Reprinted from IFPress (2002) 5(1)2,3

Discrimination Testing with Multiple Samples

Benoît Rousseau and Daniel M. Ennis

Background: The fact that variation occurs in the manufacture of consumer products leads to a challenging problem in difference testing. A product with a particular identity may be composed of different variants due to the fact that it is produced at different locations or on different machines. Difference testing designed to evaluate modifications of this product should consider the fact that the product is composed of variants. Classical difference tests such as the triangular, duo-trio or *m*-alternative forced choice methods are not capable of accounting for the possibility that a product may exist as one of several variants. In this report, an extension of the duo-trio method is used to quantify additional variation such as that caused by different production facilities.

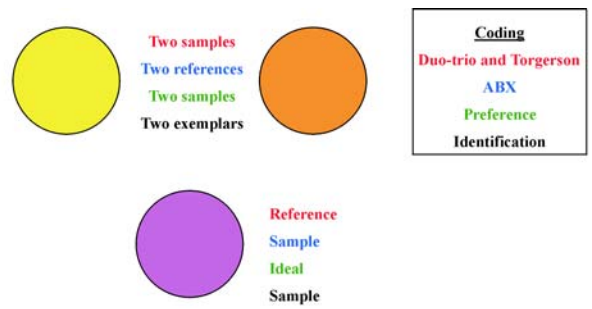
Scenario: Your supplier has proposed an ingredient change for your vanilla flavored yogurt and you would like to know what effect this change has on your product. Your inquiry is complicated by the fact that your main factory has two production lines that might produce slightly different products. The current product is labeled "A", while the reformulated product is labeled "B". Subscripts on A and B refer to different production lines. In a trial a subject is given three products, for instance $\mathbf{A_1}A_2B_1$, in which the first sample ($\mathbf{A_1}$) is a reference. The task of the subject is to select one of the two alternatives (A_2 or B_1) most similar to the reference. Each subject evaluates one of each of the 12 possible triads. This method is referred to as Torgerson's method of triads. The results are presented in Table 1.

Table 1. Number of subjects choosing the first sample of the alternative pair as most similar to the reference.

Triad	#	Total	Triad	#	Total
$A_1A_2B_1$	15	50	$A_2A_1B_1$	22	50
$A_1A_2B_2$	35	50	$A_2A_1B_2$	17	50
$A_1B_1B_2$	35	50	$A_2B_1B_2$	24	50
$B_1A_1A_2$	26	50	$B_2A_1A_2$	17	50
$B_1A_1B_2$	32	50	$B_2A_1B_1$	16	50
$B_1A_2B_2$	33	50	$B_2A_2B_1$	35	50

Torgerson's Method of Triads: It can be seen from a description of a trial given above, that Torgerson's method of triads is an extension of the duo-trio method. Unlike the duo-trio method, all three samples may be of different products. This paradigm was first proposed by Torgerson[1]. A Thurstonian model for this method is the same as that used for the ABX method, preferential choice[2] (paired preference test) and two-stimulus identification, in which two alternatives are compared to a reference point[3]. Figure 1 illustrates the similarity of these different protocols.

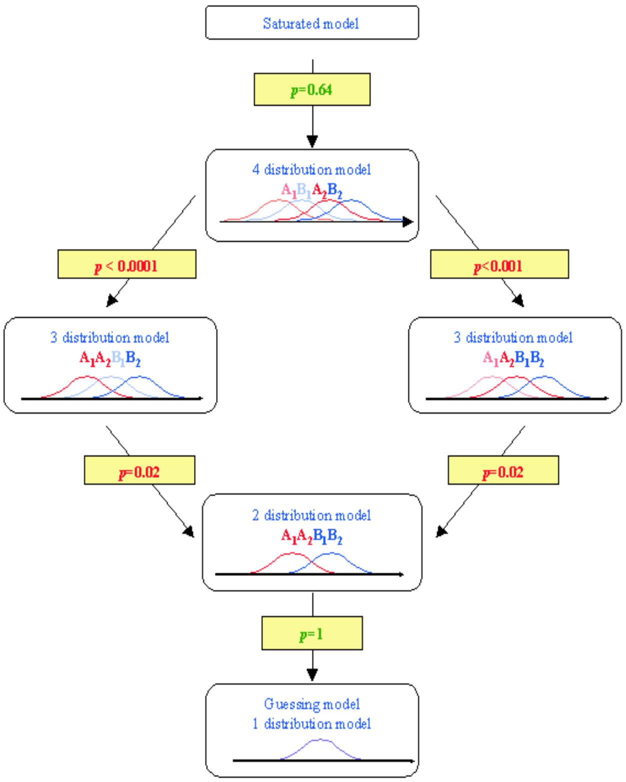
Figure 1. Similarities of the duo-trio method, Torgerson's method, the ABX method, preferential choice and two-stimulus identification.



Generalizations of other protocols also exist. These include Richardson's method of triads[4], which is a generalization of the triangular method and the multiple dual-pair[5] method, which is a generalization of the dual-pair or 4IAX procedure. A unidimensional Thurstonian model for Torgerson's method has been published[4]. In a manner similar to that used for traditional discrimination tests[6, 7], analysis of data from Torgerson's method provides scaled product similarities in units of the perceptual standard deviation. These scaled units separating products are commonly referred to as d´ values. If four products are compared, analysis of data from Torgerson's method will provide the 6 possible d´ values as well as their variances.

Fitting Hierarchical Thurstonian Models: In the scenario presented above, it is possible that all four products (2 products × 2 production lines) may be different. It is also possible that there are no differences due to the production lines and that the only difference is that due to the product modification. It is even conceivable that the modification has no effect. Which of these or other alternatives represents the best account of the data? Using the method of maximum likelihood, various models can be fit to the data. These models are presented in a hierarchical framework in Figure 2. Each node in this figure represents a model. Between the nodes there are probabilities obtained from statistical tests on differences between the models. The most complex model is presented at the top in which the number of estimated parameters equals the number of different triads. The simplest model, which is the guessing model, has no estimated parameters and is presented at the bottom. Between these extremes there are other models that require fewer parameters than the saturated model but more parameters than the guessing model. In order to find the appropriate model for your experiment, imagine that you are trying to find your way from the top of the figure to the bottom through gates that are only open if the inter-node test probability is greater than 0.05. The goal is to find the model with the fewest number

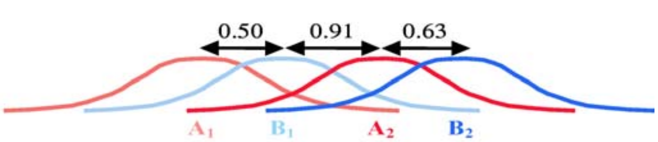
Figure 2. Hierarchical model structure for the vanilla yogurt study (the distributions given are for illustration purposes.)



of parameters that still explains the data. If there is a significant difference between the saturated model and the model immediately below it, then unidimensional Thurstonian modeling may not be adequate. This can occur, for instance, if the percepts are multidimensional. This type of model for Torgerson's method will be discussed in a future report.

Calculating d´ Values with Torgerson's Method: Based on the results presented in Table 1, you use the unidimensional Thurstonian model to analyze the results for the vanilla yogurt study. The model comparisons are given in Figure 2, while the d´ values are given in Table 2 and illustrated in Figure 3.

Figure 3. Vanilla flavored yogurts: Representation of product similarities using the unidimensional Thurstonian model.



The model that assumes four distinct distributions is found to be not significantly different from the saturated model (p=0.64), while both models with only 3 distributions are significantly different from the model with 4 distributions. This shows that the unidimensional model is sufficient to describe the product similarities and that a model with fewer parameters is not justified. In our imaginary trip down Figure 2, the gates below the 4-distribution model are closed. Taking into account the production line variation, the overall weighted average d´ value between A and B is 0.50, which is significantly different from 0. The d´ value induced by the production lines is a weighted average of 1.47. This means that, while the new ingredient significantly altered your vanilla yogurt, the extent of this change is smaller than the existing difference between your two production lines. If the difference between the two lines is considered acceptable, the product with the new ingredient may be an adequate alternative.

Table 2. Vanilla flavored yogurts: d´ values and *their variances* for all possible product pairs (distance between the product in first column and the product in first row.)

Product	A_2	B_1	B_2
A_1	1.407	0.50	2.04
	0.033	*0.029*	*0.036*
A_2	-	-0.91	0.63
		0.03	*0.029*
B_1	-	-	1.54
			0.033

Conclusion: Torgerson's method of triads is a practical way to provide simultaneous estimation of multiple product similarities. Applications include the comparison of multiple reformulations to a current standard product, or the consideration of batch-to-batch variability when investigating any kind of product change. The method has potential to be useful in product testing when simultaneous comparisons of more than two products are required.

References:

1. Torgerson, W. S. (1958). *Theory and Methods of Scaling*. New York: Wiley.
2. Ennis, D.M. (1999). Multivariate preference mapping. *IFPress*, **2(2)**, 2-3.
3. Ennis, D.M. (1993). A single multidimensional model for discrimination, identification and preferential choice. *Acta Psychologica*, **84**, 17-27.
4. Ennis, D. M., Mullen, K., & Frijters, J. E. (1988). Variants of the method of triads: Unidimensional Thurstonian models. *British Journal of Mathematical and Statistical Psychology*, **41**, 25-36.
5. Rousseau, B & Ennis, D.M. (2002). The multiple dual-pair. *Perception & Psychophysics*, *In Press.*
6. Ennis, D.M. (1998). Thurstonian scaling for difference tests. *IFPress*, **1(3)**, 2-3.
7. Rousseau, B. and Ennis, D.M. (2001). How retasting can improve the power of product testing. *IFPress*, **4(2)**, 2-3.