Differences in Preference Mapping Techniques

Renoo Yenket, Edgar Chambers IV, and Koushik Adhikari, The Sensory Analysis Center, Department of Human Nutrition,
Kansas State University, Manhattan, KS 66506, USA

Introduction

Several programs offer solutions to understand the influences and relationships between a descriptive sensory profile and consumer liking, but researchers need to know the (dis)advantages and (dis)similarities of these programs. Few studies compare co-ordinate outputs to strengths and weaknesses of software and preference mapping techniques. By contrasting the range of outputs, this study compares the advantages and disadvantages of each software/technique. The most common are: internal preference mapping, using only data of consumer acceptability (MDPREF); principal components analysis on descriptive characteristics together with consumers' mean liking scores (PCA-liking method); Autofit in SAS® (PREFMAP) either with PCA, CVA (canonical variate analysis) or GPA (generalized procrustes analysis); and partial least squares regression (PLS1 or PLS2) available with or without the passify option) for product inclusion in the correlation plot.

Method

Four statistical software programs and 7 multivariate techniques with some variants were applied to a milk study (21 descriptors, 6 samples for descriptive study, 115 consumers, and 9-point hedonic scale for consumer study), a paint study (6 descriptors, 10 samples for descriptive study, 98 consumers, and 10-point hedonic scale for consumer study), and a fragrance study (56 descriptors, 22 samples for descriptive study, 321 consumers, and 7-point hedonic scale for consumer study). The objective of this study was to investigate differences, if any, in: 1) the parameters' spatial positions, 2) Euclidean distances (proximity of the mean-liking or individual consumers to their most liked products), and 3) percent variance explained by the first two dimensions. The statistics programs used were SAS®, XLSTAT®, Unscrambler®, and Senstools®.

Results and Discussion

TABLE 1: PARTIAL CO-ORDINATES OF DESCRIPTIVE ATTRIBUTES, SAMPLES, AND CONSUMER-MEAN-LIKING VECTOR/INDIVIDUAL CONSUMER'S VECTORS FOR MILK STUDY ON COMPONENTS 1 AND 2

Method	Component	Descriptive attributes		Mean liking	Products		Consumer	
		Fat feel	Viscosity		LFA0	LFA2	C1	C2
MDPREF	1				-1.000	-0.195	1.000	0.723
	2				-0.128	0.390	0.003	-0.529
PCA	1	0.929	0.965	0.497	-1.000	0.442		
	2	0.245	0.237	-0.792	0.864	0.688		
Prefmap	1	0.939	0.952		-1.000	0.367	0.161	0.278
	2	0.317	0.260		1.000	0.743	-0.908	-0.435
CVA	1	-0.211	-0.300		1.000	-0.246	-0.248	-0.367
	2	0.468	-0.395		0.911	0.870	-0.976	-0.437
SAS® PLS1	1	0.703	0.748	0.596	-1.000	0.099		
	2	-0.655	-0.633	0.772	-0.182	-0.640		
XLSTAT® PLS1	1	0.688	0.732	0.772	-0.819	0.081		
	2	-0.636	-0.615	0.419	-0.123	-0.432		
Unscrambler® PLS1	1	0.688	0.732	0.772	-1.000	0.099		
	2	-0.636	-0.615	0.419	-0.182	-0.640		
Unscrambler® PLS1 (passified)	1	0.688	0.732	0.772	-0.819	0.081		
	2	-0.636	-0.615	0.419	-0.123	-0.432		
SAS® PLS2	1	0.889	0.930		-1.000	0.339	0.331	0.735
	2	-0.406	-0.400		-0.637	-0.691	0.879	0.643
XLSTAT® PLS2	1	0.878	0.918		-0.725	0.246	0.308	0.685
	2	-0.348	-0.343		-0.361	-0.391	0.854	0.624
Unscrambler® PLS2	1	0.878	0.918		-1.000	0.339	0.308	0.685
	2	-0.348	-0.343		-0.637	-0.691	0.854	0.624
Unscrambler® PLS2 (passified)	1	0.878	0.918		-0.725	0.246	0.308	0.685
	2	-0.348	-0.343		-0.361	-0.391	0.854	0.624
Senstools® GPA	1	-0.736	-0.802		1.000	-0.377	-0.101	-0.171
	2	0.241	0.190		0.657	0.896	-0.955	-0.460
XLSTAT® GPA	1	-0.946	-0.966		1.000	-0.379	-0.101	-0.172
	2	0.210	0.254		0.661	0.894	-0.954	-0.460

TABLE 2: PARTIAL CO-ORDINATES OF DESCRIPTIVE ATTRIBUTES, SAMPLES, AND CONSUMER-MEAN-LIKING VECTOR/INDIVIDUAL CONSUMERS' VECTORS FOR FRAGRANCE STUDY ON COMPONENTS 1 AND 2

Method	Component	Descriptive attributes		Mean liking	Products		Consumer	
		Strength	Lemon		517	621	C1	C2
MDPREF	1				0.56	0.50	0.29	0.53
	2				0.33	0.85	0.68	-0.46
PCA	1	0.27	-0.25	0.08	0.41	-0.83		
	2	0.72	-0.80	0.45	0.02	-0.10		
Prefmap	1	0.24	-0.11		-0.46	1.00	0.78	0.04
	2	0.99	-0.84		0.19	-0.61	-0.76	0.58
CVA	1	0.19	0.05		-0.06	0.67	0.50	0.16
	2	0.61	-0.15		0.33	-0.84	-0.79	0.48
SAS® PLS1	1	0.72	-0.57	0.53	0.29	-0.04		
	2	-0.55	1.00	0.83	0.37	0.92		
Unscrambler® PLS1	1	0.47	-0.37	0.73	0.29	-0.04		
	2	-0.36	0.65	0.55	0.37	0.92		
Unscrambler® PLS1 (passified)	1	0.47	-0.37	0.73	0.17	-0.02		
	2	-0.36	0.65	0.55	0.15	0.37		
SAS® PLS2	1	-0.58	0.60		-0.43	0.83	0.51	0.38
	2	0.37	-0.42		-0.09	0.28	0.39	-0.02
XLSTAT® PLS2	1	-0.48	0.50		-0.17	0.33	0.39	-0.40
	2	0.31	-0.35		-0.06	0.18	0.29	0.29
Unscrambler® PLS2	1	0.48	-0.50		0.43	-0.83	-0.39	0.40
	2	0.31	-0.35		-0.09	0.28	0.29	0.30
Unscrambler PLS2 (passified)	1	0.48	-0.50		0.17	-0.33	-0.39	0.40
	2	0.31	-0.35		-0.06	0.18	0.29	0.30
Senstools GPA	1	-0.24	0.15		0.47	-1.00	-0.78	-0.03
	2	1.00	-0.63		0.21	-0.61	-0.68	0.59
XLSTAT GPA	1	0.23	-0.11		-0.46	1.00	0.78	0.03
	2	-0.98	0.83		-0.21	0.61	0.68	-0.59

TABLE 3: PERCENTAGES REPRESENTING VECTOR(S) CLOSEST TO MOST-LIKED PRODUCTS, VARIANCE EXPLAINED AMONG DESCRIPTIVE ATTRIBUTES AND CONSUMERS

Method	%								
	Milk study			Paint study			Fragrance Study		
	Number of consumers who map closest to their most-liked products	Variance explained in descriptive attributes	Variance explained in consumers	Number of consumers who map closest to their most-liked products	Variance explained in descriptive attributes	Variance explained in consumers	Number of consumers who map closest to their most-liked products	Variance explained in descriptive attributes	Variance explained in consumers
MDPREF	56	na	61	55	na	67	32	na	29
PCA	REG3	85		593	98		513	44	
SAS® PLS1	REG3	85	78	399	98	95	237	26	83
XLSTAT® PLS1	REG3	85	78	593	98	94	na	na	na
Unscrambler® PLS1	REG3	85	78	593	98	94	237	26	83
Unscrambler® PLS1 (passified)	REG3	85	78	593	98	94	237	26	83
Prefmap	46	96	73	15	98	62	27	61	72
Senstools® GPA	42	83	81	na	na	na	26	52	80
XLSTAT® GPA	43	84	78	na	na	na	26	52	78
CVA	42	89	85	na	na	na	26	40	74
SAS® PLS2	55	86	53	56	98	60	22	44	17
XLSTAT® PLS2	55	86	54	68	98	60	27	45	17
Unscrambler® PLS2	56	86	53	44	98	60	25	45	17
Unscrambler®	55	86	53	68	98	60	27	45	17

Co-ordinate values: Separate techniques yield unequal co-ordinate values (Table 1-2). Different software yields dissimilar values in (un)passified PLS1, (un)passified PLS2, and GPA.

Parameters' spatial position: Products and their characteristics' orientations were similar. Consumer orientation and the number of matched/unmatched individual consumer positions to their "most liked" products gave some differences.

Attributes that promote liking: When the maps had a high percentage of variance explained in descriptive attributes (< 83%), methods used in this study suggested similar attributes that promote liking/acceptability. But it was hard to specify attributes when the maps had low percentages and contained individual consumers.

Variance explained by different maps: Percentage of variance explained in the descriptive attributes and consumers varied among the multivariate techniques, but did not shift across software programs for (un)passified PLS1, (un)passified PLS2, and GPA. The percent variance explained by SAS® and Unscrambler® software were similar for PLS1 or PLS2 methods regardless of whether the product positions were passified in Unscrambler® or not. As expected, MDPREF's variance explanation differed from those of the external preference mapping methods. Furthermore, among the external preference approaches, PREFMAP-PCA, PREFMAP-CVA, PLS1 and PLS2 were dissimilar in percent variance explained. The PREFMAP-PCA explained the maximum amount of variance related to descriptive characteristics, while the PLS1 maximum amount of variance related to consumer data.

Closeness of consumers to their most-liked products: Unlike internal preference mapping, all external preference mapping procedures initially place descriptive sensory attributes and product co-ordinates based on descriptive sensory data rather than consumer data. This does not guarantee consumers or the mean preference vector will be positioned close to their most-liked products. Therefore, being aware of how close individual consumer positions are to their most-liked products in a preference map will help researchers understand and use findings in the map with confidence. Among the 3 data sets, the PCA procedure on the paint study and the SAS® PLS1 procedure on the fragrance study failed to position consumer mean preference co-ordinates close to products with the highest mean-liking score. This demonstrates that those preference maps are potentially misleading when trying to understand liking vectors.

Method	Advantages	Disadvantages	Method	Advantages	Disadvantages
MDPREF	▪ Highest number of consumers who map closest to their most-liked products ▪ Product spacing on the map is different from other methods ▪ Simple and straightforward	▪ Not easy to overlay with the descriptive spacing. ▪ Difficult to identify the most-liked product when consumers do not agree on a most-liked product ▪ A variation of data explained by each component is often low (McEwan, 1996)	Unscrambler® PLS1 passified	▪ Up to standard and has a better chance than PCA of locating a mean liking vector closer to the most high-liking scored products ▪ Effortless to get the final map	▪ Tends to narrow the spacing of products, making the map very crowded in the middle ▪ Users may be unaware of extra steps programmers add to the process ▪ Maps are distorted
PCA	▪ Higher chance of mis-locating the mean liking vector to its highest-average-liking-scored product ▪ Incorporates the co-ordinate of the average consumer liking score vector so it indicates where the preferred product of this group of consumer is located. ▪ Simple, but not straight forward	▪ Co-ordinate of the average liking score vector is not necessarily correct ▪ When products differ greatly, results show in less than 50% variance explained; i.e., map not very helpful	**PLS2 in general**	▪ Shows where each consumer is located and helps identify different groups of consumer preferences; aids in positioning prototypes ▪ Difficult to define drivers of liking ▪ Presents high number of consumer co-ordinates near their most-liked products as with MDPREF	▪ Co-ordinates are widely spread when consumers do not agree
Prefmap	▪ Up to standard; maps are well spread out, and easy to read; however, maps are crowded with 321 consumers and 22 samples ▪ Products and attributes remain oriented the same on the map as PCA ▪ Shows positions of individual consumers and helps identify different clusters of consumer preference	▪ Requires familiarity with SAS®	SAS® PLS2	▪ Similar to Unscrambler® PLS2 for spatial wise.	▪ A challenging procedure; in addition to needing the ability to use SAS® language, this program requires background knowledge
CVA	▪ Up to standard ▪ Shows the map at a different angle ▪ Shows positions of individual consumers and helps to recognize different groups of consumer preference	▪ When data contains numerous samples, descriptive attributes, and consumers, they tend to crowd together ▪ Replication is required for a descriptive sensory study	XLSTAT® PLS2	▪ Co-ordinates values are similar to Unscrambler® PLS2 passified ▪ Yields a symmetry map	▪ Tends to narrow the spacing of products ▪ The descriptive data of each trained panelist must be arranged by width instead of length; risk running out of columns if not using Excel 2007 ▪ Users may be unaware of extra steps programmers add to the process
PLS1 in general	▪ Easier to identify drivers of liking	▪ Co-ordinates of an average liking score vector, generated by Unscrambler® and XLSTAT®, is not necessarily correct, e.g., in the paint study, the vector was located near 593 instead of 399 (Table 3) ▪ Suggests only one prototype that balances the intensity across all descriptive attributes to satisfy all consumers. Assumes the relationship between liking and sensory intensity for a given attribute is linear (Meullenet et al., 2007)	Unscrambler® PLS2	▪ Similar to Unscrambler® PLS2 for spatial of samples, descriptive attributes, and individual consumers.	▪ Users may be unaware of extra steps programmers add to the process ▪ Maps are distorted
SAS® PLS1	▪ Better in positioning the mean liking vector closer to products that were higher in their liking scores than XLSTAT® PLS1, Unscrambler® (with and without passified) PLS1, and PCA ▪ All co-ordinates are well spread out ▪ Easier to spot errors in the data/process ▪ Provides actual values if needed to formulate a predictive equations	▪ A challenging procedure; in addition to needing the ability to use SAS® language, this requires background knowledge	Unscrambler® PLS2 passified	▪ Co-ordinates values are similar to XLSTAT® PLS2 passified	▪ Users may be unaware of extra steps programmers add to the process ▪ Unscambler®'s maps are distorted
XLSTAT® PLS1	▪ Up to standard and has a better chance than PCA in locating a mean liking vector closer to the high-liking scored products ▪ Yields a symmetry map	▪ May not create a map (while other programs do) because of how the original data is configured ▪ The descriptive data of each trained panelist must be arranged by width instead of length; risk running out of columns if not using Excel 2007 ▪ Users may be unaware of extra steps programmers add to the process	**GPA in general**	▪ Spacing is similar to XLSTAT® and Senstools® maps ▪ XLSTAT® and Senstools® represent consumers who map closest to their most-liked products in the same distance ratio ▪ Equal in values of consumer and product co-ordinates	▪ Fewer consumers map close to their most-liked products ▪ Maps from averaged values in XLSTAT®/Senstools® programs are clearer than maps manually calculated from averaged values, because the program map tends to stretch the attribute space ▪ Replication is required for descriptive studies
Unscrambler® PLS1	▪ Up to standard and has a better chance than PCA in locating a mean liking vector closer to the high-liking scored products	▪ Users may be unaware of extra steps programmers add to the process ▪ Maps are distorted			

Conclusions

There were similarities and differences when the external preference maps from various techniques were compared. Products and their characteristics' orientations were similar, but consumer orientation and the number of matched/unmatched consumer positions to their highest scored product showed some differences. The choices of preference mapping techniques are affected by the user's expertise, time line, product knowledge, focus on consumers/products, size of data set, and product optimization. If researchers want to identify new product spaces for marketing, the internal preference map is sufficient. However, research involving marketing and development of products should consider MDPREF and one external preference mapping technique. When time is an issue, using: 1) PCA, 2) PLS1 with any software, or 3) PLS2 from XLSTAT®, Unscrambler®, or Senstools are wise choices.

References: MCEWAN, J.A. 1996. Preference mapping for product optimization. In *Multivariate Analysis of Data in Sensory Science* (T. Næs and E. Risvik, eds.) pp 71-102. Elsevier, NJ.

MEULLENET, J-F, XIONG, R., and FINDLAY, C. 2007. *Multivariate and Probabilistic Analyses of Sensory Science Problems.* pp.111-126. Blackwell Publishing, Ames, IA.