COLOR SHADE INSTRUMENTATION CORRELATION STUDY

STATISTICAL ANALYSIS

REPORT DL827T1 (REV. 1)

Alan R. Metelko
Luisa DeMorais
Melanie King
Rachel Matuszek
Jack J. Vandenberghe
Nathaniel J. Wurst

LMI

MARCH 2011

NOTICE:

THE VIEWS, OPINIONS, AND FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF LMI AND SHOULD NOT BE CONSTRUED AS AN OFFI-CIAL AGENCY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER OFFICIAL DOCUMENTATION.

Executive Summary

LMI assessed the use of spectrophotometers as an aid to military shade evaluators in the evaluation of military fabrics. Spectrophotometers detect and quantify color. Military fabric suppliers currently use spectrophotometers as part of their quality control process; however, when the suppliers ship their fabric to clothing manufacturers, the procedure for determining whether the fabric meets the military shade standard depends solely on visual evaluation. Visual evaluation by trained shade evaluators is widely considered superior to instrumentation because the human eye is more sensitive to shade variation. This is especially important in the evaluation of multicolored camouflage patterns.

The feedback from visual evaluations that is provided to fabric suppliers is often general and vague. Without more useful feedback, the fabric suppliers have difficulty making the proper adjustments to their printing and dying processes to meet the given shade standard for a specific fabric in future production. Spectrophotometer measurements provide more quantitative feedback for shade evaluations, but concerns regarding the consistency of spectrophotometer measurements have raised questions about the use of these measurements in government shade evaluations and the feedback they provide fabric suppliers.

This study addressed two main questions:

- Do spectrophotometer measurements collected at different locations produce consistent color measurements?
- Given that the instruments do not have the same sensitivity to shade as the human eye, to what extent do human visual evaluations agree with spectrophotometer evaluation?

The study was conducted in three phases. Phase 1 was an exploratory study to evaluate the feasibility of our data collection strategy and validate our methodology. In Phase 2, spectrophotometer measurements of color tiles and fabrics were collected from 19 test sites to see if the measurements were consistent with one another, both within and across test sites. With only a few exceptions, these measurements were broadly consistent.

Phase 3 involved the following:

- Recruiting a panel of 10 people to evaluate whether 100 fabric samples met designated color shade pass/fail criteria.
- Using a spectrophotometer to generate pass/fail decisions on these same fabrics.
- Obtaining an “official determination” from experts on color shades to ascertain whether the samples should have received pass/fail designations.
- Comparing the pass/fail designations of the 10 people with the pass/fail decisions generated by the spectrophotometer in terms of how often these decisions agreed with the official determination.

On average, the pass/fail decisions generated by the spectrophotometer were as accurate as the pass/fail decisions generated by the 10 people. In other words, a spectrophotometer performs just as well as a human with good color vision when judging the color fabric shades.

Contents

Figures

Tables

Acknowledgments

LMI and Army Natick Soldier Research, Development, and Engineering Center (RDEC) would like to acknowledge and thank the working group whose participation was vital in the successful completion of this study. We would specifically like to mention the contribution of the following:

Julie Tsao, Defense Logistics Agency (DLA) Research and Development Office, who is the Program Manager for the DLA Customer Driven Uniform Manufacturing II Program (CDUM II), which sponsored the study.

Luisa DeMorais, Melanie King, and Rachel Matuszek, Army Natick Soldier RDEC Textile Evaluation Team, who

- developed the data collection protocol,
- conducted the government and industry site visits,
- provided samples for use in data collection, and
- offered their expertise in the collection and use of visual evaluation and instrumentation data.

Jamie Hieber, Defense Supply Center Columbus (DSCC) Product Testing Center, who provided expertise in the visual evaluation of military uniform shade, offered insight from the government perspective, and provided samples for use in data collection.

Larry Griffin, Defense Supply Center Philadelphia (DSCP), who provided expertise and guidance on DSCP quality assurance and insight from the government perspective.

Duane Cook and Ron Pollack, Army Product Manager—Clothing and Individual Equipment, who provided expertise from the quality perspective and Shade Integrated Process Team.

Bill Gerrow (Performance Textiles), Michael Mansch (Pennsylvania Apparel), and Frank Montie (Brookwood Companies) who were instrumental in setting up industry site visits and provided the industry perspective.

Alan Metelko, Jack Vandenberghe, and Nathaniel Wurst, LMI who coordinated the study under the DLA CDUM II Program, participated in site visit data collection, and conducted the statistical analysis.

Chapter 1
Introduction

The purpose of our study was to assess the viability of using spectrophotometers to evaluate shade conformance of military fabrics. Spectrophotometers are instruments that measure color in terms of light reflected or absorbed in different portions of the spectrum and in different amounts and provide a numerical value specific to that color. Many military fabric suppliers currently use spectrophotometers as part of their quality control process; however, current shade acceptance of military textile materials to a shade standard is strictly based on visual assessment methods. Visual evaluation by trained shade evaluators is widely considered superior to instrumentation because the human eye is more sensitive to shade variation, which is especially important in the evaluation of multicolored camouflage patterns.

The feedback from visual evaluations that is provided to fabric suppliers can, at times, be vague and not provide sufficient information to correct the shade in the proper direction. Without more useful feedback, the fabric suppliers have difficulty making the proper adjustments to their printing and dying processes to meet the given shade standard for a specific fabric in future production. Spectrophotometer measurements provide more quantitative feedback for shade evaluations, but concerns regarding the consistency of spectrophotometer measurements have raised questions about the use of these measurements in government shade evaluations and the feedback they provide fabric suppliers.

The study addressed two main questions:

- Do spectrophotometer measurements collected at different locations produce consistent color measurements?
- Given that the instruments do not have the same sensitivity to shade as the human eye, to what extent do human visual evaluations agree with spectrophotometer evaluation?

If it can be shown that evaluations by spectrophotometers agree with the results of a government visual evaluation at least as often as evaluations by other human shade evaluators, a plausible case will have been presented for incorporating the use of spectrophotometer measurements somewhere in the military shade evaluation and feedback process.

In addition to the primary study, an ancillary study was conducted to collect information about the physical attributes and instrumentation used in the various shade rooms of government and industry facilities that conduct visual and

instrumental shade evaluation. The purpose was to gain insight on the amount of variation in characteristics and instrumentation used throughout the industry.

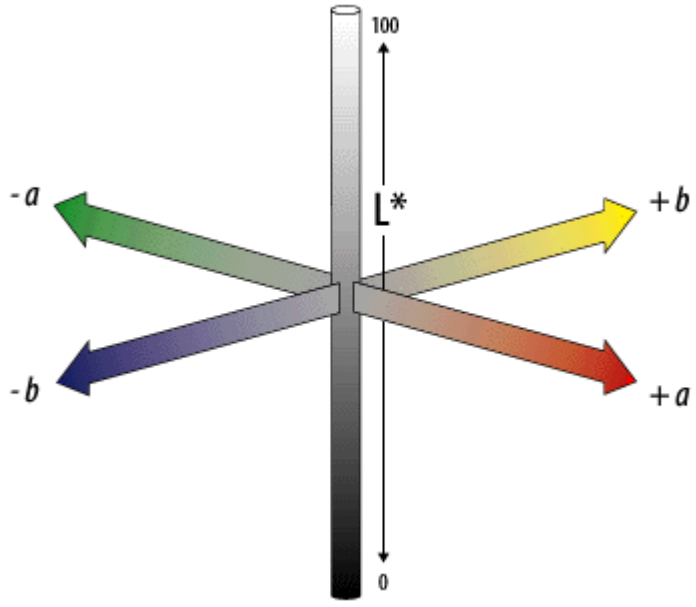
STANDARDS

The spectrophotometer measurements of interest in this study were based on standards commonly used by government and industry. These standards are based on the International Commission on Illumination (CIE, or *Commission Internationale d'Eclairage*), which specifies colors for instrumental shade evaluations using three axes (see Figure 1-1):

- L* = light-dark axis. Measured on a scale of 0 to 100, where 0 is at the darkest end of the scale, and 100 is at the lightest.
- a* = red-green axis. Scale with midpoint of 0 and no upper or lower limit. Negatives scores on this scale are more green, and positive scores are more red. The shades included in this report vary from −26.64 to 9.26 on this scale.
- b* = yellow-blue axis. Scale with midpoint of 0, with no upper or lower limit. Negative scores on this scale are more blue, and positive scores are more yellow. The shades included in this report vary from −20.56 to 18.43 on this scale.

Collectively, these dimensions are referred to as the CIELAB (CIE L*a*b*).

Figure 1-1. CIELAB Coordinate System



During the course of the study, two sets of measurements were collected based on the CIELAB standard. The first set comprised absolute measurements along each of the three color dimensions in the color space designated as L*, a* and b*. An absolute measure defines the color coordinates as a single point in the three-dimensional axis for the sample evaluated. The second set of measurements identified how far the shade of a sample differed from the shade standard for that sample's fabric type. Designated as ΔL*, Δa*, and Δb*, this set of measurements defines a vector in the three-dimensional axis that specifies the color difference between the sample and the designated shade standard evaluated. For example, measurements from the sample designated as the standard for the color shade Black 472 are

- 17.04 for L*,
- 0.45 for a*, and
- −1.17 for b*.

If another sample of Black 472 was measured at 16.62 for L*, 0.44 for a*, and −0.96 for b*, then its values would be

- −0.42 for ΔL*, or 16.62 − 17.04;
- −0.01 for Δa*, or 0.44 − 0.45; and
- 0.21 for Δb*, or −0.96 − (−1.17).

PURPOSE

The purpose of the study was a twofold assessment of whether there is a correlation between instrument shade measurements collected at government and industry sites, and whether there is a correlation between instrument pass/fail results and human visual inspection pass/fail results for a representative set of military fabrics. To make this assessment, three hypotheses were tested:

- *Hypothesis 1*: There is correlation among government and industry shade instrumentation in measuring the absolute value of a color.
- *Hypothesis 2*: There is correlation among government and industry shade instrumentation in measuring the color differences between samples.
- *Hypothesis 3*: There is correlation between instrumental measurement and human observer in making pass/fail acceptance decisions.

The first two hypotheses deal with whether instrumentation at different locations produce consistent shade measurements. This issue is addressed in this report by seeking the answers to the following two questions:

- To what extent do different locations produce consistent absolute values of L*, a*, and b* for different color shade samples?
- To what extent do different locations produce consistent color difference measurements between shade samples and a designated standard for a given fabric? That is, to what extent do ΔL^*, Δa^*, and Δb^* vary among locations?

The third hypothesis deals with the extent to which spectrophotometer evaluations agree with human evaluations. This issue is addressed by seeking the answers to the following two additional questions:

- To what extent do decisions made by human evaluators on a given set of fabric samples as compared to a shade standards (pass/fail decisions) agree with one another and with the official pass/fail determination?
- To what extent do spectrophotometer measurement decisions on a given set of fabric samples as compared to the shade standard agree with pass/fail decisions by human evaluators and with the official pass/fail determination?

STUDY APPROACH

We assembled a team to design a study and then monitor and provide feedback on its conduct. The study team comprised representatives from the following organizations:

- Defense Logistics Agency (DLA) Headquarters (HQ), Customer-Driven Uniform Manufacturing II (CDUM II) Program
- Defense Supply Center Philadelphia (DSCP) Clothing and Textiles (C&T) Quality
- Defense Supply Center Columbus (DSCC) Product Testing Center (PTC)
- Army Natick Soldier Research, Development, and Engineering Center (RDEC)
- Government/Industry Shade Integrated Product Team (IPT)
- Pennsylvania Apparel
- Brookwood Companies

- Performance Textiles
- LMI.

The study was conducted in three phases.

Phase 1 Overview

In Phase 1, CIELAB readings were taken from five sample sets of tiles and fabrics using on-site instrumentation at three textile printing and finishing plants, a color shade instrument manufacturer, and Army Natick Soldier RDEC.

Phase 1 had three objectives:

- Evaluate the ability to collect the instrumentation data at the five sites.
- Determine if the data collected are statistically sufficient for assessing whether spectrophotometers used in different locations produce consistent color measurements, modifying the data collection method if necessary.
- Explore the potential for additional data collection.

Phase 2 Overview

In Phase 2, data was collected at 19 facilities (the original 5 plus 14 additional sites) using the modified data collection method test plan derived during Phase 1 on tile and fabric set samples. The goal of Phase 2 was to determine whether different locations produce consistent color measurements.

Phase 3 Overview

Phase 3 determined the extent that pass/fail decisions based on spectrophotometer measurements agree with the official pass/fail determination and with human shade evaluators.

Ten sets of solid military fabric samples from 10 different fabrics were evaluated against their corresponding shade standard. Shade evaluators at Army Natick Soldier RDEC and DSCC PTC provided the official visual pass/fail determination for each set of samples when evaluated against the corresponding shade standard. A comparison was then made between the pass/fail determinations of human evaluators and the pass/fail determination generated by a spectrophotometer. The official pass/fail determination was used to compare the accuracy of the two sets of evaluations.

REPORT ORGANIZATION

The remainder of this report is organized as follows:

- Chapter 2 details the methodology, procedures, and findings of Phase 1.
- Chapter 3 details the methodology, data analysis, and findings of Phase 2.
- Chapter 4 details the methodology, data analysis, and findings of Phase 3.
- Appendix A reports the finding of an ancillary study conducted to determine if factors other than instrumentation were causing variability in readings of fabric shade parameters.
- Appendix B provides the data collection protocols for ceramic tiles, shade fabric samples, solid shade on different fabric samples, fabric pairs, and camouflage fabric samples.
- Appendix C defines abbreviations in the report.

Chapter 2
Phase 1

During Phase 1, CIELAB measurements were collected using the data collection protocol that was developed by textile technologists from Army Natick Soldier RDEC (provided in Appendix B). Army Natick Soldier RDEC used the American Association of Textile Chemists and Colorists (AATCC) Evaluation Procedure 6, *Instrumental Color Measurement*, Section 2, "Measurement of Color by Reflectance Methods," as a guideline for the color measurement of the samples.

Five sites—representing government, industry, and a spectrophotometer manufacturer—were selected as data collection targets. Instrumentation measurements were collected from these five sites, which are listed in Table 2-1.

Table 2-1. Phase 1 Data Collection Sites

Name	Location	Description
Army Natick Soldier RDEC	Natick, MA	Government shade evaluation lab
HunterLab	Reston, VA	Instrumentation vendor
Kenyon Industries	Kenyon, RI	Industry lab
Polartec LLC	Lawrence, MA	Industry lab
Duro Mills	Fall River, MA	Industry lab

On-site personnel followed protocol procedures to capture the spectrophotometer measurements of the samples provided. A textile technologist from Army Natick Soldier RDEC was present at each site to ensure the spectrophotometers were properly calibrated and samples were properly conditioned, and to observe the instrumentation data collection for each set of tile and fabric samples.

PHASE 1 SAMPLES

Spectrophotometer measurements were collected on the following sample sets at each site:

- A standardized set of color ceramic tiles
- A set of different shade fabric samples
- A solid shade set of different fabrics

- A set of fabric pairs
- A set of Army combat uniform (ACU) and Air Force airman battle uniform (ABU) camouflage samples.

Table 2-2 lists the material and source of each sample measured during Phase 1.

Table 2-2. Phase 1 Study Samples

Type	Material	Description	Source
Standardized color tiles	Ceramic	Deep Blue	Natick
	Ceramic	Deep Grey	Natick
	Ceramic	White	Natick
	Ceramic	Mid Grey	Natick
	Ceramic	Green	Natick
	Ceramic	Pale Grey	Natick
Single shade (Blue 450.4)—different fabric substrates	Polyester/wool elastique	Blue 450.4	DSCC PTC
	Polyester	Blue 450.4	DSCC PTC
	Polyester/wool tropical	Blue 450.4	DSCC PTC
	Polyester/wool serge	Blue 450.4	DSCC PTC
Fabric pairs	NA	Black 472	DSCC PTC
	NA	Forest Green 504	DSCC PTC
	NA	Purple 3905	DSCC PTC
	NA	CG 483	DSCC PTC
Army ACU camouflage samples	Nylon/cotton twill	Army ACU A	Natick
	Nylon/cotton twill	Army ACU B	Natick
	Nylon/cotton twill	Army ACU C	Natick
	Nylon/cotton twill	Army ACU D	Natick
	Nylon/cotton twill	Army ACU E	Natick
	Nylon/cotton twill	Army ACU F	Natick
Air Force ABU camouflage samples	Nylon/cotton twill	Air Force ABU A	DSCC PTC
	Nylon/cotton twill	Air Force ABU B	DSCC PTC
	Nylon/cotton twill	Air Force ABU C	DSCC PTC
	Nylon/cotton twill	Air Force ABU D	DSCC PTC
	Nylon/cotton twill	Air Force ABU E	DSCC PTC
	Nylon/cotton twill	Air Force ABU F	DSCC PTC

Note: Natick = Army Natick Soldier RDEC.

PHASE 1 FINDINGS

The data generated from Phase 1 showed that the data collection methodology required a redesign. Those findings were presented to the study working group on October 29, 2008. A set of recommendations was put forth for the redesigned methodology to be followed in Phase 2.

Observations from Phase 1 Data

- Only basic graphical analysis (line graphs of averages for each color fabric and tile combination by site) of the ACU and ABU samples was possible with the existing data.
- Spectrophotometer readings appeared to vary substantively from site to site.
- The observations of any given color (fabric or tile) combination were too few to determine if the variation in measurements was site-specific factors or random measurement error. More than 30 observations of a color-fabric or tile combination are desirable, and a minimum of 18 observations was determined to be necessary to be able to sort out these factors.
- The samples of color (fabric or tiles) were too many to test sufficiently for systematic variation in any given sample set.

Recommendations

- Increase the number of observations of a color (fabric or tile) sample to 18 per site.
- Reduce the number of color (fabric or tile) combinations to a core set of colors that are commonly used by the armed forces.
- Use graphical analysis, t-tests, analysis of variance (ANOVA), and regression testing by LMI statistical staff to examine Hypotheses 1 and 2.
- Consolidate the number of color (fabric or tile) combinations:
  - L*, a*, and b* testing—six tiles and three solid fabrics
  - ΔL*, Δa*, and Δb* testing—two sets of fabrics, one solid and one camouflage pattern.

Chapter 3
Phase 2

Based upon the findings and data collection during Phase 1, the data collection procedures and sample sets were modified for Phase 2. Although the procedures remained relatively unchanged, the number of readings and samples increased.

Eighteen (rather than the initial three) readings were taken for each tile and solid fabric sample during Phase 2. The number of solid color fabric samples was also changed. Due to data collection time constraints that would be encountered with the additional 15 measurements per sample, only three of the fabrics examined during Phase 1[1] were selected from which to collect instrument measurements:

- Coyote Green 498
- Desert Sand 503
- Army Green 491.

The data collection protocol was modified to collect measurements of 2 sample sets with 19 fabric samples each. The data collection procedure was modified with the following changes:

- One measurement of the "standard" and one measurement of each of the 18 fabric swatches within a sample set were taken.
- The fabric used in the sample sets were Black 472 and Air Force ABU Camouflage (Grey).
- The standard and each sample were measured to determine the values of L*, a*, b*, ΔL*, Δa*, and Δb*.

PHASE 2 SAMPLES

Table 3-1 lists the material and source of each sample measured during Phase 2.

Table 3-1. Phase 2 Study Samples

Sample #	Type	Description	Source
1	Ceramic tile	Deep Blue	Natick
2	Ceramic tile	Deep Grey	Natick
3	Ceramic tile	White	Natick
4	Ceramic tile	Mid Grey	Natick
5	Ceramic tile	Green	Natick
6	Ceramic tile	Pale Grey	Natick

[1] Colors such as Forest Green 504, Purple 3905, CG483, and Blue 450 were eliminated.

Table 3-1. Phase 2 Study Samples

Sample #	Type	Description	Source
7	Solid shade fabric	Coyote Green 498	DSCC PTC
8	Solid shade fabric	Army Green 491	DSCC PTC
9	Solid shade fabric	Desert Sand 503	DSCC PTC
10	Black 472 fabric	Black 472 “Standard”	DSCC PTC
11–28	Black 472 fabric	Black 472 samples 1–18	DSCC PTC
29	Air Force ABU Camouflage (Grey)	ABU “Standard”	DSCC PTC
30–47	Air Force ABU Camouflage (Grey)	ABU samples 1–18	DSCC PTC

Note: Natick = Army Natick Soldier RDEC.

PHASE 2 DATA COLLECTION

The data collected at the 19 tests sites are listed in Table 3-2. Those 19 sites were selected as representative of government, spectrophotometer makers, and the industry that produces uniform fabric for the U.S. military. With 19 sites, there was sufficient variation in site-relevant factors to study systematic inter-site variation and site-specific factors (e.g., maintenance of spectrophotometer and calibration settings).

Table 3-2. Phase 2 Data Collection Sites

Name	Location	Description
Army Natick Soldier RDEC	MA	Government shade evaluation lab
DSCC Product Testing Center	PA	Government shade evaluation lab
U.S. Coast Guard Shade Lab	MA	Government shade evaluation lab
U.S. Navy C&T Research Lab	MA	Government shade evaluation lab
Data Color	NJ	Instrumentation vendor
HunterLab	VA	Instrumentation vendor
X Rite	NC	Instrumentation vendor
Bondcote	VA	Industry lab
Brittany	MA	Industry lab
Burlington Finishing Plant	NC	Industry lab
Burlington Raeford	NC	Industry lab
Carlisle	SC	Industry lab
Crystal Springs	GA	Industry lab
Duro Mills	MA	Industry lab
Kenyon Industries	RI	Industry lab
Milliken	SC	Industry lab
Mount Vernon	GA	Industry lab
Polartec LLC	MA	Industry lab
W.L. Gore	MD	Industry lab

When presenting actual data or analyzed data specific to a site, a Greek letter was randomly assigned (Alpha through Zeta) in place of the actual site name. This was done to preserve the anonymity of the site providing the data.

ADDITIONAL COROLLARIES

The primary purpose of Phase 2 was to determine if CIELAB scores from different spectrophotometers at different sites were correlated. This study also looked into how CIELAB readings varied by the size of aperture (large or small) used in capturing spectrophotometer measurements. This led to the following corollaries to test in the analysis:

- Corollary 1.a—Spectrophotometers produced by different manufacturers will produce different values of L*, a*, and b*.
- Corollary 1.b—Spectrophotometers produced by different manufacturers will produce different values of ΔL*, Δa*, and Δb*.
- Corollary 2.a—Spectrophotometers with different lens aperture sizes will produce readings of L*, a*, and b* that correlate differently.
- Corollary 2.b—Spectrophotometers with different lens aperture sizes will produce values of ΔL*, Δa*, and Δb that correlate differently.

To again preserve the anonymity of the spectrophotometer manufacturers, each manufacturer is represented by a letter from the Hebrew alphabet: Aleph, Beth, Gimel, or Daleth.

ANALYTICAL METHODS

We analyzed the data collected from the 19 sites using Microsoft Excel 2003 and Statistical Package for the Social Sciences (SPSS) version 17. We employed a variety of statistical analysis methods to examine the two hypotheses and four corollaries. We then produced tables and graphs to illustrate the variations in mean and standard deviation for sample tiles and fabrics, sites, and spectrophotometers. We also generated correlation matrices in order to examine the strength and patterns of correlations for CIELAB L* a* and b* readings between test sites.

DATA ANALYSIS

Hypothesis 1

We examine Hypothesis 1 first:

> *Hypothesis 1*: There is correlation among government and industry shade evaluation instrumentation values of military fabric for the absolute value of color shades—L*, a*, and b*.

CENTRAL TENDENCY OF ABSOLUTE MEASURES (L*, a*, AND b*) OF SOLID COLOR TILES AND FABRICS

When we wish to summarize any distribution of numbers, the one characteristic of the distribution that we are often interested in is its central tendency. The central tendency of a distribution is its central or most characteristic value. The most widely used measures of central tendency are the mean (or average value), the median (or the middle value), and the mode (the value that occurs most often). For the purposes of this study, we used the mean as our measure of central tendency.

Table 3-3 shows the mean values of L*, a*, and b* for each type of material sample obtained from 18 of the 19 sites. One site is not included here because we did not have data from that site for all of the sample sets.

Table 3-3. L, a*, and b* by Type of Material Sampled*

Type of material sampled	L*	a*	b*
Deep Blue Tile	28.3981	8.4949	−19.3166
Green Tile	55.7443	−26.1381	13.3189
Pale Grey Tile	83.7194	−0.3753	0.5739
Mid Grey Tile	59.9678	−0.4884	−0.1982
Deep Grey Tile	36.1034	−0.1510	0.2656
White Tile	95.6810	−0.2501	2.6837
Army Green 491	24.2242	−3.6765	1.5410
Coyote 498	41.1680	5.4887	17.3420
Desert Sand 503	67.4196	2.2124	10.3663

Note: Each number is an average of 324 observations—18 observations at 18 of 19 sites.

Table 3-3 is useful because it shows how the scores vary for the different colors. For example, the deep blue tile has the largest reading on the blue end of the b* (yellow-blue) scale, exactly as we would expect. Similarly, the white tile has the largest reading on the L*(dark-light) scale with a score near the largest possible value of 100.

We also used the scores in Table 3-3 as benchmarks in assessing inter-site dispersion of color measurements. If the color measurements from an individual site vary significantly from the mean measurements from all sites, we assessed whether the measurements from that site are divergent.

DISPERSION OF ABSOLUTE MEASURES (L*, a*, AND b*) FOR SOLID COLOR TILES AND FABRICS

In addition to central tendency, the other characteristic of a distribution we are often interested in is dispersion. In addition to a distribution's most characteristic value, we also wish to know how tightly or sparsely the other values are spread around that characteristic value. This gives us a sense of how the other numbers in that distribution vary from that characteristic value. In the present study, the real issue for Phase 2 is not what the absolute measurements L*, a*, and b* are, but to what extent they vary from site to site and even within sites themselves. Our first step was to examine overall variation in the scores to see if some types of samples had more variation than other types of samples. There are three measures of dispersion that we use in this study: the range, the standard deviation, and a comparison between the subgroup mean and the overall mean for a given measurement. In this context, the subgroups are the individual test sites.

The range is simply the distance between the lowest score and the highest score. The range is useful because it is easy to understand and because it gives a sense of how much total variation exists in a sample. The disadvantage of the range is that outliers can give a distorted sense of how much variation exists in a sample. To complement the range, we use the standard deviation. The standard deviation is somewhat complicated,[2] but the idea is to measure how much, on average, each observation varies from the mean of the sample. The standard deviation is the most widely used indicator of dispersion among statisticians.

When comparing subgroup means to overall means, we conclude that there is dispersion among subgroups if the absolute value of the difference between these means is greater than some given value. For example, if the overall mean is 30.3 and we use 0.5 as our given value, we might conclude that a subgroup mean was divergent if it was either below 29.8 or above 30.8.

To show how much observations vary by sample type, we used the standard deviation. For variation within a site, we used the range because even a single aberrant reading from a given site is important and should be noted. For inter-site variation, we used the difference between the site mean and the mean for all sites.

[2] The formula for the sample standard deviation is the square root of the sum of the squared deviations from the mean divided by the number of observations minus one.

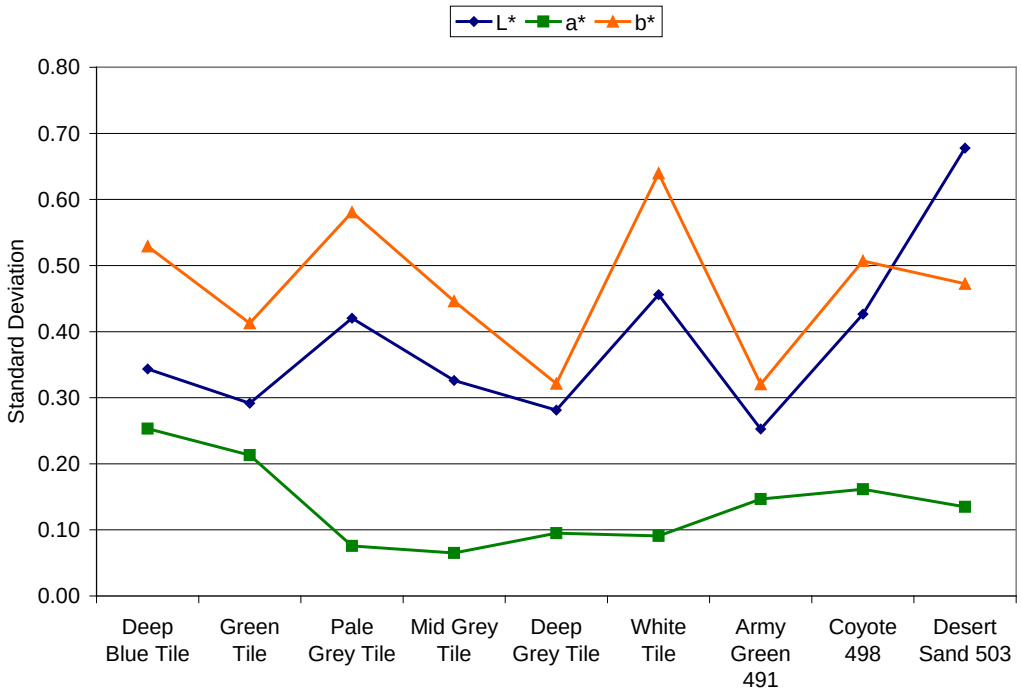
$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n-1}}$$

Figure 3-1 shows standard deviations for L*, a*, and b* by type of color sample.[3] Among the L* measurements Desert Sand 503 has the highest level of dispersion. Also, note that the three color samples with the highest measurements of L* on Table 3-3 also show the highest level of dispersion in Figure 3-1. It appears that colors with a higher absolute value of L* also tend to have more divergent L* measurements.

The a* measurements have relatively little dispersion overall. We found the most dispersion in a* measurements for the deep blue tile and the green tile. These tiles also have the highest and lowest a* values in Table 3-3.

Among the b* measurements the white and pale grey tiles had the highest levels of dispersion in their measurements. There is also little evidence of a relationship between the absolute values of b* in Table 3-3 and the dispersion of b* measurements shown in Figure 3-1, with the possible exception of the deep blue tile. The deep blue tile has the lowest b* measurement on Table 3-3, and the third-highest level of dispersion for b* in Figure 3-1.

Figure 3-1. Standard Deviations for L, a*, and b* by Type of Sample*



[3] L*, a*, and b* are shown on the same graph only for convenience. Since the scales are different, the standard deviations across scales are not comparable.

INTER-SITE DISPERSION OF ABSOLUTE MEASURES (L*, a*, AND B*) FOR SOLID COLOR TILES AND FABRICS

Table 3-4 gives summary results of cross-site dispersion for the six color tile samples and the three solid color fabric samples in this study for which values of L*, a* and b* are available. Keep in mind the actual site name has been replaced with a random letter of the Greek alphabet to maintain anonymity. Table 3-4 indicates which color samples at each site had readings that were more than 0.5 above or below the mean values of readings recorded from all sites.

Table 3-4. Color Samples with L, a*, and b* Readings at an Individual Site More Than 0.5 Above or Below Mean for All Sites*

Site	L*	a*	b*
Alpha	Desert Sand 503		
Beta			
Chi			Army Green 491
Delta			
Epsilon			
Iota			
Kappa			
Lambda			
Mu	Desert Sand 503		
Omega	Desert Sand 503 Deep Blue	Deep Blue Tile	Deep Blue Tile
Omicron	Coyote 498 Desert Sand 503		
Psi	Desert Sand 503		
Rho	All Samples		All Samples
Sigma			
Tau	Desert Sand 503 Deep Grey Tile Deep Blue Tile	Deep Blue Tile	Coyote 498 Deep Grey Tile Deep Blue Tile
Theta			
Upsilon	Desert Sand 503 Coyote 498	Green Tile	Coyote 498 Deep Blue Tile Green Tile
Xi			
Zeta			

Several things stand out in Table 3-4:

- L* and b* measurements from site Rho differ significantly from all other sites for all color samples in the study. No other site had divergent measurements for more than three colors; site Rho had divergent L* and b* measurements for all nine color samples.
- There is less evidence of dispersion overall for a* measurements than for L* and b* measurements.
- Some color samples showed evidence of more dispersion than others. Just as we saw in Figure 3-1, Desert Sand 503 in particular appears to have a great deal of dispersion in its L* readings. Coyote 498 and the deep blue tile also show more dispersion than the other color samples throughout all of the color dimensions. Overall, there appears to be less dispersion on readings of color tiles than fabrics.

WITHIN-SITE DISPERSION OF ABSOLUTE MEASURES (L*, a*, AND b*)

Table 3-5 gives summary results of within-site dispersion by site for the six color tile samples and the three solid color fabric samples in this study for which values of L*, a* and b* are available.

Table 3-5. Sites with L, a*, and b* Readings with a Range 0.3 or Larger*

Site	L*	a*	b*
Alpha			Desert Sand 503
Beta			
Chi			
Delta			Desert Sand 503
Epsilon		Army Green 491	
Iota			
Kappa			
Lambda			Desert Sand 503
Mu			
Omega			
Omicron			Desert Sand 503
Psi			
Rho			
Sigma		Green Tile	
Tau	Deep Blue Tile		
Theta			
Upsilon			
Xi			
Zeta			

The table indicates which site reported readings of a given sample type that had a range of 0.3 or larger. Recall that each site reported 18 different L*, a*, and b* readings for the 6 color tile samples and the 3 solid color fabric samples. The range is the distance between the largest and the smallest of the 18 readings. Overall, we find very little inter-site dispersion for these samples. The one major exception is b* readings of Desert Sand 503; four different sites reported sets of readings with a range of 0.3 or greater. This is somewhat surprising given that we did not observe an unusually large standard deviation for b* readings of Desert Sand 503 in Figure 1-1. This points to the possibility that the dispersion in the b* readings of Desert Sand 503 is caused by a few outliers. (We revisit this issue at the end of this chapter.)

SUMMARY OF CENTRAL TENDENCY AND DISPERSION OF ABSOLUTE MEASURES (L*, a*, AND b*)

Three main points should be taken from this section of our analysis:

- In general, spectrophotometers at the 19 sites generated similar absolute measures (L*, a*, and b*) for the color tile and fabric samples included in this study. The one exception was site Rho, which consistently produced L* and b* readings that diverged significantly from the readings produced at the other 18 sites. We therefore find that hypothesis 1 is, for the most part, confirmed.

- With one major exception, there was relatively little within-site dispersion of the absolute measures (L*, a* and b*). That exception was the Desert Sand 503 fabric. Four sites reported divergent sets of b* readings for that fabric. With this one exception, we have high confidence that the sites are generating consistent readings.

CORRELATIONS OF ABSOLUTE MEASURES (L*, a*, AND b*) FOR SOLID COLOR TILES AND FABRICS

Having examined central tendency and dispersion in the absolute measures (L*, a*, and b*), we next examined to what extent these measures are correlated with one another. The correlation coefficient "r" indicates to what extent two sets of observations are related to one another. This coefficient has a range of possible values from −1 to 1. The standard that we will use in this study to interpret the correlation coefficient is as follows:[4]

- −1.0 to −0.5, strong negative association
- −0.49 to −0.3, moderate negative association
- −0.29 to −0.1, small negative association

[4] Jessica Steele, *Choosing the Correct Statistical Test,* www.radford.edu/~jcsteele/Choosing the Correct Statistical Test, 2006.

- −0.09 to 0.09, no association
- 0.1 to 0.29, small positive association
- 0.3 to 0.49, weak positive association
- 0.5 to 1.0 strong positive association.

Figure 3-2, Figure 3-3, and Figure 3-4 show the correlations matrices for the mean site L*, a*, and b* tile measurements. All three of the matrices indicate an extremely high degree of correlation among the mean measurements from all sites. This means that even if there are small differences in reported measurements between the sites, these differences remain highly stable across color tile samples.

An illustration of the high degree of correlation across color tile samples can be seen in Figure 3-5 by examining the graph of the mean L* color tile measurements from sites Alpha and Rho. Recall from Table 3-4, which showed that L* measurements from site Rho were consistently different from measurements from other sites across color sample type. Despite the discrepancy, there is still a very close correspondence between the mean L* measurements from these sites.

Figure 3-6, Figure 3-7, and Figure 3-8 show the correlation matrices for the mean L*, a*, and b* site measurements for the color fabric samples Army Green 491, Coyote 498, and Desert Sand 503. Once again, we see an extremely high degree of correlation for all these measurements across sites. We can conclude this section of our analysis by stating that despite a limited amount of across-site and within-site dispersion in absolute L*, a*, and b* of six color tile samples and three fabric samples, we still find a high degree of correlation across sites for these measurements.

Hypothesis 1—Findings

We find there is a high degree of correlation among industry color shade evaluation instrumentation values of military fabric for the absolute value of color shades L*, a*, and b*, as shown in Figures 3-2–3-8.

Figure 3-2. Correlation Matrix for L Site Means of Color Tiles*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.999999	1																	
Chi	0.999998	0.999999	1																
Delta	0.999999	0.999998	0.999997	1															
Epsilon	0.999999	0.999998	0.999996	0.999999	1														
Iota	0.999997	0.999998	0.999997	0.999999	0.999996	1													
Kappa	0.999998	1	0.999998	0.999999	0.999998	0.999999	1												
Lambda	0.999982	0.999981	0.999978	0.999988	0.999981	0.99999	0.999984	1											
Mu	0.999999	1	0.999999	0.999999	0.999998	0.999999	1	0.999984	1										
Omega	0.999977	0.99998	0.999981	0.99997	0.999977	0.999967	0.999977	0.999923	0.999977	1									
Omicron	0.99999	0.999994	0.999996	0.999988	0.999985	0.999992	0.999992	0.999971	0.999993	0.999976	1								
Psi	0.999999	0.999998	0.999996	0.999999	1	0.999996	0.999998	0.99998	0.999998	0.999978	0.999985	1							
Rho	1	0.999999	0.999998	0.999998	0.999999	0.999996	0.999999	0.999979	0.999999	0.999982	0.99999	0.999999	1						
Sigma	0.999998	0.999999	0.999999	0.999999	0.999996	0.999999	0.999999	0.999986	1	0.999973	0.999995	0.999996	0.999998	1					
Tau	0.999984	0.999985	0.999983	0.99999	0.999983	0.999993	0.999987	0.999999	0.999987	0.99993	0.999978	0.999982	0.999981	0.99999	1				
Theta	0.999998	1	0.999999	0.999997	0.999996	0.999998	0.999999	0.999979	0.999999	0.999981	0.999996	0.999996	0.999998	0.999999	0.999983	1			
Upsilon	0.999992	0.999994	0.999995	0.999993	0.999987	0.999996	0.999993	0.999986	0.999994	0.99996	0.999996	0.999987	0.99999	0.999997	0.999991	0.999994	1		
Xi	0.999997	0.999997	0.999996	0.999999	0.999996	0.999999	0.999998	0.999993	0.999998	0.999962	0.999989	0.999996	0.999996	0.999999	0.999995	0.999996	0.999996	1	
Zeta	0.999999	1	0.999999	0.999998	0.999998	0.999999	1	0.999982	1	0.999979	0.999994	0.999998	0.999999	1	0.999986	1	0.999994	0.999997	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
171	0	0
Median Correlation		Mean Correlation
0.999996		0.999992

small weak strong association

Figure 3-3. Correlation Matrix for a Means of Color Tiles*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.999999	1																	
Chi	0.999978	0.999975	1																
Delta	0.999996	0.999992	0.999979	1															
Epsilon	0.99999	0.999989	0.999967	0.999994	1														
Iota	0.999993	0.999994	0.999986	0.999992	0.999993	1													
Kappa	0.999994	0.999997	0.999983	0.999986	0.999984	0.999997	1												
Lambda	0.999989	0.999991	0.999982	0.999976	0.999968	0.999988	0.999997	1											
Mu	0.99999	0.999992	0.99998	0.99999	0.999996	0.999999	0.999993	0.999981	1										
Omega	0.999788	0.999773	0.99967	0.999773	0.999738	0.999715	0.99974	0.99975	0.999705	1									
Omicron	0.999996	0.999997	0.999986	0.999988	0.999981	0.999994	0.999998	0.999997	0.999989	0.999757	1								
Psi	0.999993	0.999991	0.99995	0.999985	0.999977	0.999975	0.999982	0.999978	0.999972	0.999855	0.999984	1							
Rho	0.999809	0.999806	0.999756	0.999768	0.999719	0.999755	0.999802	0.999839	0.999729	0.999876	0.999819	0.99985	1						
Sigma	0.999992	0.999995	0.999975	0.99998	0.999976	0.999991	0.999998	0.999998	0.999986	0.999764	0.999997	0.999985	0.999832	1					
Tau	0.999942	0.999944	0.999969	0.99995	0.999962	0.999972	0.999955	0.999938	0.999975	0.999518	0.999948	0.999897	0.99958	0.999938	1				
Theta	0.999991	0.99999	0.999961	0.999977	0.999964	0.999976	0.999988	0.999991	0.999968	0.999833	0.999991	0.999994	0.999881	0.999993	0.999898	1			
Upsilon	0.999959	0.99996	0.999969	0.999969	0.99998	0.99998	0.999963	0.999942	0.999985	0.999581	0.999957	0.999922	0.999602	0.999946	0.999995	0.999915	1		
Xi	0.999989	0.999994	0.999979	0.999976	0.999972	0.99999	0.999998	0.999997	0.999985	0.999729	0.999997	0.999976	0.999816	0.999998	0.999946	0.999987	0.999952	1	
Zeta	0.999999	0.999998	0.999977	0.999994	0.999986	0.999991	0.999992	0.999987	0.999987	0.999783	0.999996	0.999991	0.999809	0.99999	0.999939	0.999989	0.999957	0.999989	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
171	0	0
Median Correlation		Mean Correlation
0.99998		0.999934

small weak strong association

Figure 3-4. Correlation Matrix for b Means of Color Tiles*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.999867	1																	
Chi	0.999881	0.999941	1																
Delta	0.999992	0.999862	0.99987	1															
Epsilon	0.999976	0.99991	0.999894	0.999991	1														
Iota	0.999905	0.999988	0.999931	0.999914	0.999956	1													
Kappa	0.999885	0.999993	0.999935	0.999894	0.999941	0.999999	1												
Lambda	0.999942	0.99997	0.999929	0.999952	0.999981	0.999994	0.999988	1											
Mu	0.999919	0.999988	0.999932	0.999923	0.99996	0.999999	0.999996	0.999995	1										
Omega	0.999913	0.99959	0.999653	0.999879	0.999808	0.999644	0.999606	0.999717	0.999674	1									
Omicron	0.999986	0.999933	0.999909	0.999975	0.999977	0.999954	0.99994	0.999974	0.999965	0.99985	1								
Psi	0.999981	0.999899	0.999916	0.999993	0.999995	0.999943	0.999928	0.999972	0.999948	0.999827	0.999973	1							
Rho	0.997594	0.998557	0.998239	0.99754	0.997766	0.998348	0.998424	0.998153	0.998324	0.99682	0.997937	0.997705	1						
Sigma	0.999974	0.999955	0.999931	0.999974	0.999988	0.999978	0.999968	0.999992	0.999984	0.999798	0.999994	0.999983	0.998019	1					
Tau	0.999969	0.999936	0.999879	0.999954	0.999961	0.999953	0.999939	0.999967	0.999965	0.999826	0.999995	0.999949	0.99803	0.999985	1				
Theta	0.999918	0.999973	0.999979	0.99991	0.999937	0.999974	0.999973	0.999974	0.999977	0.999698	0.999955	0.999944	0.998295	0.999971	0.99994	1			
Upsilon	0.999921	0.999705	0.999716	0.999869	0.999817	0.999722	0.999693	0.999769	0.999752	0.999959	0.999898	0.999826	0.997333	0.999845	0.999894	0.999761	1		
Xi	0.999985	0.999937	0.999923	0.999985	0.999992	0.999966	0.999953	0.999986	0.999973	0.999826	0.999995	0.99999	0.997905	0.999998	0.999983	0.999962	0.999859	1	
Zeta	0.999943	0.99997	0.999934	0.999951	0.99998	0.999993	0.999986	0.999999	0.999995	0.999723	0.999974	0.999972	0.998158	0.999992	0.999967	0.999979	0.999771	0.999986	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
171	0	0

Median Correlation	Mean Correlation
0.999942	0.999713

small weak strong association

Figure 3-5. L Means for Color Tile Samples from Sites Alpha and Rho*

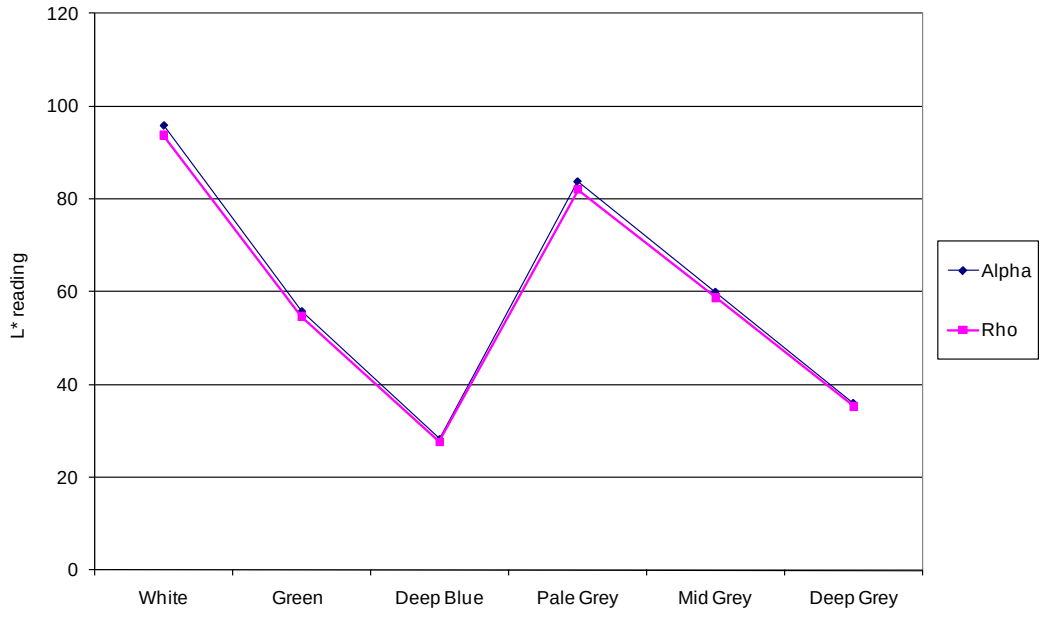


Figure 3-6. Correlation Matrix for L Means of Color Fabrics*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Zeta
Alpha	1																	
Beta	0.999999	1																
Chi	0.999998	0.999999	1															
Delta	0.999999	0.999998	0.999997	1														
Epsilon	0.999999	0.999998	0.999996	0.999999	1													
Iota	0.999997	0.999998	0.999997	0.999999	0.999996	1												
Kappa	0.999998	1	0.999998	0.999999	0.999998	0.999999	1											
Lambda	0.999982	0.999981	0.999978	0.999988	0.999981	0.99999	0.999984	1										
Mu	0.999999	1	0.999999	0.999999	0.999998	0.999999	1	0.999984	1									
Omega	0.999977	0.99998	0.999981	0.99997	0.999977	0.999967	0.999977	0.999923	0.999977	1								
Omicron	0.99999	0.999994	0.999996	0.999988	0.999985	0.999992	0.999992	0.999971	0.999993	0.999976	1							
Psi	0.999999	0.999998	0.999996	0.999999	1	0.999996	0.999998	0.99998	0.999998	0.999978	0.999985	1						
Rho	1	0.999999	0.999998	0.999998	0.999999	0.999996	0.999999	0.999979	0.999999	0.999982	0.99999	0.999999	1					
Sigma	0.999998	0.999999	0.999999	0.999999	0.999996	0.999999	0.999999	0.999986	1	0.999973	0.999995	0.999996	0.999998	1				
Tau	0.999984	0.999985	0.999983	0.99999	0.999983	0.999993	0.999987	0.999999	0.999987	0.99993	0.999978	0.999982	0.999981	0.99999	1			
Theta	0.999998	1	0.999999	0.999997	0.999996	0.999998	0.999999	0.999979	0.999999	0.999981	0.999996	0.999996	0.999998	0.999999	0.999983	1		
Upsilon	0.999992	0.999994	0.999995	0.999993	0.999987	0.999996	0.999993	0.999986	0.999994	0.99996	0.999996	0.999987	0.99999	0.999997	0.999991	0.999994	1	
Zeta	0.999997	0.999997	0.999996	0.999999	0.999996	0.999999	0.999998	0.999993	0.999998	0.999962	0.999989	0.999996	0.999996	0.999999	0.999995	0.999996	0.999996	1

Strong Associations (>.5)		Weak Associations (.3 to .5)		Small Assocations (.1 to .3)	
153		0		0	
Median Correlation			Mean Correlation		
0.999996			0.999991		

small weak strong association

Figure 3-7. Correlation Matrix for a Means of Color Fabrics*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Zeta
Alpha	1																	
Beta	0.999963	1																
Chi	0.99997	1	1															
Delta	0.999996	0.999936	0.999945	1														
Epsilon	0.999987	0.999905	0.999917	0.999997	1													
Iota	0.999982	0.999893	0.999906	0.999995	1	1												
Kappa	0.999989	0.999992	0.999995	0.999973	0.999952	0.999943	1											
Lambda	0.999992	0.999919	0.99993	0.999999	0.999999	0.999998	0.999962	1										
Mu	0.999999	0.999953	0.999961	0.999999	0.999992	0.999988	0.999983	0.999996	1									
Omega	0.9997	0.999451	0.99948	0.999762	0.999812	0.999828	0.999573	0.999791	0.999726	1								
Omicron	0.99981	0.999605	0.999629	0.999859	0.999897	0.999909	0.999708	0.999881	0.999831	0.999987	1							
Psi	1	0.99997	0.999977	0.999993	0.999981	0.999976	0.999993	0.999987	0.999998	0.999676	0.999791	1						
Rho	0.99994	0.999997	0.999995	0.999907	0.99987	0.999857	0.99998	0.999887	0.999927	0.999371	0.999537	0.99995	1					
Sigma	0.999922	0.999777	0.999795	0.999952	0.999973	0.999979	0.999852	0.999964	0.999935	0.999928	0.999976	0.99991	0.999725	1				
Tau	0.999919	0.999992	0.999987	0.999881	0.99984	0.999825	0.999968	0.999859	0.999904	0.999305	0.99948	0.99993	0.999998	0.999681	1			
Theta	0.99998	0.999997	0.999999	0.999959	0.999934	0.999924	0.999999	0.999946	0.999972	0.999524	0.999667	0.999986	0.999989	0.999823	0.999979	1		
Upsilon	0.999981	0.999997	0.999999	0.999961	0.999936	0.999927	0.999999	0.999948	0.999974	0.99953	0.999671	0.999986	0.999988	0.999826	0.999978	1	1	
Zeta	0.999917	0.999991	0.999987	0.999878	0.999837	0.999822	0.999966	0.999856	0.999902	0.999301	0.999476	0.999929	0.999998	0.999678	1	0.999979	0.999977	1

Strong Associations (>.5)		Weak Associations (.3 to .5)		Small Assocations (.1 to .3)	
153		0		0	
Median Correlation			Mean Correlation		
0.999948			0.999888		

small weak strong association

Figure 3-8. Correlation Matrix for b Means of Color Fabrics*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Zeta
Alpha	1																	
Beta	0.999572	1																
Chi	0.997396	0.999078	1															
Delta	0.999903	0.999883	0.998305	1														
Epsilon	0.999998	0.999625	0.997529	0.999927	1													
Iota	0.999754	0.999975	0.998751	0.999966	0.999793	1												
Kappa	0.999796	0.999959	0.99865	0.99998	0.999832	0.999998	1											
Lambda	0.999991	0.999689	0.997697	0.999953	0.999997	0.99984	0.999873	1										
Mu	0.999998	0.999518	0.997265	0.999876	0.999993	0.999712	0.999758	0.999981	1									
Omega	0.999386	0.997934	0.994256	0.998799	0.999319	0.998362	0.998473	0.999226	0.999447	1								
Omicron	0.999915	0.999869	0.998252	1	0.999937	0.999958	0.999974	0.999962	0.99989	0.998844	1							
Psi	0.999997	0.999636	0.997556	0.999932	1	0.999801	0.999839	0.999998	0.999992	0.999304	0.999942	1						
Rho	0.999244	0.997679	0.993837	0.998603	0.999169	0.998135	0.998253	0.999067	0.999312	0.999993	0.998651	0.999153	1					
Sigma	0.998861	0.999829	0.999701	0.99943	0.998948	0.999674	0.999621	0.999057	0.998774	0.996575	0.999398	0.998966	0.99625	1				
Tau	0.999832	0.998868	0.995906	0.999478	0.999796	0.999179	0.999257	0.999744	0.999863	0.99986	0.999508	0.999788	0.999789	0.997818	1			
Theta	0.999893	0.999893	0.998343	1	0.999919	0.999971	0.999984	0.999947	0.999866	0.998767	0.999999	0.999924	0.998569	0.999451	0.999457	1		
Upsilon	0.999786	0.998753	0.995689	0.999399	0.999745	0.99908	0.999163	0.999687	0.999821	0.999897	0.999431	0.999736	0.999834	0.997659	0.999997	0.999377	1	
Zeta	0.999884	0.999902	0.998378	0.999999	0.999911	0.999976	0.999988	0.99994	0.999855	0.998736	0.999998	0.999916	0.998536	0.999472	0.999437	1	0.999355	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
153	0	0

Median Correlation	Mean Correlation
0.999712	0.999267

small weak strong association

Hypothesis 2

It is now appropriate to examine Hypothesis 2:

> *Hypothesis 2*: There is correlation among government and industry color shade evaluation instrumentation measurements of the difference of color shade values—ΔL*, Δa*, and Δb*—of different military fabric samples in comparison to a "standard" sample.

In order to test Hypothesis 2, we conducted additional analysis using the Air Force Grey shade of the Air Force ABU camouflage pattern and Black 472 solid shade. The data for these sample sets were collected in a different manner. Two data sets (one for each fabric type) consisted of a single "standard" fabric sample and 18 other samples. Rather than taking 18 measurements of each fabric sample, we collected a single measurement for each of the 19 fabric samples contained in the two sample sets. In addition, since the grey shade of the ABU camouflage pattern did not allow for the use of a large aperture view in collecting its instrumentation measurements, all sites collected their ABU grey measurements using a small aperture opening. Replicating the statistical analyses performed for the nine samples above could help us understand what a "normal" set of data might look like and the implications that would have for Hypothesis 1.

Mean Values of Absolute Measures (L*, a*, and b*) and Difference Measures (ΔL*, Δa*, and Δb*) for Air Force Grey and Black 472

Table 3-6 gives the means and standard deviations of the absolute measures (L*, a*, and b*) and the difference measures (ΔL*, Δa*, and Δb) for the samples of Air Force Grey and Black 472 used in this study. Notice the standard deviations for the measurements of the Black 472 samples are consistently smaller than

the standard deviations for the measurements of the Air Force Grey samples. This indicates there is more variation among the Air Force Grey samples than there is among the Black 472 samples. For this reason alone, differences among the Air Force Grey samples should be easier to detect than differences among the Black 472 samples. The Δa* scores for the Black 472 samples appear to be particularly problematic. Table 3-6 indicates that the average difference between the Δa* designated as the standard and all of the other Black 472 samples is 0.001 with a standard deviation of only 0.086. With differences this small, we may run into problems finding correlations between Δa* scores among the different sites. Correlation essentially means that one set of scores tends to change in direct proportion to the change in a second set of scores. If a sample does not exhibit a sufficient amount of variation, then it becomes more difficult to demonstrate correlation in that sample.

Table 3-6. L, a*, b*, ΔL*, Δa*, and Δb* Means and Standard Deviations for Air Force Grey and Black 472*

Material sampled and statistic	L*	a*	b*	ΔL*	Δa*	Δb*
Air Force Grey Mean	50.27	0.15	5.75	0.234	0.234	0.694
Air Force Grey Standard Deviation	1.45	0.44	0.75	1.414	0.455	0.599
Black 472 Mean	16.36	0.36	−1.51	−0.5273	0.001	0.119
Black 472 Standard Deviation	0.52	0.12	0.51	0.484	0.086	0.163

Table 3-7 displays mean ΔL*, Δa*, and Δb* scores for the Air Force Grey and Black 472 samples by site. Means shaded in yellow were more than 0.5 above or below the mean difference scores for all sites. Means shaded in red were more than 1 above or below the mean difference scores for all sites. Mean site ΔL* scores for Air Force Grey samples were more likely than other measure to diverge from the mean scores of other sites. This is consistent with Table 3-6, where the ΔL* scores for Air Force Grey had the highest standard deviation. Two sites also reported divergent Δb* scores for Air Force Grey. This is again consistent with Table 3-6, where Δb* scores for Air Force Grey had the second largest standard deviation.

The ΔL* and Δb* scores from two sites in particular, Omega and Omicron, were much more divergent from the mean difference scores reported by other sites. Recall from Table 3-4 that these two sites had some divergent values for some of the absolute measures (L*, a*, and b*) of the solid color samples as well. Surprisingly, ΔL*, Δa*, and Δb* scores from site Rho differed little from the ΔL*, Δa*, and Δb* reported from other sites, despite the fact that absolute measures L*, a*, and b* for solid color samples from site Rho had shown a large degree of divergence.

Table 3-7. Difference Scores for Air Force Grey and Black 472

	Air Force Grey			Black 472		
Site	ΔL*	Δa*	Δb*	ΔL*	Δa*	Δb*
Alpha	0.03	0.29	0.56	−0.47	−0.02	0.14
Beta	−0.39	0.23	0.45	−0.66	0.04	0.06
Chi	−0.46	0.08	0.45	−0.77	0.05	0.03
Delta	−0.39	−0.03	0.54	−0.37	0.02	0.22
Epsilon	0.04	−0.17	0.77	−0.53	0.03	0.20
Iota	0.06	0.31	0.63	−0.64	0.04	0.06
Kappa	0.37	0.26	0.63	−0.28	−0.05	0.13
Lambda	1.56	0.13	0.86	−0.62	0.03	0.02
Mu	0.03	0.21	0.60	−0.82	0.00	0.16
Omega	1.47	0.11	1.44	−0.02	0.02	0.08
Omicron	1.90	0.72	1.45	−0.36	−0.03	0.14
Psi	0.01	0.26	0.52	−0.75	0.06	0.13
Rho	0.00	0.28	0.61	−0.59	−0.01	0.12
Sigma	0.07	0.23	0.56	−0.55	−0.08	0.18
Tau	−0.05	0.23	0.56	−0.52	−0.01	0.17
Theta	0.04	0.28	0.61	−0.66	0.01	0.12
Upsilon	0.02	0.39	0.69	−0.20	−0.07	0.09
Xi	−0.03	0.32	0.58	−0.83	−0.01	0.10
Zeta	0.16	0.33	0.69	−0.52	−0.03	0.12
Mean	0.23	0.23	0.69	−0.53	0.00	0.12

Note: Each number is an average of 18 observations per site, except Upsilon, which has 17 observations. Values shaded in red are more than 1.0 above or below the average of other sites. Values shaded in yellow are more than 0.5 above or below the average of other sites.

Correlations of Difference Measures (ΔL*, Δa*, and Δb*) for Air Force Camouflage Grey and Solid Black 472 Samples

Figure 3-9, Figure 3-10, and Figure 3-11 display the correlation matrices for the difference measures ΔL*, Δa*, and Δb* of the Air Force Grey samples. Just as in the correlation matrices for the solid color tiles and samples, the measurements from the various sites of the Air Force Grey samples tended to be highly correlated with one another.

While the correlations represented in Figure 3-9, Figure 3-10, and Figure 3-11 are nearly all strong, they are not nearly as strong as the correlations for the measurements of the solid color tiles and fabrics shown earlier. This is due to the differing research design used to gather shade measurement differences versus absolute shade measurements.[5]

One thing that we do not notice in Figure 3-9, Figure 3-10, or Figure 3-11 are low correlations between difference measures ΔL*, Δa*, and Δb* reported from sites Omega and Omicron and the difference measures reported from the other sites. Even though ΔL* and Δb* scores of Air Force Grey samples from sites Omega and Omicron diverged from the scores of the rest of the sites, changes in the scores from sites Omicron and Omega still tended to track closely with the changes reported by other sites.

Figure 3-9. Correlation Matrix of ΔL Scores for Air Force Grey*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.904845	1																	
Chi	0.930347	0.970237	1																
Delta	0.809341	0.86528	0.895085	1															
Epsilon	0.874522	0.926784	0.94908	0.864829	1														
Iota	0.94853	0.980559	0.986712	0.891009	0.949238	1													
Kappa	0.925608	0.964786	0.984982	0.881373	0.958767	0.985056	1												
Lambda	0.910178	0.938153	0.919723	0.814507	0.864608	0.948363	0.929791	1											
Mu	0.936141	0.968225	0.985779	0.891616	0.952847	0.991397	0.976652	0.92462	1										
Omega	0.903036	0.881361	0.908661	0.783704	0.933096	0.919886	0.913902	0.852604	0.933604	1									
Omicron	0.814587	0.800477	0.794836	0.803541	0.659313	0.811357	0.742302	0.783866	0.785927	0.628226	1								
Psi	0.92597	0.977973	0.976591	0.882115	0.928097	0.987893	0.978017	0.90887	0.980678	0.893769	0.794177	1							
Rho	0.902986	0.977154	0.970344	0.865813	0.933689	0.969056	0.960901	0.907261	0.95195	0.85442	0.794564	0.964166	1						
Sigma	0.942314	0.982924	0.98985	0.893897	0.952033	0.998863	0.986864	0.938726	0.99245	0.921112	0.805261	0.990443	0.969994	1					
Tau	0.942999	0.975681	0.982342	0.878231	0.9347	0.995799	0.979294	0.945608	0.986817	0.901738	0.81826	0.98479	0.964221	0.993893	1				
Theta	0.940987	0.979517	0.980566	0.875989	0.943444	0.996795	0.982212	0.933001	0.987398	0.913793	0.808039	0.993464	0.967013	0.9965	0.994547	1			
Upsilon	0.957855	0.958415	0.969074	0.853425	0.94106	0.987187	0.970394	0.938539	0.977115	0.928565	0.781595	0.971784	0.949771	0.985696	0.977269	0.984144	1		
Xi	0.954296	0.972152	0.976425	0.887892	0.933069	0.996786	0.977921	0.94374	0.982906	0.909172	0.828627	0.988187	0.959336	0.994473	0.993439	0.996455	0.987578	1	
Zeta	0.944695	0.971731	0.986361	0.895053	0.945602	0.996949	0.985358	0.936911	0.992668	0.921387	0.806185	0.98868	0.961479	0.995701	0.995011	0.995708	0.978379	0.993937	1

Strong Associations (>.5)		Weak Associations (.3 to .5)		Small Assocations (.1 to .3)	
171		0		0	
Median Correlation		Mean Correlation			
0.968225		0.939517			

small weak strong association

[5] Absolute shade measurement data collection of the solid color tiles and fabrics captured 18 measurements each of a single sample from each solid color type. Because we took so many measurements of each sample type, any within-site variation of measurements was most likely going to be averaged out. Shade difference measurement data collection, on the other hand, used 1 standard sample and 18 other samples for each of the sample types that were only measured once. This means divergent individual measurements would not be averaged out. Also, since each site was measuring a total of 19 different samples, just 1 divergent measurement out of 19 could have a significant effect on the correlation of that site's measurements with those of other sites. Given this increased potential for error in the difference measures, we should not expect the near-perfect correlations of around 0.999 that we saw in shade absolute measurements of the solid-color samples.

Figure 3-10. Correlation Matrix of Δa Scores for Air Force Grey*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.959803	1																	
Chi	0.960632	0.979317	1																
Delta	0.555623	0.568305	0.578983	1															
Epsilon	0.652768	0.686115	0.679542	0.590692	1														
Iota	0.967569	0.988949	0.981736	0.569213	0.688895	1													
Kappa	0.969275	0.970604	0.983175	0.639193	0.680909	0.973942	1												
Lambda	0.940333	0.966382	0.970043	0.60173	0.678461	0.975363	0.947498	1											
Mu	0.971493	0.992487	0.984361	0.586227	0.695021	0.997989	0.979776	0.97106	1										
Omega	0.561058	0.621518	0.572576	0.242716	0.563156	0.587395	0.507411	0.574318	0.579526	1									
Omicron	0.784583	0.784267	0.724907	0.223667	0.651672	0.801569	0.738313	0.709544	0.792354	0.605394	1								
Psi	0.962048	0.982634	0.982158	0.573693	0.7092	0.992132	0.969608	0.981762	0.988829	0.61207	0.776892	1							
Rho	0.95272	0.984867	0.982048	0.648114	0.724782	0.9864	0.983385	0.972156	0.988171	0.591586	0.747545	0.987675	1						
Sigma	0.968674	0.988437	0.971873	0.531627	0.683832	0.994989	0.971238	0.960635	0.995018	0.579777	0.832386	0.982403	0.98023	1					
Tau	0.866061	0.882335	0.879348	0.425086	0.557916	0.900345	0.879112	0.869976	0.900201	0.466047	0.778897	0.880982	0.873972	0.911177	1				
Theta	0.972254	0.985823	0.977859	0.59156	0.686954	0.995836	0.975608	0.966416	0.997842	0.562554	0.800604	0.983323	0.981471	0.992988	0.910204	1			
Upsilon	0.959841	0.966667	0.971923	0.611555	0.668499	0.973057	0.966583	0.94842	0.974274	0.58121	0.731826	0.963857	0.968435	0.963822	0.830732	0.970247	1		
Xi	0.963737	0.986631	0.976872	0.571307	0.684613	0.997663	0.973234	0.969625	0.995376	0.585976	0.803333	0.988813	0.987227	0.994922	0.891844	0.993246	0.975667	1	
Zeta	0.974399	0.985614	0.982372	0.587141	0.679489	0.996693	0.981643	0.967245	0.99646	0.57629	0.798835	0.986407	0.987133	0.993746	0.900846	0.996471	0.979452	0.996759	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
167	2	2
Median Correlation		Mean Correlation
0.960632		0.838751

small weak strong association

Figure 3-11. Correlation Matrix of Δb Scores for Air Force Grey*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.959281	1																	
Chi	0.94885	0.971595	1																
Delta	0.820653	0.828152	0.777369	1															
Epsilon	0.546472	0.588836	0.617929	0.686504	1														
Iota	0.981792	0.982768	0.980747	0.8502	0.615882	1													
Kappa	0.945584	0.962106	0.960729	0.818075	0.60114	0.970941	1												
Lambda	0.926757	0.950305	0.919215	0.825488	0.542379	0.938878	0.96415	1											
Mu	0.985753	0.981803	0.974255	0.843735	0.599455	0.996938	0.971771	0.942275	1										
Omega	0.845773	0.782867	0.777172	0.726703	0.668834	0.830735	0.763856	0.744428	0.84363	1									
Omicron	0.82582	0.781751	0.755814	0.807006	0.366158	0.8155	0.727518	0.760827	0.806576	0.636051	1								
Psi	0.975025	0.979562	0.974672	0.830222	0.587168	0.98607	0.962372	0.937434	0.98347	0.81535	0.78262	1							
Rho	0.969788	0.973841	0.948818	0.864221	0.593802	0.984356	0.965722	0.947196	0.984416	0.825031	0.800143	0.972671	1						
Sigma	0.983797	0.981021	0.972688	0.809638	0.59111	0.992703	0.965725	0.92803	0.995452	0.84012	0.794464	0.981753	0.97902	1					
Tau	0.976073	0.981914	0.984536	0.839751	0.611532	0.994534	0.961713	0.930902	0.992684	0.829132	0.813603	0.990947	0.973741	0.989431	1				
Theta	0.978524	0.972587	0.979625	0.844641	0.606067	0.992695	0.964221	0.940543	0.993748	0.837679	0.817175	0.978133	0.970755	0.983309	0.990127	1			
Upsilon	0.980519	0.985849	0.965077	0.849531	0.573791	0.990318	0.964999	0.944768	0.993298	0.811444	0.811173	0.984313	0.977413	0.990401	0.989202	0.982017	1		
Xi	0.976674	0.978838	0.983412	0.845261	0.619249	0.997391	0.970127	0.934763	0.993473	0.820741	0.802512	0.980938	0.983405	0.987009	0.990857	0.992542	0.982867	1	
Zeta	0.972559	0.977471	0.985713	0.846522	0.609544	0.994943	0.959567	0.930943	0.989007	0.808658	0.82989	0.983965	0.974235	0.982016	0.995425	0.991335	0.981431	0.994753	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
170	1	0
Median Correlation		Mean Correlation
0.960729		0.886589

small weak strong association

Figure 3-12 shows the correlation matrix for the ΔL* scores reported by the sites for the samples of the solid Black 472 fabrics. All of these correlations are strong with the exception of the correlation between ΔL* scores from site Omicron and ΔL* scores from the other sites. This is surprising since the mean ΔL* score from site Omicron shown in Table 3-7 was not significantly different from the mean ΔL* score reported from the other sites.

Figure 3-12. Correlation Matrix for ΔL Scores of Black 472*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.959224	1																	
Chi	0.789031	0.792858	1																
Delta	0.958025	0.955999	0.829243	1															
Epsilon	0.891062	0.858462	0.712071	0.859851	1														
Iota	0.845244	0.870244	0.937983	0.863047	0.765754	1													
Kappa	0.899429	0.895097	0.861339	0.896206	0.767838	0.908471	1												
Lambda	0.898153	0.916095	0.901531	0.91996	0.815438	0.957915	0.897542	1											
Mu	0.777911	0.809198	0.931922	0.796079	0.707889	0.957011	0.822698	0.914117	1										
Omega	0.732083	0.727149	0.643652	0.758024	0.642094	0.65304	0.817443	0.737658	0.517571	1									
Omicron	0.259639	0.258651	0.464355	0.358238	-0.01049	0.46722	0.3893	0.368114	0.485331	0.1205	1								
Psi	0.902939	0.900164	0.927226	0.894409	0.790445	0.969523	0.939953	0.965677	0.941326	0.710178	0.404199	1							
Rho	0.964641	0.965196	0.819571	0.95708	0.831795	0.897915	0.913032	0.92961	0.802839	0.756136	0.332921	0.925476	1						
Sigma	0.836397	0.869456	0.925975	0.830841	0.772872	0.966225	0.881291	0.959416	0.959951	0.642679	0.356229	0.968747	0.86773	1					
Tau	0.750946	0.791981	0.921624	0.831851	0.706259	0.914814	0.805877	0.918018	0.895746	0.658675	0.415844	0.897046	0.813547	0.896437	1				
Theta	0.824027	0.852736	0.943798	0.862388	0.738905	0.985073	0.904543	0.960749	0.94495	0.688207	0.462002	0.962679	0.885666	0.967151	0.91011	1			
Upsilon	0.985863	0.963891	0.785272	0.962942	0.883212	0.844768	0.894083	0.893189	0.770291	0.718477	0.3025	0.882515	0.961009	0.824919	0.750723	0.818883	1		
Xi	0.887434	0.874108	0.936234	0.881593	0.80028	0.973609	0.928655	0.953853	0.912153	0.718533	0.42809	0.977527	0.923447	0.950044	0.908856	0.95668	0.874573	1	
Zeta	0.934662	0.939701	0.876868	0.952433	0.822926	0.933269	0.943092	0.97262	0.884384	0.773088	0.407531	0.962824	0.945896	0.919668	0.859188	0.933071	0.933411	0.941479	1

Strong Associations (>.5)		Weak Associations (.3 to .5)		Small Assocations (.1 to .3)	
153		14		3	
Median Correlation			Mean Correlation		
0.895097			0.831739		

small weak strong association

The reason for the low correlations for site Omicron can be seen in Table 3-8. There was one aberrant measurement from site Omicron that had a disproportionate effect on the correlations between the L* scores from site Omicron and the Black 472 L* scores from the other sites. This measurement of 21.18 is circled in Table 3-8. If this single measurement is replaced by the average measurement of that sample from the other sites, the recalculated correlations between the Black 472 L* scores from Omicron and those from the other sites all become strong.

Table 3-8. L Measurements of Black 472 Samples*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Standard	16.84	17.01	17.24	16.38	16.67	17	16.51	18.11	17.29	16.73	17.04	17.01	16.32	17.15	16.73	17.02	16.55	17.77	16.76
Sample 1	16.25	16.36	16.64	15.93	15.98	16.39	16.13	17.57	16.67	16.67	16.37	16.35	15.63	16.8	16.5	16.45	16.12	16.96	16.19
Sample 2	16.49	16.45	16.35	16.03	16.28	16.51	16.40	17.65	16.61	16.64	16.55	16.44	15.82	16.76	16.37	16.45	16.39	17.08	16.39
Sample 3	15.94	16.02	16.08	15.59	15.66	16.06	15.74	17.18	16.15	16.36	15.94	15.93	15.54	16.24	16.02	16.07	15.84	16.67	15.83
Sample 4	16.28	16.31	16.25	15.93	16.09	16.24	16.23	17.34	16.19	16.97	16.36	16.07	15.74	16.49	15.91	16.32	16.32	16.81	16.07
Sample 5	16.43	16.36	16.50	15.98	15.9	16.33	16.26	17.61	16.36	16.78	16.55	16.29	15.81	16.68	16.21	16.36	16.42	16.99	16.26
Sample 6	16.55	16.46	16.78	16.24	16.22	16.61	16.50	17.67	16.64	16.94	16.77	16.49	15.97	16.75	16.35	16.67	16.66	17.14	16.45
Sample 7	16.05	16.10	15.90	15.50	15.78	15.98	15.93	16.99	16.02	16.12	16.03	15.9	15.45	16.21	15.36	15.92	16.01	16.63	15.87
Sample 8	16.04	16.08	16.22	15.74	15.69	15.95	15.89	17.13	16.19	16.23	16.13	15.94	15.40	16.31	15.79	16.02	16.03	16.61	16
Sample 9	15.95	15.84	15.89	15.56	15.72	15.79	15.75	16.97	15.95	16.32	15.85	15.74	15.16	15.99	15.47	15.82	15.88	16.39	15.79
Sample 10	16.43	16.26	16.45	16.08	16.2	16.28	16.36	17.48	16.38	17.21	16.39	16.34	15.78	16.55	16.23	16.34	16.38	16.98	16.29
Sample 11	16.54	16.47	16.42	16.23	16.33	16.27	16.28	17.47	16.39	16.82	16.42	16.22	15.79	16.55	16.18	16.33	16.40	16.92	16.25
Sample 12	16.88	16.79	17.00	16.56	16.68	16.79	16.80	18	16.75	18.24	16.85	16.66	16.19	16.98	16.89	16.76	16.88	17.43	16.76
Sample 13	16.06	16.22	16.32	15.90	16.02	16.16	16.03	17.19	16.23	16.22	16.13	15.97	15.50	16.31	16.23	16.10	16.15	16.7	15.94
Sample 14	17.61	17.33	17.09	17.14	17.34	17.01	16.82	18.32	17.14	17.34	17.07	16.94	16.72	17.25	16.94	16.93	17.66	17.57	17.04
Sample 15	16.28	16.48	16.51	16.09	16.2	16.52	16.25	17.76	16.71	16.80	16.66	16.33	15.75	16.81	16.36	16.60	16.34	16.93	16.44
Sample 16	16.13	16.01	16.61	15.71	16.35	16.42	16.00	17.45	16.56	16.08	16.38	16.2	15.48	16.67	16.29	16.40	16.07	17.02	16.05
Sample 17	16.40	16.37	16.78	16.22	15.75	16.63	16.40	17.65	16.78	16.55	21.18	16.43	15.86	16.73	16.54	16.62	16.46	17.15	16.44
Sample 18	16.39	16.36	16.67	15.83	16.06	16.48	16.39	17.43	16.72	16.54	16.62	16.38	15.68	16.77	16.15	16.45	16.34	16.99	16.2

Figure 3-13 shows the correlation matrix for the Δa* scores reported by the sites for the samples of the Black 472 fabrics. Unlike all of the other correlation matrices that we have seen up to this point, this correlation matrix shows a large number of correlations that are weak, small, or even nonexistent. Furthermore, even the strong correlations in Figure 3-13 are not generally as strong as the correlations we saw in Figure 3-12 for the ΔL* scores of the black 472 samples. The mean and median correlations for the Δa* scores are only 0.553 and 0.518, compared to 0.895 and 0.832 for the ΔL* scores. A likely cause for the correlations could be the relative

lack of variation of the Δa* scores of the Black 472 fabric samples. Recall from Table 3-6 that the Δa* varied within a very narrow range around a mean of 0.001. It appears the range of variation of the Δa* scores is so narrow as to be close to the limits of what the spectrophotometers can detect. Confirmation of this hypothesis can be seen in Table 3-9. The sites in Figure 3-13 that have the lowest correlations also tend to have the highest standard deviations in Table 3-9. It appears that, at least for sites Epsilon and Lambda, it is difficult to distinguish actual variation in the measurements from measurement error.

Figure 3-13. Correlation Matrix for Δa Scores for Black 472*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.475482	1																	
Chi	0.574446	0.537655	1																
Delta	0.635369	0.66778	0.375979	1															
Epsilon	0.223177	0.082317	0.215766	0.115875	1														
Iota	0.562219	0.541512	0.654358	0.544343	0.045791	1													
Kappa	0.490127	0.465478	0.377588	0.524361	0.081356	0.679359	1												
Lambda	0.28694	0.236629	0.172083	0.313216	0.204306	0.141484	0.027744	1											
Mu	0.607217	0.738338	0.670364	0.763561	0.183161	0.717176	0.773994	0.38479	1										
Omega	0.219015	0.42474	0.410873	0.422086	-0.02451	0.618945	0.284093	0.471242	0.574515	1									
Omicron	0.345228	0.234464	0.444423	0.349192	-0.03876	0.33626	0.572569	-0.00883	0.597611	0.152387	1								
Psi	0.462553	0.663437	0.664709	0.60926	-0.11415	0.64587	0.505741	0.30752	0.741579	0.523241	0.486771	1							
Rho	0.60996	0.73495	0.631987	0.6307	0.079858	0.652847	0.574449	0.368493	0.829716	0.567211	0.557561	0.644967	1						
Sigma	0.323162	0.19186	-0.0327	0.494066	-0.00799	0.494448	0.417886	0.008242	0.255349	0.400439	-0.15123	0.258155	0.146715	1					
Tau	0.544065	0.507734	0.331121	0.750599	0.255373	0.752502	0.711004	0.275734	0.695412	0.423896	0.276277	0.594159	0.464661	0.590257	1				
Theta	0.655515	0.610307	0.546462	0.683759	-0.02389	0.719207	0.62946	0.263101	0.761223	0.443806	0.215759	0.752206	0.595665	0.42486	0.717819	1			
Upsilon	0.467911	0.512597	0.004088	0.715525	0.010359	0.523338	0.666149	0.311456	0.563599	0.343931	0.161777	0.477902	0.408621	0.618333	0.863252	0.615719	1		
Xi	0.437132	0.552624	0.30642	0.654655	0.094195	0.63073	0.534112	0.341272	0.658713	0.626898	0.065549	0.575136	0.361369	0.539192	0.759526	0.756051	0.721477	1	
Zeta	0.551234	0.695677	0.681293	0.745329	0.025797	0.759709	0.678278	0.357298	0.894921	0.693002	0.562127	0.870508	0.792089	0.402563	0.708815	0.757673	0.553669	0.661899	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
89	40	23
Median Correlation	Mean Correlation	
0.552624	0.51837	

small weak strong association

Table 3-9. Standard Deviations of Δa Scores of Black 472 Samples by Site*

Site	Standard deviation
Alpha	0.075
Beta	0.047
Chi	0.057
Delta	0.063
Epsilon	0.139
Iota	0.040
Kappa	0.042
Lambda	0.120
Mu	0.046
Omega	0.147

Site	Standard deviation
Omicron	0.067
Psi	0.056
Rho	0.045
Sigma	0.093
Tau	0.048
Theta	0.075
Upsilon	0.066
Xi	0.051
Zeta	0.059

Figure 3-14 show the correlation matrix for the Δb* scores reported by the sites for the samples of the Black 472 fabrics. These correlations are quite high with the exception of the correlations between the Δb* scores from site Omega and the scores reported from the other sites. The low correlations for site Omega are

somewhat puzzling. There is no indication from Table 3-7 of anything out of the ordinary about the Δb* scores of the Black 472 fabric samples from site Omega. One possibility is within-site variation of the Δb* scores in site Omega.

Figure 3-14. Correlation Matrix for Δb Scores for Black 472*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.933015	1																	
Chi	0.86476	0.887908	1																
Delta	0.943366	0.969791	0.872336	1															
Epsilon	0.738525	0.714502	0.795976	0.679506	1														
Iota	0.931365	0.924057	0.925579	0.954333	0.757267	1													
Kappa	0.893199	0.970823	0.907416	0.947212	0.736237	0.915482	1												
Lambda	0.8141	0.780039	0.82053	0.764297	0.819568	0.790199	0.733316	1											
Mu	0.895851	0.92176	0.952754	0.937216	0.703253	0.96266	0.932454	0.806618	1										
Omega	0.413027	0.347344	0.227571	0.347794	0.272102	0.2738	0.268587	0.390469	0.303647	1									
Omicron	0.932967	0.963887	0.92402	0.960963	0.734514	0.957625	0.962953	0.790357	0.973687	0.32442	1								
Psi	0.912687	0.961386	0.924095	0.944014	0.768442	0.922551	0.960322	0.79498	0.93476	0.33177	0.958172	1							
Rho	0.911428	0.968582	0.835139	0.959355	0.686688	0.931092	0.926913	0.77028	0.909644	0.350038	0.95607	0.91459	1						
Sigma	0.670383	0.664377	0.67237	0.703248	0.517797	0.752987	0.711528	0.52886	0.769575	0.328789	0.745459	0.666043	0.726945	1					
Tau	0.936208	0.958166	0.92504	0.9642	0.763852	0.959618	0.964788	0.813019	0.970363	0.340044	0.981737	0.950559	0.945076	0.739332	1				
Theta	0.713926	0.725387	0.80469	0.790325	0.475676	0.793594	0.792138	0.582355	0.867465	0.21728	0.814674	0.777424	0.713807	0.839441	0.787146	1			
Upsilon	0.942704	0.934669	0.830764	0.926537	0.730086	0.908586	0.91944	0.714153	0.872897	0.38021	0.934145	0.943264	0.921799	0.666096	0.929746	0.694701	1		
Xi	0.902405	0.928411	0.934803	0.933412	0.797626	0.952098	0.929894	0.863646	0.97313	0.328905	0.964478	0.949658	0.915242	0.733078	0.977084	0.791398	0.896651	1	
Zeta	0.919833	0.95779	0.928404	0.945604	0.756928	0.932716	0.974584	0.746876	0.943901	0.238372	0.973484	0.946075	0.923086	0.720835	0.974014	0.788268	0.9127	0.940323	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
152	13	6
Median Correlation	Mean Correlation	
0.915242	0.818749	

small weak strong association

Table 3-10 shows the standard deviations of the Δb* scores of the Black 472 samples by site. The standard deviations from sites Sigma and Omega are higher than those from the other sites, yet the correlations of scores from site Sigma and the scores from the other sites are still high. We have no other explanation for these low correlations from site Omega for the Δb* scores.

Table 3-10. Standard Deviations of Δb Scores of Black 472 Samples by Site*

Site	Standard deviation
Alpha	0.140
Beta	0.153
Chi	0.127
Delta	0.162
Epsilon	0.199
Iota	0.137
Kappa	0.149
Lambda	0.152
Mu	0.145
Omega	0.214

Site	Standard deviation
Omicron	0.153
Psi	0.149
Rho	0.150
Sigma	0.215
Tau	0.144
Theta	0.165
Upsilon	0.136
Xi	0.129
Zeta	0.149

Hypothesis 2—Findings

Overall, we find that spectrophotometer machines at different sites tend to report ΔL*, Δa*, and Δb* readings that are reasonably close to one another. The caveats to this overall conclusion are as follows:

- ΔL* and Δb* readings of Air Force Grey from sites Omega and Omicron tended to diverge significantly from the ΔL* and Δb* readings recorded at other sites.
- Readings of ΔL*, Δa*, and Δb* across sites were more highly correlated for the Air Force Grey Fabrics than they were for the Black 472 fabrics. We found the weakest correlation among the Δa* scores of the Black 472 fabrics. This lack of correlation appeared to be related to a lack of variation in the Δa* scores of the Black 472 fabrics.

Corollary 1.a

Having tested the hypotheses 1 and 2, we now turn to the corollaries. The first corollary we examine is Corollary 1.a:

> *Corollary 1.a*: Spectrophotometers produced by different manufacturers will produce different values of L*, a*, and b*.

Central Tendency and Dispersion of Absolute Measures (L*, a*, and b*) of Solid Color Tiles and Fabrics by Manufacturer

Table 3-11 gives summary results by device manufacturer for the six color tile samples and the three solid color fabric samples in this study for which values of L*, a*, and b* are available.

Table 3-11. Color Samples by Manufacturer That Varied by More Than 0.3 from the Mean Readings from All Manufacturers

Spectrophotometer brand	L*	a*	b*
Aleph			
Daleth	Coyote 498 Army Green 491	Army Green 491 Desert Sand 503	Deep Blue Tile
Gimel			
Beth	Deep Blue Tile Deep Grey Tile Coyote 498 Desert Sand 503		White Tile Deep Blue Tile Pale Grey Tile Mid Grey Tile Deep Grey Tile Coyote 498

Table 3-11 shows the number of color samples for which the mean measurements from a particular spectrophotometer manufacturer were more than 0.3 above or below the mean values reported by all manufacturers. We use the value of 0.3 rather than 0.5 because, in no instance, were the reported mean values from particular spectrophotometer manufacturers more than 0.5 above or below the mean for all models for either L*, a*, or b*. Highlights of this table are as follows:

- Machines manufactured by Gimel and Aleph tend to have mean readings closest to one another.[6]

- The Beth machines have the most divergent readings for L* and b*. In particular, the Beth machines had divergent b* readings for five of the six tile color samples. Further manual inspection of the data (not shown) revealed that these divergent readings came mostly from the Beth machine at site Tau. Therefore, it may be site-specific factors at site Tau rather than the Beth machine that is causing the divergent readings.

- The Daleth machine had two out of nine divergent readings for L* and a*. In both cases, these were readings from fabric samples.

We will now look at how the standard deviations of the scores for L*, a*, and b* vary by the brand of spectrophotometer (see Figure 3-15, Figure 3-16, and Figure 3-17).[7]

[6] Sixteen of the 19 spectrophotometers in this study were either Gimel or Aleph models. Since the readings from these two models tend to agree with one another, the readings from these two models have a disproportionate effect on the overall mean for all manufacturers. The fact that the Daleth machine and the two Beth machines produce readings that diverge from this mean does not necessarily imply anything about the relative accuracy among the different brands.

[7] There was only one Daleth spectrophotometer. This reduces the potential for variation on the Dalet brand significantly below that of other brands. As a result, no inferences should be drawn about the former Daleth's brand's performance relative to the other brands based upon these graphs.

Figure 3-15. Standard Deviations of Absolute L by Spectrophotometer Maker and Type of Sample*

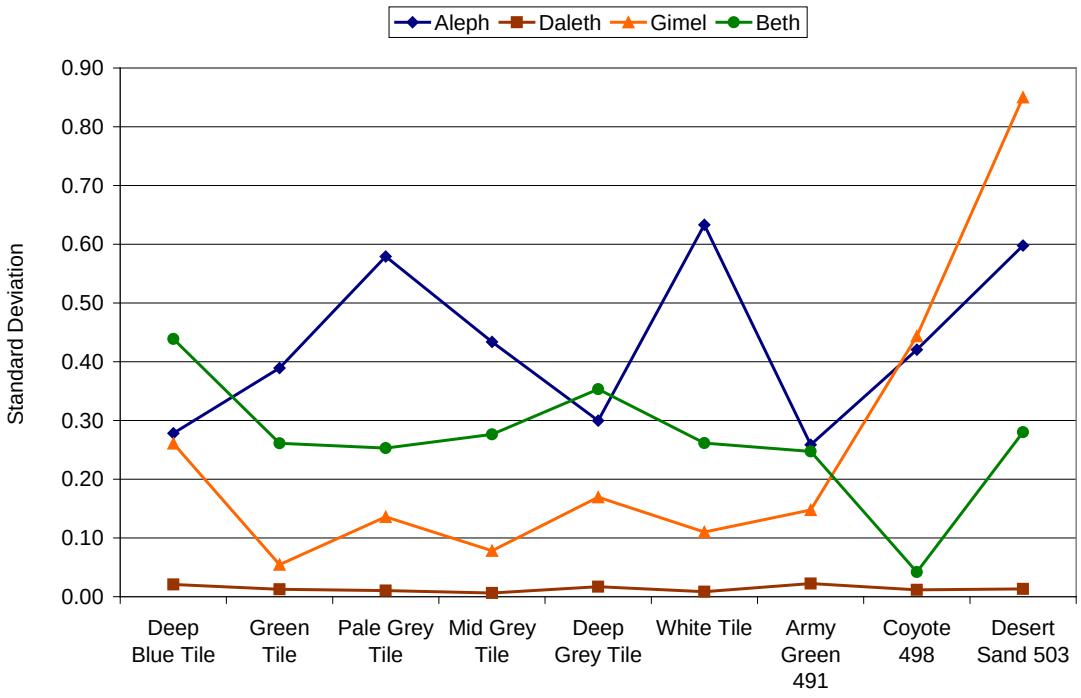


Figure 3-16. Standard Deviations for Absolute a by Spectrophotometer Maker and Type of Sample*

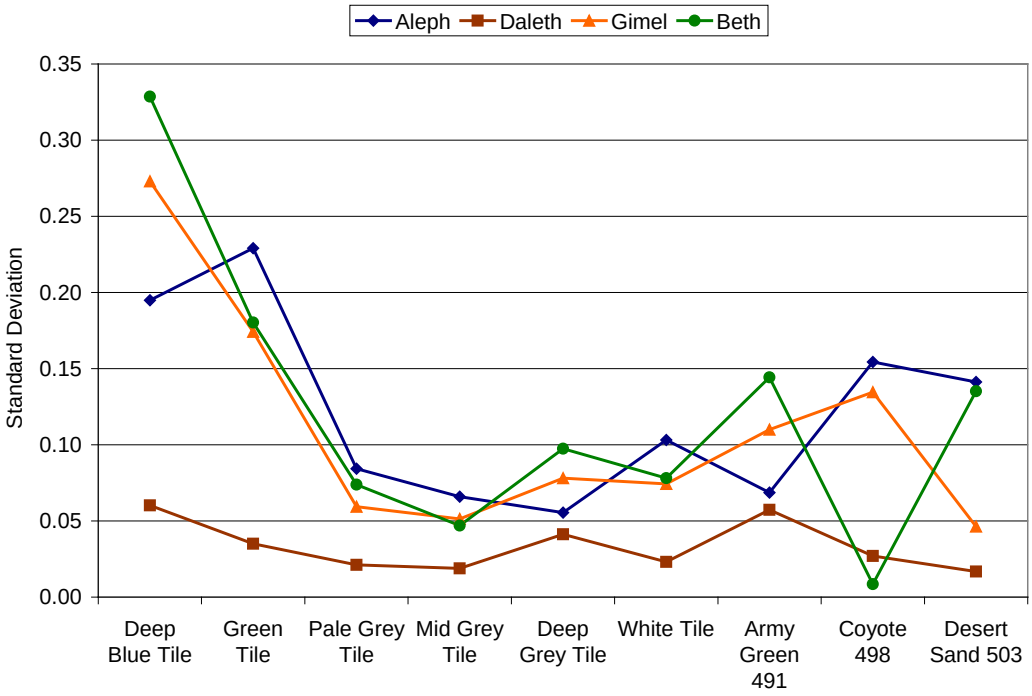
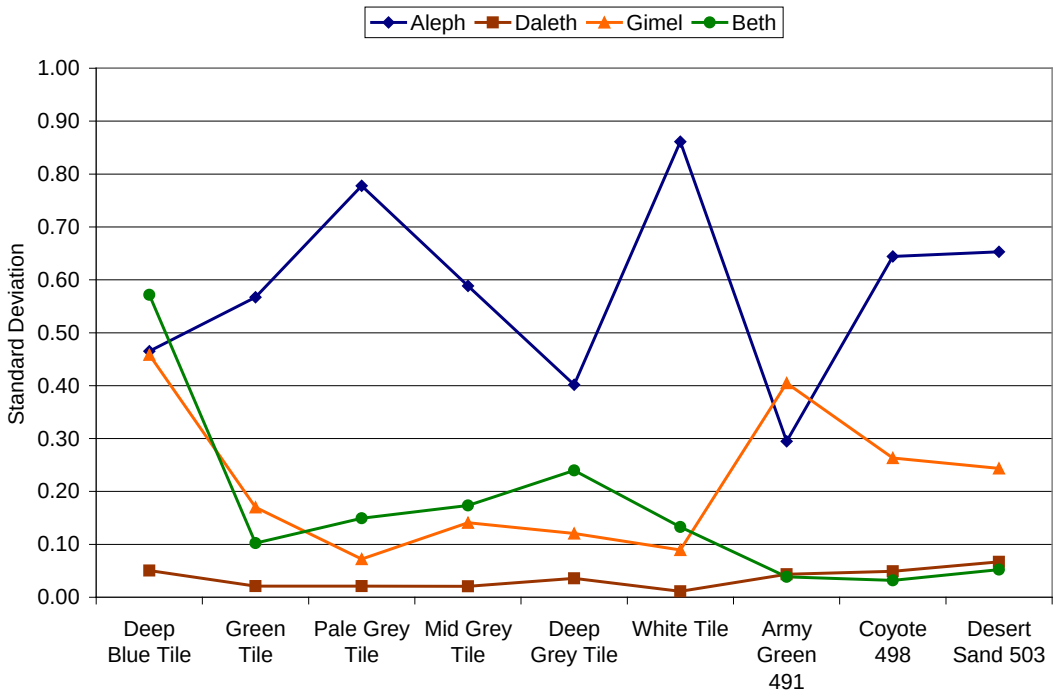


Figure 3-17. Standard Deviations of Absolute b by Spectrophotometer Maker and Type of Sample*



Highlights from Figure 3-15, Figure 3-16, and Figure 3-17 include the following:

- Sites with Aleph spectrophotometers showed the most dispersion in L* and b* measurements. The high level of dispersion of L* and b* readings of the White and Pale Grey tiles we observed in Figure 3-1 appears to be caused by the divergent readings from the Aleph machines. The greater divergence in L* measurements from Aleph machines was limited to colors with the highest L* scores (White and Pale Grey). The greater divergence in b* measurements from Aleph machines was seen for seven of the nine colors.[8] An analysis of variance (not shown) comparing the dispersion of measurements within Aleph sites to the dispersion of measurement across Aleph sites indicates that nearly all of the variation observed for measurements from Aleph machines was between Aleph sites. In other words, readings from a particular Aleph machine showed little dispersion, while readings from different Aleph machines had a tendency to diverge from one another for L* and b* measurements.

[8] This finding of greater dispersion for L* and b* scores from Aleph machines may seem to contradict our earlier finding from Table 3-11 that average readings from Aleph and Gimel machines were the most similar to average readings of all the machines. While the individual Aleph machines are producing divergent readings, when these divergent readings are averaged, the resulting average tends to match the average of the readings from the Gimel machines. This is a good example of the importance of looking at measures of central tendency (means) as well as dispersion (standard deviation). Measures of central tendency give us a sense of what value the scores tend to cluster around, while measures of dispersion give us a sense of how tightly or loosely those values cluster around that mean value.

- Sites with Gimel spectrophotometers generally showed a similar pattern of dispersion in their CIELAB measurements of color tiles to those of the Beth machines. On the other hand, sites with Gimel machines generally reported more dispersed L* and b* readings of color fabrics than did sites with Beth machines. Further inspection of the data (not shown) indicated that the unusually high standard deviation for Gimel L* scores of Desert Sand 503 was caused primarily by divergent readings from site Omicron.

- The White tile and the three shades of grey tiles all showed low levels of dispersion in a* readings from all sites. These four color shades also had a* values closest to zero on Table 3-3.

Central Tendency and Dispersion of Absolute Measures (L*, a*, and b*) of Black 472 and Air Force Grey by Manufacturer

Table 3-12 shows the means and standard deviations for the samples of Black 472 and Air Force Grey by spectrophotometer manufacturer. None of the means in the table is more than 0.3 above or below the overall means for Air Force Grey or Black 472 shown in Table 3-6. One surprising result in Table 3-12 is that the standard deviations for CIELAB scores of Black 472 and Air Force Grey from Gimel machines tend to be higher than the standard deviations for Aleph or Beth machines. Gimel machines for the most part showed no such tendency in the measurements of the solid color samples.

Table 3-12. Means and Standard Deviations of Air Force Grey and Black 472 by Manufacturer

	Air Force Grey			Black 472		
Scanner maker	L*	a*	b*	L*	a*	b*
Aleph mean	50.30	0.17	5.49	16.29	0.35	−1.49
Aleph standard deviation	1.28	0.39	0.76	0.40	0.13	0.33
Gimel mean	50.21	0.11	6.03	16.38	0.35	−1.59
Gimel standard deviation	1.70	0.49	0.71	0.61	0.12	0.71
Beth mean	50.34	0.16	5.81	16.58	0.46	−1.32
Beth standard deviation	1.26	0.42	0.47	0.51	0.06	0.15

Note: Each number is an average of observations taken by spectrophotometers of that brand. There were 144 observations for Aleph, 126 for Gimel, and 36 for Beth. There are no L*, a*, or b* measurements for the site with the Daleth machine. No means are shaded because there are no significant divergences in readings between the different models of machines.

Corollary 1.a—Findings

We find limited support[9] for Corollary 1.a. Readings from spectrophotometers produced by different manufacturers do diverge from one another, but this divergence is never over 0.5 on any of the CIELAB measurements. Perhaps the most important difference among the machines is the dispersion in their readings. Aleph machines at different sites showed a tendency to produce divergent readings of the solid color samples. Gimel machines showed the greatest tendency to produce divergent readings of Air Force Grey and Black 472.

Corollary 1.b

> *Corollary 1.b*: Spectrophotometers produced by different manufacturers will produce different values of ΔL*, Δa*, and Δb*.

Table 3-13 displays means and standard deviations for the Air Force Grey and Black 472 samples by spectrophotometer manufacturer. Our earlier findings for absolute CIELAB scores cannot be extended to the difference scores of ΔL*, Δa*, and Δb*. The only significantly divergent readings we see in this table are the ΔL* scores for Air Force Grey from the Daleth machine, which was the only mean difference score to be greater than 1.0. Overall, we find limited support for Corollary 1.b.

Table 3-13. ΔL, Δa*, and Δb* by Spectrophotometer Brand and Type of Material Sampled*

	Air Force Grey			Black 472		
Scanner maker	ΔL*	Δa*	Δb*	ΔL*	Δa*	Δb*
Aleph mean	0.041	0.279	0.041	−0.531	−0.015	0.115
Aleph standard deviation	1.232	0.387	1.232	0.397	0.072	0.158
Daleth mean	1.556	0.130	0.862	−0.618	0.028	0.019
Daleth standard deviation	1.343	0.408	0.511	0.342	0.120	0.152
Gimel mean	0.372	0.180	0.817	−0.469	0.020	0.133
Gimel standard deviation	1.571	0.542	0.734	0.607	0.100	0.174
Beth mean	−0.038	0.274	0.569	−0.673	−0.009	0.136
Beth standard deviation	1.262	0.415	0.474	0.383	0.049	0.140

Note: Each number is an average of observations taken by spectrophotometers of that brand. There were 161 observations for Aleph, 18 for Daleth, 126 for Gimel, and 36 for Beth. Means are shaded in yellow if absolute value of the average difference value was greater than 1.

[9] When we test a hypothesis or corollary, we are seeking evidence to determine whether or not this hypothesis is true or not true. If our evidence leads us to believe that the hypothesis is true, we say that the evidence supports the hypothesis. If our evidence does not lead us to believe that the hypothesis is true, then we say that evidence does not support the hypothesis.

Corollary 1.b—Findings

Overall, we find very little support for Corollary 1.b.

Corollary 2.a

> *Corollary 2.a*: Spectrophotometers with different lens aperture sizes will produce readings of L*, a*, and b* that correlate differently.

The primary purpose of our data collection from sample readings of Air Force Grey and Black 472 was to test Hypothesis 2. However, this data collection also provided us with an opportunity to test Corollaries 2.a and 2.b because these samples were measured using two different types of aperture size. The Air Force Grey samples were measured using a small aperture size and the Black 472 samples were measured using a large aperture size. This was not a perfect test of Corollaries 2.a and 2.b because we cannot rule out the possibility that different levels of correlation among these sets of reasons are affected by differences in the samples themselves. We have already found, for example, that the standard deviations of CIELAB scores were higher for Air Force Grey samples than for Black 472 samples.

Figure 3-18, Figure 3-19, Figure 3-20, Figure 3-21, Figure 3-22, and Figure 3-23 show the correlation matrices for L*, a*, and b* CIELAB scores for the Air Force Grey and Black 472 fabric samples. These correlation matrices are similar to the earlier figures that showed the difference CIELAB scores for the Air Force Grey and Black 472 fabric samples. All the CIELAB scores for Air Force Grey are more highly correlated than the CIELAB scores for Black 472.

Figure 3-18. Correlations of L Scores for Air Force Grey*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.902436	1																	
Chi	0.927269	0.969906	1																
Delta	0.808256	0.866876	0.895894	1															
Epsilon	0.880036	0.92185	0.942491	0.86975	1														
Iota	0.948653	0.976765	0.98163	0.888203	0.953399	1													
Kappa	0.924293	0.955444	0.973666	0.872358	0.955432	0.983497	1												
Lambda	0.880009	0.882919	0.861423	0.765882	0.845655	0.918578	0.912666	1											
Mu	0.949505	0.975316	0.98478	0.880888	0.955868	0.998066	0.982722	0.898429	1										
Omega	0.876873	0.834004	0.857663	0.742017	0.90516	0.894605	0.900273	0.862147	0.893333	1									
Omicron	0.766273	0.722755	0.712876	0.728348	0.638185	0.764935	0.71692	0.800646	0.754001	0.65694	1								
Psi	0.925823	0.975284	0.973021	0.880161	0.927385	0.98743	0.9757	0.876251	0.990298	0.866896	0.744105	1							
Rho	0.903453	0.974515	0.966386	0.865476	0.926908	0.968862	0.958433	0.875188	0.971314	0.828132	0.745505	0.963988	1						
Sigma	0.942833	0.979152	0.985178	0.891412	0.953774	0.998687	0.985217	0.908542	0.997867	0.896913	0.758397	0.98995	0.969578	1					
Tau	0.943389	0.973373	0.979386	0.877447	0.946829	0.995501	0.976248	0.909691	0.995194	0.871772	0.764629	0.984361	0.963778	0.993305	1				
Theta	0.941635	0.976234	0.976676	0.873957	0.948379	0.996778	0.980402	0.900918	0.9979	0.887343	0.759667	0.993369	0.966907	0.996313	0.99435	1			
Upsilon	0.958147	0.955439	0.965276	0.852011	0.936397	0.987089	0.968235	0.906573	0.984346	0.900804	0.734009	0.971326	0.949525	0.985682	0.976806	0.984011	1		
Xi	0.954793	0.969924	0.973471	0.886904	0.939525	0.996738	0.975397	0.909319	0.99497	0.880992	0.775912	0.98773	0.959583	0.994198	0.993479	0.996457	0.98747	1	
Zeta	0.944896	0.966353	0.979946	0.890874	0.953566	0.996757	0.984668	0.911362	0.996714	0.900544	0.766163	0.988047	0.960958	0.995374	0.994244	0.99532	0.977929	0.993462	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
171	0	0
Median Correlation		Mean Correlation
0.949525		0.920152

small weak strong association

Figure 3-19. Correlations of a Scores for Air Force Grey*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.961593	1																	
Chi	0.953298	0.975239	1																
Delta	0.543126	0.562158	0.577389	1															
Epsilon	0.63724	0.676383	0.679028	0.598459	1														
Iota	0.969828	0.987907	0.971849	0.559707	0.66873	1													
Kappa	0.969851	0.9702	0.976764	0.627882	0.666456	0.974928	1												
Lambda	0.938092	0.965015	0.968714	0.599674	0.676691	0.969258	0.945217	1											
Mu	0.971287	0.992488	0.980822	0.575636	0.681725	0.996906	0.979529	0.969425	1										
Omega	0.564951	0.623672	0.572552	0.242869	0.563025	0.589308	0.510255	0.577007	0.582701	1									
Omicron	0.790028	0.782795	0.704632	0.210146	0.604244	0.807602	0.743927	0.698608	0.791933	0.59754	1								
Psi	0.963753	0.982988	0.976636	0.563748	0.695894	0.991833	0.97128	0.97882	0.98862	0.615177	0.77876	1							
Rho	0.953252	0.984615	0.974656	0.640845	0.713448	0.986245	0.982824	0.969248	0.987661	0.591907	0.752252	0.986946	1						
Sigma	0.910698	0.93867	0.946955	0.525755	0.687654	0.92874	0.914232	0.929024	0.946478	0.558949	0.728008	0.925421	0.923349	1					
Tau	0.970254	0.986669	0.96861	0.538435	0.678777	0.996676	0.972081	0.964898	0.994105	0.595476	0.818005	0.988796	0.982765	0.938279	1				
Theta	0.97416	0.985727	0.970667	0.582289	0.676072	0.995986	0.975966	0.963805	0.997309	0.564991	0.804233	0.983616	0.981646	0.933119	0.991227	1			
Upsilon	0.96713	0.967637	0.951306	0.615366	0.671025	0.973899	0.977311	0.931532	0.972277	0.562561	0.772626	0.964662	0.974406	0.881817	0.975296	0.971824	1		
Xi	0.96446	0.985795	0.967506	0.561606	0.664689	0.997413	0.974127	0.963932	0.994304	0.58707	0.808181	0.98829	0.987508	0.928387	0.996899	0.992532	0.978498	1	
Zeta	0.975295	0.983973	0.970636	0.573174	0.657378	0.997069	0.981509	0.960366	0.994583	0.577113	0.807061	0.985912	0.986212	0.922043	0.994043	0.995551	0.981818	0.996803	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
169	0	2
Median Correlation		Mean Correlation
0.960366		0.841496

small weak strong association

Figure 3-20. Correlations of b Scores for Air Force Grey*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.959586	1																	
Chi	0.948528	0.97211	1																
Delta	0.822892	0.832951	0.783304	1															
Epsilon	0.575482	0.613882	0.637078	0.703541	1														
Iota	0.980717	0.979177	0.974479	0.847309	0.648469	1													
Kappa	0.947827	0.962619	0.959574	0.820394	0.632215	0.972923	1												
Lambda	0.926214	0.940292	0.907876	0.813383	0.580589	0.943333	0.961339	1											
Mu	0.985367	0.980306	0.970581	0.844301	0.631585	0.997187	0.973951	0.944259	1										
Omega	0.833428	0.7664	0.754015	0.696288	0.676308	0.837485	0.768472	0.783329	0.841049	1									
Omicron	0.805669	0.749416	0.719572	0.74377	0.426971	0.811153	0.727178	0.791799	0.796891	0.733618	1								
Psi	0.975702	0.979801	0.974304	0.832633	0.617438	0.985572	0.964456	0.935379	0.983679	0.807154	0.766887	1							
Rho	0.969392	0.972884	0.945935	0.863307	0.622855	0.985326	0.968379	0.949189	0.985294	0.823277	0.793532	0.973796	1						
Sigma	0.98343	0.981213	0.971657	0.812019	0.620999	0.992406	0.968998	0.9292	0.995772	0.830637	0.777344	0.983035	0.980035	1					
Tau	0.97643	0.981046	0.982545	0.838868	0.637615	0.99385	0.964549	0.930763	0.99259	0.822911	0.795083	0.990991	0.974574	0.990021	1				
Theta	0.979326	0.970853	0.975528	0.843559	0.637	0.993048	0.96638	0.943104	0.993903	0.838609	0.807858	0.979632	0.972327	0.984417	0.9903	1			
Upsilon	0.980566	0.982302	0.960625	0.846818	0.609383	0.99132	0.966647	0.94738	0.993704	0.818146	0.807487	0.983997	0.978727	0.990248	0.988869	0.983514	1		
Xi	0.977284	0.978383	0.980468	0.845265	0.648412	0.996943	0.972312	0.937451	0.994041	0.82043	0.79198	0.982114	0.983972	0.988068	0.991354	0.993494	0.983669	1	
Zeta	0.974058	0.976697	0.982684	0.846963	0.637176	0.994304	0.962385	0.932568	0.989544	0.808952	0.812545	0.98491	0.9757	0.983167	0.995714	0.992016	0.982528	0.995322	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
170	1	0
Median Correlation		Mean Correlation
0.960625		0.888741

small weak strong association

Figure 3-21. Correlations of L Scores for Black 472*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.95178	1																	
Chi	0.795896	0.832181	1																
Delta	0.958213	0.935623	0.812208	1															
Epsilon	0.904939	0.872077	0.742541	0.863704	1														
Iota	0.850195	0.892766	0.951288	0.847065	0.803671	1													
Kappa	0.903617	0.879706	0.839081	0.89891	0.792783	0.889692	1												
Lambda	0.902138	0.929467	0.915621	0.90927	0.843396	0.962804	0.888858	1											
Mu	0.781929	0.848809	0.945849	0.774228	0.740334	0.962131	0.791155	0.922352	1										
Omega	0.707849	0.660643	0.570015	0.742055	0.626744	0.588207	0.801734	0.681613	0.444055	1									
Omicron	0.265375	0.264201	0.443926	0.366892	0.023624	0.453111	0.393323	0.366185	0.447066	0.121075	1								
Psi	0.885662	0.91757	0.942954	0.856861	0.811375	0.973543	0.895245	0.964314	0.956783	0.614038	0.379916	1							
Rho	0.96295	0.96964	0.842481	0.945423	0.858277	0.908499	0.9041	0.938498	0.829356	0.703533	0.332533	0.926522	1						
Sigma	0.847858	0.889481	0.936863	0.8284	0.802276	0.97035	0.875509	0.965354	0.955531	0.593874	0.360018	0.965941	0.884816	1					
Tau	0.770726	0.808418	0.914049	0.838794	0.741909	0.912114	0.812669	0.920401	0.879267	0.634335	0.41983	0.883289	0.825852	0.902797	1				
Theta	0.826241	0.877862	0.956136	0.840091	0.774116	0.987123	0.879387	0.964431	0.95735	0.611907	0.445054	0.968923	0.895234	0.970024	0.90594	1			
Upsilon	0.972187	0.909816	0.754134	0.961758	0.864524	0.814545	0.906489	0.863536	0.716774	0.719623	0.309194	0.819882	0.934665	0.801996	0.745919	0.788363	1		
Xi	0.859644	0.893651	0.947555	0.830996	0.810584	0.97154	0.868492	0.946632	0.935812	0.598236	0.393157	0.982798	0.916898	0.944103	0.880984	0.960317	0.788059	1	
Zeta	0.93754	0.946074	0.889741	0.944924	0.84153	0.939636	0.936225	0.975723	0.890838	0.72508	0.405471	0.95551	0.952619	0.930257	0.870365	0.936606	0.913067	0.927997	1

Strong Associations (>.5)	Weak Associations (.3 to .5)	Small Assocations (.1 to .3)
152	15	3
Median Correlation		Mean Correlation
0.877862		0.808816

small weak strong association

Figure 3-22. Correlations of a Scores for Black 472*

	Alpha		Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.451525	1																	
Chi	0.530218	0.54112	1																
Delta	0.642969	0.680679	0.377874	1															
Epsilon	0.228623	0.111241	0.20755	0.133124	1														
Iota	0.552615	0.574334	0.634109	0.600874	0.094129	1													
Kappa	0.532829	0.392454	0.263566	0.520883	0.118105	0.607753	1												
Lambda	0.294146	0.242465	0.190402	0.309009	0.25057	0.138983	0.044607	1											
Mu	0.617688	0.710552	0.640293	0.77238	0.186948	0.726625	0.743843	0.37514	1										
Omega	0.212168	0.413383	0.40875	0.432514	0.003443	0.583729	0.297904	0.466506	0.575964	1									
Omicron	0.356971	0.208599	0.377549	0.332806	-0.05336	0.332119	0.553958	-0.01408	0.589625	0.142223	1								
Psi	0.457064	0.673139	0.657904	0.620903	-0.07675	0.644452	0.425058	0.288862	0.736561	0.512564	0.464277	1							
Rho	0.612561	0.706572	0.574238	0.631702	0.083863	0.623348	0.585226	0.336461	0.840282	0.522478	0.58077	0.593989	1						
Sigma	0.339347	0.142549	-0.07837	0.472667	0.024575	0.423025	0.497197	-0.003	0.249249	0.386307	-0.12781	0.190378	0.133501	1					
Tau	0.596083	0.610157	0.30249	0.791005	0.015832	0.727887	0.72097	0.360265	0.728281	0.571291	0.273079	0.65879	0.548695	0.612706	1				
Theta	0.665699	0.592797	0.475574	0.708378	0.003577	0.679674	0.620606	0.277451	0.781654	0.424323	0.217261	0.723122	0.602395	0.424392	0.767034	1			
Upsilon	0.484999	0.431677	-0.07893	0.689415	0.026234	0.454011	0.713934	0.277725	0.525299	0.313436	0.175767	0.401553	0.377687	0.641571	0.885301	0.614385	1		
Xi	0.438117	0.554566	0.289625	0.680944	0.120519	0.618514	0.573722	0.32737	0.679842	0.625863	0.090418	0.550599	0.375663	0.512412	0.788391	0.734101	0.716622	1	
Zeta	0.569909	0.656727	0.615163	0.740922	0.051467	0.700968	0.702305	0.349386	0.893503	0.678172	0.566421	0.813194	0.787775	0.412707	0.787964	0.75067	0.563905	0.669177	1

Strong Associations (>.5)			Weak Associations (.3 to .5)			Small Assocations (.1 to .3)		
	88			38			28	
	Median Correlation				Mean Correlation			
		0.520883				0.456751		

small weak strong association

Figure 3-23. Correlations of b Scores for Black 472*

	Alpha	Beta	Chi	Delta	Epsilon	Iota	Kappa	Lambda	Mu	Omega	Omicron	Psi	Rho	Sigma	Tau	Theta	Upsilon	Xi	Zeta
Alpha	1																		
Beta	0.95178	1																	
Chi	0.795896	0.832181	1																
Delta	0.958213	0.935623	0.812208	1															
Epsilon	0.904939	0.872077	0.742541	0.863704	1														
Iota	0.850195	0.892766	0.951288	0.847065	0.803671	1													
Kappa	0.903617	0.879706	0.839081	0.89891	0.792783	0.889692	1												
Lambda	0.902138	0.929467	0.915621	0.90927	0.843396	0.962804	0.888858	1											
Mu	0.781929	0.848809	0.945849	0.774228	0.740334	0.962131	0.791155	0.922352	1										
Omega	0.707849	0.660643	0.570015	0.742055	0.626744	0.588207	0.801734	0.681613	0.444055	1									
Omicron	0.265375	0.264201	0.443926	0.366892	0.023624	0.453111	0.393323	0.366185	0.447066	0.121075	1								
Psi	0.885662	0.91757	0.942954	0.856861	0.811375	0.973543	0.895245	0.964314	0.956783	0.614038	0.379916	1							
Rho	0.96295	0.96964	0.842481	0.945423	0.858277	0.908499	0.9041	0.938498	0.829356	0.703533	0.332533	0.926522	1						
Sigma	0.847858	0.889481	0.936863	0.8284	0.802276	0.97035	0.875509	0.965354	0.955531	0.593874	0.360018	0.965941	0.884816	1					
Tau	0.770726	0.808418	0.914049	0.838794	0.741909	0.912114	0.812669	0.920401	0.879267	0.634335	0.41983	0.883289	0.825852	0.902797	1				
Theta	0.826241	0.877862	0.956136	0.840091	0.774116	0.987123	0.879387	0.964431	0.95735	0.611907	0.445054	0.968923	0.895234	0.970024	0.90594	1			
Upsilon	0.972187	0.909816	0.754134	0.961758	0.864524	0.814545	0.906489	0.863536	0.716774	0.719623	0.309194	0.819882	0.934665	0.801996	0.745919	0.788363	1		
Xi	0.859644	0.893651	0.947555	0.830996	0.810584	0.97154	0.868492	0.946632	0.935812	0.598236	0.393157	0.982798	0.916898	0.944103	0.880984	0.960317	0.788059	1	
Zeta	0.93754	0.946074	0.889741	0.944924	0.84153	0.939636	0.936225	0.975723	0.890838	0.72508	0.405471	0.95551	0.952619	0.930257	0.870365	0.936606	0.913067	0.927997	1

Strong Associations (>.5)			Weak Associations (.3 to .5)			Small Assocations (.1 to .3)		
	152			13			6	
	Median Correlation				Mean Correlation			
		0.899				0.802902		

small weak strong association

Corollary 2.a—Findings

We find some support for Corollary 2.a. Spectrophotometers taking measurements with a small aperture size do report more highly correlated CIELAB measures than spectrophotometers taking measurements with a large aperture sizes. We must qualify our support for this corollary by noting that the small and large aperture sizes were used on two entirely different fabric sample sets.

Corollary 2.b

> *Corollary 2.b*: Spectrophotometers with different lens aperture sizes will produce readings of ΔL*, Δa*, and Δb* that correlate differently.

Our test of corollary 2.b is the same as our test of Corollary 2.a. The correlation matrices we used to test this corollary are in Figure 3-9, Figure 3-10, Figure 3-11, Figure 3-12, Figure 3-13, and Figure 3-14. These figures show the correlation matrices for ΔL*, Δa*, and Δb* readings of the Air Force Grey and Black 472 samples.

Corollary 2.b—Findings

The scores from the Air Force Grey samples were consistently more highly correlated than the scores for the Black 472 samples. Recall that the Air Force Grey samples were measured with small lens aperture, while the Black 472 samples were measured with a large lens aperture. As a result, our findings for corollary 2.b are the same as our findings for corollary 2.a.

SUPPLEMENTAL ANALYSIS—DOES TESTING CHANGE THE L*, a*, AND b* VALUES OF SAMPLES?

During our study, we found b* readings tended to decay with repeated measurements for specific samples.[10] We examined if the tendency was statistically significant and if it occurred on all CIELAB axes. We examined the CIELAB readings from the Desert Sand 503 fabric to test the hypothesis that spectrophotometer readings decay over time.

Figure 3-24 through Figure 3-26 indicate the results of repeated L*, a*, and b* measurements of the Desert Sand 503 fabric. There was little change in the L* and a* measurements, but b* measurements to appear to decay after the first measurement.

[10] The Natick specialists noticed the pattern in b* readings while testing Desert Sand 503.

Figure 3-24. Desert Sand L Readings*

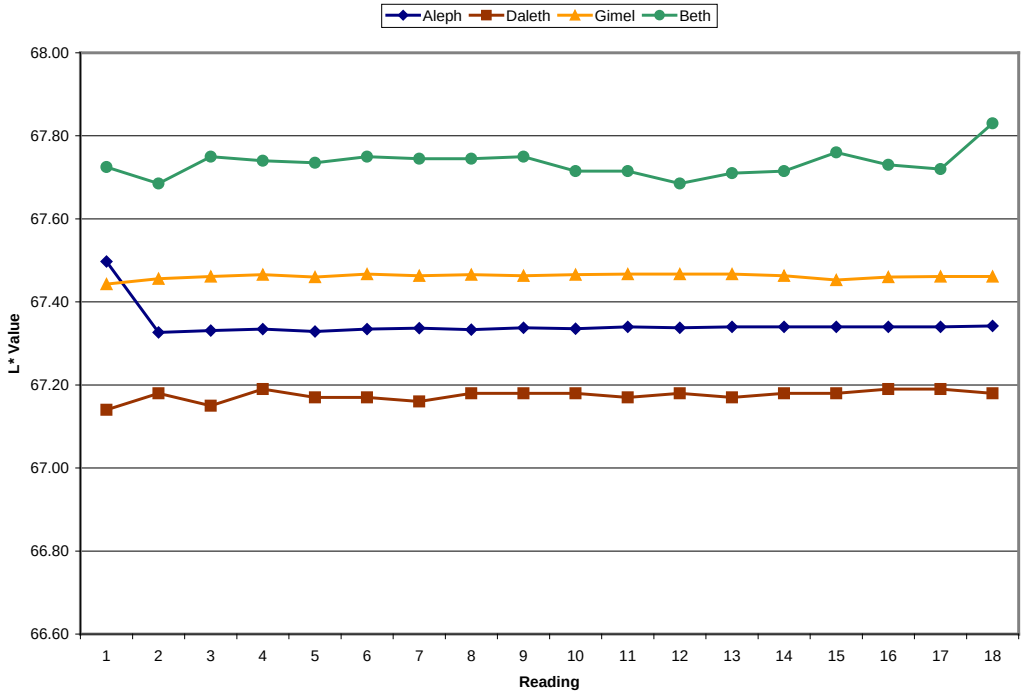


Figure 3-25. Desert Sand a Readings*

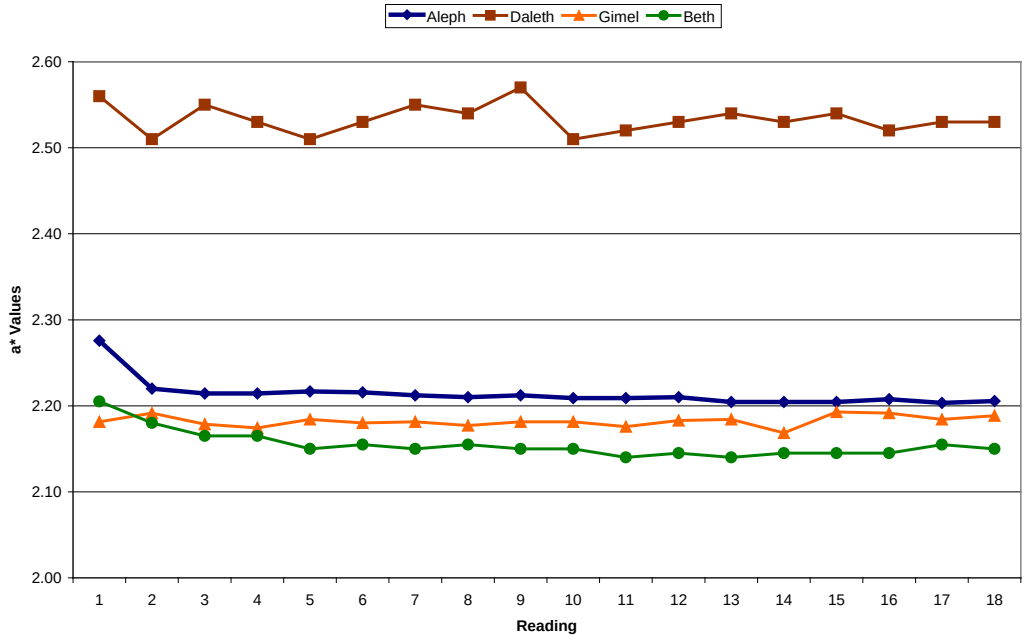
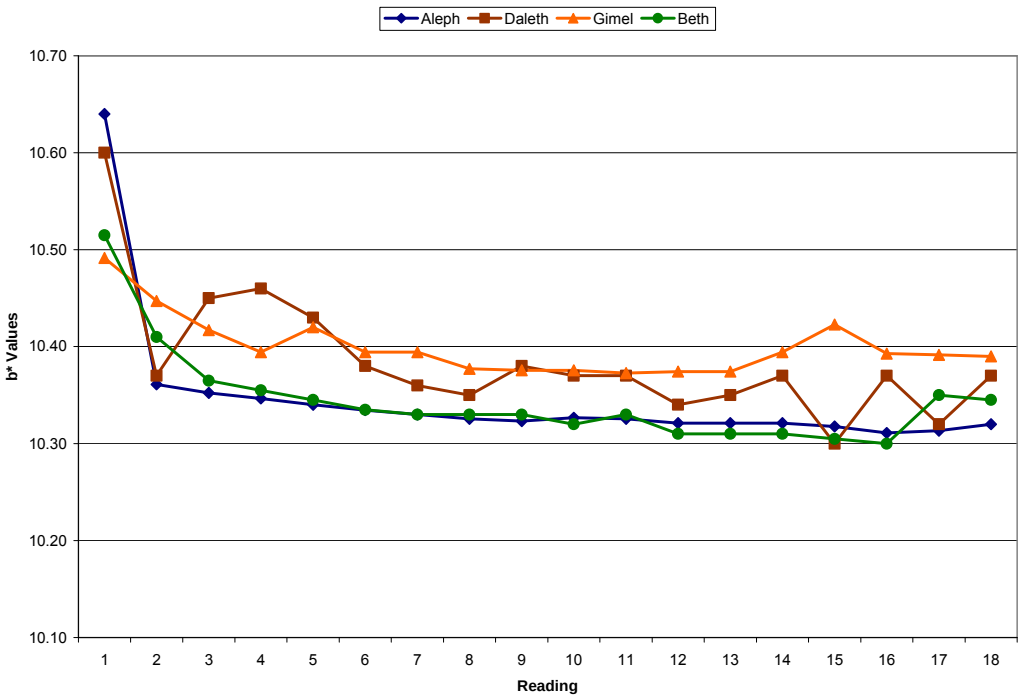


Figure 3-26. Desert Sand b Readings*



After visually inspecting the data, our next step was to test if there was an overall downward trend in L*, a*, or b* readings of the Desert Sand 503 fabric. Table 3-14 shows the correlation coefficients[11] we obtained between the number of observations and the L*, a*, and b* measurements. All of the correlation coefficients are tiny, and none reaches statistical significance.

Table 3-14. Bivariate Correlations between Observations and CIELAB Measures

	L*	a*	b*
Observation number	−0.004	−0.036	−0.059

Note: Each correlation estimated from 340 cases.

While there is no evidence of an overall trend toward decreasing L*, a*, or b* measurements, there is still the possibility that the difference between the first b* reading and subsequent b* readings could be statistically significant. As we see from Figure 3-26, this decrease appears particularly evident in readings from Aleph machines. To test the possibility of a difference between first and subsequent readings, we conducted t-tests comparing first readings and subsequent readings. We conducted separate t-tests for machines produced by different

[11] Correlation coefficients can range from −1 to 1. A correlation coefficient of 0 indicates that there is no relationship between two variables. A correlation coefficient of 1 or −1 means the value of one variable can be perfectly predicted from the value of another variable. An example of two variables with a correlation of 1 would be the same temperature recorded in Fahrenheit and Celsius.

manufacturers. The results are shown on Table 3-15. These results indicate that the Aleph, Daleth, and Beth machines all had statistically significant differences between their first b* readings and their subsequent b* readings.

Table 3-15. Mean First and Subsequent b Readings by Manufacturer*

Manufacturer	Mean first b* reading	Mean subsequent b* readings	Mean decrease
Aleph	10.640	10.329	0.311
Gimel	10.491	10.395	0.097
Daleth	10.600	10.373	0.227
Beth	10.515	10.334	0.181

Note: A total of 18 readings were taken from each machine. The first readings from two of the Aleph machines were missing from the data set. Values for mean decrease shaded in yellow were statistically significant.

SUMMARY OF FINDINGS IN PHASE 2

Overall, we find support for Hypothesis 1 and Hypothesis 2. Spectrophotometer readings tend to be consistent across sites, equipment manufacturers, and spectrophotometer aperture sizes. Other findings from Phase 2 include the following:

- Site Rho consistently reported readings for L* and b* for all color samples that diverged significant from the scores reported from all other sites; however, the ΔL*, Δa*, and Δb* readings from site Rho did not significantly differ from the ΔL*, Δa*, and Δb* readings from the other sites.

- Sites Omega and Omicron reported ΔL* and Δb* readings that were significantly different from the ΔL* and Δb* readings reported by other sites. The greater dispersion of readings from site Omicron was mostly due to a single aberrant reading.

- The brand of the machine had a modest effect on machine readings. Different brands produced readings with different dispersion profiles. None of the brands was superior overall to the other brands in terms of the consistency of their readings.

- L*, a*, b*, ΔL*, Δa*, and Δb* readings generated by a machine with a small aperture lens on camouflage fabrics are correlated at least as strongly as L*, a*, b*, ΔL*, Δa*, and Δb readings generated by a machine with a large aperture lens on camouflage fabrics.

- The first b* readings of the color Desert Sand tended to be significantly higher than subsequent b* readings. This effect was largest for the Aleph machines and smallest for the Gimel machines.

Chapter 4
Phase 3

In Phase 2 evidence was presented that spectrophotometer measurements tend to correlate highly with one another. In Phase 3, comparison was made of evaluations of military fabrics based on spectrophotometer measurements with evaluations made by human shade evaluators. The objective was to test the following hypothesis:

> *Hypothesis 3:* There is a correlation between spectrophotometer pass/fail evaluation based on instrumental measurement and human shade evaluator pass/fail evaluation based on visual methods.[1]

METHODOLOGY

Ten human shade evaluators from the Army Natick Soldier RDEC and the DSCC PTC were recruited for the study. Each evaluator was qualified using the Farnsworth Munsell 100 Hue Test. The intent was to recruit evaluators with normal color vision who were not color shade experts. The evaluators reviewed 10 sets of 10 colored fabric samples. Since each evaluator made 100 evaluations, this provided a data set of 1,000 evaluations.

In each assessment, the evaluator was asked to make a determination as to whether a given fabric passed or failed based on a given shade standard and set of shade tolerances.

Army Natick Soldier RDEC and DSCC PTC provided an "official" determination for each fabric sample as to whether the sample met the given shade standard and was acceptable within tolerance. The fabric samples included the following:

- Black 385 (Army)
- Blue 3329 (Navy)
- Blue 3346 (Navy)
- Blue 3376 (Navy)
- Blue 3386 (Navy)
- Army Green 415 (Army)
- Green 489 (Army)
- Sage Green 1590 (Air Force)
- MC 2312 (Marines)
- Pewter 2246 (Marines).

[1] Use of the industry standard criteria of dE CMC (a modification of the CIELAB equation to represent shade variation as one value) 2:1 of 1.0 is implied by the hypothesis.

PERCENTAGE AGREEMENT—HUMANS WITH "OFFICIAL" DETERMINATION

Table 4-1 shows what percentage of the time the pass/fail decisions given by each of the evaluators agreed with the "official" pass/fail decision. We obtained these percentages by adding the number of times a human evaluator and the official determination both gave a "pass" to a given fabric sample to the number of times that a human evaluator and the "official" determination both gave a "fail" to a given fabric sample. We then divided that result by the total number of pass/fail decisions (in this case 100). The results in Table 4-1 indicate a range of agreement, from a low of 65 percent to a high of 77 percent.

Table 4-1. Agreement on Pass/Fail Decision of Human Shade Evaluators with the "Official" Determination of Pass/Fail

	Agreement percentage	Phi measure
Evaluator 1	70%	0.410[a]
Evaluator 2	76%	0.519[a]
Evaluator 3	69%	0.373[a]
Evaluator 4	70%	0.369[a]
Evaluator 5	73%	0.431[a]
Evaluator 6	65%	0.285[a]
Evaluator 7	77%	0.518[a]
Evaluator 8	77%	0.526[a]
Evaluator 9	71%	0.386[a]
Evaluator 10	75%	0.476[a]
Median among evaluators	72%	0.421
Mean among evaluators	72.3%	0.429

[a] Indicates relationship is statistically significant.

These results suggest a moderate level of agreement between the human evaluators and the "official" determination, but they are not definitive. This is because the percentage of agreement is not completely reliable as a measure of the strength of the relationship between two dichotomous variables (in this case, the two sets of pass/fail decisions). The reason is the agreement percentage does not give us any information about why two parties agree. For example, suppose we were looking at the strength of the relationship of pass/fail decisions between two human evaluators. Suppose that one evaluator passed fabrics 75 percent of the time after carefully scrutinizing the fabrics while another evaluator passed all of the fabrics after only a cursory glance. The resulting agreement percentage would be 75 percent, even though there was no relationship between the standards that the two evaluators were using to judge the fabrics. In fact, if both evaluators merely flip a coin to decide whether a fabric passes, they should still agree about 50 percent of the time.

RELATIONSHIP AMONG HUMAN EVALUATOR PASS/FAIL DECISIONS USING THE PHI COEFFICIENT

To confirm our findings about agreement percentage among the human shade evaluators, we turn to the phi coefficient. The phi coefficient can vary between −1 and 1 in the same fashion as the more widely known r coefficient. If the phi coefficient between two sets of pass/fail decisions was 1, the pass/fail decisions agreed with one another every single time. If the phi coefficient between two sets of pass/fail decisions was −1, the pass/fail decisions disagreed with one another every single time. If the phi coefficient between the two sets of pass/fail decisions was zero, this would mean knowing one pass/fail decision would give you no indication whatsoever about what the other pass/fail decision was likely to be. For example, if both sets of pass/fail decisions used a coin flip to determine whether color samples met the standard, then we would expect the phi coefficient to be zero.

The standard that we used to judge the phi coefficients is as follows:[2]

- −1.0 to −0.5, strong negative association
- −0.49 to −0.3, moderate negative association
- −0.29 to −0.1, small negative association
- −0.09 to 0.09, no association
- 0.1 to 0.29, small positive association
- 0.3 to 0.49, moderate positive association
- 0.5 to 1.0, strong positive association.

The phi coefficients expressing the relationship between the pass/fail decisions of the human participants and the pass/fail decisions generated by the spectrophotometer are shown in the third column of Table 4-1. All of the coefficients are statistically significant and positive. Small associations are highlighted in blue, moderate associations are highlighted in yellow, and strong associations are highlighted in red. All but one of the associations is moderate or strong. The phi correlations track very closely with the pass/fail agreement percentages, thus confirming that the agreement percentages that show the relationship between human pass/fail decisions and spectrophotometer pass/fail decisions are valid.

[2] Jessica Steele, *Choosing the Correct Statistical Test,* www.radford.edu/~jcsteele/Choosing the Correct Statistical Test, 2006.

Overall, we find that the mean agreement between the human pass/fail decisions and the spectrophotometer pass/fail decisions is 72.3 percent. The median percentage of agreement is 72 percent. The degree of closeness between the median and the mean percent agreement tells us that the agreement percentages are evenly clustered around 72 percent. A visual inspection of the data confirms this. Five of the agreement percentages are below 72 percent and five of the agreement percentages are above 72 percent. Similarly, the mean and median phi statistics 0.429 and 0.421 are also close to one another, and the phi statistics are evenly divided above and below the mean and median phi statistics.

COMPARING THE INSTRUMENTATION AND THE "OFFICIAL" DETERMINATION

After comparing the individual human evaluator pass/fail decisions to the "official" pass/fail decisions, our next step was to see how pass/fail decisions generated by a spectrophotometer compared to the official pass/fail decisions. To generate the spectrophotometer pass/fail decision, we used a gauge based on dE CMC tolerance. If the dE CMC score of the sample fabric differed by more than 1.0 from the dE CMC score of the fabric designated as the standard, then the fabric sample was given a "fail" score.

Table 4-2 shows the comparisons of the agreement percentage scores and the phi coefficients between pass/fail decisions of the human evaluators, the spectrophotometer, and the official pass/fail determination. The agreement percentage and phi coefficient for the spectrophotometer are statistically indistinguishable from the mean and median agreement percentage and phi coefficient for the human evaluators. In other words, the spectrophotometer performed as well as the average human evaluator in terms of how well its pass/fail decision compared to the "official" pass/fail determination.

Table 4-2. Percentage of Agreement between Human Evaluators, Instrumentation, and the "Official" Determination

	Agreement percentage	Phi
dE CMC with "official" determination	71%	0.425[a]
Median human agreement with "official" determination	72%	0.421
Mean human agreement with "official" determination	72.3%	0.429

[a] Indicates relationship is statistically significant.

IS THERE A PATTERN BETWEEN "OFFICIAL" DETERMINATION AND INSTRUMENTATION-GENERATED PASS/FAIL DECISIONS?

Table 4-3 is a crosstab between the official and spectrophotometer-generated pass/fail decisions. The "official" determination was much more likely than the spectrophotometer—41 percent versus 14 percent—to fail fabric samples. In fact, there was only one instance when the spectrophotometer passed a fabric that was failed by the official determination. On the other hand, there were 28 instances when the official determination failed a fabric that was passed by the spectrophotometer.

Table 4-3. "Official" Determination and Spectrophotometer Pass/Fail Decisions

Official decision	Spectrophotometer decision		Total
	Fail	Pass	
Fail	13	28	41
Pass	1	58	59
Total	14	86	100

When the principal military shade evaluators at Army Natick Soldier RDEC and the DSCC PTC made the official pass/fail decisions, they also provided a reason for why the fabric sample was not acceptable. This gives us an opportunity to test if there is a relationship between the reason for the fail decision and the spectrophotometer's pass/fail decision. For example, perhaps fabric samples official evaluators judged "too full" were more likely to be passed by the spectrophotometer than were fabric samples judged "too red."

Table 4-4 shows the reason for the official fail decision and the instrumentation pass/fail decision. The chi-square for this crosstab is 0.407, indicating that we have no verification to suppose that there is any evidence for a relationship. In order to conclude that we have evidence of a relationship, the chi-square value would need to be below 0.05. As a result, we have no basis for concluding that instrumentation will agree with officials for some types of fail decision more often than for other types.

Table 4-4. Reasons Given by Officials for Rejecting Fabric Samples by Instrumentation Pass/Fail Decision

Official reason for rejection	Spectrophotometer decision		Total
	Pass	Fail	
Too full	7	3	10
Too thin	1	0	1
Too yellow	3	3	6
Too blue	3	0	3
Too red	0	1	1
Too green	3	0	3
Thin blue	0	1	1
Thin red	2	2	4
Thin green	1	0	1
Full yellow	2	2	4
Full blue	2	0	2
Full red	4	1	5
Total	28	13	41

SUMMARY OF FINDINGS IN PHASE 3

We find there is correlation between spectrophotometer pass/fail decisions based on instrumental measurement and human evaluator pass/fail decisions concerning whether solid fabrics meet acceptable shade standards.

- When individual human shade evaluators made pass/fail decisions concerning whether solid color fabrics met shade standards, they agreed with the official determination between 65 percent and 77 percent of the time.
- The pass/fail determination generated by spectrophotometers agreed with the official determination 71 percent of the time. This rate is comparable to the success rates for the human evaluators. (We validated the above percentages with the appropriate statistical procedures.)
- The official pass/fail visual determination was more likely than the spectrophotometer to determine which samples did not meet pass criteria.
- There is no evidence of a relationship between reasons given by officials for determining that a sample did not meet standards and whether or not the spectrophotometer generated a failing score for that sample.

Appendix A
Ancillary Study Report

This report describes the purpose, process, and findings of the ancillary study portion of the Customer Drive Uniform Manufacture (CDUM) program. The purpose of the ancillary study was to provide insight on the variability of shade evaluation labs used in both industry and government facilities. The ancillary study was conducted to determine if factors other than instrumentation were causing variability in readings of fabric shade parameters. Questions related to ancillary factors, including illumination, environment, sample submittal, and spectrophotometer use, were used to collect data from each lab. The data was then compiled into table format for comparison. The tables illustrated the level of variation for each ancillary factor.

Some industry and government facilities have multiple labs on site. The labs observed in this study are used for the purpose of evaluating fabric or material for lot acceptance or rejection and adherence to shade specifications.

PROCESS

The ancillary study was conducted by observing laboratory conditions, measuring lab dimensions, and interviewing lab and production personnel. Representatives from LMI and the U.S. Army Soldier Research, Development and Engineering Center at Natick, MA, conducted site visits between September 2008 and April 2009 to 14 labs. The labs visited were as follows:

- Duro
- Mount Vernon
- Crystal Springs
- Defense Supply Center, Philadelphia
- Gore
- U.S. Navy
- Brittany
- Kenyon
- Polartec

- Bondcote
- Burlington Finishing Plant
- Burlington Raeford
- Milliken
- Carlisle.

During the site visits, LMI and Army Natick Soldier RDEC representatives observed the conditions of the labs used for shade evaluation, and production facilities were toured where applicable. Each site visit started with a meeting between the LMI representative, Army Natick Soldier RDEC representative, and the site's host (usually a member of management). After introductions, the purpose of the visit was discussed. The host outlined the agenda for the visit, and then the lab and production areas were toured.

The tours usually started in the shade evaluation lab. The LMI representative asked the lab employees several questions related to the ancillary factors. The questions and factors are listed in the questionnaire shown as Figure A-1. The questionnaire was used during all site visits.

The lab employees answered the questions, and the LMI representative documented their answers. When the lab employees could not answer a question, a coworker or member of management was called in to assist. Employees of the production area were sometimes called in to assist with answering questions related to sample submittal. In some cases, answers to questions were not known by either the individuals being interviewed or any of their peers. In these cases, the answers were left blank on the questionnaire and tables.

The condition of the lab was assessed by the LMI representative without input from any of the other parties.

After gathering the necessary data, the LMI representative, the Army Natick Soldier RDEC representative, and site host toured the production area.

Figure A-1. Questions and Factors Asked to Lab Personnel

Ancillary Questions

Illumination Conditions

1) What illuminant is being used? (i.e., Daylight 75)
2) What is the brand and make of the illuminant?
3) How long since illuminant was changed?
4) When was the illuminant simulator last tested or calibrated? How often?

Viewing Environment

5) Is a shade table or lightbox being used?
6) What are the dimensions of the room? (length, width, height)
7) What are the colors of the walls and flooring?
8) If gray, is it equivalent to Munsell N5 thru N8?
9) Is the environment climate controlled? (i.e., central air, localized air, climate control)
10) Is option A, B, C, or D from AATCC evaluation procedure 9 being used?
11) What is the angle of the viewing table?
12) What is the distance of the table to the illuminant source?
13) What is the condition of the lab? (i.e., overall cleanliness)

Sample Submittal

14) How are lot samples generated for selection?
15) How are they stored/handled until selection?
16) What is the time period from sample generation to selection?
17) How are selected lot samples handled after selection?
18) How long between lot sample selection and shipment to evaluation lab?
19) How are selected lot samples sent to evaluation lab?

Other

20) How are spectrophotometers used in the manufacturing process?

FINDINGS

The three primary factors observed for the study were illumination, view environment, and sample submittal. Descriptions of each factor and the data collected are contained below. Information on spectrophotometer brand and software is also provided.

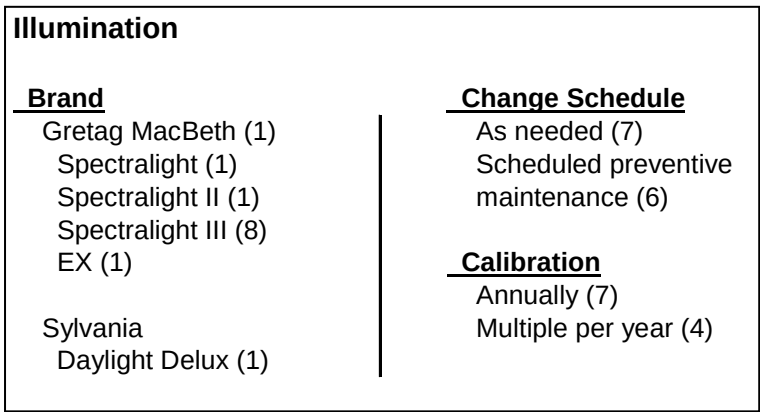
Discrete values are displayed for some factors; a range is shown for others. The distinction between discrete value and range was based upon the most effective way to display the data.

Illumination

Illumination is the light source used to illuminate the fabric while it is evaluated for adherence to specifications by the human eye. The brand, change, and calibration schedules were observed for the illuminants.

Figure A-2 displays the observations for illumination. Note that the numbers in parentheses following the descriptors indicate the number of labs for which that descriptor applies. For example, eight labs use the Spectralight III brand of illuminant.

Figure A-2. Observations of Illumination



View Environment

View environment describes the type of device being used to evaluate fabric (a shade table or lightbox), the square footage of the lab, the colors of the lab walls and flooring, the type of climate control the lab utilizes, which American Association of Textile Chemists and Colorists, or AATCC, option the lab follows when evaluating fabric, the angle of the table on which the fabric rests for evaluation, lab cleanliness, and the distance from the table's surface to the illuminant.

Figure A-3 displays the different options available for shade evaluation under AATCC Evaluation Procedure 9. Figure A-4 displays the observations for view environment.

Figure A-3. Options for Shade Evaluation under AATCC Evaluation Procedure 9

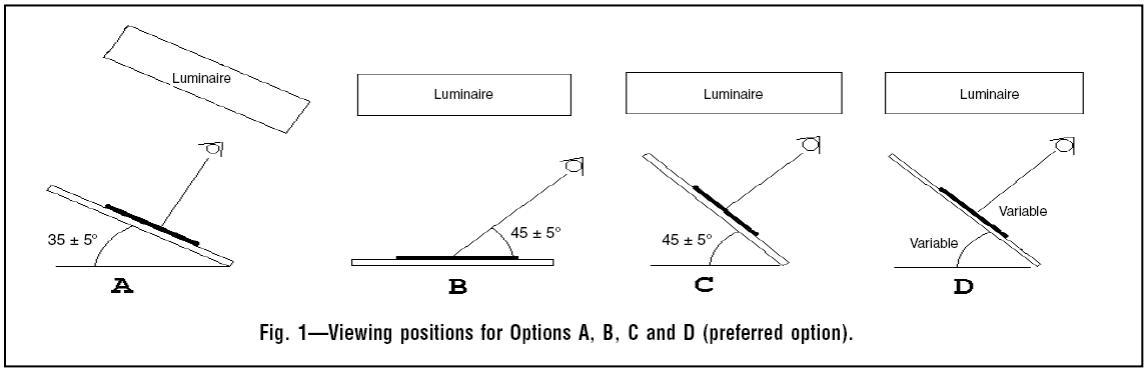


Fig. 1—Viewing positions for Options A, B, C and D (preferred option).

Figure A-4. Observations for View Environment

View Environment*

Device
Shade table (14)

Square footage (approx.)
57 to 1056

Wall/flooring
Gray (8)
Gray/white (1)
Tan/green/black (1)
Tan/gray (1)
Gray/red (1)
Gray/brown (2)

Climate Control
Central A/C (10)
Local A/C (2)

AATCC procedure 9
Option A (10)
Option B (1)
Option D (2)
Procedure similar to AATCC (1)

Table angle (approx.)
0 to 45 degrees

Distance to illuminant (approx.)
29.5" to 49"

*We observed cleanliness ranging from very clean, average to cluttered

SAMPLE SUBMITTAL

Sample submittal describes the type of packaging used to hold the fabric samples and the time between when a sample is generated and when it is selected for review by the government representative. Figure A-5 displays the observations for sample submittal.

Figure A-5. Observations for Sample Submittal

Sample Submittal

Packaging
- Cardboard box (3)
- Plastic bag (5)
- Envelope (3)
- Rubber band (1)

Sample generation to selection
- 1 day to 1 week

SPECTROPHOTOMETER

A spectrophotometer is an instrument used by lab personnel to evaluate fabric shade against a given standard. Some labs use spectrophotometers in conjunction with human evaluation to judge whether fabric meets shade specifications. The spectrophotometer brands and associated software documented observed during the ancillary study are listed in Figure A-6.

Figure A-6. Spectrophotometer Brands

Spectrophotometer

Brand	Software
Datacolor	Datacolor tools 1.1.0 (1)
	Colortools QC 3.95 (1)
	Colortools QC (1)
	Datacolor tools 1.0.2 (1)
	Datacolor tools 1.3 (1)
	Colortools QC 3.1 (2)
	Chroma QC (1)
	Colortools QC 3.0 (1)
Hunterlab	Easy Match QC 3.72.00 (2)
	Easy Match QC 3.84.00 (1)
	Color icontrol 4.11.002 (1)
	Easy Match QC 4.03.00 (1)
	Universal 4.01 (1)
	Easy Match QC 3.90.00 (1)
X Rite	X Rite Colormaster 8.0.3 (1)
	Color iQC v.6.0 (1)
Gretag MacBeth	Gretag Quality Control System Version 329 (1)

Appendix B
Data Collection Protocol

DATA COLLECTION PROTOCOL FOR CERAMIC TILES

- The ceramic tiles were clean and free of lint and fingerprints.
- Using the green tile, the spectrophotometer was calibrated to the manufacturer's specifications before measurements were taken.
- The following parameters were used for measuring the ceramic tiles:
  - Illuminant: D65
  - Area view: Large
  - Observer: 10°
  - Mode: Specular included—Reflectance
  - Ultraviolet (UV) filter: Set to OUT
  - UV lamp: Set to OFF
  - Scale: CIELAB Absolute values.
- The tiles were as follows:
  - Pale Grey
  - Mid Grey
  - Deep Grey
  - Green
  - White
  - Blue.
- Each tile was measured in the same area, with the same orientation, to collect the L*, a*, and b* values.
- The absolute L*, a*, and b* values were recorded for each tile sample.

DATA COLLECTION PROTOCOL FOR DIFFERENT SHADE FABRIC SAMPLES

We used AATCC Evaluation Procedure 6, *Instrumental Color Measurement*, Section 2, "Measurement of Color by Reflectance Methods," as a guideline for the different shade fabric samples.

- The samples were clean and free of lint, and were allowed to condition if necessary.
- Using the green tile, the spectrophotometer was properly calibrated to the manufacturer's specifications before measurements were taken.
- The following parameters were used for measuring the fabric samples:
  - Illuminant: D65
  - Area view: Large
  - Observer: 10°
  - Mode: Specular Included—Reflectance
  - UV filter: Set to OUT
  - UV lamp: Set to OFF
  - Scale: CIELAB Absolute values.
- The samples were single-layer samples with a black tile as the backing.
- The samples were as follows:
  - Coyote 498
  - Desert Sand 503
  - Army Green 491.
- Each sample was measured in the same area, with the same orientation, to collect the L*, a*, and b* values.
- The absolute L*, a*, and b* values were recorded for each fabric sample.

DATA COLLECTION PROTOCOL FOR SOLID SHADE ON DIFFERENT FABRIC SAMPLES

We used AATCC Evaluation Procedure 6, *Instrumental Color Measurement*, Section 2, "Measurement of Color by Reflectance Methods" as a guideline for the same shade on different fabrics.

- The samples were clean and free of lint, and were allowed to condition if necessary.
- Using the green tile, the spectrophotometer was properly calibrated to the manufacturer's specifications before measurements were taken.
- The following parameters were used for measuring the fabric pairs:
  - Illuminant: D65
  - Area view: Large
  - Observer: 10°
  - Mode: Specular included—Reflectance
  - UV filter: Set to OUT
  - UV lamp: Set to OFF
  - Scale: CIELAB Absolute values.
- The samples were single-layer samples with a black tile as the backing.
- The samples were as follows:
  - Blue 450 p/w serge
  - Blue 450 poly
  - Blue 450 p/w elastique
  - Blue 450 p/w tropical.
- Each sample was measured in the same area, with the same orientation, to collect the L*, a*, and b* values.
- The absolute L*, a*, and b* values for each fabric sample were recorded.

DATA COLLECTION PROTOCOL FOR FABRIC PAIRS

We used AATCC Evaluation Procedure 6, *Instrumental Color Measurement*, Section 2, "Measurement of Color by Reflectance Methods," and the AATCC Test Method 173-2005 CMC, *Calculation of Small Color Differences for Acceptability*, as guidelines for the color measurement of fabric pairs.

- The samples were clean and free of lint, and were allowed to condition if necessary.
- Using the green tile, the spectrophotometer was properly calibrated to the manufacturer's specifications before measurements were taken.
- The following parameters were used for measuring the fabric pairs:
  - Illuminant: D65
  - Area view: Large
  - Observer: 10°
  - Mode: Specular included—Reflectance
  - UV filter: Set to OUT
  - UV lamp: Set to OFF
  - Scale: CIELAB and CMC ratio 2:1.
- The samples were single-layer samples with a black tile as the backing.
- The samples were as follows:
  - Blue 3329 standard and sample
  - Blue 3372 standard and sample
  - Purple 3905 standard and sample
  - Black 472 standard and sample
  - CG 483 standard and sample
  - FG 504 standard and sample.
- Each "standard" and sample was measured in the same area, with the same orientation, to collect the de CMC,[1] ΔL*, Δa*, and Δb* values.
- The de CMC, ΔL*, Δa*, and Δb* values for each color were recorded.

[1] de CMC is a summary of the L*, a*, and b* data.

DATA COLLECTION PROTOCOL FOR CAMOUFLAGE FABRIC SAMPLES

We used AATCC Evaluation Procedure 6, *Instrumental Color Measurement*, Section 2, "Measurement of Color by Reflectance Methods," as a guideline for the color measurement of camouflage fabric samples (six Air Force ABU samples and six Army ACU samples).

- The samples were clean and free of lint, and were allowed to condition if necessary.
- Using the green tile, the spectrophotometer was properly calibrated to manufacturer's specifications before measurements were taken.
- The following parameters were used for measuring the camouflage fabric samples:
  - Illuminant: D65
  - Area view: Small
  - Observer: 10°
  - Mode: Specular included—Reflectance
  - UV filter: Set to OUT
  - UV lamp: Set to OFF
  - Scale: CIELAB Absolute values.
- The samples were single-layer samples with a black tile as the backing.
- The samples were identified as follows:
  - Army ACU A
  - Army ACU B
  - Army ACU C
  - Army ACU D
  - Army ACU E
  - Army ACU F
  - Air Force ABU A
  - Air Force ABU B
  - Air Force ABU C
  - Air Force ABU D
  - Air Force ABU E
  - Air Force ABU F.

Each sample was measured in the same area, with the same orientation, to collect the L*, a*, and b* values.

Appendix C
Abbreviations

AATCC	American Association of Textile Chemists and Colorists
ABU	airman battle uniform
ACU	Army combat uniform
ANOVA	analysis of variance
C&T	Clothing and Textiles
CDUM	Customer Driven Uniform Manufacturing
CIE	*Commission Internationale d'Eclairage* (International Commission on Illumination)
CIELAB	CIE L*a*b*
DLA	Defense Logistics Agency
DSCC	Defense Supply Center Columbus
DSCP	Defense Supply Center Philadelphia
HQ	Headquarters
IPT	Integrated Product Team
MC	Marine Corps
PTC	Product Testing Center
RDEC	Research, Development, and Engineering Center
SPSS	Statistical Package for the Social Sciences

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
03-2011	Final	

4. TITLE AND SUBTITLE
Color Shade Instrumentation Correlation Study: Statistical Analysis (Revision)

5a. CONTRACT NUMBER
BPA SP4701-07-A-0002

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Metelko, Alan R.; Author
DeMorais, Luisa; Author
King, Melanie; Author
Matuszek, Rachel; Author
Vandenberghe, Jack ; Author
Wurst, Nathaniel J.; Author

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
LMI
2000 Corporate Ridge
McLean, VA 22102-7805

8. PERFORMING ORGANIZATION REPORT NUMBER
DL827T1(Rev. 1)

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Logistics Agency
8725 John J. Kingman
Ft. Belvoir, VA 22060

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
A Approved for public release. Distribute to DTIC and defense libraries, NTIS, or other appropriate government report databases

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Military fabric suppliers currently use spectrophotometers as part of their quality control process; however, when the suppliers ship their fabric to clothing manufacturers, the procedure for determining whether the fabric meets the military shade standard depends solely on visual evaluation. Visual evaluation by trained shade evaluators is widely considered superior to instrumentation because the human eye is more sensitive to shade variation. The feedback from visual evaluations that is provided to fabric suppliers is often general and vague, and suppliers have difficulty making the proper adjustments to their printing and dying processes. Spectrophotometer measurements, on the other hand, provide more quantitative feedback, but there have been concerns regarding the consistency of measurements. LMI and Army Natick Soldier RDEC assessed the use of spectrophotometers as an aid to military shade evaluators, asking two main questions: Do spectrophotometer measurements collected at different locations pro-duce consistent color measurements? To what extent do human visual evaluations agree with spectrophotometer evaluation?

15. SUBJECT TERMS
Spectrophotometers; shade evaluation; L*A*B* measurement

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE	Unclassified Unlimited	76	Nancy E. Handy
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED			19b. TELEPHONE NUMBER *(include area code)* 703-917-7249