Setting significance levels in Consensus Sensory Profiling

C. Barnagaud[1], P. Mehring[1], S. Ferris[1] and V. Mialon[2] | MMR Research Worldwide, United Kingdom[1] /USA[2]
c.barnagaud@mmr-research.com

With consensus profiling, members of a trained sensory panel work in concert to agree magnitude ratings for each sensory attribute rather than providing independent ratings in duplicate or triplicate (Meilgaard, 2006). This makes it more rapid, aligning it better with the fast pace of commercial product development.

However, the credibility of this process is often questioned because lack of replication precludes attribute-by-attribute testing of statistical significance across products.

One potential solution might be to obtain an independent set of ratings from each assessor for each attribute of each product, prior to obtaining the group consensus, and using this to calculate a Least Significant Difference across the product set for each attribute (albeit devoid of intra-assessor variability).

Method

To explore this possibility, one of MMR's Reading (UK)-based expert sensory panels conducted consensus sensory profiling of 8 brands of chocolate flavoured milk drinks.

Summary of consensus profiling approach:

1. Vocabulary development across all products (yielded 53 attributes)
2. Attribute rating training – to promote common qualitative and quantitative interpretation
3. Panellists independently profile each product – averaged across panellists for each attribute of each product
4. Final scores for each attribute reached via discussion:
    - When panel agreement is high - the average is taken as final score
    - When panel agreement is low - an amended score is agreed via consensus and highlighted (see Fig. 1)

Across all products and attributes (8 x 53 = 424) there was initial disagreement in 31 instances (7%) and only 18/53 attributes were affected. Thus, LSDs across the product set could be calculated for 95% LOC for all attributes.
MMR's second expert panel profiled the same 8 chocolate milk drinks using conventional duplicate assessment.

Analyses & Results

Only 'presence of bubbles' was not significant in the consensus profiling and all attributes were significant in the conventional duplicate profiling. Principal Component Analysis (PCA) subsequently conducted on both data sets (Figures 2 and 3).

The PCA plots from both approaches show similar patterns. Only the positions of Slim Fast and Rice Dream relative to Alpro Soya show some differences.

To further explore the sensitivities of both methods, the profiles of two relatively different products (Yazoo vs. Tesco Ultra Slim – Figures 4 & 5) and two relatively similar products (Frijj vs. Yazoo – Figures 6 & 7) are compared.

Across the two sensorially different products, the pattern of discrimination is very similar for both methods.

And the pattern of discrimination between the two relatively similar products is also broadly the same across methods, especially with respect to appearance and mouthfeel attributes. However, some of the more subtle differences in appearance, odor and flavor differed across the two approaches.

Conclusions

Using single replicate data to calculate LSDs across assessors, as part of a consensus profiling approach, gave a reliable read on key significant differences between products when benchmarked against conventional duplicated assessments.

This consensus approach provides sensory scientists with a reliable method for quickly and accurately profiling a large number of products when time or product availability is a constraint and/or when the sample set is inherently difficult / complex and quantification in isolation would thus be very difficult (i.e. samples that are very similar, show a lot of variability or that represent sensory challenges).

Consensus adds the ability to do some 'side by side' assessment during profiling to understand not just the differences but the similarities among the sample set.

Fig. 1 - Consensus ratings for 8 chocolate milks - highlighted scores agreed via consensus

	Yazoo	Frijj	Galaxy	Mars	Slimfast	Tesco Ultra Slim	Rice dream	Alpro Soya
Dry/powdery	26	23	18	16	18	18	20	17
Cocoa	54	49	45	35	8	10	30	24
Artificial chocolate	0	0	0	0	0	32	0	3
Dairy/milk	45	40	42	40	11	5	11	10
Rice milk	0	0	0	0	0	0	30	0
Soya milk	0	0	0	0	18	0	3	30
Buttery	0	0	10	18	0	0	0	1
Vanilla	15	10	20	26	0	8	7	8
Sweet	51	45	59	79	28	65	41	30
Sour/acid	0	4	0	2	16	21	2	3
Caramel/toffee	0	6	28	39	0	0	3	0
Over ripe fruit/banana	0	0	0	0	0	53	0	0
Wheat	0	8	0	0	0	0	29	22
Nutty	0	0	0	0	10	0	2	11
Mint	0	0	0	0	0	8	0	0
Vitamin B	0	0	8	8	42	0	0	7

Fig. 2 PCA - Consensus profiling

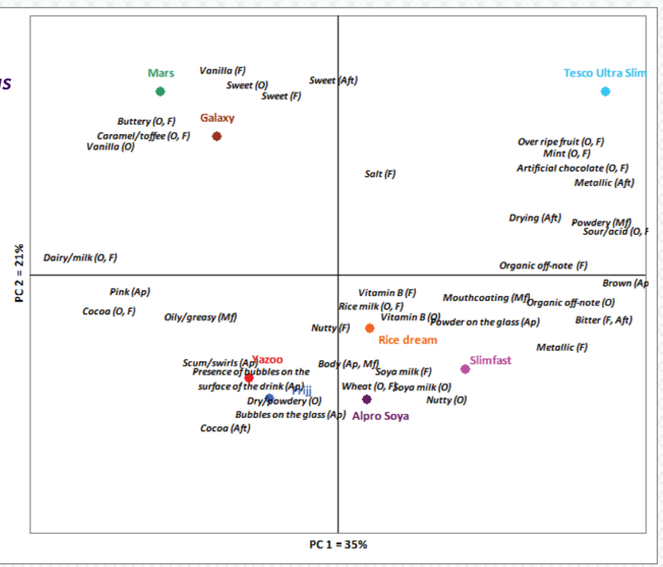


Fig. 3 PCA - Conventional duplicate profiling



Fig. 4 - Yazoo vs. Tesco Ultra Slim - consensus profiling

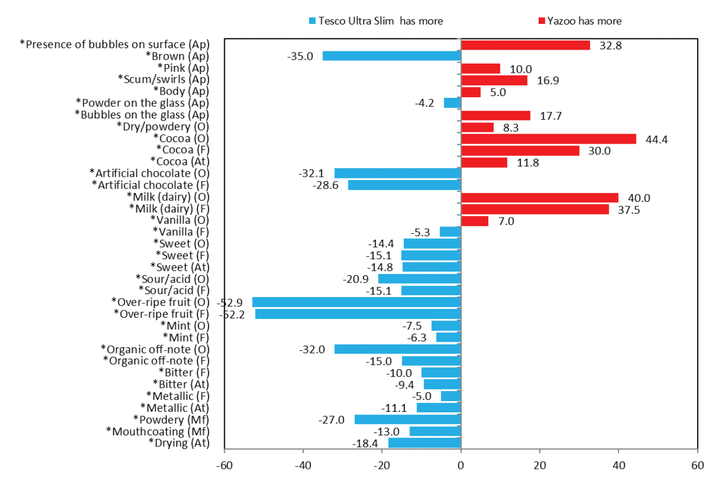


Fig. 5 - Yazoo vs. Tesco Ultra Slim - conventional duplicate profiling

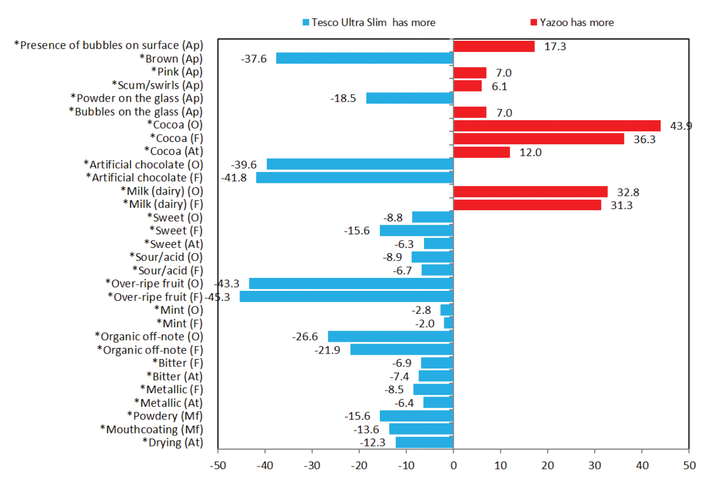


Fig. 6 - Frijj vs. Yazoo - consensus profiling

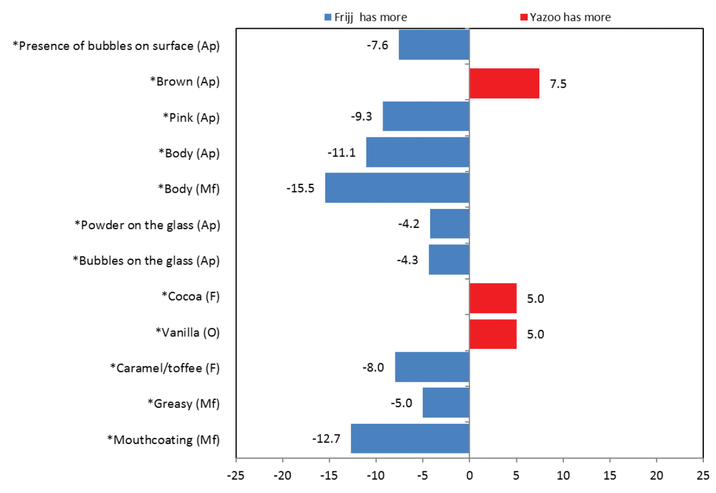
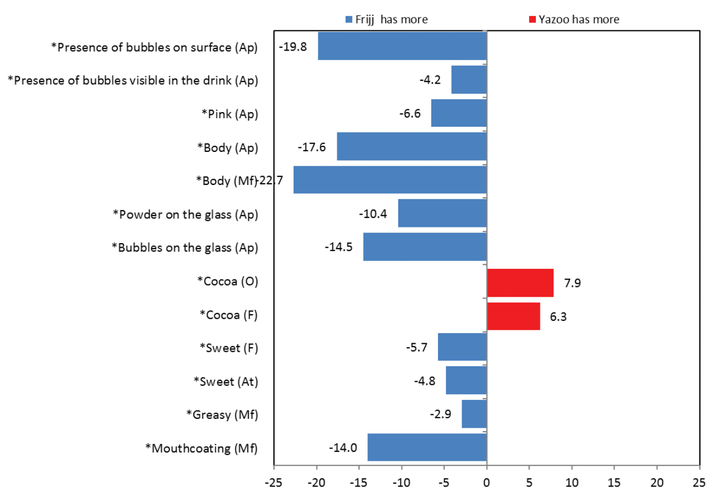


Fig. 7 - Frijj vs. Yazoo - conventional duplicate profiling



References

- Morten C. Meilgaard, Gail Vance Civille, B. Thomas Carr (2006). Sensory Evaluation Techniques, Fourth Edition. New York: Taylor & Francis