Accuracy and Reliability of Fabric's Hand Subjective Evaluation

Loreta VALATKIENĖ*, **Eugenija STRAZDIENĖ**

Department of Clothing and Polymer Product Technology, Faculty of Design and Technologies, Kaunas University of Technology, Studentu 56, LT-51424 Kaunas, Lithuania

Received 01 February 2006; accepted 03 July 2006

Subjective assessments treat fabric hand as a psychological reaction obtained from the sense of touch, based on experience and sensitivity of humans. In this paper we identify the effect of judge panel's training number and methodology upon the accuracy and agreement level between the results of fabric's hand subjective evaluation. The objects of subjective evaluation are three groups of 100 % cotton fabric, treated by commercial stiffener. The concentration of stiffener in each group differed by 3 ml/l, by 5 ml/l, by 7 ml/l, respectively. The experiments are performed with the panel of ten qualified judges. The samples of assessed fabric are ranked individually for such textile subjective characteristics as: hardness, stiffness, flexibility, softness, roughness, smoothness, stretchability and resiliency. The discussion of agreement level changes on the basis of Kendall's coefficient of concordance W is presented for total hand evaluation, as well as for individually assessed fabric's subjective characteristics. The reliability of trained judge panel is evaluated.

Keywords: fabric hand, judge panel, subjective evaluation, agreement level.

INTRODUCTION

The successful application of fabric objective measurement (FOM) depends as much on establishing reliable methods for quantifying subjective judgments, and on establishing equations that accurately predict such judgments from the chosen objective measurements, as it does on the selection and precision of the objective measurement themselves. For this reason, sensory methods for the evaluation of fabric handle are discussed in some detail before any of the instrumental methods more usually associated with FOM are described. Fabric evaluations carried out by people (subjects) are usually called subjective evaluations, whereas evaluations made by using instruments (objects) are called objective measurements [1].

The term fabric "hand" or "handle" has been defined as the quality of a fabric or yarn assessed by the reaction obtained from the sense of touch or the sum total of the sensations expressed when a textile fabric is handled by touching, flexing of the fingers and so on [1]. It implies the ability of the fingers to make a sensitive and discriminating assessment, and of the mind to integrate and express the results in a single-valued judgment [2].

W. S. Howorth and P. H. Oliver studied the subjective assessment of fabric hand. They used a panel of 25 people with no special experience in handling fabrics to rank 27 samples of worsted suiting fabrics. Fabrics were ranked according to hand by the method of comparison in pairs, and each judge was asked to state for accepting or rejecting a particular fabric from a pair [3, 4].

Through the development of new web formation technologies, new fibers, and new finishing procedures, fabrics have made significant gains from the standpoint of their aesthetic acceptability. R. H. Brand is one of several researches, who commented on differences between vocabularies of experts and untrained judges of textile hand. He stated that, "…aesthetic concepts are basically people's preferences and should be evaluated subjectively by people" [5]. This differentiation has initiated much research focused on how to model subjective fabric hand objectively.

C. L. Hui *et al.*, in order to assess the reliability of the fourteen significant bipolar pairs of sensory attributes of fabric hand conducted a test-retest reliability study at the completion of the panelists' training. Panelists were trained to understand the definitions of these sensory attributes [6].

K. L. Yick, K. P. S. Chang, Y. L. How studied subjective handle assessment and used a panel of 199 judges. They were divided into two groups based on their academic and industrial experiences in the textile and clothing industries: people who had less than five years of experience and people who had five or more years experience in the clothing industry. More experienced judges exhibited a higher percentage of significance and gave a higher level of overall agreement [7].

In the work of V. A. Cardello and C. Winterhalter a standardized hand evaluation methodology was checked for its sensitivity and reliability and used to characterize military fabrics. 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 (HSDA). In order to assess the reliability and sensitivity of this method, they conducted a test-retest reliability study at the completion of training and to assess long-term reliability, two fabrics were tested again six month later. They concluded that in conjunction with the panel training program, result in a sensory hand evaluation method is highly sensitive and reliable over extended period of time [8]. The later this method provides useful information, notwithstanding the fact that it expresses the individual assessor's experience and knowledge.

*Corresponding author. Tel.: +370-37-300205.; fax: +370-37-353989.
E-mail address: *loreta.valatkiene@stud.ktu.lt* (L. Valatkienė)

The comparison of methods for the evaluation of woven fabric hand was performed in Kaunas University of Technology. The blind respondents carried out subjective assessment: in terms best characterizing textile hand [9].

The aim of this research is to define the effect of judge panel's training number and methodology upon the accuracy and agreement level between the results of fabric's hand subjective evaluation and to assess the reliability of such evaluations.

EXPERIMENTAL

Investigations were performed with 100 % cotton plane weave fabric. Samples for fabric hand subjective evaluation were prepared in such a way: they were soaked for 15 min in stiffener (PVA dispersion) solutions of 30 ±2 ºC temperature and dried at 20 ±3 ºC temperature. In order to determine the agreement level of fabric's hand subjective evaluation and judge panel's sensitivity limit, cotton samples were divided into three groups. The concentration of stiffener solution in each group differed by 3 ml/l, by 5 ml/l and by 7 ml/l. Thus stiffener concentrations for following groups of samples were 0; 3; 6; 9 ml/l, 0; 5; 10; 15 ml/l and 0; 7; 14; 21 ml/l, respectively. With the purpose to assess judge panel's reliability the fourth group of samples was prepared, i.e. into the second group of samples random sample from the third group (concentration of stiffener – 7 ml/l) was added. So, stiffener concentrations for this group were 0; 5; 7; 10; 15 ml/l, respectively. Before testing specimens were kept in standard atmosphere conditions (temperature – 20 ±2 ºC, humidity – 65 ±2 %) not less than for 24 h.

Judge panel consisting of ten experts (researches and students from the textile and clothing sectors) was chosen for fabric hand subjective evaluation. Experience of the other researches shows that it is important to control the climatic conditions where the subjective evaluation is carried out [1 – 4]. For this reason, subjective evaluation was performed in standard atmosphere conditions with all groups of samples. The samples for testing were 300 mm × 300 mm. Definitions and assessment techniques of fabric hand attributes are given in Table 1 [1].

In order to minimize the influence of fabric appearance on the perception of other attributes the "black box" with two hand holes was used. In such a way blind subjective evaluation was performed.

After establishing the assessment methodology judges were trained individually to use the prescribed techniques. They were also provided with explanatory and visual information how to assess these features. Descriptions how to assess each fabric attribute are given in Table 1. Later samples were given for the experts one after the other in mixed order and they were asked to rank the fabrics (e.g. from less hand to more hand).

The experts individually evaluated and scored each

Table 1. Definitions and assessment techniques of fabric hand attributes

Attribute	Definition	Assessment technique
Hardness	Non-resistance / resistance to compression or bending.	To put down the fabric on the base of "black box" and to compress it for three times. Afterward to take between the fingers and to flex the corners.
Stiffness	Flexible (not stiff) / stiff to bending.	Fabric sample is taken in to the palm where it is clenched and unclenched for three times.
Stiffness*		Fabric sample is held between two fingers in one hand and swept from top to bottom with the palm of the other hand.
Flexibility	Non-resistance / resistance to bending. The more floppy it is and the closer it follows the line of the knuckles, the more flexible it is.	To get the sample between the thumb and index finger so that it "drapes" down across the knuckles. "Flip" it to and fro from one side of the hand to the other.
Softness	Resistance / non-resistance to compression or bending.	To put down the fabric on the base of "black box" and to press it with the palms for three times and to bent the corners.
Softness*		Fabric sample is held between two fingers in one hand and swept from top to bottom with the palm of the other hand.
Roughness	Large / small amount of small particles rise on fabric's surface.	To put down the fabric on the base of "black box" and with the light pressure to move the palm of the hand across the surface of the sample.
Smoothness	The surface of a smooth fabric will offer little resistance to slipping when rubbed.	Fabric is taken between two fingers of both hands and it is pulled by one hand so that it would slide between two fingers.
Smoothness*		Fabric sample is held between two fingers in one hand and swept from top to bottom with the palm of the other hand.
Stretchability (in warp and in weft)	Degree to which a fabric stretches without tearing from its original shape.	The edges of the sample are held with both hands then stretched for three times in the same direction.
Resiliency (in warp and in weft)	Rate at which a fabric returns to its original position after deformation was removed.	The edges of the sample are held with both hands then stretched for three times in the same direction and left to return back to initial position.

* Fabric's hand attributes with special assessment technique.

fabric attribute at their own speed; the time of each evaluation session was limited to one hour, because hands become less sensible if the test is too long.

Fabric properties differ in different directions. Therefore experts were instructed to evaluate "stretchability" and "resiliency" in warp and weft directions.

Kendall's coefficient of concordance W [10] was used to determine the level of agreement between experts:

$$W = \frac{12\left[\sum_{j=1}^{n}\left(Rj - \overline{R}\right)^2\right]}{r^2 n(n-1)(n+1)},$$

where R_j is the sum of ranks given to each fabric sample; $\overline{R}$ is the mean values of rank sums; R is the number of experts; N is the number of samples.

RESULTS AND DISCUSSION

The ability of judges to discriminate between fabrics clearly depends on the particular attribute being assessed. It is therefore necessary to analyze separately the data generated for each fabric attribute descriptor.

The assessment of fabric hand subjective evaluation can be performed by a twofold interpretation. The first is done by analyzing the agreement level changes on the basis of Kendall's coefficient of concordance W for individual fabric's hand attributes. Whereas the second consists of area comparison of radial diagrams for total hand evaluation (if the area is larger, then the agreement is better).

In order to observe the rating tendency in each group of samples, the average values of each fabric attribute were used to analyze the accuracy in subjective hand evaluation and Kendall's coefficient of concordance was computed.

The results of subjective hand evaluation are illustrated by radial diagrams with 11 axes corresponding to each individual fabric hand attribute. Radial diagrams for three training sessions when concentration of stiffener differed by 3 ml/l, by 5 ml/l and by 7 ml/l are presented in Fig. 1, a – c). Diagram illustrating the reliability of judge panel's upon training number when concentration of stiffener differed by 5 ml/l is given in Fig. 1, d. As stated above the fourth group of samples was prepared in such a way: the specimen with 7 m/l concentration of stiffener was added to the second group of samples (difference of stiffener concentration – 5 ml/l).

The results indicated that for the first group (difference of stiffener concentration – 3 ml/l), during the first training experts were in good agreement evaluating fabric "hardness" ($W = 0.85$), "stiffness*" ($W = 0.79$), "softness" ($W = 0.78$) (see Fig. 1, a). While agreement level for "stretchability" and "resiliency" in both principal directions was poor ($W = 0.13 \div 0.63$). During the second training experts showed higher agreement evaluating all fabric attributes ($W = 0.68 \div 1.0$). The best agreement level was reached evaluating "softness" ($W = 1.0$). During the third training session the best agreement level was reached evaluating "stiffness*" ($W = 1.0$) and highest disagreement was reached evaluating "stretchability" and "resiliency" in warp direction ($W = 0.74 \div 0.78$).

For the second group (difference of stiffener concentration – 5 ml/l) during the first training experts were in good agreement evaluating "hardness" ($W = 0.94$), "stiffness" ($W = 0.89$), "stiffness*" ($W = 0.82$), "flexibility" ($W = 0.78$), "softness" ($W = 0.87$), "softness*" ($W = 0.89$) (see Fig. 1, b). The experts' agreement level evaluating "roughness", "smoothness", "stretchability" in warp and "resiliency" in warp and in weft directions was at lower level ($W = 0.42 \div 0.65$). During the second training experts showed higher agreement evaluating all 11 fabric hand attributes ($W = 0.76 \div 1.0$). The third training showed that agreement between the experts was the highest. The best agreement level was reached evaluating "hardness", "stiffness", "stiffness*", "flexibility", "softness", "softness*" ($W = 1.0$), evaluating other fabric attributes – $W = 0.76 \div 0.96$.

In order to establish the reliability and sensitivity limit of subjective evaluation judge panel was provided with samples differing by 7 ml/l concentration of stiffener (see Fig. 1, c). During the first training session judge panel showed good agreement evaluating only "hardness", "stiffness", "stiffness*", "flexibility", "softness", "softness*" ($W = 0.87 \div 1.0$). After the second session we observe that they were in a good agreement evaluating all fabric attributes ($W = 0.8 \div 1.0$). During the third session the results were slightly higher ($W = 0.86 \div 1.0$).

So, during the first training the judge panel was not sensitive evaluating the first group of samples but after the training program the results showed that subjective evaluation was highly sensitive.

With the purpose to assess the reliability of judge panel's evaluation the fourth training session was performed. The comparison of four trainings for the group of samples with 5 ml/l difference showed that the judges were trusty. Hence, more experienced judges gave a higher level of overall agreement and are not confused even when new sample was added.

For the general perception the area of radial diagrams was determined, which presented the total hand subjective evaluation for all fabric attributes particularly for each training session (Table 2).

Table 2. The changes of total hand rate

Training	Changes of stiffener concentration		
	by 3 ml/l	by 5 ml/l	by 7 ml/l
1	0.54	0.73	0.52
2	0.84	0.86	0.94
3	0.88	0.93	0.96
4		0.93	

Results showed that after the first training the handle was better evaluated for the group where stiffener concentration differed by 5 ml/l: the total hand rate was 0.73. Evaluating the third group of fabrics (difference – 7 ml/l) experts had difficulties on perception of "roughness", "smoothness", "smoothness*", "stretchability" and "resiliency" so the total hand rate was 0.52 and it is 2 % less compared with the first group of samples (difference of stiffener concentration 3 ml/l). For the third group (difference of stiffener concentration – 7 ml/l) after the second training the total hand rate increased up to 0.94.

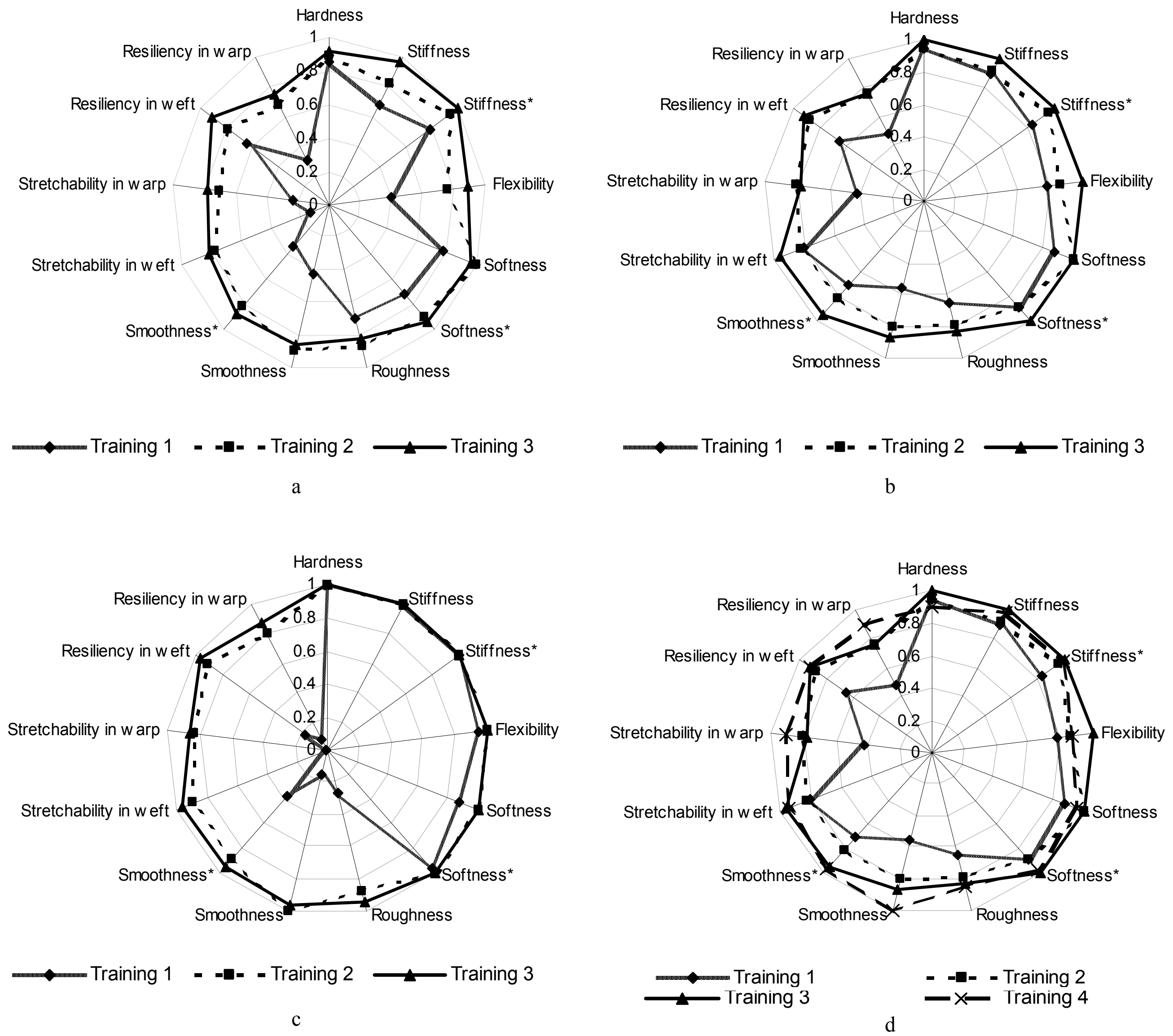

Fig. 1. Radial diagrams illustrating fabric's total hand subjective evaluation for three training sessions: a – difference of stiffener concentration 3 ml/l; b – difference of stiffener concentration 5 ml/l; c – difference of stiffener concentration 7 ml/l; d – reliability of judge panel's evaluation

The results showed that after third training session the areas of radial diagrams for all groups of samples were slightly higher compared to the second training: for the third group of samples the total hand rate increased up to 0.96.

These results led to important conclusions concerning the limits of subjective discrimination between fabrics having different characteristics. The experts' agreement level increased with the increase of difference between stiffener's concentrations in the groups of samples.

Assessing the reliability of fabric's hand subjective evaluation we can state that judge panel is trustworthy. During the fourth training the difference between samples with 5 ml/l and 7 ml/l concentration of stiffener was only 2 ml/l and it was expected that agreement between the experts will be lower, but results showed that total hand rate was the same as during the third training – 0.93.

In order to evaluate the quality of the assessment technique subjective evaluation of such fabric attributes as "stiffness", "softness" and "smoothness" was performed in two ways (Table 1). The more acceptable method for experts evaluating all three fabric attributes was – fabric sample was held between two fingers in one hand and swept from top to bottom with the palm of the other hand, compared to "stiffness*", "softness*" and "smoothness*". This method is reliable because it was unopposed between the experts, in this instance they felt more than only one attribute and tactile sensation was deeper.

Meantime the judges had difficulties evaluating "roughness" and "smoothness", because, supposedly, the definition was not clear and the assessment technique was non-effective. The fabric's "smoothness" or "roughness" was compared with it "stiffness": the roughest specimen was taken as the stiffest one and otherwise. The evaluation of "stretchability" and "resiliency" for judges was questionable. These fabric attributes were compared with fabric "stiffness", too. In their opinion the most resilient specimen is the stiffest one and otherwise.

The fabric "stiffness" apparently distracted the judges' ability to concentrate on "smoothness", "roughness", "stretchability" and "resiliency" as a single attribute. This tend to confirm that the subjective perception of each fabric attribute such as "roughness" or "resiliency", can be represented by a combination of objectively measurable properties and that this combination of properties can vary with fabric construction or end-use. This also implies that the interpretation of each fabric attribute descriptor can be different for fabrics having markedly different constructions or end-uses.

So, to proof the effect of judge panel's reliability upon training number the relationships between agreement level changes for individual fabric's hand attributes in 4 trainings for the group of samples differed by 5 ml/l were determined (see Fig. 2).

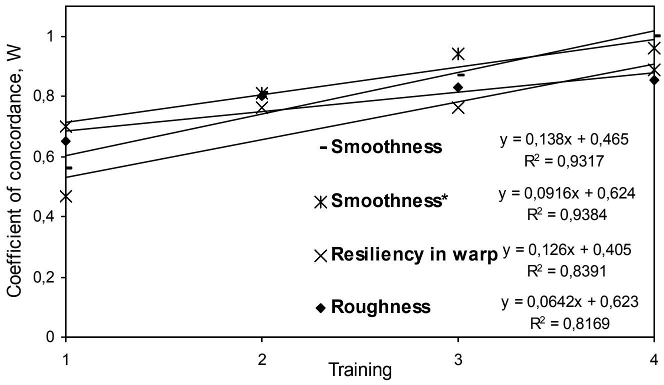

Fig. 2. Agreement level changes versus training sessions for individual fabric's hand attributes

The dependencies between agreement level and the number of training sessions can be described by linear equation $y = a + bx$. The coefficient of determination defined the distribution of agreement level between the experts evaluating each fabric attribute at different trainings. Subjective evaluation of fabric attributes as "smoothness", "smoothness*", "roughness" and "resiliency" in warp direction during the first training was very complex therefore during the following trainings experts' agreement level increased and coefficient of determination R^2 evaluating these fabric attributes varying in the range between 0.8169 and 0.9384. Other coefficients of determination of fabric's hand attributes in 4 trainings are given in Table 3.

Table 3. Coefficients of determination

Hardness	$R^2 = 0.05$	$y = -0.01x + 0.96$
Stiffness	$R^2 = 0.68$	$y = 0.03x + 0.87$
Stiffness*	$R^2 = 0.76$	$y = 0.05x + 0.79$
Flexibility	$R^2 = 0.36$	$y = 0.04x + 0.77$
Softness	$R^2 = 0.31$	$y = 0.03x + 0.89$
Softness*	$R^2 = 0.63$	$y = 0.03x + 0.85$
Stretchability (in warp)	$R^2 = 0.75$	$y = 0.14x + 0.37$
Stretchability (in weft)	$R^2 = 0.78$	$y = 0.05x + 0.75$
Resiliency (in weft)	$R^2 = 0.74$	$y = 0.09x + 0.63$

CONCLUSIONS

In the present work the analysis of subjective assessment was established.

Significant or non-significant agreement between the experts depends on more (difference between stiffener concentration 7 ml/l) or less (3 ml/l) perceptible differences between tested fabrics. Selected techniques for fabric's hand assessment showed high reliability after four trainings. Therefore the whole panel is trustful for total hand evaluation.

The experts' agreement level increased with the increase of difference between concentrations of stiffener in the groups of samples. Comparing the first and the fourth training sessions when concentration of stiffener differed by 5 ml/l the total hand rate increased up to 0.93. The more noticeable total hand rate difference between the first and last trainings was 44 % when the change of stiffener concentration differed by 7 ml/l.

The judges had difficulties evaluating "smoothness", "roughness", "stretchability" and "resiliency" as a single attribute because these fabric attributes they compared with the fabric "stiffness".

The dependencies between agreement level and training number can be described by linear equation $y = a + bx$. The coefficients of determination R^2 evaluating "smoothness", "smoothness*", "roughness" and "resiliency" in warp direction varies in the range between 0.8169 and 0.9384.

REFERENCES

1. **Bishop, D. P.** Fabrics: Sensory and Mechanical Properties *Journal of the Textile Institute* 26 (3) 1996: pp. 1 – 63.
2. **Ellis, B. C., Gransworthy, R. K.** A Review of Techniques for the Assessment of Hand *Textile Research Journal* 50 (4) 1980: pp. 231 – 238.
3. **Howorth, W. S., Oliver, P. H.** The Application of Multiple Factor Analysis to the Assessment of Fabric Handle *Journal of the Textile Institute* 49 (11) 1958: pp. T540 – T553.
4. **Howorth, W. S.** The Handle of Suiting, Lingerie, and Dress Fabrics *Journal of the Textile Institute* 55 (4) 1964: pp. T251 – T256.
5. **Brand, R. H.** Measurement of Fabric Aesthetics: Analysis of Aesthetic Components *Textile Research Journal* 34 1994: pp. 791 – 804.
6. **Hui, C. L., Lau, T. W., Ng, S. F., Chan, K. C. C.** Neural Network Prediction of Human Psychological Perceptions of Fabric Hand *Textile Research Journal* 74 (5) 2004: pp. 375 – 383.
7. **Yick, K. L., Cheng, K. P. S., How, Y. L.** Subjective and Objective Evaluation of Men's Shirting Fabrics *International Journal of Clothing Science and Technology* 7 (4) 1995: pp. 17 – 29.
8. **Cardello, V. A., Winterhalter, C., Schutz, H. G.** Predicting the Handle and Comfort of Military Clothing Fabrics from Sensory and Instrumental Data: Development and Application of New Psychophysical Methods *Textile Research Journal* 73 (3) 2003: pp. 221 – 237.
9. **Grinevičiūtė, D., Gutauskas, M.** The Comparison of Methods for the Evaluation of Woven Fabric Hand *Materials Science* (*Medžiagotyra*) 10 (1) 2004: pp. 97 – 100.
10. **Hartman, K., Lezki, E., Schafer, W.** Experiment Planning and Evaluation of Technological Processes. Moscow, 1977: 552 p. (in Russian).