CONSUMER DIVERSITY IN SENSORY EVALUATION DATA

John C. Castura

Compusense Inc.
Guelph, Ontario, Canada
jcastura@compusense.com

Abstract

Ralph Waldo Emerson once cautioned that "a foolish consistency is the hobgoblin of little minds." In this talk, this quote is playfully considered in light of consumer diversity as manifested in sensory evaluation studies.

The talk begins with a focus on clustering of consumers based on hedonic data. This type of data is interesting. An individual consumer might have equal liking for dissimilar samples, and two consumers with the same preferences might assign different liking responses to the samples. Four types of studies are introduced:

- A study in which all consumers evaluate all products without replication. An experimental design that balances position and carryover effects seems appropriate, but there is a danger that consumer clusters will be strongly influenced by serving order, especially when the number of products evaluated per consumer is low. It will be discussed why this can happen, as well as potential solutions, such as subtracting the estimated contribution of order effects prior to cluster analysis.
- A study in which all consumers evaluate all products without replication, but give repeated measures. Examples of such data include consumer chewing gum evaluations, but could apply equally to other product categories in which liking responses are given at multiple predetermined time points. This type of study has all the challenges of the first type of data mentioned, but additional problems are posed by the repeated (autocorrelated) hedonic measurements.
- A study in which consumers evaluate only a subset of the products, without replication. This problem comes up often in product categories that do not lend themselves to more than a small number of evaluations by consumers in a single sitting, usually because the product type is fatiguing, potentially intoxicating (e.g. wine, beer), stimulating (coffee), or is particularly sating or exhausts finite resources (e.g. grooming with a razor removes the hair that is required for a razor evaluation). For 25 years the standard advice has been to use a balanced incomplete block (BIB) design, with the goal of balancing position and carryover effects, and this solution is perfect if the objective is to estimate average consumer liking. What if the objective is to understand consumer segments? Sensory-informed designs are presented as an alternative to a conventional balanced-incomplete-block design. These designs create subsets of samples that span the sensory space, exposing consumers to a higher level of product sensory variability than would otherwise be found in a BIB design. Clustering can be conducted in which consumers are allocated to groups, and imputation of values for samples not presented is conditional on group membership.
- A study in which consumers evaluate products with replication, for example in multiple sittings. These data are interesting consumers are often inconsistent with their liking responses. The problem of working with these data will be discussed in light of scenarios previously mentioned.

The second topic is the clustering of consumers based on sensory response data. Consumers' product characterizations have been of interest in recent years, and open the possibility of co-investigating relationships between consumer perception and liking. This time two types of studies will be introduced:

- A study in which consumers are presented with check-all-that-apply (CATA) questions, in which they can select all the terms from a list to describe the sample being evaluated. Consumer clustering is then possible based on CATA data, which enables an exploration of whether consumers differ in their responses across all response data. Consumer clusters are obtained and investigated.
- A study in which consumers evaluate products using temporal check-all-that-apply (TCATA). During each evaluation, consumers will indicate and continuously describe the sample by selecting as many attributes as are appropriate from a list. Starting from liking data collected in the same study, we perform conventional clustering then show how product perception differs across liking clusters.

Next we turn attention to clustering consumers based on TCATA data, and identify perception clusters.

A recurring theme is that the objectives of a sensory study need to guide its design, and that slavishly following standard conventions for design and analysis can be the foolish consistency which is the hobgoblin of little sensory scientists.

Keywords **-** Consumer heterogeneity; Unsupervised clustering; Experimental design; Liking; Temporal check-all-that-apply (TCATA)

INTRODUCTION

In the 1970 edition of The Structure of Scientific Revolutions, historian and philosopher of science Thomas S. Kuhn includes a postscript in which he states: "To the extent, of course, the individuals belong to the same group and thus share education, language, experience, and culture, we have good reason to suppose that their sensations are the same" (p. 193). Kuhn was a was not a pioneer of sensory evaluation, but the publication of his words seem to endorse a phenomenon that was just getting underway in the United States: the adoption in the 1970s and 1980s of a powerful new sensory evaluation method called quantitative descriptive analysis by food and beverage industries. A quantitative descriptive analysis panel is comprised of a small group of assessors who undergo an appropriate common training experience and who align their vocabularies can describe and quantify the sensations elicited by specific foods and beverages in an objective manner. The idea is that that panel will function as an analytical instrument. A key assumption underlies this method: product profiles obtained from the panel are assumed to be aligned with the human sensory experiences of some larger, untrained population, even if the individuals in this population cannot express these experiences due to lack of adequate sensory training.

But there is various evidence that foods might be experienced differently by different consumers in a non-laboratory setting due to complex factors. Consumers might differ in their genetics, receptor expression, age, dentition, microbial communities, salivation, habits for processing foods orally, strength of masticatory musculature, pregnancy status, neurodevelopmental disorder status, mental illness status, medication, sensory education, willingness to eat new foods, thermal tasting status, and other factors (see, e.g., [1-12]). What happens if we cannot assume that there is an average or representative set of sensations that go along with any particular product? And how does this affect the design and analysis of studies in which consumer diversity is a salient feature to be investigated? Ralph Waldo Emerson once cautioned that "a foolish consistency is the hobgoblin of little minds" [13]. In this talk, this quote is considered in light of consumer diversity as it manifests in sensory evaluation studies.

CONSUMER DIVERSITY IN HEDONIC SENSORY STUDIES

Three examples of sensory studies in which consumers provide hedonic responses are given.

First, there is the complete block design. This is a study in which **all consumers evaluate all products** without replication. An experimental design that balances position and carryover effects seems appropriate in this case. This approach will balance the serving order bias across all samples, which allow for proper estimation of average hedonic responses. But an average value is of interest when "average" is in some way typical. Now suppose that data arise from a bimodal population with diametrically opposed preferences: here the average value is no longer of interest. Cluster analysis is often conducted to determine whether data arise from different populations. But when there is a sample order effect and the number of samples is small, it has been reported that cluster analysis might produce clusters that are strongly influenced by the serving order [14]. It will be discussed why this can happen. The merits of potential solutions, such as subtracting the estimated contribution of order effects prior to cluster analysis, will be discussed.

Second, there is a complete block design with replication, but in which the protocol involves repeated measures. This is a study in which **all consumers evaluate all products at multiple predetermined time points**. Examples of such a study include consumer chewing gum evaluations. Evaluations of products that are worn or used for an extended period of time also fall into this category, as do dietary interventions that are intended to promote satiety. This type of study has all the challenges of the first type of data mentioned, but additional problems are posed by the repeated (autocorrelated) hedonic measurements.

Third, there is an incomplete block design without replication. This is a study in which **consumers evaluate only a subset of the total number of products**, without replication. This problem comes up often in product categories that do not lend themselves to more than a small number of evaluations by consumers in a single sitting, usually because the product type is fatiguing, potentially intoxicating (e.g. wine, beer), stimulating (coffee), or is particularly sating or exhausts finite resources (e.g. grooming with a razor removes the hair that is required for a razor evaluation). For 25 years the standard advice has been to use a balanced incomplete block design, with the goal of balancing position and carryover effects. This solution is perfect if the objective is to estimate average consumer liking. What if the objective is to understand consumer segments? This talk will discuss why there is a better solution than a conventional balanced-incomplete-block design. The sensory-informed design [15, 16] provides a key strategy for addressing this type of problem.

Fourth, there is a complete block design with replication. This is a study in which **consumers evaluate all products with replication**. This problem comes up when consumers are asked on multiple occasions to given their responses to products. In such a study we have both position effects and replicate effects to consider.

CONSUMER DIVERSITY IN PERCEPTION AND PERCEPTION DYNAMICS

Two examples of sensory studies in which consumers provide sensory response data are given. Consumers also provide liking data, which makes it possible to coinvestigate sensory and hedonic responses [17].

First, there is a consumer study in which consumers are presented with check-all-that-apply (CATA) questions, in which they can select all the terms from a list to describe the sample being evaluated. Consumer clustering is then possible based on CATA data, which enables an exploration of whether consumers differ in their responses across all response data. Consumer clusters are obtained and investigated based on several studies in which the product of interest was orange juice [18], strawberries [19, 20], and whole grain bread [21].

Second, there is a consumer study in which consumers evaluate products using temporal check-all-that-apply (TCATA). During each evaluation, consumers will indicate and continuously describe the sample by selecting as many attributes as are appropriate from a list. Starting from liking data collected in the same study, we perform conventional clustering then show how product perception differs across liking clusters [22]. Additional approaches are being investigated actively to enable consumer perception clusters to be obtained from TCATA data.

CONCLUSIONS

Consumers can be heterogeneous with respect to liking or with respect to perception, or both. It is possible to investigate conventional consumer sensory test data to better understand consumer diversity. This talk focuses on sensory data arising two types of studies: consumer hedonic data, and consumer perception (and perception dynamics) data. A recurring theme is that the objectives of a sensory study need to guide its design, and that slavishly following standard conventions for design and analysis can be the foolish consistency which is the hobgoblin of little sensory scientists.

REFERENCES

[1] Pickering, G. J., & Kvas, R. (2016). Thermal Tasting and Difference Thresholds for Prototypical Tastes in Wine. Chemosensory Perception, 9, 37-46.

[2] Jaeger, S. R., McRae, J. F., Bava, C. M., Beresford, M. K., Hunter, D., Jia, Y., et al. (2013). A Mendelian trait for olfactory sensitivity affects odor experience and food selection. Current Biology, 23, 1601-1605.

[3] Ghanizadeh, A., Bahrani, M., Miri, R., & Sahraian, A. (2012). Smell identification function in children with attention deficit hyperactivity disorder. Psychiatry Investigation, 9, 150-153.

[4] Ikebe, K., Matsuda, K., Kagawa, R., Enoki, K., Yoshida, M., Maeda, Y., & Nokubi, T. (2011). Association of masticatory performance with age, gender, number of teeth, occlusal force and salivary flow in Japanese older adults: is ageing a risk factor for masticatory dysfunction. Archives of Oral Biology, 56, 991-996.

[5] Foster, K. D., Grigor, J. M. V., Cheong, J. N., Yoo, M. J. Y., Bronlund, J. E., & Morgenstern, M. P. (2011). The role of oral processing in dynamic sensory perception. Journal of Food Science, 76, R49-R61.

[6] Quandt, S. A., Chen, H., Bell, R., Savoca, M. R., Anderson, A. M., Leng, X., Kohrman, T., Gilbert, G. H., & Arcury, T. A. (2010). Food avoidance and food modification practices of oral rural adults: association with oral health status and implications for service provision. Gerontologist, 50, 100-101.

[7] Mustonen, S., & Tuoria, H. (2010). Sensory education decreases food neophobia score and encourages trying unfamiliar foods in 8-12-year-old children. Food Quality and Preference, 21, 353-360.

[8] Douglass, R., & Heckman, G. (2010). Drug-related taste disturbance: a contributing factor in geriatric syndromes. Canadian Family Physician, 56, 1142-1147.

[9] Hayes, J. E., Bartoshuk, L. M., Kidd, J. R., & Duffy, V. B. (2008). Supertasting and PROP bitterness depends on more than the TAS2R38 gene. Chemical Senses, 33, 255-265.

[10] Bergdahl, M., & Bergdahl, J. (2002). Perceived taste disturbance in adults: prevalence and association with oral and psychological factors and medication. Clinical Oral Investigations, 6, 145-149.

[11] Kölble, N., Hummel, T., von Mering, R., Huch, A., & Huch, R. (2001). Gustatory and olfactory function in the first trimester of pregnancy. European Journal of Obstetrics & Gynecology and Reproductive Biology, 99, 179-183.

[12] Lavin, J. G., & Lawless, H. T. (1998). Effects of color and odor on judgments of sweetness among children and adults. Food Quality and Preference, 9, 283-289.

[13] Emerson, R. W. (1841/1993). “Self-Reliance.” In Self-Reliance and Other Essays. Mineola, NY: Dover Publications.

[14] Hottenstein, A. W., Taylor, R., & Carr, B. T. (2008). Preference segments: A deeper understanding of consumer acceptance or a serving order effect? Food Quality and Preference, 19, 711-718.

[15] Franczak, B. C., Browne, R. P., McNicholas, P. D., & Findlay, C. J. (2015). Product selection for liking studies: The sensory informed design, Food Quality and Preference, 44, 36-43.

[16] Franczak, B. C., Castura, J. C., Browne, R. P., Findlay, C. J., & McNicholas, P.D. (2016). Handling missing data in consumer hedonic tests arising from direct scaling: Imputation techniques for consumer hedonic tests. Journal of Sensory Studies, 31, 514–523.

[17] Castura, J. C. (2018). Dynamics of consumer perception. In G. Ares & P. Varela (eds.). Methods in Consumer Research, Volume 1 (pp. 211-240). Boca Raton, FL: Woodhead Publishing.

[18] Lee, Y., Findlay, C. J., & Meullenet, J.-F. (2013). Experimental consideration for the use of check-all-that-apply questions to describe the sensory properties of orange juices. International Journal of Food Science and Technology, 48, 215-219.

[19] Tang, Y., Browne, R. P., & McNicholas, P. B. (2015). Model based clustering of high-dimensional binary data. Computational Statistics & Data Analysis, 87, 84-101.

[20] Meyners, M., & Castura, J. C. (2014). Check-all-that-apply questions. In: P. Varela and G. Ares (eds.): Novel Techniques in Sensory Characterization and Consumer Profiling (pp. 271-306). Boca Raton, FL: CRC Press.

[21] Meyners, M., Castura, J. C., & Carr, B. T. (2013). Existing and new approaches for the analysis of CATA data. Food Quality and Preference, 30, 309-319.

[22] Castura, J. C., King, S. K., & Findlay, C. J. (2017). Do teenaged chocolate-flavored cereal consumers go cuckoo trying to do TCATA while eating Cocoa Puffs? 12th Pangborn Sensory Science Symposium. 20-24 August. Providence, RI, USA. (Data Snapshot / Poster).