Applying STRESS to Evaluate Variabilities in Consumer Responses in the Textile Industry

Zhenhua Luo, Jiajun Liu, Jiufeng Liu, Lori Rothenberg, and Renzo Shamey*
North Carolina State University, USA
*correspondence: rshamey@ncsu.edu

Keywords: STRESS, multiple dimensional scaling, consumer

Background and objective. Textile products and apparel are often evaluated by consumers subjectively. Assessments include the appearance (including color), hand, fragrance, and texture, among other characteristics. The subjective nature of human sensory observations, however, necessitates repeated trials to obtain statistically sound responses. Repeated trials by an observer in multiple assessments or responses among a group of observers are subject to different levels of variability, which must be addressed to improve confidence in reported outcomes. MultiDimensional Scaling (MDS) is a multivariate analytical technique, which was originally developed within the domain of psychometrics for measuring human preference and perceptions. Applications of MDS expanded in a wide range of fields, including travel, migration, health behavior (Dunfield, 1996), food science, color science, hazard perceptions, resource management, retail and residential choices, and market research. The basic idea of MDS involves presenting n objects by n corresponding points in a geometric space, where the inter-point distances approximate dissimilarities between pairs of objects. The objectives of MDS are to determine a spatial configuration in a certain dimensionality whose distances fit the dissimilarities best. A metric index denoted STRESS (STandardized REsidual Sum of Squares), was proposed by Kruskal (Kruskal, 1964), to evaluate the goodness of fit between the configuration and the dissimilarities.

A modified form of this metric is currently employed in the color science domain to test the performance of various color-difference formulas, a popular topic with noteworthy needs in many industries, including textile, automotive, and paper industries, among others. It has been shown that STRESS outperforms many other conventional metrics (ρ, γ, PF/3, PF/4, etc.) in various aspects, especially in its ability to indicate the statistical significance of the difference between two different datasets. STRESS is similar to a coefficient of variation, and its values are confined between 0-1 (or 0-100), where smaller values indicate stronger agreement between datasets, and as such the index presents a simple means of comparison between datasets. Moreover, STRESS benefits from having a symmetry property, which means the same results are achieved independently of the order of the two lists of input values. As such, the usage of the index has been extended to determining the goodness-of-fit between any two sets of results (García, Huertas, Melgosa, & Cui, 2007). STRESS model can be represented by Equation 1, where X and Y constitute two sets of data under consideration, and N indicates the number of data pairs.

$$STRESS = 100\left(\frac{\sum_{i=1}^{N}\left(X_i - F_1\ Y_i\right)^2}{\sum_{i=1}^{N} F_1^2 Y_i^2}\right)^{1/2} \quad \text{where } F_1 = \frac{\sum_{i=1}^{N} X_i^2}{\sum_{i=1}^{N} X_i Y_i} \qquad \text{(Eq. 1)}$$

More recently STRESS was recommended as a metric for determination of the degree of inter-individual variability (deviation between each individual's responses from the mean results obtained from a panel, also known as reproducibility) and the intra-individual variability (deviation among results of an individual in repeated trials in an experiment, also known as repeatability) in various fields related to human sensory responses (Shamey, Cárdenas, Hinks, & Woodard, 2010; García et al., 2011; Barcenas, Pérez Elortondo, & Albisu, 2004). Despite its advantages, however, only very few studies in the textile and apparel industry report the application of the STRESS index for evaluation of variability among subjects.

The objectives of this work are twofold. First, we report the application of STRESS for evaluation of inter-and intra-observer variability for 78 observers' preferences for apparel colors. Secondly, we aim at recommending STRESS as a useful general metric for evaluation of variability among consumers' behavioral and perceptual observations in the textile and apparel industries.

Case studies. Volunteer observers between the ages of 18 and 45 participated in a study that involved assessing preference for the color of t-shirts worn by simulated models (see Figure 1) that were displayed on a color-calibrated monitor. Assessments were conducted over three independent trials and included examinations of ten different colors. The main objective of this work is the examination of the utility of STRESS for assessment of the level of variability among repeated trials by the same observer and the extent of variation between mean responses of a given observer against grand mean values from all observations.

Assessment results for the case study described above indicate that the highest agreement in observers' preference ratings was in the selection of black colors. In terms of intra-observer variability, mean STRESS values for participants' repeatability between Trials 1&2 and Trials 2&3 were 19.58 and 15.79, respectively. These results indicate that preferences for garment color remained consistent over different assessment sessions. The reduction in the magnitude of STRESS from Trials 1&2 to Trials 2&3 may also suggest a slight training effect. The analysis of responses from all observers showed the repetition of the visual assessment resulted in negative ΔSTRESS for the majority of observers for comparisons between Trials 2 and 3 against those from Trials 1 and 2 (as shown in Figure 2). A similar analysis could be employed to screen inconsistent observer responses from overall assessments in different studies. In terms of inter-observer variability, the experimental results in the case study showed average STRESS values of 24.38, 23.46, and 23.11 for the three trials, respectively, which indicate consistent responses among individual preferences.

Figure 1. A simulated model dressed in a red t-shirt.

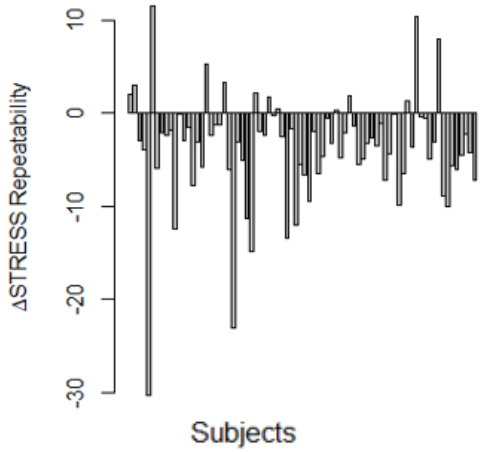

Figure 2. Subject repeatability evaluated by the difference in STRESS repeatability for Trial 2 and 3 minus STRESS repeatability for Trial 1 and 2.

Several additional case studies may be used to provide additional insight into STRESS utilization and its potential application as a metric in the apparel and textiles domains. Shamey et al. (Shamey et al., 2010) applied STRESS to compare variability in the visual assessment of small color differences of textiles. The study involved assessing 27 pairs of colored textile samples using 25 novice color assessors, who had no previous commercial color difference assessment experience, and 25 expert assessors who routinely engaged in the shade matching of textiles industry. Their results showed that the mean STRESS values for three trials conducted by the novice participants were 37, 36, and 34. Meanwhile, the experts exhibited a comparable average STRESS value of 36. The authors argued that based on STRESS results, naïve observers could be utilized for quality control of color in the supply chain. In another case study, Barcenas et al. (Barcenas et al., 2004) investigated the similarities/dissimilarities between a trained panel and naïve consumers in a sensory assessment of ewes milk cheeses. The mean STRESS value for the untrained panelists for three replicates was 29.4, while the average value obtained from the expert group was moderately lower of 27.2. The results revealed a slightly lower homogeneity degree for the consumer group than the trained panel. It can be argued, however, that the significance of the difference between the mean results of the two groups should be tested to determine whether the two observer panels were significantly different.

Conclusion. These studies and others show that STRESS function can be employed to analyze and interpret repeatability and reproducibility of sensory and perceptual evaluations in different

domains and is thus proposed as a useful general metric for assessment of variability in sensory evaluation of products in the apparel and textile industries.

References.

Barcenas, P., Pérez Elortondo, F. J., & Albisu, M. (2004). Projective mapping in sensory analysis of ewes milk cheeses: A study on consumers and trained panel performance. *Food Research International*, *37*(7), 723–729.

Dunfield, J. F. (1996). Consumer perceptions of health care quality and the utilization of non-conventional therapy. *Social Science and Medicine*, *43*(2), 149–161.

García, P. A., Gómez-Robledo, L., Melgosa, M., Shamey, R., Luo, M. R., Hinks, D., & Cui, G. (2011). Notes on the application of the standardized residual sum of squares index for the assessment of intra- and inter-observer variability in color-difference experiments. *Journal of the Optical Society of America A*, *28*(5), 949.

García, P. a, Huertas, R., Melgosa, M., & Cui, G. (2007). Measurement of the relationship between perceived and computed color differences. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *24*(7), 1823–1829.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*(1), 1-27.

Shamey, R., Cárdenas, L. M., Hinks, D., & Woodard, R. (2010). Comparison of naïve and expert subjects in the assessment of small color differences. *Journal of the Optical Society of America A*, *27*(6), 1482.