R03-6

Predicting the Handle and Comfort of Military Clothing Fabrics from Sensory and Instrumental Data: Development and Application of New Psychophysical Methods

ARMAND V. CARDELLO[1] AND CAROLE WINTERHALTER

U.S. Army Natick Soldier Center, Natick, Massachusetts 01760, U.S.A.

HOWARD G. SCHUTZ

Department of Consumer Sciences, University of California, Davis, California 95616, U.S.A.

ABSTRACT

Sensory and instrumental fabric characteristics that contribute to military clothing comfort are analyzed in a series of studies. A standardized hand evaluation methodology is checked for its sensitivity and reliability and used to characterize thirteen military fabrics. A labeled magnitude scale of comfort is developed using consumer magnitude estimates of the semantic meaning of verbal phrases denoting different levels of comfort/discomfort. The sensitivity and reliability of this "CALM" scale is assessed in two studies, and the scale is then used by consumers to rate the handle and comfort of the thirteen test fabrics. The descriptive sensory data and comfort data are combined with Kawabata data obtained on a subset of the fabrics, and the data are analyzed with a principal components analysis. Multiple regression analyses are performed using the component scores to predict consumer comfort from the sensory and instrumental data. The results show a high degree of predictability of comfort responses from a combination of sensory and Kawabata parameters.

The United States Department of Defense (DoD) procures over 1.1 billion dollars of clothing and individual equipment each year. A large portion of these expenditures goes toward the purchase of battle dress uniforms (BDU), the two-piece, camouflage uniforms worn by troops in combat, training, and garrison situations. While the comfort of these garments has been a major consideration in their design and development, much of the research to date has focused on the *thermal* comfort of the garments, because thermal stress is a major factor contributing to human performance degradation. More recently, focus has turned toward the less studied area of *tactile* comfort. This refocusing has been precipitated both by the knowledge that the BDU is worn on a daily basis in garrison situations, where heat stress is less of an issue than in combat, and the fact that procurement policy changes have moved DoD away from specifications of fabric composition and toward specifications based on functional or performance characteristics, *e.g.*, durability and comfort criteria. In order to better understand and quantify the tactile comfort of military clothing and to determine predictive relationships between fabric properties, sensory perceptions, and consumer comfort, a research program has been undertaken to identify and define the critical factors contributing to the tactile comfort of military fabrics.

[redacted]

Scientific studies of perceptual and affective responses to clothing originated in the early years of the past century, when investigators such as Binns [3], Pierce [33], Houghton and Yaglou [22], Winslow [54,55], and others began systematic analyses of subjective responses to textiles and clothing. From these early efforts evolved the conceptual bases for the study of fabric "handle" and the analysis of the determinants of sensory, thermal, and overall clothing comfort. While the next fifty years produced an expanding literature on these topics, studies of human responses to clothing materials suffered from a lack of theoretical models to guide research in the field. As a result, the field was plagued by idiosyncratic and undefined terminology, a lack of operational constructs, confusion over the kinds of panelists to use, failure to adopt modern psychophysical techniques, and general confusion in communication about fabric attributes and qualities [5, 10, 53, 56].

Beginning with the work of Fourt and Hollies [13], a better conceptualization of clothing comfort began to emerge, which focused on three important components of clothing comfort: the person, the clothing, and the environment. Subsequent theoretical work by Slater [42, 43], Rohles [38, 39], Pontrelli [37], and Sontag [44] drew finer conceptual and empirical distinctions between the physical factors of both the garment and the environment, the physiological and sensory responses of the individual, psychological "filters" by which these latter responses are modified prior to conscious awareness, and the final affective response that we call comfort (see Branson and Sweeney [7] for a detailed review of these theoretical developments). Within the context of this evolving theoretical framework, it became possible to better isolate the variables contributing to clothing comfort and to begin to refine techniques for measuring both these antecedent variables and the primary dependent variable of clothing comfort itself.

There are two fundamental psychological dimensions that comprise all sensations produced by contact of clothing fabrics with the skin. The first is *qualitative* (descriptive) and relates to the specific sensory quality or attribute that is being perceived, *e.g.*, roughness, stiffness, etc. The second is *quantitative* (intensive) in nature and relates to the perceived magnitude of that sensation, *e.g.*, *very* rough, *slightly* stiff, etc. Both dimensions of experience are involved in the perception of fabrics on the skin, and the psychophysical methodologies used to identify and define these dimensions are critical factors determining the validity of the data and the conclusions that can be drawn from them.

Recent Developments in Sensory Hand Analysis

Civille and Dus [10] reviewed the published studies on developing sensory hand attributes, terminology, and systems. Confirming earlier analyses [5, 51, 53], they concluded that there were significant deficiencies in the existing methods in terms of the development of primary (discrete and independent) tactile characteristics, the operationalization of terminology and evaluation procedures, proper scaling methodology, subject/panelist training, and established test protocols and controls. In response to this lack of standardization, Civille and Dus [10] developed the Handfeel Spectrum Descriptive Analysis (HSDA) method as a more analytical, comprehensive, controlled approach to sensory analysis of woven and nonwoven fabrics. This method is modeled after similar, highly successful, descriptive methods used for sensory analysis of consumer products, *e.g.*, foods, perfumes, and skin care products [6, 41]. The attribute terms and protocols for the HSDA method have been reviewed and refined by the Other Senses Task Group (E18.02.06.03) of ASTM Committee E-18, and this method for the descriptive analysis of textiles has been reported previously [36, 37].

The HSDA method significantly enhanced the ability to define and study the qualitative aspects of sensory handle by establishing operationally defined attribute terminology that is free of affective (good/bad) associations. Furthermore, by avoiding idiosyncratic terminology and the unnatural separation of the visual component of handle [5, 9, 28], the method minimizes differences between trained panelist ratings and consumer perceptions, significantly improving the likelihood of developing predictive relationships with consumer comfort. Since a major goal of the HSDA methodology is interlaboratory standardization, the psychophysical scaling method uses physical fabric standards as reference points along a fifteen-point intensity scale for each hand attribute. Thus, the intensity scale for fabric "stiffness" is anchored at the upper end by a cotton organdy standard with a stiffness rating of 14.0, and at the lower end by a 50/50% polyester/cotton single knit fabric with a stiffness rating of 1.3. Other fabrics define intermediate points on the continuum. Similar sets of fabrics define the intensity scales for other attributes [9, 25]. Such stimulus-referenced or "learned" rating scales are widely used in commercial sensory evaluation and are particularly effective in helping to conceptualize and define the stimulus dimension of interest. In addition, they reduce intersubject variability [52] and can be easily transferred from one subject group to another, thereby ensuring high interlaboratory reliability.

Psychophysical Scaling of Comfort

Although a valid and reliable system for quantifying the descriptive hand attributes of fabrics is a logical prerequisite for identifying the fabric attributes that contribute to clothing comfort, no less important is a reliable and valid measure of comfort itself. Unlike tactile attributes, comfort is not a sensory dimension, because it is not associated directly with any single human sense organ. Rather, it is an evaluative or affective dimension, analogous to liking. Thus, there is no underlying physical dimension of the stimulus that varies continuously and is monotonic with the perception of comfort. The same stimulus can elicit quite different comfort responses from different individuals. As a result, it is not possible to define a comfort scale based on physical standards that is valid for all users. In addition, since comfort is an affective dimension, it is appropriately judged only by untrained consumers. This requires a method for scaling comfort that is simple and unencumbered by the necessity for training or complex instructions.

The most common kind of subjective scale for rating comfort is a "category scale," which is characterized by a series of verbally and/or numerically labeled points or categories. Individuals rate their subjective sensations by placing them into one of several descriptive categories. Since less than five categories can result in a loss of discrimination sensitivity, the number of categories is typically around seven to nine [8], but can be greater [31]. Several of the best known category scales for evaluating clothing sensations and/or comfort are Hollies' subjective comfort rating chart [20, 21], which uses both a category scale of intensity (partially, mildly, definitely, totally) and the thirteen-point McGinnis category scale of comfort, and Gagge *et al.*'s, [14] scale of comfort sensation ("comfortable, slightly uncomfortable, uncomfortable, very uncomfortable"). The reasons for the widespread use of category scales to measure subjective comfort and other psychological dimensions is their simplicity, versatility, ease of use by subjects, and high reliability.

In spite of these advantages, there are significant problems associated with the use of category scales. Although it is often assumed that the points on a numbered category scale represent equal subjective intervals, this is not the case [49]. On labeled category scales, subjects attend primarily to the word labels and not to the numbers [15]. In these cases, unless the verbal labels are chosen on the basis of extensive testing to verify that such differences as those between "slightly comfortable" and "moderately comfortable" are the same as those between "moderately comfortable" and "extremely comfortable," then the scale cannot be considered to be an interval scale, but merely an ordinal scale. This has implications for the kind of statistics to be applied to the data (*e.g.*, nonparametric rather than parametric). In addition, both the range and frequency of stimuli to be evaluated can significantly influence category scale ratings [32, 35].

Another common problem with category scales is that subjects tend not to use the end categories [18, 49]. This "category end effect" results in seven-point category scales being functionally reduced to five-point scales, five-point scales to three-point scales, etc. A further complication occurs in those cases where the category scale is bi-directional and uses a "neutral" or null category. Such categories have been shown to encourage subjects to be noncommittal in their responses, *i.e.*, they overly rely on this "safe" category [17, 23].

An alternative scaling approach that avoids these problems, while providing ratio (rather than ordinal or interval) level data, was proposed by S.S. Stevens [47]. Stevens developed a scaling method in which subjects assign their own internal numbers to represent the magnitude of their sensations. He named the method "magnitude estimation" [47, 48], and it avoids the major problems of category scaling by providing an unbounded upper limit for ratings. In addition, because magnitude estimation uses a true zero point of sensation and because all judgments are made relative to one another in a ratio manner (*e.g.*, stimulus X is three times (one-half, etc.) as stiff (comfortable, etc.) as stimulus Y), the resultant data provide a ratio scale of the subjective dimension being evaluated, allowing for valid parametric analyses of the data.

In several studies examining the human sensory and comfort responses to clothing and textiles, magnitude estimation has been successfully used as a ratio scale measure of tactile responses [1, 11, 12, 50]. Although this technique significantly increases the ability to accurately quantify subjective sensations, magnitude estimation requires that sensations be directly compared to one another, thereby precluding judgments that must be made over extended time periods. In addition, magnitude estimation requires detailed instruction for proper use and time-consuming normalization of the data prior to statistical analysis. More recently, these practical limitations were eliminated with the development of *labeled* magnitude scales (sometimes called "semantic" or "category" ratio scales). These scales take the form of a visual analogue or "line" scale, but they are anchored with verbal labels that define a ratio scale of sensory magnitude. This stands in contrast to unlabeled visual analogue scales, *e.g.*, [29], which rely on the instructional set to create the ratio aspects of the scale. The first published labeled magnitude scale was the "Borg" scale of perceived exertion [4]. However, similar labeled magnitude scales have been developed recently for both sensory [16] and affective [40] continua.

Study Objectives

The recent developments in psychophysical methodology that enable better quantification of both the descriptive aspects of hand sensations and the scaling of the affective dimension of handle open the possibility of a more well-grounded psychophysical approach to studying the sensory and comfort characteristics of clothing fabrics. Combining these new sensory methodologies with established instrumental measures of fabric characterization, *e.g.*, the Kawabata system [24–26] now makes it possible to develop better predictive relationships between sensory, instrumental, and comfort measures of fabrics.

With this in mind, we have initiated a multiphase research program to develop a sensitive, reliable, standardized method for assessing the sensory tactile characteristics of military fabrics, to develop a sensitive and reliable labeled affective magnitude scale for rating fabric/clothing comfort, to use these methods to characterize a variety of military fabrics, and to develop predictive relationships between the tactile attributes of the fabrics, their instrumental properties, and their perceived comfort.

Experiment 1: Establishing a Sensitive and Reliable Sensory Hand Method

EXPERIMENTAL

Fifteen panelists (ten females, five males) were selected from volunteer employees at Natick, chosen on the basis of interest, availability, and successful completion of a screening test to establish minimum tactile acuity [10]. The latter was necessary because tactile acuity/sensitivity has been shown to vary significantly as a function of age [46], degree of skin hydration/wettedness [19], dermatitis, and other factors.

Panelists participated in a six-month training program that consisted of training in the basic methodology and operational (manual) evaluation techniques employed in the Handfeel Spectrum Descriptive Analysis method [10], repeated practice with the attribute definitions and fabric intensity scales for each of seventeen different sensory hand attributes (four related to surface geometry, ten to mechanical properties, and two to sound properties), and tailoring the seventeen-attribute definition, operational techniques, and physical reference standards to the specific military clothing fabrics to be used in testing. Table I lists the seventeen sensory attributes employed in testing. We will provide the operational techniques and physical standards of intensity upon request.

In order to assess the reliability and sensitivity of the HSDA method, we conducted a test-retest reliability study at the completion of training. We selected three fabrics for evaluation—a jersey knit, a polyester/wool serge (MIL-C-823), and a Tencel® ripstop poplin—to represent a range of tactile attributes that might be encountered in testing and to include both similar and dissimilar fabrics. The three test fabrics were evaluated by the hand panel on two different occasions, separated by a two-week interval. In addition, two of the test fabrics (Tencel ripstop poplin and polyester/wool serge) were tested again six months later to assess long-term reliability. All testing was conducted in a textile conditioning room at a temperature of 70 $\pm$1.4 F and at 65% $\pm$1.3 RH at large open tables with smooth, black, stone-top surfaces. Panelists evaluated test samples on their "face" (labeled) surface, independently, and in random order. All fabrics were laundered five times to remove nondurable sewing lubricants or softeners that could influence the tactile characteristics. The laundering was in accordance with American Association of Textile Chemists and Colorists (AATCC) test method #96, test condition IIIc, tumble dry (option A). After laundering, the fabrics were cut into 30 × 30 cm swatches, with edges parallel to the fabric warp and filling directions. All edges were serrated to prevent raveling.

RESULTS AND DISCUSSION

Figure 1 shows the average panel data for all three fabrics. Looking at the fabric profiles, we observe significant differences in the attribute profiles between fabrics (Figure 1d), but a high degree of similarity in the

TABLE I. Definitions of fabric hand attributes used for descriptive analysis.

Attribute	Definition
Grainy	amount of small, round particles in the surface of the sample
Gritty	amount of small, abrasive, picky particles in the surface of the sample
Fuzziness	amount of pile, fiber, fuzz on the surface of the sample
Thickness	perceived distance between the thumb and index finger (when the sample is placed between the two)
Tensile stretch	degree to which the sample stretches from its original shape
Hand friction	force required to move the palm of the hand across the surface of the sample
Fabric-fabric friction	force required to move the fabric over itself
Depression depth	amount that the sample depresses when downward force is applied
Springiness	rate at which the sample returns to its original position after the downward force is released
Force to gather	amount of force required to compress the gathered sample into the palm
Stiffness	degree to which the sample feels pointed, ridged, and cracked; not pliable
Force to compress	amount of force required to compress the gathered sample into the palm
Fullness/volume	amount of material felt in the hand
Compression resilience intensity	perceived force with which the sample exerts resistive pressure against the cupped hands
Compression resilience rate	rate at which the sample returns to its original shape or rate at which the sample opens after compression
Noise intensity	loudness of the noise
Noise pitch	pitch (frequency) of the noise

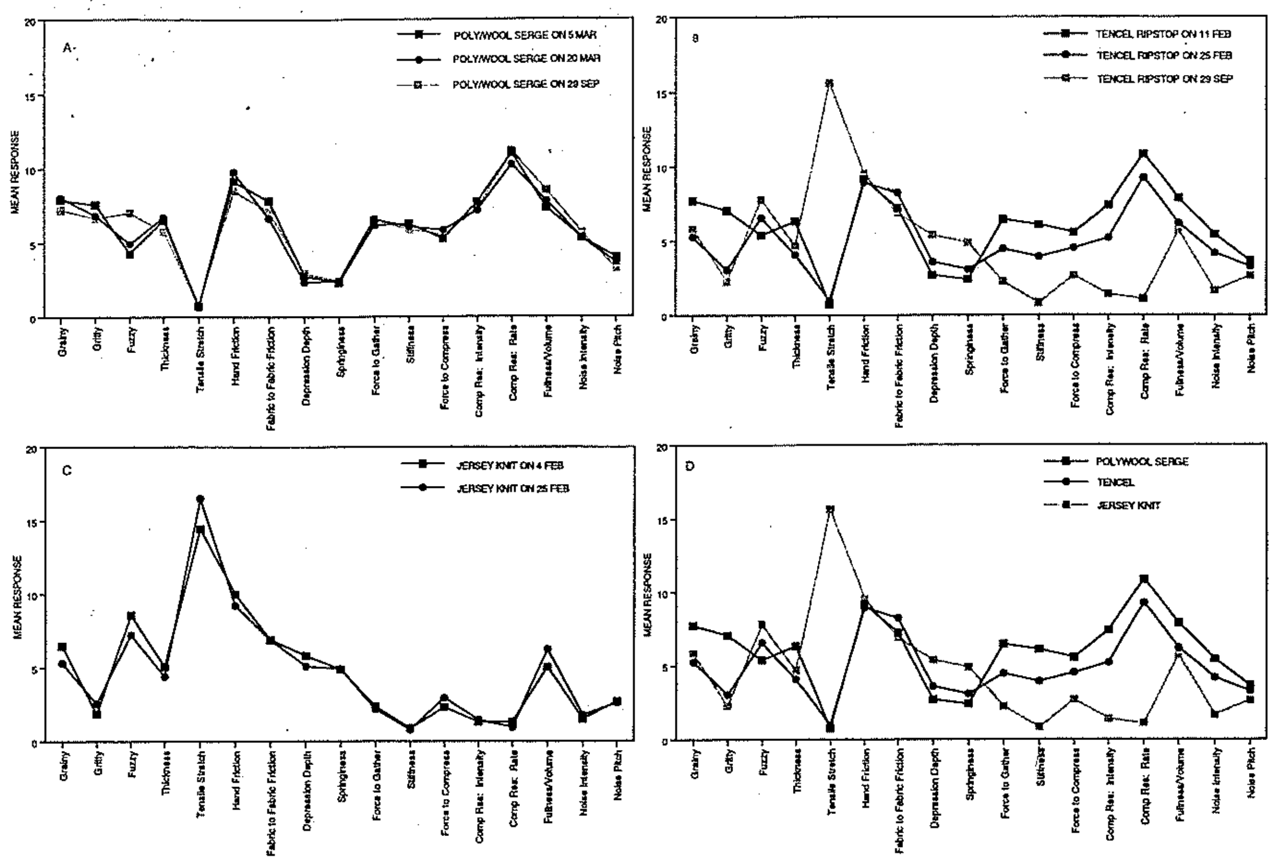

FIGURE 1. Mean panel ratings of hand attributes obtained during different test sessions for poly/wool serge (a), Tencel ripstop poplin (b), and a jersey knit (c) fabric. Panel d shows the ratings for each fabric averaged over all test sessions.

profiles for the same fabrics obtained on different testing dates (Figures 1a, b, c). For example, Figure 1d shows that the poly/wool serge fabric differed greatly from the Tencel ripstop on such attributes as "grainy," "gritty," "thickness," "force to gather," "stiffness," and "intensity of compressive resistance." There were even larger differences when comparing the jersey knit to the other two fabrics, yet ratings of attributes for each fabric evaluated on multiple occasions (Figures 1a, b, c) were very similar. Pearson product-moment correlations were calculated across mean attribute ratings for each fabric rated on the different test days. The correlation coefficients between fabrics tested two weeks apart were 0.98 (poly/wool serge), 0.93 (Tencel ripstop), and 0.98 (jersey knit). Correlations of panel ratings for the same fabrics by attribute ranged from 0.93–0.98, depending upon the attribute examined.

For the two fabrics tested again six months later, the correlation coefficients between each of the first two sessions and the third were 0.94 and 0.95 (poly/wool serge) and 0.89 and 0.93 (Tencel ripstop), indicating only a minor drop in test-retest reliability over the six-month period.

From these data we concluded that the HSDA methods, in conjunction with the panel training program, result in a sensory hand evaluation method that is highly sensitive and reliable over extended periods of time.

Experiment 2: Descriptive Analysis of Military Fabrics

EXPERIMENTAL

In order to quantify the sensory hand attributes for a wide range of military fabrics, thirteen fabrics used in U.S., British, Canadian, and Australian military garments (Table II) were evaluated by the sensory hand panel. These fabrics were chosen to represent a wide range of tactile (and likely comfort) characteristics to be found in U.S. and foreign military uniforms. Of these thirteen fabrics, eight (see asterisked fabrics in Table II) were down-selected for subsequent evaluation of their mechanical properties using the Kawabata (KES-F) system of

TABLE II. Military fabrics used in experiments 2 and 6.

Test fabric	Sample code
50%/50% Nylon/combed cotton, ripstop poplin weave[a]	10R
50%/50% Nylon/polyester, oxford weave (Australian)[a]	11A
50%/50% Nylon/cotton, twill weave[a]	12T
92%/5%/3% Nomex, Kevlar, P140, plain weave[a]	13P
100% Cotton, twill weave (former flame retardant treated)[a]	14N
77%/33% Cotton sheath/synthetic core, twill (U.K.)[a]	15B
100% Combed cotton, ripstop poplin (former hot weather BDU)[a]	16C
65%/35% Wool/polyester, plain weave (Canada-unlaundered)[a]	17C
65%/35% Wool/polyester, plain weave (Canada-laundered)	18L
92%/5%/3% Nomex, Kevlar, P140, oxford weave	19N
Carded cotton sheath/nylon core, plain weave (Canada)	20J
100% Pima cotton ripstop poplin (experimental)	124
50%/50% Nylon carded cotton ripstop poplin weave	176

[a] Fabrics for which Kawabata data were also obtained (experiment 7).

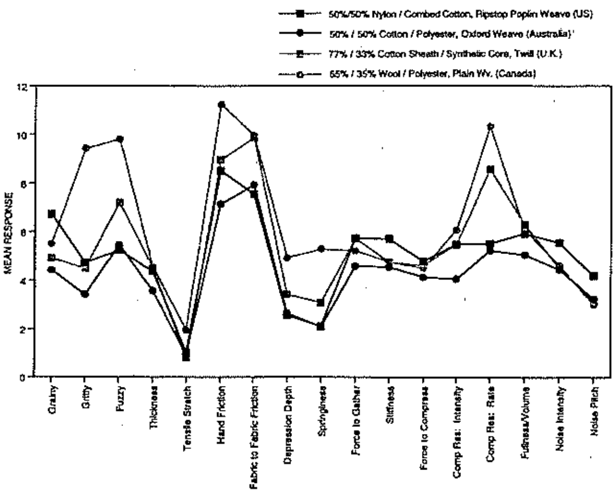

FIGURE 3. Mean panel ratings of hand attributes averaged over three replicates for four additional fabrics tested in experiment 2.

fabric testing. At most, four fabrics were evaluated during any panel session. Each fabric evaluation was replicated three times, using the same testing procedures and test conditions described in Experiment 1.

RESULTS AND DISCUSSION

Figures 2 and 3 show the sensory hand profiles for the eight fabrics down-selected for further testing. Figure 2 shows four of the eight fabrics. One of the fabrics is currently used in the U.S. Army Aircrew BDU (black circles), one is used in the Temperate BDU (black squares), while the other two fabrics are materials recently used in U.S. Navy coveralls (gray squares) and in the U.S. Army Hot Weather BDU (gray circles).

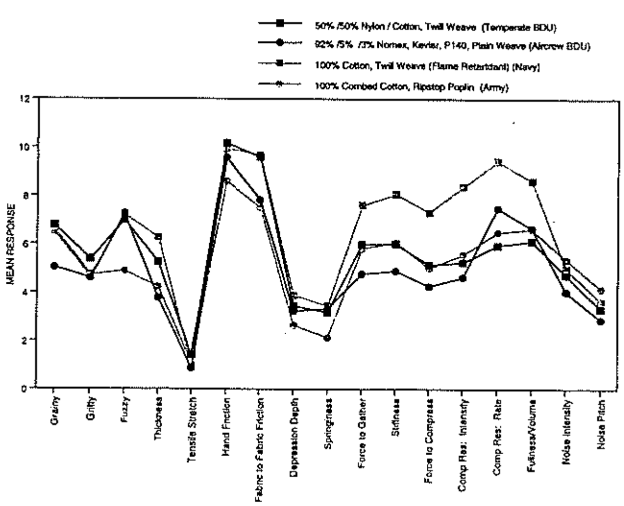

FIGURE 2. Mean panel ratings of hand attributes averaged over three replicates for four of the eight fabrics tested in experiment 2.

As these figures show, the sensory differences between the Army Aircrew and the Temperate BDU fabrics (black circles/squares) are relatively small. The Army Hot Weather BDU fabric (gray circles) is somewhat similar, but differs greatly from the former two in "fuzziness" and tends to be lower on several other attributes, *e.g.*, "hand friction," "depression depth," and "springiness". On the other hand, the Navy fabric (gray squares) is quite different in its hand characteristics. In particular, it is "thicker," has greater "force to gather," "stiffness," "compressive resilience," and "fullness/volume" than any of the other fabrics. The Army flame-resistant fabric exhibits some similar sensory properties, *e.g.*, in terms of "fuzziness," "tensile stretch," "hand friction," "depression depth," and "springiness," but is a thinner, much smoother (less grainy) fabric, and has lower "force to gather," "stiffness," and "compressive resistance" characteristics than the Navy material.

Figure 3 shows the other four fabrics, including the three non-U.S. fabrics. Again, there are large differences in the hand profiles for the fabrics. Table III shows the results of ANOVAS for each hand attribute for the fabrics shown in Figures 2 and 3, along with the number of statistically significant subsets of samples (based on the Newman-Keuls test of differences between means). As we see from the highly significant F values, all of the seventeen hand attributes discriminated between the test fabrics. Several of the attributes, such as "hand friction," "force to compress," and both the "intensity and rate of compression resilience" significantly differentiated the fabrics into as many as five distinct subsets. Several other

TABLE III. F values and the number of statistically significant subsets discriminated by each hand attribute (from Newman-Kuels post-hoc tests) for the eight fabrics tested in experiment 2 and asterisked in Table II.

Attributes	F value[a]	Number of significant subsets, $p < 0.05$
Grainy	11.47	3
Gritty	61.20	3
Fuzziness	42.51	3
Thickness	36.65	4
Tensile stretch	8.47	3
Hand friction	18.44	5
Fabric to fabric friction	18.47	2
Depression depth	27.91	4
Springiness	22.27	3
Force to gather	38.09	4
Stiffness	45.76	3
Force to compress	39.72	5
Compression resil: int.	50.14	5
Compression resil: rate	33.66	5
Fullness/volume	19.53	3
Noise intensity	8.28	3
Noise pitch	6.58	2

[a] All F-values are significant at $p < 0.01$.

attributes differentiated three or four subsets. Of the three attributes with somewhat lower F values, "tensile stretch," "noise intensity," and "noise pitch," an examination of Figures 2 and 3 reveals that few of the fabrics showed any tensile stretch. In contrast, the intensities of the sound attributes for these fabrics, although low, were as high or higher than other attributes that showed better discrimination between the fabrics, *e.g.*, depression depth.

Pearson product-moment correlation coefficients were calculated for all possible pairs of the seventeen hand attributes (136 coefficients). Of these 136 pairings, twelve had coefficients greater than 0.90. An examination of these highly associated attributes revealed several distinct and logical groupings. For example, there was a highly significant association between the attributes "springy," "fuzzy," and "depression depth" (all r values > 0.96, $p < 0.05$), an association that is logically consistent with a fuzzy surface texture giving way to slight finger pressure and then springing back after the pressure is removed. Similarly, "force to gather," "force to compress," and "compression resilience intensity" were all highly correlated (r values > 0.91, $p < 0.05$) and logically consistent with the operational technique of gathering the fabric in the hand, compressing it, and perceiving the resistance to compression. The third grouping of highly associated attributes that emerged ("force to gather," "force to compress," "thickness," and "stiffness") is logically consistent with the fact that thicker and stiffer fabrics require greater force to gather and to compress in the hand. Last, there was a high correlation between "noise intensity" and "noise pitch" ($r = 0.96$, $p < 0.05$), which is consistent with the physics of sound production, since more abrasive surface textures produce louder sounds with higher frequency when rubbed together. Only one association between attributes—"gritty" and "tensile stretch"—was highly significant ($r = 0.96$, $p > 0.05$) but defied a logical explanation in terms of the definitions and techniques involved in the evaluation process. The large differences between fabrics in Figures 2 and 3, combined with the demonstrated sensitivity (Table III) and reliability (Figures 1a, b, c) of the HSDA methodology, establish a strong empirical basis upon which to examine both the comfort of these fabrics and their mechanical parameters, so that the relationships between hand attributes, comfort, and instrumental properties can be determined.

Experiment 3: Developing a Labeled Magnitude Scale for Measuring Comfort

EXPERIMENTAL

In order to develop a sensitive, reliable, and valid labeled magnitude scale of comfort, thirty-five Natick employees, none of whom were members of the descriptive hand panel, were recruited from a random list of volunteers. Word adjectives that could be used to modify the terms "comfortable" and "uncomfortable" to reflect intensity differences were compiled from previous scaling literature and from standard English language resources. The adjectives "greatest imaginable" and "greatest possible" were included to define scale values commensurate with a common fixed end-point of positive and negative affective experience, as used in previously developed labeled magnitude scales [4, 16, 40]. These adjectives were used to create forty-one word phrases, which in combination with two nonpolar terms ("neutral" and "neither comfortable nor uncomfortable"), resulted in a total of forty-three phrases to be used in scale development. These phrases appear in the left-hand column of Table IV.

The forty-three phrases were printed on separate pages and assembled in random order into testing booklets. Before testing, subjects were provided with written instructions on the procedure to be used in scaling the semantic meaning of the phrases. Oral instructions with an example were also provided. Subjects sequentially rated each of the phrases to index the magnitude of comfort or discomfort connoted by the phrase, using a modulus-free magnitude estimation procedure. In this procedure, subjects assign an arbitrary number to indicate the magnitude of comfort or discomfort reflected by the first phrase (positive numbers used for comfort, neg-

TABLE IV. Word phrases, geometric mean magnitude estimates (n = 35), standard errors, and standard errors divided by the geometric mean for the data obtained in experiment 3.

Comfort/discomfort word phrases	Geom. mean mag. est.	Standard error	Standard error/G.M.
Greatest imaginable comfort	366.72	34.88	0.10
Greatest possible comfort	345.28	28.76	0.08
Exceptionally comfortable	280.20	16.03	0.06
Superior comfort	279.71	19.27	0.07
Intensely comfortable	268.44	19.82	0.07
Extremely comfortable	260.75	23.51	0.09
Highly comfortable	224.01	15.80	0.07
Very comfortable	203.99	13.96	0.07
Terribly comfortable	135.93	48.72	0.36
Moderately comfortable	130.18	10.51	0.08
Comfortable[a]	109.22	10.81	0.10
Satisfactory comfort	86.11	11.68	0.14
Fairly comfortable	85.16	8.62	0.10
Average comfort	77.58	17.30	0.22
Acceptable comfort	72.17	8.85	0.12
Somewhat comfortable	59.98	9.07	0.15
Slightly comfortable	38.26	9.96	0.06
A little comfortable	28.77	7.82	0.27
Mediocre comfort	22.63	9.60	0.42
Barely comfortable	15.42	4.77	0.31
Neutral	0	0	N.A.
Neither comfortable nor uncomfortable	0	0	N.A.
Barely uncomfortable	−27.61	4.38	0.16
A little uncomfortable	−40.90	5.05	0.12
Slightly uncomfortable[a]	−52.95	5.73	0.11
Somewhat uncomfortable	−71.56	6.74	0.09
Average discomfort	−76.64	13.55	0.18
Mediocre discomfort	−79.56	10.96	0.14
Uncomfortable[a]	−96.34	8.21	0.09
Fairly uncomfortable	−99.38	10.07	0.10
Moderately uncomfortable	−145.63	7.23	0.05
Very uncomfortable[a]	−209.86	11.00	0.05
Awfully uncomfortable	−228.96	10.71	0.05
Highly uncomfortable	−231.80	11.42	0.05
Terribly uncomfortable	−257.78	14.51	0.06
Exceptionally uncomfortable	−272.76	12.41	0.05
Intensely uncomfortable	−274.34	18.28	0.07
Oppressively uncomfortable	−279.70	15.71	0.06
Horribly uncomfortable	−283.88	22.86	0.08
Extremely uncomfortable	−290.84	15.57	0.05
Unbearably uncomfortable	−298.44	21.79	0.07
Greatest possible discomfort	−345.82	24.29	0.07
Greatest imaginable discomfort	−350.67	35.85	0.10

[a] Word phrases used in Gagge *et al.*'s comfort sensation scale[14].

ative numbers for discomfort). Subjects then make all subsequent judgments relative to the first, so that if the second phrase denotes twice as much comfort as the first, a number twice as large is assigned; if it denotes one-third as much comfort, a number one-third as large as the first is assigned, etc. All ratings were made in spaces provided in the testing booklet.

RESULTS AND DISCUSSION

The geometric means and standard errors of the assigned magnitude estimates were calculated for each of the comfort/discomfort phrases after applying an equalization procedure [27]. These data are shown in Table IV. Geometric means were used because magnitude estimates have been shown to be log-normally distributed [45]. As we can see, the geometric mean magnitude estimates ranged from −351 for "greatest imaginable discomfort" to +367 for "greatest imaginable comfort," with the other phrases distributed between these two extremes. (The phrases "neutral" and "neither comfortable nor uncomfortable" were assigned zero ratings by all subjects.)

Examination of the data in Table IV reveals the geometric mean ratings to have construct validity, because the rank order of mean values corresponds to the generally understood and accepted semantic meaning of the phrases. Also, in keeping with previous findings on the non-equivalence of intervals between the labeled points on category scales, the data in Table IV clearly demonstrate that the phrases used in Gagge *et al.*'s [30] comfort sensation scale (asterisked in Table IV) are *not* perceptually equivalent. For example, while the interval between the phrases "uncomfortable" and "very uncomfortable" is 113 units, the interval between the phrases "uncomfortable" and "slightly uncomfortable" is only forty-three units. The data also reveal a slight asymmetry between the ratings of comfort and discomfort. Examining common adjective phrases above and below the "neutral" and "neither comfortable nor uncomfortable" categories in Table IV reveals that discomfort initially grows more quickly than comfort, *i.e.*, "barely comfortable" = 15.42, "barely uncomfortable" = −27.61, "a little comfortable" = 28.77, "a little uncomfortable" = −40.90, "somewhat comfortable" = 59.98, and "somewhat uncomfortable" = −71.56. With some exceptions, this difference can be observed throughout the scale. It is only at the highest levels of comfort/discomfort, *i.e.*, "greatest possible" and "greatest imaginable," that comfort ratings achieve the same levels of magnitude as ratings of discomfort.

Based on the data in Table IV, we chose a subset of phrases to construct a labeled magnitude scale of comfort. The criteria for selecting terms were low variability in perceived semantic meaning, parallelism in the terms used to describe comfort and discomfort, and selection of an equal number of comfortable and uncomfortable phrases (a decision based on evidence from the preference scaling literature showing that balanced scales are better for differentiating products).

Examination of the standard errors of the geometric means for each of the phrases (Table IV) led to the elimination of several phrases (*e.g.*, "mediocre comfort," "barely comfortable," "a little comfortable") due to their variable semantic meaning to the subjects. Other phrases were eliminated because of a lack of suitable parallelism in terminology for the purpose of establishing bipolarity (*e.g.*, "superior comfort," "oppressively uncomfortable"). Applying the remaining criterion to the phrases resulted in the selection of eleven phrases for use in the scale: five associated with comfort, five associated with discomfort, and one neutral term ("neither comfortable nor uncomfortable") to define the zero point. The geometric mean magnitude estimates of the positive and negative phrases were transformed to range from 0 to +100 (positive phrases) and 0 to −100 (negative phrases). The phrases were then placed along a 100-mm vertical analogue line scale in accordance with their transformed values. The resulting labeled affective magnitude scale of comfort is shown in Figure 4.

The comfort affect labeled magnitude (CALM) scale shown in Figure 4 has several advantages over other comfort scales commonly used in the literature. With this scale, the level of comfort or discomfort experienced by an individual can be readily indexed by simply placing a mark somewhere on the line. This stands in contrast to the difficulty often encountered by subjects using magnitude estimation procedures. However, by having positioned the phrases of comfort/discomfort along the analogue line scale at points representing the magnitude of their semantic meaning as determined by a magnitude estimation procedure, it becomes possible to treat the measured distances along the scale as ratio level data. This stands in contrast to category scales of comfort, which provide only ordinal data. The ratio nature of the CALM scale enables statements to be made about whether

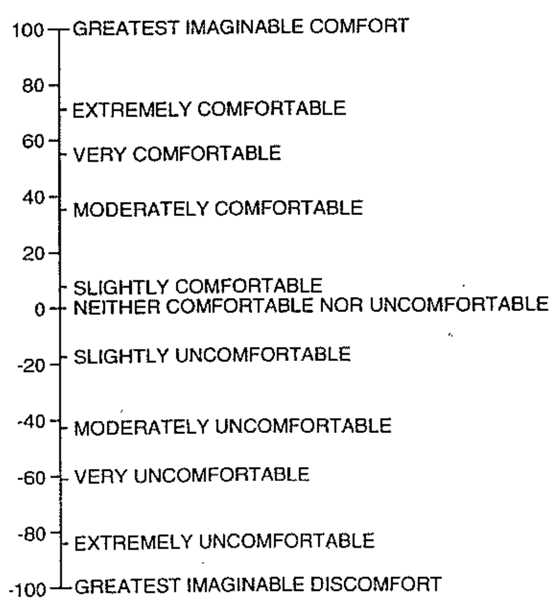

FIGURE 4. The CALM (comfort affective labeled magnitude) scale.

a particular sample is 20%, 40%, three times, etc., as comfortable (or uncomfortable) as another sample. In addition, it does not require that the data be normalized, as is the case with magnitude estimates. Last, by using the "greatest imaginable" comfort (or discomfort) as end-points on the scale, the scale enables better discrimination between samples/conditions that are either very high or very low in comfort/discomfort and establishes a common ruler by which comfort/discomfort ratings of different subjects can be compared.

Experiment 4: Reliability and Sensitivity of the CALM Scale (Image-Based Stimuli)

EXPERIMENTAL

In order to evaluate the reliability, validity, and sensitivity of the newly developed CALM scale, we conducted a study in which subjects used the scale to index the comfort/discomfort associated with several image-based clothing and environmental stimuli. The use of image-based stimuli in psychophysical scaling has been shown to produce data patterns similar to those of actual stimuli [2] and is a convenient approach for testing scale properties of validity, sensitivity, and reliability.

Twenty-seven Natick volunteer employees served as subjects, all drawn from the same general subject pool as used in experiment 3. In order to establish a clear and unambiguous set of comfort levels for testing the sensitivity of the scale, written comfort scenarios were developed describing a wide range of clothing and environmental conditions, using clothing type, ambient temperature, humidity, wind speed, and the activity of the subject as test variables. Each scenario described a particular fabric type (for a shirt or blouse) and a set of environmental/activity conditions in which the garment would be worn. The purpose of the scenarios was to create realistic, image-based stimuli that would be associated with discrete and distinct levels of perceived comfort/discomfort for all subjects. A valid comfort scale should discriminate between the levels of comfort/discomfort represented by the image-based stimuli and should be reliable from one judgment time to the next.

Subjects were tested in individual consumer testing booths. Each subject was given a self-administered questionnaire that included written instructions and a set of eight stimulus/response sheets (in random order) with the six scenarios (plus two repeated scenarios to obtain a measure of reliability) printed on them. Subjects were asked to rate the comfort or discomfort associated with each scenario by placing a mark somewhere on the labeled magnitude (CALM) scale (Figure 4). However, since previous research has shown that the *numerical* labels commonly accompanying the verbal labels on labeled magnitude scales do not affect ratings (subjects attend to the verbal labels and extrapolate between them) [16, 40], the scale points were rescaled to range from 0 ("greatest imaginable discomfort") to +100 ("greatest imaginable comfort"). This ensured that subjects would not be unduly influenced to assign negative ratings to negatively valenced scenarios and positive ratings to positively valenced scenarios without due consideration to the comfort/discomfort levels evoked by the scenarios and the semantic differences reflected in the verbal scale labels.

RESULTS AND DISCUSSION

Data were analyzed by measuring the distances of the marks from the zero point along the rating scale. Frequency distributions for each scenario, analysis of variance (with Newman-Kuels post-hoc tests) across scenarios, and correlation coefficients between the replicated scenarios were computed.

Examining the frequency distributions for the different scenarios revealed no unusual or unexpected distribution of values for any of the stimuli. The mean comfort ratings of subjects differed significantly across scenarios ($F = 83.77$, $df = 7{,}175$, $p < 0.001$), ranging from 15.7 or "very uncomfortable" (wearing denim at 100°F/60%RH) to 75.5 or "moderately comfortable" (wearing cotton at 72°F/30%RH). Mean comfort ratings for the two replicated scenarios (denim at 100°F/60%RH and wool at 0°F/20%RH) were nearly identical. In addition, the Pearson product-moment correlation coefficient calculated across subjects for the two "denim" scenarios was 0.84 and for the two "wool" scenarios was 0.94, both significantly different from zero at $p < 0.0001$.

Considering both the mean ratings and the r values between replicated scenarios, we can conclude that there was a wide range of comfort ratings assigned to the different comfort scenarios and these ratings were consistent with the logically expected levels of comfort defined by the scenario, there were statistically significant differences between pairs of scenarios that would be expected to differ, and the correlations between the replicated scenarios were very high. Taken together, these data show the high degree of sensitivity of the CALM scale to image-based scenarios, a high degree of construct validity because the mean comfort ratings of the scenarios are logically ordered, and good test-retest reliability.

Experiment 5: Reliability and Sensitivity of the CALM Scale (Clothing Stimuli)

EXPERIMENTAL

Thirty-seven volunteer consumer panelists served as subjects, coming from the same pool of subjects described previously. All were naïve to testing of the hand and comfort properties of clothing and fabrics. The stimuli consisted of three gloves differing in both fabric composition and construction. Gloves were chosen as stimuli to mimic comfort responses similar to those involved in handling fabric swatches. The first glove was an 8-ounce jersey fabric glove with a knit wrist, clute cut (Dickey brand, general utility Williamson-Dickey Mfg Co.), the second was an 8-oz blended canvas glove with a knit wrist (Wells-Lamont "Basics" work glove, 65% polyster, 35% cotton, Wells-Lamont, Inc. Niles, IL), and the third was a U.S. military glove insert made of 70% wool and 30% nylon.

All testing was done in the same consumer testing booths (70°F) used previously, in order to avoid influences of temperature on comfort ratings [19, 30]. Glove samples were presented twice in a restricted random order of two series (the same glove could not be presented sequentially in the two series). Subjects were instructed to place each glove on their preferred hand (determined in advance) and to rate its comfort after clenching the fist three times. Subjects were specifically instructed to ignore fit in their evaluation. Comfort ratings were made using the CALM scale (labeled −100 to +100). After the comfort evaluation of each glove, a 60-second interstimulus interval elapsed before presentation of the next glove. After the first series of evaluations was complete, ratings of the same gloves were repeated in a second test series.

RESULTS AND DISCUSSION

The mean (+ standard error) of the comfort ratings for the two replicates of the jersey glove were 65.0 (2.6) and 66.9 (2.6). For the canvas glove, the ratings were 37.3 (4.5) and 36.3 (4.3), and for the wool glove, they were 6.0 (6.4) and −7.8 (7.1). An analysis of variance of the data showed a significant main effect of glove type ($F = 52.23$, $df = 2{,}72$, $p < 0.001$). In addition, there was a significant session (replication) effect ($F = 13.17$, $df = 1{,}36$, $p < 0.001$) and a significant session × glove effect ($F = 17.96$, $df = 2{,}72$, $p < 0.001$). The latter effects can be attributed entirely to the difference in mean comfort ratings for the wool glove between replicates. This effect may be due to a greater variability in the comfort sensation around the neutral point (neither comfortable nor uncomfortable), as reflected in the mean comfort ratings and associated standard errors for this glove. Surprisingly, in spite of this session effect, the Pearson correlation coefficient across subjects for the two replicates of the wool glove was 0.93 ($p < 0.01$). The correlation coefficient between replicates for the jersey glove was 0.88 ($p < 0.01$), and for the canvas glove it was 0.91 ($p < 0.01$). The results of this study show the CALM scale to be a sensitive measure of the perceived comfort of fabrics/clothing worn on the hand. The correlation coefficients between replicates also show good reliability of the scale for this purpose, although the reliability may be reduced when comfort ratings fall near the neutral point.

Experiment 6: Comfort Scaling of Military Fabrics

EXPERIMENTAL

In order to examine the relationship of sensory hand attributes to clothing comfort, the same thirteen fabrics for which descriptive hand data were obtained in experiment 2 were evaluated for their hand comfort by naïve consumers. Forty civilian employees of Natick who had no formal training in textiles served as subjects. The same 30 × 30 cm swatches of each military fabric previously tested (see Table II) were used as stimuli. All samples were stored under controlled climatic conditions (70° ±1.4 F/65 ±1.3% RH) until just prior to testing, which occurred in the same temperature controlled, individual sensory testing booths used previously.

All thirteen samples were presented in random order during a single test session. Each sample was evaluated for comfort using the CALM scale shown in Figure 4. Subjects were instructed that they could "hold, touch, feel or squeeze the material in any manner" so long as they only felt and evaluated the coded (face) side of the fabric. After evaluating each sample, the swatch and rating form were returned, and the next sample was presented after a 60-second interstimulus interval. Testing was repeated in the same manner five days after initial testing in order to assess the reliability of subject ratings.

RESULTS AND DISCUSSION

Table V shows the mean comfort ratings for each of the thirteen test fabrics and the results of the ANOVA and post-hoc tests conducted on the mean comfort ratings. It is evident from Table V that the CALM scale was used effectively to differentiate between the comfort levels of the fabrics. In terms of absolute comfort levels, the fabrics had a range of perceived comfort/discomfort that varied from slight discomfort (−9.8) to above moderately comfortable (47.2). (Note that these fabrics represented materials used in military garments, so very uncomfortable fabrics were not part of the stimulus set.)

A Pearson product-moment correlation coefficient between comfort ratings obtained during the initial test session and the replication had a value of 0.68 ($p < 0.01$). Although not as high as the test-retest correlation coefficients for image-based stimuli and gloves, judgments in those latter experiments were repeated within a single session, whereas in this experiment, judgments were separated by a five-day interval.

TABLE V. Mean comfort ratings (n = 45) for the 13 fabrics tested in experiment 6. Means with different letter superscripts are significantly different at $p < 0.05$.

Fabric	Mean comfort rating	Standard deviation
18L	-9.8^{a}	44.8
17C	-1.4^{ab}	40.3
176	2.4^{ab}	29.4
124	9.8^{bc}	25.0
20J	10.9^{bcd}	31.0
16C	22.0^{cde}	26.2
12T	23.6^{cde}	27.1
14N	24.2^{cde}	30.8
19N	28.5^{cdef}	36.1
10R	28.9^{def}	25.7
13P	37.4^{ef}	25.3
15B	46.4^{f}	22.5
11A	47.2^{f}	27.8

The sensory descriptive data collected previously on the samples were used to correlate with the consumer comfort data collected here. In addition, for each fabric, the mean of the descriptive attribute intensity ratings across all attributes was calculated to serve as an index of the overall salience of the fabric's handle. This was done in order to test the hypothesis that clothing comfort is related to the *absence of tactile sensation*. This hypothesis derives from such studies as Gwosdow *et al.* [19], which showed that increased perception of fabric texture significantly decreased fabric acceptability.

Table VI shows the Pearson product-moment correlation coefficients between comfort ratings, individual sensory hand attributes, and the mean intensity rating across all

TABLE VI. Pearson product-moment correlation coefficients for the associations of each descriptive hand attribute and the mean intensity across all attributes with the judged comfort of the fabrics: * $p < 0.05$, ** $p < 0.01$.

Hand attribute	r with comfort
Grainy	−0.41
Gritty	−0.92**
Fuzziness	−0.60
Thickness	−0.32
Tensile stretch	−0.92**
Hand friction	−0.77*
Fabric to fabric friction	−0.36
Depression depth	−0.71*
Springiness	−0.72*
Force to gather	−0.17
Stiffness	−0.17
Force to compress	−0.17
Compression resilience/intensity	−0.42
Compression resilience/rate	−0.53
Fullness/volume	−0.17
Noise intensity	−0.25
Noise pitch	−0.03
Mean intensity over all attributes	−0.70

attributes. Many of the individual sensory attributes are significantly correlated with consumer comfort ratings. These include "gritty," "tensile stretch," "hand friction," "depression depth," and "springiness." Also evident is the fact that all seventeen of the descriptive hand attributes are *negatively* correlated with comfort, suggesting that the higher the salience of any fabric attribute, the lower the perceived comfort. Given these high negative correlations, it is not surprising that the mean intensity rating across all attributes is also negative and accounts for about 50% of the variance in comfort responses ($r = -0.70$).

Experiment 7: Correlations of Sensory, Instrumental (Kawabata), and Comfort Data

EXPERIMENTAL

Logically, the sensory hand attributes of a fabric should be a better predictor of its comfort than any mechanical measure, because the human observer can only base his/her comfort judgment on perceptual experience. Of course, mechanical measures are extremely convenient, and it would be desirable to find one or more that correlate well with either perceived sensory experience or comfort. One instrumental technique that has achieved particular popularity is the Kawabata evaluation system for fabrics (KES-F) [24–26]. This technique consists of a set of mechanical parameters of fabrics that can be combined, using established regression formulas, to predict hand attributes. The methodology generates predictions of the hand attributes of "stiffness," "anti-drape stiffness," "crispness," "fullness and softness," "smoothness," and "total hand value."

To assess the relationship between Kawabata parameters and the descriptive handle and comfort data that we collected, the eight fabrics asterisked in Table II were submitted for Kawabata mechanical testing by Milliken Research, Corp., under standardized textile testing conditions. Table VII lists the Kawabata parameters that were tested and their associated units of measure. All testing was conducted on the reverse side of the fabric swatches. Due to the large number of mechanical properties tested (seventeen) and the equally large number of descriptive hand attributes in the HSDA method (seventeen), it is impractical to examine correlations between individual mechanical and hand properties. However, it is feasible to correlate Kawabata predicted hand values and the total hand value (as they apply to men's winter and summer suit fabrics) to HSDA hand attributes.

TABLE VII. Kawabata parameters and associated units of measure (experiment 7).

Blocked property	Kawabata mechanical parameters: Symbol	Characteristic value	Unit
Tensile	KEMT	extensibility	gf - dimensionless
	K_LT	linearity	dimensionless
	K_WT	tensile energy	gf - cm/cm^2
	K_RT	resilience	%
Bending	K_B	bending rigidity	gf-cm^2/cm
	K_HB	hysteresis	gf-cm^2/cm
Shearing	K_G	shear stiffness	gf/cm - degree
	K_HG	hysteresis at $0 = 0.5°$	gf/cm
	K_HG5	hysteresis at $0 = 5°$	gf/cm
Compression	K-LC	linearity	dimensionless
	K_WC	compressional energy	gf - cm/cm^2
	K_RC	resilience	%
Surface	K_MIU	coefficient of friction	dimensionless
	K_MMD	mean deviation of MIU	dimensionless
	K_SMD	geometrical roughness	micron
Weight & thickness	K_W	weight per unit	mg/cm^2
	K_T	thickness at 0.5 gf/cm^2	mm

TABLE VIII. Pearson correlation coefficients greater than 0.50 for the associations between Kawabata hand values and HSDA hand attributes, * $p < 0.05$.

Hand properties	r value
Stiffness	
Force to compress	0.83*
Stiffness	0.80*
Force to gather	0.79*
Compression resilience intensity	0.71
Thickness	0.68
Fullness/volume	0.63
Anti-drape stiffness	
Force to compress	0.87*
Stiffness	0.84*
Force to gather	0.80*
Compression resilience intensity	0.73*
Fullness/volume	0.71*
Thickness	0.67
Fullness/softness	
Springiness	0.87*
Depression depth	0.85*
Fuzziness	0.85*
Hand friction	0.77*
Gritty	0.76*
Tensile stretch	0.67
Smoothness	
Fuzziness	0.55
Fabric to fabric friction	0.50
Crispness	
No correlation > 0.50	

RESULTS AND DISCUSSION

Table VIII shows Pearson product-moment correlation coefficients between the Kawabata hand parameter predictions and the HSDA hand attribute ratings. Only correlation coefficients greater than 0.50 are listed, and those that are statistically significant are so indicated in the table.

Examination of the HSDA hand attributes that correlate best with each primary hand expression reveals good conceptual agreement between the methods. For example, both Kawabata "stiffness" and "anti-drape stiffness" are correlated with the same six HSDA hand attributes, and both are correlated very highly with HSDA "stiffness" ($r = 0.80, 0.84$). Similarly, the HSDA hand attributes that correlate highly with Kawabata "fullness/softness" (defined as "bulky," "rich," and "springy" sensations) are those of "depression depth," "fuzziness," and "springiness." Last, Kawabata smoothness (defined as "limber" and "soft" like "cashmere fiber") is correlated most highly with HSDA "fuzziness" and "fabric to fabric friction," both attributes that would be expected to be positively associated with softer pile fabrics.

While the Kawabata method predicts hand values from independent mechanical properties of fabrics, they are based on predictive equations derived from quite different fabrics than we tested here. A more direct approach to reduce the number of Kawabata mechanical properties to a manageable number for the purpose of correlation with sensory or comfort data is to use principal component analysis (PCA) to derive the component (factor) structure in the data. Such an approach can also be used to reduce redundancy in the sensory hand data.

We conducted a PCA analysis of the Kawabata data obtained on the eight fabrics. A Varimax rotation with Kaiser normalization and an Eigen value criterion of 1.0 to stop extracting factors resulted in a five-component solution. An analysis of the variable loadings on each component resulted in the interpretation of these components as being related to "shear properties," "bending properties," "compression/friction," "tensile properties," and "surface roughness." These five components accounted for 98% of the variance in the instrumental data. We conducted a similar PCA on the sensory descriptive hand data, which resulted in a three-component solution accounting for 92% of the variance in the data set. Analysis of the attribute loadings on each component identified the three components as "surface texture/depth," "volume," and "noise."

In order to assess the relationship of the sensory hand attributes with comfort ratings, we conducted a third PCA using both the sensory attributes and "comfort" ratings. The results of this PCA are shown in Table IX. Analysis of the sensory attribute loadings revealed the same pattern as obtained previously, resulting in three components: "surface texture/depth" (component 1), "volume" (component 2), and "noise" (component 3). As we can see, comfort ratings loaded negatively on "surface texture/depth," suggesting that the comfort of these fabrics is inversely related to their perceived surface texture/depth. This is consistent with the negative correlations found previously (Table VI) between comfort and hand attribute intensities.

TABLE IX. Rotated component matrix resulting from a PCA of the HSDA hand attributes plus comfort. Rotation converged in 8 iterations.

	Components		
	1	2	3
Gritty	.990		
Tensile stretch	.937		
Comfort	−.919		
Springiness	.901		
Depression depth	.879		
Hand friction	.862		
Fuzziness	.835		−.527
Compression resilience: rate	.660		
Fabric to fabric friction	.589		
Force to compress		.975	
Force to gather		.962	
Fullness/volume		.931	
Thickness		.917	
Compression resilience: intensity		.915	
Stiffness		.908	
Noise intensity			.923
Noise pitch			.883
Grainy			.780

Last, we conducted a PCA using both the Kawabata data and the comfort ratings, and the results are shown in Table X. In this PCA, which accounted for 98% of the variance in the data, comfort is positively loaded on component 1 ("shear") but negatively loaded on component 4 ("compression/friction"). The other components are "bending properties" (component 2), "tensile properties" (component 3), and "surface roughness" (component 5).

TABLE X. Rotated component matrix resulting from a PCA of the Kawabata mechanical parameters plus comfort. Rotation converged in 8 iterations.

	Components				
	1	2	3	4	5
Hysterisis at 0 = 0.5°	.939				
Resilience	−.923				
Linearity	.898				
Hysteresis at 0 = 5°	.864				
Shear stiffness	.835				
Compressional energy	−.828				
Thickness of 0.5 gf/cm^2	−.786				
Resilience	−.749			.592	
Weight per unit area		.918			
Bending rigidity		.881			
Hysteresis		.848			
Extensibility			.961		
Tensile energy			.932		
Comfort	.594		−.748		
Linearity				.957	
Coefficient of friction				.926	
Mean deviation of MIU					.927
Geometric roughness					.819

With independent component scores established from the principal component analyses for both the Kawabata and sensory hand data, it was possible to predict perceived comfort from either the sensory component scores, the Kawabata component scores, or both. First, by including all three sensory components in a multiple regression of component scores on comfort ratings, we developed the following regression equation:

$$\text{Comfort} = -15.6\ (\text{surface texture/depth}) - 1.07\ (\text{volume}) - 7.67\ (\text{noise}) + 27.5\ (\text{constant}) \quad (R = 0.96,\ R_{adj}^2 = 0.87) \quad . \quad (1)$$

The weightings of the components in this regression model support the findings of the PCA of sensory attributes plus comfort, which show the comfort variable to be loaded highly on component 1 (surface texture/depth). In addition, the fact that all the component weightings are negative supports the notion put forth earlier that comfort is inversely related to the average perceived intensity across all hand attributes.

A multiple regression using the Kawabata component scores to predict comfort produced the following regression equation:

$$\text{Comfort} = 11.8\ (\text{shear}) - 3.1\ (\text{bending}) - 0.3\ (\text{compression/friction}) - 11.9\ (\text{tensile}) + 0.4\ (\text{surface roughness}) + 27.5\ (\text{constant}) \quad (R = 0.94,\ R_{adj}^2 = 0.60 \quad . \quad (2)$$

While producing a predictive model as good as that obtained with the sensory component scores, the Kawabata regression model used all five factors. Since there were only eight fabrics in the data set, it was not as compelling or as useful a predictive model. In addition, none of the components were statistically significant at the $p < 0.05$ level, as compared with the sensory regression model in which only component 2 (volume) was not statistically significant.

Combining both the sensory component scores and the Kawabata component scores into a stepwise multiple regression model to predict comfort resulted in the following equation:

$$\text{Comfort} = -16.3\ (\text{sensory surface texture/depth}) - 8.7\ (\text{sensory noise}) - 4.3\ (\text{Kawabata surface texture}) + 27.5\ (\text{constant}) \quad (R = 0.99,\ R_{Adj}^2 = 0.96) \quad , \quad (3)$$

where all three components contributed significantly ($p < 0.05$) to the model. We chose this three-variable solution because including more variables resulted in solutions that were overdetermined. As might well be expected, both sensory surface texture and Kawabata surface texture factors were important predictors of comfort in Equation 3. In addition, the noise factor was an important sensory predictor of comfort due to its unique contribution to the model variance.

The reader may have noted that we obtained all of our sensory hand data on the face surface of the test fabrics, whereas the Kawabata data came from the reverse side of the fabrics. Although there were no obvious tactile differences between the face and reverse surfaces of these fabrics, in order to ensure that any possible differences would not alter the basic findings of the research, the descriptive hand panel evaluated the reverse surfaces of all eight test fabrics used in Kawabata comparisons for eight attributes that could potentially vary between face and reverse surfaces. Of the 64 *possible* differences between face and reverse (eight fabrics × eight attributes), there were only seven significant differences. The absolute magnitude of these differences was small (the largest was equal to one-half a scale point) and did not change the overall tactile profile for any fabric. Note that for any other fabric types, differences between face and reverse surfaces could contribute more importantly to the interpretation of such data.

General Discussion and Conclusions

The research reported here establishes the basis for a standardized approach to characterizing the hand properties of clothing fabrics and analyzing the contributions of their sensory and mechanical properties to perceived comfort. Although developed and applied for characterizing and analyzing military fabrics, our approach and techniques can be used for any clothing fabrics. The approach is predicated on the use of sound psychophysical principles for assessing both the qualitative and quantitative aspects of sensory handle and comfort experience.

Adapting the HSDA method of hand analysis to military fabrics constitutes a significant advance in enabling the sensory characterization of fabrics with a set of well-defined, independent attributes, each with a detailed operational technique for its evaluation. These standardized operational techniques enable ready transfer of the methodology to other laboratories. This fact, combined with the use of stimulus-referenced intensity scales for each attribute, establishes a unique, standard protocol for interlaboratory studies or for establishing functional performance-based specifications for military (or other) clothing fabrics. In addition, the extremely high reliability of the method ensures that data collected over long periods of time, *e.g.*, during storage trials, can be readily compared.

The experiments reported here also establish the validity, reliability, and sensitivity of a new psychophysical scale for assessing comfort. The CALM (comfort affective labeled magnitude) scale developed here has several important advantages over simple category scales of comfort. First, the CALM scale enables statements to be made about the ratios of perceived comfort between samples (*e.g.*, sample X is twice as comfortable as sample Y). However, it avoids a major disadvantage of magnitude estimation—the inability to index and compare absolute levels of liking among different individuals. The CALM scale avoids this problem by using the word phrases "greatest imaginable comfort" and "greatest imaginable discomfort" to anchor the scales to a common ruler of perceptual experience [4]. Another related advantage of the CALM scale is its potentially greater sensitivity to differences between very comfortable (or uncomfortable) stimuli. This is a logical consequence of the CALM scale end-points ("greatest imaginable liking/disliking"), which enable more extreme ratings than "extremely comfortable (or uncomfortable)". These end-point labels serve not only to anchor different subject ratings to a common scale, but also to foster better discrimination of very comfortable (or uncomfortable) fabrics or items of clothing. This can be an important advantage, because in most product development applications, the samples being tested are near optimal comfort. The CALM scale has the potential to enable better discrimination between fabrics and clothing items that fall in this "near optimal" category.

From a sensitivity and reliability standpoint, the data show the CALM scale to be sensitive to a wide variety of comfort-related stimuli, including image-based stimuli, fabrics, and garments (gloves), and to have good reliability both within and across test sessions. In addition, a practical aspect of the scale is that a simple arithmetic mean can be used as a measure of central tendency. This stands in contrast to magnitude estimations, where medians, geometric means, or log transformations of the data must be calculated to arrive at a measure of central tendency. Also, because the scale produces ratio level data, standard parametric statistics can be used to analyze these data.

Finally, of some practical importance is the fact that the specific numerical labels on the CALM scale are somewhat arbitrary. Previous research showed that a scale with no numbers produces data equivalent to scales labeled with numbers ranging from 0 to 100 or -100 to $+100$ [40]. Subjects pay relatively little attention to the numbers on the scale, as suggested previously by Green *et al.* [16]. It may well be the case that no numbering is the best option in certain cases. This is particularly true if the data from the scale are to be compared by users who differ significantly in their knowledge or use of numbers (children versus adults) or where cultural or practical concerns may make the use of numbers a distraction. When no numbers are used, data from the scale can simply be transcribed from measurements with a ruler on the 100-mm analogue line scale and then transformed to a -100 to $+100$ scale. Of course, if it is desirable to make ratio statements about liking/disliking, the scale must conform to the numerical values that were originally used to locate the semantic labels (Figure 4), or to a multiplicative transformation of these values.

The approach to uncovering sensory instrumental-comfort relationships outlined here is a valuable approach for understanding the complex factors that contribute to the perceived comfort of fabrics and clothing. By reducing the large array of sensory and mechanical properties that can be measured for fabrics to a small number of independent components, it is possible to derive regression models to predict the perceived comfort of the fabrics. The results of this research show that a judicious combination of sensory *and* instrumental factors can be used to predict the handle and comfort of military fabrics, while accounting for >95% of the variance in the comfort ratings.

Further research is now being conducted in our laboratory to determine the extent to which the sensory, instrumental, and hand-comfort data collected here can be used to predict the dynamic comfort of users wearing garments constructed from these same fabrics in controlled wear trials.

ACKNOWLEDGMENTS

We would like to thank Mr. Jim Whitworth of Milliken Research Corporation, Spartanburg, SC, for conducting the Kawabata analyses, Mr. Larry Lesher of Geo-Centers, Inc. for help with the statistical analysis of the data, and the Industry/Government Working Group for performance criteria and the Technical Cooperative Panel for their technical input on the British, Australian, and Canadian military fabrics.

Literature Cited

1. Ali, S. I., and Begum, S., Fabric Softeners and Softness Perception, *Ergonomics* **37** (5), 801–806 (1994).
2. Baird, J. C., and Harder, K. A., The Psychophysics of Imagery, *Percept. Psychophys.* **62** (1), 113–126 (2000).
3. Binns, H., The Discrimination of Wool Fabrics by the Sense Touch, *Br. J. Psychol.* **16**, 237–247 (1926).
4. Borg, G., A Category Scale with Ratio Properties for Intermodal and Interindividual Comparisons, in "Psychophysical Judgment and the Process of Perception," H-G.

Geissler, and P. Petxoid, Eds., VEB Deutxcher Veriag der Wissenschaften, Berlin, 1982, pp. 25–34.

5. Brand, R. H., Measurement of Fabric Aesthetics, *Textile Res. J.* **34,** 791–804 (1964).
6. Brandt, M. A., Skinner, E. Z., and Coleman, J. A., Texture Profile Method, *J. Food Sci.* **28,** 404–409 (1963).
7. Branson, D. H., and Sweeney, M., Conceptualization and Measurement of Clothing Comfort: Toward a Metatheory, in "Critical Linkages in Textiles and Clothing: Theory, Method and Practice," S. Kaiser and M. L. Damhorst, Eds., International Textile and Apparel Association, Monument, CO, 1991, pp. 94–105.
8. Burgard, D. R., and Kuznicki, J. T., "Chemometrics: Chemical and Sensory Data," CRC Press, Boca Raton, FL, 1990.
9. Burns, L. D., Brown, D. M., Cameron, B., Chandler, J., Dallas, M. J., and Kaiser, S. B., Sensory Interaction and Descriptions of Fabric Hand, *Percept. Mot. Skills* **81,** 120–122 (1995).
10. Civille, G. V., and Dus, C. A., Development of Terminology to Describe the Handfeel Properties of Paper and Fabrics, *J. Sens. Stud.* **5,** 19–32 (1990).
11. Elder, H. M., Fisher, S., Armstrong, K., and Hutchinson, G., Fabric Softness, Handle and Compression, *J. Textile Inst.* **75,** 37–46 (1984).
12. Elder, H. M., Fisher, S., Armstrong, K., and Hutchinson, G., Fabric Stiffness, Handle and Flexion, *J. Textile Inst.* **75,** 99–106 (1984).
13. Fourt, L., and Hollies, N. R. S., "Clothing: Comfort and Function," Marcel Dekker, NY, 1970.
14. Gagge, A. P., Stolwijk, J. A. J., and Hardy, J. D., Comfort and Thermal Sensations and Associated Physiological Responses at Various Ambient Temperatures, *Environ. Res.* **1,** 1–20 (1967).
15. Green, B. G., Dalton, P., Cowart, B., Shaffer, G., Rankin, K., and Higgins, J., Evaluating the "Labeled Magnitude Scale" for Measuring Sensations of Taste and Smell, *Chem. Senses* **21** (3), 323–335 (1996).
16. Green, B. G., Shaffer, G. S., and Gilmore, M. M., Derivation and Evaluation of a Semantic Scale of Oral Sensation Magnitude with Apparent Ratio Properties, *Chem. Senses* **18,** 683–702 (1993).
17. Gridgeman, N. T., A Comparison of Some Test Methods, *J. Food Sci.* **26,** 171–177 (1961).
18. Guilford, J. P., and Dingman, H. F., A Modification of the Method of Equal-appearing Intervals, *Am. J. Psychol.* **68,** 450–454 (1955).
19. Gwosdow, A. R., Stevens, J. C., Berglund, L. G., and Stolwijk, J. A. J., Skin Friction and Fabric Sensations in Neutral and Warm Environments, *Textile Res. J.* **56** (9), 574–580 (1986).
20. Hollies, N. R., Investigation of the Factors Influencing Comfort in Cotton Apparel Fabrics, Contract 12-14-100-7183 (72), US Dept. of Agriculture, New Orleans, (referenced in Hollies *et al.,* 1979), 1965.
21. Hollies, N. R., Custer, A. G., Morin, C. J., and Howard, M. E., A Human Perception Analysis Approach to Clothing Comfort, *Textile Res. J.* **49,** 557–564 (1979).
22. Houghton, F. C., and Yaglou, C. P., Determining Lines of Equal Comfort, *ASHRAE Trans.* **28,** 163–176, 361–384 (1923).
23. Jones, L. V., Peryam, D. R., and Thurstone, L. L., Development of a Scale for Measuring Soldiers' Food Preferences, *Food Res.* **20,** 512–520 (1955).
24. Kawabata, S., "Standardization and Analysis of Hand Evaluation," 2nd ed., The Textile Machinery Society of Japan, Osaka, 1980.
25. Kawabata, S., and Niwa, M., Analysis of Hand Evaluation of Wool Fabrics for Men's Suits Using Data of a Thousand Samples and Computation of Hand from the Physical Properties, in "Proc. 5th Int. Wool Textile Research Conference," vol. 5, 1975, pp. 413–424.
26. Kawabata, S., Postle, R., and Niwa, M., Eds. "Objective Specification of Fabric Quality, Mechanical Properties and Performance," The Textile Machinery Society of Japan, Osaka, 1982.
27. Lane, H. L., Catania, A. C., and Stevens, S. S., Voice Level: Autophonic Scale, Perceived Loudness and Effect of Side Tone, *J. Acoust. Soc. Am.* **33,** 160–167 (1961).
28. Laughlin, J., Perception of Texture, Visually and Tactually: An Exploratory Study Using Multidimensional Scaling Analysis, *Int. J. Cloth. Sci. Technol.* **3** (1), 28–36 (1991).
29. Mackay, C., Effect of Laundering on the Sensory and Mechanical Properties of 1 × 1 Rib Knitwear Fabrics, Masters thesis, Bolton Institute of Higher Education, Bolton, England, p. 63, 1992.
30. Marks, L. E., and Gonzales, R. R., Skin Temperature Modifies the Pleasantness of Thermal Stumuli, *Nature* **247** (5441), 473–475 (1974).
31. Meilgaard, M. C., Civille, G. V., and Carr, B. T., "Sensory Evaluation Techniques," 2nd ed., CRC Press, Boca Raton, FL, 1991.
32. Parducci, A., Contextual Effects: A Range-Frequency Analysis, in "Handbook of perception" vol. 3, E. C. Carterette and M. P. Friedman, Eds., Academic Press, 1974.
33. Pierce, F. T., The "Handle" of Cloth as a Measurable Quantity, *J. Textile Inst.* **21,** T377–416 (1930).
34. Pontrelli, G. J., Partial Analysis of Comfort's Gestalt, in "Clothing Comfort: Interaction of Thermal, Ventilation, Construction and Assessment Factors," N. R. S. Hollies and R. F. Goldman, Eds., Ann Arbor Science Publishers Inc., Ann Arbor, MI, 1977, pp. 71–80.
35. Riskey, D. R., Uses and Abuses of Category Scales in Sensory Measurement, *J. Sens. Stud.* **1,** 217–236 (1986).
36. Robinson, K. J., Chambers, E., and Gatewood, B. M., Influence of Pattern Design and Fabric Type on the Hand Characteristics of Pigment Prints, *Textile Res. J.* **67** (11), 837–845 (1997).
37. Robinson, K. J., Gatewood, B. M., and Chambers, E., Influence of Domestic Fabric Softeners on the Appearance, Soil Release, Absorbency, and Hand of Cotton Fabrics, in "Book of Papers, 1994 International Conference and Ex-

hibition," American Association of Textile Chemists and Colorists, Charlotte, NC, 1994, pp. 58–66.
38. Rohles, F. H., Comfort and the Man-Environment System, in "Proc. Clothing and Energy Resources Workshop", Michigan State University, East Lansing, MI, 1978, pp. 22–35.
39. Rohles, F. H., Psychological Aspects of Thermal Comfort, *ASHRE J.* **13,** 86–90 (1971).
40. Schutz, H. G., and Cardello, A. V., A Labeled Affective Magnitude (LAM) Scale for Assessing Food Liking/Disliking, *J. Sens. Stud.* **16,** 117–159 (2001).
41. Schwartz, N. O., Adaptation of the Sensory Texture Profile Method to Skin Care Products, *J. Texture Stud.* **6,** 33–42 (1975).
42. Slater, K., Comfort Properties of Textiles, *Textile Prog.* **9,** 1–71 (1977).
43. Slater, K., "Human Comfort," Charles C. Thomas, Springfield, IL, 1985.
44. Sontag, M. S., Comfort Dimensions of Actual and Ideal Insulative Clothing for Older Women, *Cloth. Textile Res. J.* **4,** 9–17 (1985–1986).
45. Stevens, J. C., A Comparison of Ratio Scales for the Loudness of White Noise and the Brightness of White Light, Doctoral dissertation, Harvard University, 1957.
46. Stevens, J. C., Aging and Spatial Acuity of Touch, *J. Gerontol.* **47** (1) 35–40 (1992).
47. Stevens, S. S., On the Psychophysical Law, *Psychol. Rev.* **64,** 153 (1957).
48. Stevens, S. S., "Psychophysics," Wiley, NY, 1975.
49. Stevens, S. S., and Galanter, E. H., Ratio Scales and Category Scales for a Dozen Perceptual Continua, *J. Exp. Psychol.* **54,** 377–411 (1957).
50. Sweeney, M. M., and Branson, D. H., Sensorial Comfort, Part II: A Magnitude Estimation Approach for Assessing Moisture Sensation, *Textile Res. J.* **60,** 477–452 (1990).
51. Vaughn, E. A., and Kim, C. J., Definition and Assessment of Fabric Hand, in "Book of Papers, AATCC Conference," American Association of Textile Chemists and Colorists, Research Triangle Park, NC, 1975, pp. 66–77.
52. West, R. L., Ward, L. M., and Khosla, R., Constrained Scaling—The Effect of Learned Psychophysical Scales on the Idiosyncratic Response Bias, *Percept. Psychophys.* **62** (1), 137–151 (2000).
53. Winakor, G., Kim, C. J., and Wolins, L., Fabric Hand: Tactile Sensory Assessment, *Textile Res. J.* **50,** 601–610 (1980).
54. Winslow, C-E. A., Herrington, L. P., and Gagge, A. P., Physiological Reactions of the Human Body to Varying Environmental Temperature, *Am. J. Physiol.* **120,** 1–20 (1937).
55. Winslow, C-E. A., Herrington, L. P., and Gagge, A. P., Relations Between Atmospheric Conditions, Physiological Reactions and Sensations of Pleasantness, *Am. J. Hyg.* **26,** 103–115 (1937).
56. Yick, K. L., Cheng, K. P. S., and How, Y. L., Subjective and Objective Evaluation of Men's Shirting Fabrics, *Int. J. Cloth. Sci. Technol.* **7** (4), 17–29 (1995).

Manuscript received December 17, 2001; accepted August 21, 2002.