Must not be too costly for industrial acceptance.

The principal conclusion from Phase I was that all three instruments are remarkably stable and repeatable. A few minor optical and computational modifications were identified in each case. These are discussed in Sections 4 and 5.

b. **Phase II. Design of System.** The major challenges in this phase were to design a system that can meet all the above criteria by reasonable modification of existing equipment and software. During this phase it became increasingly apparent that the system must also be able to measure infrared reflectance as well as provide colorimetric data. Major elements in the design focussed on

(1) Minor optical changes

(2) Development of acceptability criteria

(3) Development of a system-wide calibration

(4) Development of a detailed specification for procurement of a prototype system.

These objectives were accomplished and are discussed in detail in Section 5.

c. **Phase III. Assembly of System.** The purpose of Phase III is to acquire and evaluate a prototype system based on the knowledge derived from Phases I and II and to evaluate the design in a realistic situation. Specifications for procurement of a two-unit system were developed on the basis of contract effort with Rensselaer Polytechnic Institute and guidance from the NRC committee and in-house studies. These specifications consist mostly of elaboration and quantification of the criteria listed above. In addition, a few other requirements were added.

(1) The instrument must be able to measure and print out reflectances 400 to 900 nm.

(2) Provision was made for mass storage of data and for subsequent analytical and statistical computation.

11

(3) Communication capabilities between units is required.

A two-unit system has been procured; the evaluation aspect is under way. One of the two units is located at DPSC, mainly for acquisition and processing of primary data and for transmission of these data to NLABS. The data will be processed at DPSC to the degree needed to determine acceptability. In Phase III such data are being used to determine whether the objectively structured method is valid.

The second unit is located at NLABS. This unit is being used for several purposes. It must first be determined that the two units agree and can function as a single system. Second, both units are being used to determine whether further modifications of either equipment or methodology are required. The NLABS unit is also being used to perform various analytical work related to the system calibration and extension of the procedure for defining acceptability to a wider range of shades. The NLABS unit will also be used to perform the correlation studies between units and between visual and instrumental judgments.

Finally, the two-unit system will be tested in a large procurement to determine whether it can provide the basis for reliable decisions on acceptability. Only when the system has earned our confidence will the program proceed to Phase IV.

d. **Phase IV. System Trial.** Following the operational trial in Phase III, the system will be used in the procurement of shade for which all data have been developed. Decisions on acceptability will be based on the trial's results but will be checked by visual observations.

4. Phase I. Survey of Commercial Equipment

The practicability of using color-measuring instruments to determine shade acceptability of textile samples has long been known and occasionally practiced in many industries concerned with color. The resulting data usually serve as a supplement to visual observations, despite evidence published as early as 1969 attesting to the ability of instruments to reliably discriminate color differences smaller than those discerned by observers, even when viewing conditions were ideal. Several studies on the comparative performance of color-measuring instruments were conducted by the Rensselaer Color Measurement Laboratory on instruments commercially available at the time.[2-6] Recent publications dealing with textile samples have shown the

[2] F. W. Billmeyer, Jr., Comparative performance of color measuring instruments, Appl. Optics 8, 775–783 (1969).

[3] F. W. Billmeyer, Jr., E. D. Campbell, and R. T. Marcus, Comparative performance of color-measuring instruments; Second report, Appl. Optics 13, 1510–1518 (1974); ibid. Authors' reply to comments, Appl. Optics 14, 265 (1975).

[4] R. T. Marcus and F. W. Billmeyer, Jr., Statistical study of color-measurement instrumentation, Appl. Optics 13, 1519–1530 (1974).

[5] F. W. Billmeyer, Jr. and P. J. Alessi, Assessment of color-measuring instruments, Color Res. Appl., 6, 195–202 (1981).

[6] F. W. Billmeyer, Jr. and P. J. Alessi, Assessment of color-measuring instruments for objective textile acceptability judgment, NATICK/TR–79/044, US Army Natick R&D Laboratories, Natick, MA, March 1979. (AD A081 231)

results from instrumental shade passing are more consistent that the results obtained visually for the same samples.[7-10] A study of the precision, accuracy and reproducibility of current color-measuring instruments for use in the measurement of textile samples for the Army and the Department of Defense was the objective of the contract awarded Rensselaer Polytechnic Institute. Data resulting from this study have contributed to defining criteria for instruments that are suitable for use in the color measurement program.

a. Instruments

The newly developed instruments listed in 3a. are three of the five originally scheduled for evaluation by Rensselaer, the remaining two being a Diano-Hardy II Spectrophotometer and the NLABS Hunter D38D Colorimeter. The latter two instruments were soon eliminated from consideration: the Diano-Hardy II because of its high cost and relative slowness of operation; the Hunter D38D, a colorimeter, because a spectrophotometer was required for determining spectral reflectance factor data on metameric pairs and also in the infrared. The three new spectrophotometers, all stand-alone units with their own computational capability, were marketed as having both higher precision and reproducibility than older instruments plus the advantages of speed and ease of operation.[11,12] The moderate cost of all three instruments contributes to their attractiveness. The Macbeth MS—2000 is an abridged double-beam spectrophotometer (Figure 3) having a pulsed xenon flash-tube as its light source.[13] The Hunter D54P—5, a single-beam scanning spectrophotometer (Figure 4), uses a quartz-halogen tungsten-filament lamp as source.[14] The most conventional instrument of the three is the

[7]S. M. Jaekel, Utility of color-difference formulas for match-acceptability decisions, Appl. Optics 12, 1299–1316 (1973).

[8]K. Jeltsch and X. Fink, Analysis of errors in acceptability experiments, J. Soc. Dyers Colourists 92, 227–232 (1976).

[9]K. McLaren, An introduction to instrumental shade passing and sorting and a review of recent developments, J. Soc. Dyers Colourists 92, 317–326 (1976).

[10]S. M. Jaekel and C. D. Ward, Practical instrumental colour quality control-the Hatra experience, J. Soc. Dyers Colourists 92, 353–363 (1976).

[11]F. W. Billmeyer, Jr. and D. C. Rich, Color measurement in the computer age, Plastics Eng. 34, 35–39 (1978).

[12]D. C. Rich and F. W. Billmeyer, Jr., Practical aspects of current color measurement instrumentation for coatings technology, J. Coatings Technol. 51, 45–47 (1979).

[13]S. J. Kishner, A pulsed-xenon spectrophotometer with parallel wavelength sensing, in F. W. Billmeyer, Jr. and G. Wyszecki, Eds., Color 77, Adam Hilger Bristol, 305–308 (1978).

[14]R. Stanziola, B. Momiroff, and H. Hemmendinger, The Spectro Sensor-A new generation spectrophotometer, in F. W. Billmeyer, Jr. and G. Wyszecki, Eds. Color 77, Adam Hilger, Bristol, 313–315 (1978), and Color Res. Appl. 4, 157–163 (1979).

Diano Match-Scan, a double-beam spectrophotometer (Figure 5), which also uses a quartz-halogen tungsten filament lamp as source. All three utilize integrating-sphere geometry, diffuse polychromatic irradiation and near-normal viewing, optically filtering their light sources to simulate CIE Standard Illuminant D_{65}. The Diano Match-Scan also has the option of 8° monochromatic irradiation and diffuse viewing. Both the Macbeth and Diano instruments employ grating monochromators in their designs, the former a fixed grating and the latter a single-pass grating. The Hunter instrument uses a rotating interference-wedge monochromator. The detectors differ for each unit, a photomultiplier tube being used by the Diano, a silicon-photodiode by the Hunter and an array of 17 silicon-photodiodes, each one simultaneously receiving a 20 nm interval projection from the monochromator, by the Macbeth. The Hunter instrument is optically identical to the Spectro Sensor, a unit available from Applied Color Systems.[15]

b. Samples

To assess properly the capability of an instrument to measure a sample reproducibly, it first must be established that the samples have been carefully prepared and are themselves reproducible. Many materials other than textiles have met the criteria, e.g., ceramic tiles, transmitting filters, and carefully prepared painted papers. Reproducibility of measurements of textile samples is more uncertain, owing to the differences and possible instability of surface characteristics and weave among fabrics. The feasibility of using textile samples for this study was investigated, and the specific areas of application were determined.

Various sets of samples were used to evaluate instrument criteria; the same set was used in the investigation of several parameters. Calibrated standards such as the NBS Standard Reference Materials 2101–2105, a set of transmitting filters, and the British Ceramic Research Association (BCRA) Reflecting Tiles, calibrated at the Hemmendinger Color Laboratory, were used to test accuracy of color measurement.[16-18] Also used were six sets of tiles produced by NLABS and previously used in instrument studies by Rensselaer. Each set of ceramic tiles contained a standard, a high-limit sample and a low-limit sample for a different Army shade.[19,20]

[15] See reference 14.

[16] H. J. Keegan, J. C. Schleter, and D. B. Judd, Glass filters for checking the performance of spectrophotometer-integrator systems of color measurement. J. Res. Nat'l Bur. Stand. 66A, 203–211 (1962).

[17] K. L. Eckerle and W. H. Venable, Jr., 1976 Remeasurement of NBS spectrophotometer integrator filters, Color Res. Appl. 2, 137–141 (1977).

[18] F. J. J. Clarke, Ceramic colour standards-an aid for industrial colour control, Printing Technology 13, 101–113 (1969).

[19] See reference 2.

[20] See reference 3.

Figure 3.
Macbeth MS–200
Spectrophotometer

Figure 4.
Hunter D54P–5
Spectrophotometer
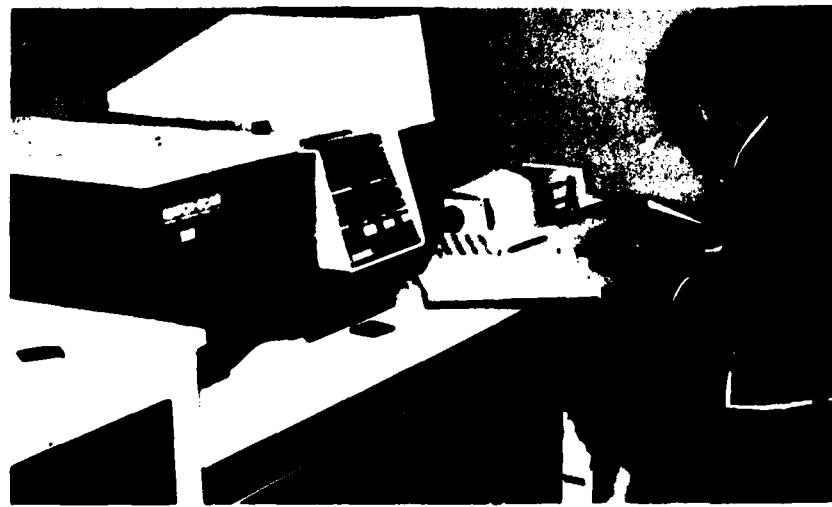
Figure 5.
Diano Match-Scan
Spectrophotometer

15

The short-term repeatability of each instrument was evaluated by measurements on the NBS SRM 2101—2105 filters, the BCRA tiles, a set of varied-colored highly fluorescent samples, and the NLABS textile samples. The latter are a collection of sixteen sets, each set a different military shade or fabric type, composed of a standard plus eight tolerance limits. The tolerance limits had been established visually. The textile samples, BCRA tiles, and ten "Carrara" glass tiles were used to determine long-term repeatability of the instruments.

Sensitivities to the following parameters were also investigated:

(1) Precision required of measurements to be used in computing color differences was determined by measuring the NLABS porcelain-enamel tiles and the sets of textile samples.

(2) Effect of weave orientation on the measurements was determined by comparison of the measurement data from the same sample area after rotation of the sample, in its own plane, 45° and 90° from the reference position. Specific samples from each of the NLABS textile sets were used.

(3) Effectiveness of each instrument in rejecting the specular component of reflectance was determined by measurement of the glossy "Carrara" glass tiles and the mainly matte-surfaced textile samples with specular both included and excluded.

(4) Differences among photometric scales were evaluated relatively, because of the non-availability of calibrated nonselective neutral samples that would be required for tests of absolute accuracy. Eleven Johnson gray tiles were measured on each instrument and the scaled CIE Y (D_{65}) values were compared.[21]

(5) Within-piece variation of textiles was investigated by a comparison of the color differences calculated between the center and each of the four corners of a specimen, using the samples in the NLABS textile sets.

(6) Sensitivity of the instruments to point-to-point variations within a specimen was determined by measurements on the BCRA tiles. Data were collected and analyzed in a similar manner to (5).

c. Test Procedure

Rensselaer used white ceramic tiles, calibrated to the 1977 NBS scale of absolute reflectance, as the preferred white standards for the test measurements.[22] Pressed barium sulfate standards were used only when necessary, as in inter-instrument comparisons. The CIE 1976 L*a* b* (CIELAB) color coordinates and color differences, computed for Illuminant D_{65}

[21] A. R. Robertson and W. D. Wright, An international comparison of working standards for colorimetry, J. Opt. Soc. Am. 55, 694—706 (1965).

[22] See reference 16.

and the 1931 Standard Observer were used in all instances except where comparisons were made to literature values.[23,24] In the latter cases, Illuminant C/1931 Standard Observer, and FMC–2 color differences were computed.[25] All Rensselaer data reported were the result of multiple measurements of the specimens.[26] For example, short-term repeatability studies involved five repeat measurements per sample. The NBS filters were removed and replaced between measurements, whereas the BCRA tiles and the fluorescent samples remained in place for repeat measurements. Long-term repeatability studies involved measurement of the samples once a week for eight weeks. The combined short and long-term repeatability studies were a combination of the two data-gathering procedures. In the majority of cases, the data were reported as color differences, having been calculated between the tristimulus values for each measurement and the mean tristimulus values for the multiple measurements.

d. Conclusions

The final report of the Rensselaer evaluation led to the following conclusions.[27]

(1) Each of the three color-measuring instruments tested had adequate short-term and long-term repeatability and adequate accuracy to carry out the necessary measurements for NLABS' proposed program of objective textile acceptability judgements.

(2) However, none of the instruments evaluated was found to be fully satisfactory for the proposed program from operational and computational points of view. Modifications to each instrument would be required to produce a completely suitable configuration.

(3) With appropriate modifications, any one of the three instruments tested should be suitable for NLABS' requirements.

The data generated from Rensselaer's test studies indicate that the repeatability of color measurement with all three instruments was within 0.1 CIELAB color-difference unit. The data in Table 1 show each instrument's precision of measurement of the same area of a sample;

[23] Recommendations of Uniform Color Spaces, Color-Difference Equations, and Psychometric Color Terms, Supplement No. 2 to CIE Publication No. 15, Colorimetry, E–1.3.1 (1971), Bureau Central de la CIE, Paris (1978).

[24] Publications CIE No. 15 (E–1.3.1 1971), Colorimetry Bureau Central de la CIE, Paris 1971; Supplement 1, 1972; Available from the U.S. National Committee, CIE, National Bureau of Standards, Washington, D.C. 20234.

[25] K. D. Chickering, FMC color difference formulas: clarification concerning usage, J. Opt. Soc. Am. 61, 118–121 (1971).

[26] See reference 6.

[27] See reference 6.

i.e., repeat measurements on precisely the same area. The results are virtually identical for non-fluorescent and fluorescent samples, on a short-term basis. Examination of long-term means suggest some differences of performance among the three instruments. These are not satistically significant, because all of the means fall within one standard deviation of one another. Data on point-to-point variations over a sample are listed in Table 2. In no case is the mean color-difference from a mean greater than 0.1 CIELAB unit in all the tests performed, both on a short-term and long-term basis.

To determine whether the repeatability of instrumental color measurement as found was of significance in the measurement of Army shades, the data were compared with the sizes of the color differences to be measured. Rensselaer averaged the mean color differences between the standard and each of the eight limit standards per textile shade, using the data resulting from the measurement of these samples at 0° orientation. These data are listed in Table 3.

Rensselaer's conclusion that any one of the three spectrophotometers evaluated is suitable for performing the objective textile acceptability judgements is based on the following three findings:

(1) The repeatability of color measurement for each instrument is better than 0.1 CIELAB unit.

(2) The mean color differences exhibited in Table 3 are always much larger than 0.1 CIELAB unit.

(3) The performance of the three instruments is satistically indistinguisable.

Selection of one of the three instruments for use in the program depended on factors other than their overall performance in terms of repeatability of color and color-difference measurements. Rensselaer indentified the operational computational modifications they felt were required of each instrument to obtain a fully satisfactory system. In addition, personnel at NLABS identified other design modifications. These are discussed in Section 5.

5. Phase II. Design of System

The findings and recommendations from Phase I served as the basis for actions in Phase II, the major steps of which are listed in 3.b. These culminated in the design of a prototype system.

a. Optical Modifications

(1) Infrared requirement. Each of the three instruments can be modified by straightforward means. The Diano Match-Scan requires only additional software and a filter to reject the second-order spectrum from the grating for infrared measurements. This has been done commercially for at least three customers. The Hunter D54P—5 requires similar software additions, replacement of the monochromating wedge by another commercially available wedge and relocation of the D_{65} filter to the filter wheel. In addition to added

18

Table 1. Precision of Measurement of the Same Area of a Specimen (Color Differences in CIELAB Units for Illuminant D_{65})

	Repeatability		
Instrument	Short-Term Fluorescent	Short-Term Non-Fluorescent	Long-Term Carrara Glasses
Macbeth MS—2000	0.03 ± 0.018	0.02 ± 0.010	0.06 ± 0.019
Hunter D54P—5	0.03 ± 0.025	0.03 ± 0.014	0.06 ± 0.053
Diano Match-Scan	0.03 ± 0.010	0.03 ± 0.007	0.08 ± 0.042

Table 2. Short-Term and Long-Term Repeatability of Color Measurement When Within-Sample Variation is Included (Color Differences in CIELAB Units for Illuminant D_{65})

	Repeatability		
Instrument	NBS SRM 2101–2105 Short-Term	BCRA Tiles Long-Term	Textile Standards Long-Term
Macbeth MS—2000	0.07 ± 0.033	0.05 ± 0.026	0.08 ± 0.043
Hunter D54P—5	0.04 ± 0.035	0.07 ± 0.077	0.07 ± 0.028
Diano Match-Scan	0.05 ± 0.035	0.07 ± 0.038	0.09 ± 0.038

19

Table 3. Average Color Differences for Textile Acceptability

NLABS Textile Set	NLABS MS–2000	NLABS Hunter D54P–5	NLABS Match-Scan	Grand Mean over 3 Instruments
AG 344 P/W Gab.	0.60 ± 0.315	0.56 ± 0.281	0.52 ± 0.271	0.56 ± 0.287
OG 106 Ox./Nyl.	2.49 ± 1.411	2.43 ± 1.346	2.45 ± 1.318	2.46 ± 1.340
OG 107 Nyco. Pop.	0.92 ± 0.300	0.86 ± 0.307	0.91 ± 0.299	0.90 ± 0.299
OG 107 Ctn. B'loon	1.32 ± 0.490	1.56 ± 0.518	1.61 ± 0.516	1.49 ± 0.516
OG 107 Ctn. Sat.	1.24 ± 0.319	1.27 ± 0.216	1.22 ± 0.285	1.24 ± 0.274
OD7 Ctn. Duck MRWRP	1.24 ± 0.815	1.31 ± 0.834	1.32 ± 0.800	1.29 ± 0.802
Tan 46 Ctn. Pop.	1.54 ± 0.581	1.48 ± 0.608	1.54 ± 0.576	1.52 ± 0.580
AG 344 P/W Trop.	1.54 ± 0.668	1.72 ± 0.672	1.58 ± 0.656	1.61 ± 0.661
Tan 445 Twill poly./cot.	0.88 ± 0.793	0.85 ± 0.724	0.84 ± 0.722	0.86 ± 0.737
Blue 150 Gab./Wool	0.29 ± 0.127	0.35 ± 0.080	0.31 ± 0.105	0.32 ± 0.108
OD7 Ctn. Duck UTRD	1.61 ± 0.734	1.72 ± 0.853	1.63 ± 0.799	1.65 ± 0.787
AG 44 Wl. Serge	0.36 ± 0.083	0.38 ± 0.070	0.39 ± 0.077	0.38 ± 0.077
Tan M1 Cl. P/W Trop.	0.59 ± 0.342	0.55 ± 0.334	0.58 ± 0.322	0.57 ± 0.329
Blue 150 Trop. Wl.	0.48 ± 0.174	0.43 ± 0.143	0.42 ± 0.168	0.45 ± 0.162
OG 108 W/N Fl. Shirt	1.09 ± 0.325	1.08 ± 0.353	1.13 ± 0.388	1.10 ± 0.352
Blue 151 Wool Trop.	0.71 ± 0.263	0.71 ± 0.239	0.72 ± 0.249	0.71 ± 0.247

software, the Macbeth MS–2000 requires a different grating and detector array or other provisions for accomplishing the same results, perhaps by rotation of the existing grating for the infrared measurements; other options for extending the range are available. In no case are these changes considered to be especially difficult.

(2) Polarization. Both grating instruments, the Macbeth MS–2000 and the Diano Match-Scan, displayed some polarization of the beam. This was adequately corrected by use of a retardation plate with the Diano Match-Scan, whereas the Macbeth MS–2000 was found to be within allowable limits established for the procurement of this prototype system. The requirement provides for a variation of not more than four percent of measured spectral transmittance of a polarizing filter as it is rotated through 360 degrees. Textile specimens are expected to display far less effect on reflectance due to sample orientation.

(3) Illumination. Because the system must handle both fluorescence and metamerism, it is necessary that the instruments provide both a simulated daylight illumination as well as tungsten illumination. Source A colorimetric data cannot be calculated for a fluorescent specimen, if measurements are acquired with a lamp that simulates Illuminant D_{65}. A slight modification of two of the instruments is required to meet this requirement; it is not certain whether the Macbeth instrument can be modified appropriately.

The Illuminant D_{65} simulation is achieved by using selected filters to modulate the primary source, whether it be a xenon flash or tungsten-halogen lamp. Because some of the specimens to be measured may be flourescent, it is important that the spectral power distribution of the "daylight" illumination closely resemble that of CIE Illuminant D_{65} in both the ultraviolet and visible regions of the spectrum. It was concluded that all three instruments comply with the latest CIE suggestions for spectral power distribution sufficiently well for purposes of this development.

b. Acceptability Criteria

Small color differences have been determined for many years. The well-trained, careful operator with a good spectrophotometer and integrator or with a good colorimeter has been able to obtain useful data. Modern instrumentation makes possible even more reliable results.

The major problem in applying any of a number of color difference equations to the NLABS development is that they are based on equal perceptible differences in color. It is well known but not very well documented that acceptability is often biased in terms of perceptual color-difference data. Thus, conventionally obtained colorimetric data must be skewed in some manner to fit the visually acquired acceptability criteria.

The principles for expressing acceptability in colorimetric terms have been developed and applied to three shades under contract with Lehigh University.[28]

[28] E. Allen and B. Yuhas, Investigations to define acceptability tolerance ranges in various regions of color space, Final Technical Report on Contract DAAK60–78–C–0084, US Army Natick R&D Labs, Natick, MA 01760, Feb 80, AD A094 163.

Three shades were studied: Olive Green 107, Tan, and Blue 150. From the colorimetric data on over 200 specimens for each of the three shades, pairs of samples differing only in hue, chroma or lightness were chosen. For each of the three shades, twelve pairs of samples were chosen (four color difference levels in three directions). The smallest color difference should pass most of the time and the largest should fail most of the time. The two intermediate color differences would be chosen close to the tolerance limits. The reference color was not necessarily the same for more than one series, nor was it the official standard for the shade. Plus and minus directions in color space were obtained by reversing the standard and sample.

A panel of six observers viewed each sample pair ten times in a random order and made a judgment as to the acceptability of the color match. The method of logistic functions was then applied to the acceptance data to determine the color difference that would be accepted 50% of the time.

Data for each of the six directions about the reference were analyzed for each shade. The major conclusions of the study are summarized as follows:

(1) Acceptability thresholds were defined by ellipsoids in CIELAB space.

(2) For Olive Green and Tan, orientation of the ellipses in the chromaticity plane was in the direction of an iso-hue line. In the case of the Blue 150 dress fabric, the orientation of the ellipse was more along an iso-chroma line.

(3) The ratio of the lengths of the three axes in CIELAB units was approximately 3:2:1 in lightness, chroma, and hue for Olive Green and Tan shades.

(4) The over-all sizes of the ellipses depended on the observers' knowledge of the ultimate end use of the fabric. In decreasing order, the sizes of the ellipses were for the olive green, tan and blue.

(5) The size and shape of each ellipsoid is defined by three constants. A fourth constant is needed to define the orientation of the ellipse in the chromaticity plane.

(6) From the evaluation of twelve current Army standards and their limit samples, it was found that the standard was not centered in the locus of acceptable shades. Some bias was noted; e.g., a deviation in the green direction might be more objectionable on a tan shade than a red deviation.

(7) The end results of the analysis are equations for an acceptability index for each of the three shades. These equations are in the form

$$\Delta A = [g_{11}(\Delta a^*)^2 + 2g_{12}\Delta a^* \Delta b^* + g_{22}(\Delta b^*)^2 + g_{33}(\Delta L^*)^2]^{1/2}$$

(8) Constants have been determined to be different for each of the three shades in order for the acceptability index to equal unity for barely acceptable specimens. These constants are programmed into the computer in appropriate files for each of those shades.

(9) The procedure for determining similar constants for other shades, including the necessary computer software for analysis, has been developed. Data are being acquired on the four colors of the Woodland Pattern to compare results by the visual and analytical methods.

c. System-Wide Calibration

The calibration requirements for a multi-unit system are understandably more rigourous than for a single unit that operates alone. In a multi-unit system each unit must not only agree with itself over a long period of time but each unit must agree with others in the system. The two-unit system that has been procured (see d. below) was furnished with a system calibration procedure and standards by the manufacturer. Under contract, Clemson University has reported on the design of such a calibration procedure and provided the necessary standards. Both procedures have at least one feature in common, fail-safe operation.[29] The calibration procedures are tied into the software in such a manner that the instrument cannot operate unless all steps in the calibration procedure have been carried out and the measured values reproduce stored values within predetermined, acceptable limits.

The Clemson study considered procedures for both the Match-Scan and Hunter instruments because the decision as to which instrument would be used in the system had not been made at that time. The two procedures were similar in concept and material standards. They differed primarily in operating details related to optical differences between the two instruments. It is not intended here to reproduce the procedures in detail but only to provide an outline of the logic used. Reference 29 gives a full description, complete with experimental results, computer programs, and description of material standards. Several materials were selected to assure proper calibration for one or more parameters of operation.

(1) A white opal glass is used to check nominal 100 percent reflectance throughout the spectral range of interest. This glass has previously been calibrated and its data are stored in the computer. Correction factors for the 100 percent line are calculated and applied to subsequent measurements, provided that the values measured at the time of calibration are within acceptable limits.

(2) A Corning 3307 Signal Yellow filter is used to check photometric linearity at low and high values, at wavelengths were the transmittance curve is relatively flat. By choosing a particular wavelength where the rising transmittance curve is linear, the wavelength of the instruments may be calibrated.

(3) A gray porcelain tile is used to check the photometric linearity at a mid-range reflectance level.

[29] F. T. Simon and J. H. Lubar, Standardization procedure for two instruments for color measurement, NATICK/TR—82/024. US Army Natick R&D Laboratories, Natick, MA 01760, Sept 1981.

(4) A pair of tan polyester gelcoat plaques of known color difference are used to check the ability of the unit to measure repeatably both differences in spectral reflectance and to compute color differences.

d. Procurement of Prototype System

The above results were included in procurement documents for an expandable two-unit prototype system. As indicated earlier, one unit is now located at NLABS, the second at DPSC. The two units are being evaluated thoroughly to assure that they work reproducibly together as a system and meet the criteria set forth in 3.a. When confidence in the prototype system has been demonstrated, the system will be expanded by locating a few units in contractors' plants. That event will signify the beginning of the verification step of Phase III. It should not be surprising if a few minor modifications, especially in operating procedure, may be made at that point.

6. Phase III. Assembly of System

a. Overview of System Capabilities

The instrument procured through competitive bidding was the Applied Color System (ACS) Spectro Sensor, which is essentially the same as the Hunter D54P—5 studied earlier, with the exception of a modification to scan into the near infrared. Instruments are present at both NLABS and DPSC, but the associated computers have different capabilities. The computer at NLABS is a Digital Equipment Corporation (DEC) PDP 11/23 with 124K of memory, a DEC VT100 video terminal, two DEC LA120 Decwriter Hard Copy Terminals (one for use as a line printer and one as an additional tie-in to the computer), two DEC RL02 disk drives and a modem interface that connects to DPSC. The computer at DPSC is a DEC PDP 11/03 with 32K of memory, a DEC VT100 video terminal, a DEC LA120 for use as a line printer, two DEC RX02 floppy disk drives and a modem interface to transmit data to NLABS. Figures 6 and 7 show the two systems. The main differences in the two computers is that the one at DPSC has a RT—11 single-user operating system and the one at NLABS has an RSX—11M multi-user operating system and more memory.

The software to run the instrument was written by ACS especially for NLABS. The complete program package, called NLASS (Natick Labs Assessment and Sample Storage System), includes programs for measurement, acceptability testing, file maintenance and data transmission and reception. Table 4 shows which programs are available at each installation, and the following paragraphs briefly describe each program.

Table 4. Summary of Programs Available at Each Installation

DPSC	NLABS
NFILX	NFILX
NLAM	NLAM
NLMM	NLMM
NXMIT	HNRCV
	NLFILE
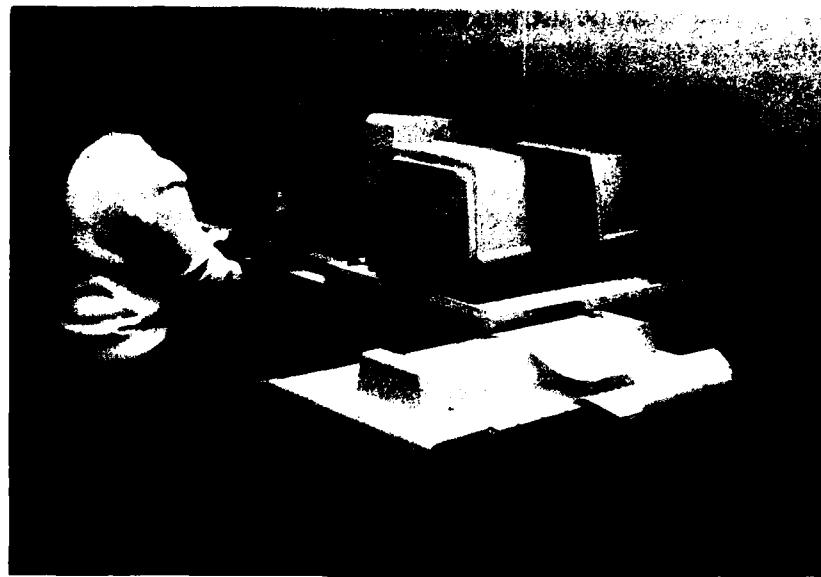
Figure 6. The prototype unit located at U3 Army Natick R&D Laboratories
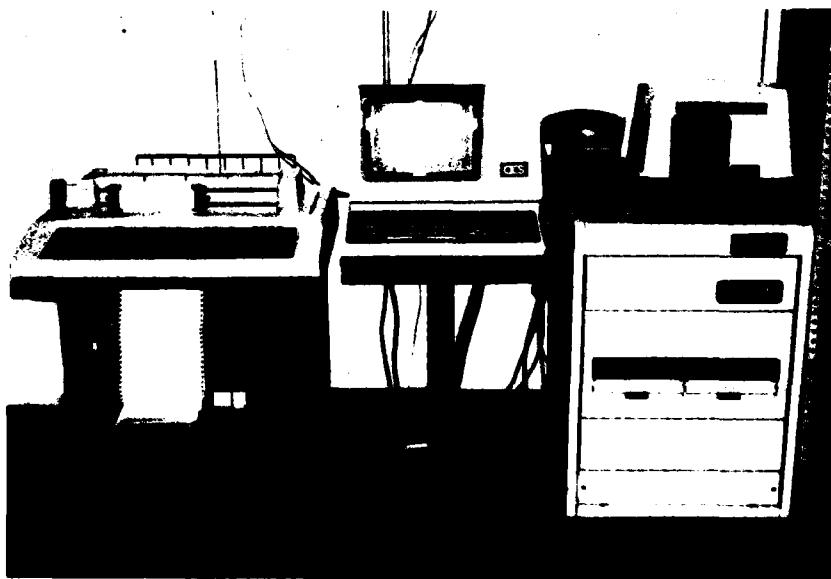
Figure 7. The prototype unit located at the Defense Personnel Support Center

The NFILX program can be used to create, store, list, edit and delete standard files. These files contain the acceptability parameters for the standard shades that the Army has. This program will also create, list or delete from a sample file.

The NLAM program measures a sample to be evaluated for shade and compares it to the values stored by NFILX for the standard. The program gives a pass/fail judgement and prints the acceptability factor. This data can then be stored, if the operator chooses that option.

The NLMM program is similar to the NLAM program, but it only measures samples; it does not evaluate them for acceptability. Both of these programs use a fail-safe calibration procedure to insure that the instrument is working properly. This calibration procedure measures a black trap to set the zero, a white tile to calibrate to 100%, and, because the Spectro Sensor is a single-beam instrument, a gray tile to determine the sphere correction factor. The program checks to see that the tiles were measured in the correct order and that the light source the user has chosen for calibration is the same as that for which the instrument is actually set up. If the gray tile cannot be measured repeatably, the calibration fails and must be redone. A green tile is measured after the gray tile and compared to stored values, to check photometric, wavelength and colorimetric accuracy. If it passes, measurement is enabled; otherwise the calibration procedure must be repeated, after whatever was wrong has been corrected. Almost invariably, simple correction of operation order or easy cleaning of one of the tiles is sufficient to correct the trouble and to enable the spectrophotometer.

The program NXMIT is used by DPSC to transfer the data they have stored with NLAM to NLABS over telephone lines. This transfer can be done by record number, by query (where the computer displays each record and asks if it is to be transferred) or automatically, where all of the stored records are transferred.

Using HNRCV, the computer at NLABS receives the data from DPSC. Because it is a multi-user system, the NLABS computer can be used for other purposes at the same time that data are being transmitted. When the transfer is complete, a message is printed on the NLABS line printer indicating the number of records that have been transmitted.

When DPSC transfers data to NLABS, the data goes into a sample file. The records must then be transferred to permanent library files by using the file maintenance program, NLFILE. These library files are generally on a separate disk, which will allow NLABS to store more efficiently the large amounts of data that will be received from DPSC. By use of the NLFILE program, NLABS can also replicate the colorimetric and acceptability calculations of records transferred by DPSC.

b. Testing the System

Upon installation of the two units, completion of preliminary training, and familiarization with the equipment, testing of the units for reliability began. It is necessary that both units are shown to produce repeatable results. It is also necessary to show that they agree with each other. At the present time, the instrument at NLABS is being tested for short and long-term

reproducibility with two sets of porcelain-enamel tiles and fabric samples. After the performance of the NLABS instrument has been determined, the samples will be taken to DPSC to establish the performance of their instrument and to evaluate how well the two instruments agree.

Parallel with this work, data are being accumulated on the four shades of the Woodland Camouflage pattern in order to implement the Allen-Yuhas method for determining acceptability parameters.

7. Conclusions

Three commercial spectrophotometers were evaluated under Phase I. All three were found to be capable of meeting the requirements with minor modifications.

Under Phase II, a two-unit prototype system was designed and procured. Each unit consists of a spectrophotometer, computer, and specialized software. A method of quantifying acceptability tolerances is incorporated into the software. Two special calibration methods are available, which will be evaluated in Phase III.

Phase III work has begun. This will establish the validity of the overall method and identify any necessary adaptations in design or method. Two elements of this phase are of primary significance. The two units must be self-consistent and they must also agree with each other. Second, data banks on acceptability tolerances must be entered into the computer and the resulting color measurements on actual textile specimens satisfactorily compared with visual observations by the current method. Phase III will be reported when the work is completed for three or four shades.

Actual implementation of the system in Phase IV is expected to begin by 1984.

REFERENCES

Allen, E. and B. Yuhas, Investigations to define acceptability tolerance ranges in various regions of color space. Final Technical Report on Contract DAAK60—78—C—0084, US Army Natick Research and Development Laboratories, Natick, MA 01760, Feb 80. AD A094 163.

Billmeyer, Jr., F. W. Comparative performance of color measuring instruments, Appl Optics 8, 775—783 (1969).

Billmeyer, Jr. and P. J. Alessi, Assessment of color-measuring instruments for objective textile acceptability judgment, NATICK/TR—79/044, US Army Natick R&D Laboratories, Natick, MA, March 1979. (AD A081 231)

Billmeyer, Jr., F. W. and P. J. Alessi, Assessment of color-measuring instruments, Color Res. Appl., 6, 195—202 (1981).

Billmeyer, F. W., Jr., E. D. Campbell, and R. T. Marcus, Comparative performance of color-measuring instruments; Second report, Appl. Optics 13, 1510—1518 (1974); ibid. Authors' reply to comments, Appl. Optics 14, 265 (1975).

Billmeyer, Jr., F. W. and D. C. Rich, Color measurement in the computer age, Plastic Eng. 34, 35—39 (1978).

Chickering, K. D., FMC color difference formulas: clarification concerning usage, J. Opt. Soc. Am. 61, 118—121 (1971).

Clarke, F. J. J., Ceramic colour standards-an aid for industrial colour control, Printing Technology 13, 101—113 (1969).

Eckerle, K. L., and W. H. Venable, Jr., 1976 Remeasurement of NBS spectrophotometer integrator filters, Color Res. Appl 2, 137—141 (1977).

Federal Standard, CCC—T—191b, Test Method 9010.

Jaekel, S. M., Utility of color-difference formulas for match-acceptability decisions, Appl. Optics 12, 1299—1316 (1973).

Jaekel, S. M. and C. D. Ward, Practical instrumental colour quality control-the Hatra experience, J. Soc. Dyers Colourists 92, 353—363 (1976).

Jeltsch, K. and X. Fink, Analysis of error in acceptability experiments, J. Soc Dyers Colourists 92, 227—232 (1976).

Keegan, H. J., J. C. Schleter, and D. B. Judd, Glass filters for checking the performance of spectrophotometer-integrator systems of color measurement. J. Res. Nat'l Bus. Stand. 66A, 203—211 (1962).

Kishner, S. J., A pulsed-xenon spectrophotometer with parallel wavelength sensing, in F. W. Billmeyer, Jr., and G. Wyszecki, Eds., Color 77, Adam Hilger Bristol, 305–308 (1978).

Marcus, R. T. and F. W. Billmeyer, Jr., Statistical study of color-measurement instrumentation, Appl. Optics 13, 1519–1530 (1974).

McLaren, K., An introduction to instrumental shade passing and sorting and a review of recent developments, J. Soc Dyers Colourists 92, 317–326 (1976).

Publications CIE No. 15 (E–1.3.1 1971), Colorimetry Bureau Central de la CIE, Paris 1971; Supplement 1, 1972; Available from the U.S. National Committee, CIE, National Bureau of Standards, Washington, D.C. 20234.

Recommendations of Uniform Color Spaces, Color-Difference Equations, and Psychometric Color Terms, Supplement No. 2 to CIE Publication No. 15, Colorimetry, E–1.3.1 (1971) Bureau Central de la CIE, Paris (1978).

Rich, D. C. and F. W. Billmeyer, Jr., Practical aspects of current color measurement instrumentation for coatings technology, J. Coatings Technol. 51, 45–47 (1979).

Robertson, A. R and W. D. Wright, An international comparison of working standards for colorimetry, J. Opt. Soc. Am. 55, 694–706 (1965).

Simon, F. T. and J. H. Lubar, Standardization procedure for two instruments for color measurement, NATICK/TR–82/024. US Army Natick R&D Laboratories, Natick, MA 01760, Sep 1981.

Stanziola, R., B. Momiroff, and H. Hemmendinger, The Spectro Sensor-A new generation spectrophotometer, in F. W. Billmeyer, Jr., and G. Wyszecki, Eds. Color 77, Adam Hilger, Bristol, 313–315 (1978), and Color Res. Appl. 4, 157–163 (1979).