# So close, yet so far? A methodological investigation of the potential of and optimal sample sizes for the application of napping as rapid sensory method in marketing contexts

**Kathrin Heim**
Austrian Marketing University of Applied Sciences, Campus Wieselburg der Fachhochschule Wiener Neustadt GmbH
**Robert Fina**
Austrian Marketing University of Applied Sciences, Campus Wieselburg der Fachhochschule Wiener Neustadt GmbH
**Roswitha Enzelberger**
Fachhochschule Wr. Neustadt GmbH, Campus Wieselburg
**Elisabeth Steiner**
Austrian Marketing University of Applied Sciences, Campus Wieselburg der Fachhochschule Wiener Neustadt GmbH
**Alexander Höftberger**
Fachhochschule Wr. Neustadt GmbH, Campus Wieselburg

Cite as:
Heim Kathrin, Fina Robert, Enzelberger Roswitha, Steiner Elisabeth, Höftberger Alexander (2020), So close, yet so far? A methodological investigation of the potential of and optimal sample sizes for the application of napping as rapid sensory method in marketing contexts. *Proceedings of the European Marketing Academy*, 11th, (84947)

Paper presented at the 11th Regional EMAC Regional Conference, Zagreb (online), September 16-19,2020

# So close, yet so far? A methodological investigation of the potential of and optimal sample sizes for the application of napping as rapid sensory method in marketing contexts

**Abstract**:

**This paper introduces a large-scale Napping study, serving the purpose of identifying the optimal sample size for a scientific employment of Napping in marketing contexts. A napping with 15 yogurt samples (n = 104) for the properties appearance, taste and texture was arranged. For the evaluation of the napping data, random samples were evaluated using MFA and GPA. Tucker´s congruence coefficient (rc) was used to interprete the results. In contrast to current standards recommending sample sizes of 9-15 respondents, our results suggest Napping sample sizes of at least 20, ideally 33 untrained subjects in order to get statistically sound results.**

*Keywords: Napping, MFA, sensory marketing*

# 1 Introduction of paper

Sensory profiling, serving the purpose of defining and quantifying sensory characteristics on which products differ, has been an established and essential tool for food scientists and food manufacturers for a long time. The numerous applications of traditional descriptive profiling techniques include classical “sensory” tasks, such as product development, product improvement and quality control (Valentin, Chollet, Lelièvre, and Abdi, 2012), however, also expand to the fields of marketing and consumer science, involving advertising claim substantiation (Lawless & Heymann, 2010) and the understanding of consumer preferences (Greenhof & MacFie, 1994).

Though being referred to as “one of the most powerful, sophisticated and most extensively used tools in sensory science” (Varela & Ares, 2012), sensory profiling, as conducted with trained assessor panels, is time-consuming and cost-intensive. Consequently, despite of the relevance of traditional profiling, several alternative methods have evolved in recent years, following industrial demand for faster and more cost-effective methods. The resulting rapid sensory methods may – as novel methods for product characterization – not only be regarded as efficient alternatives to traditional methods in sensory science, but expand their range of application to new fields of use (Delarue, Lawlor, and Rogeaux, 2014): Particularly the suitability of these rapid sensory methods for sensory product characterization with consumers (rather than trained assessors) substantiates their noteworthy relevance for the field of marketing, even inducing researchers to refer to these novel methods as the “blurred line between sensory and consumer science” (Varela & Ares, 2012).

While three categories of rapid sensory methods may be distinguished, namely verbal-based methods (flash profile and check-all-that-apply), reference-based methods (polarised sensory positioning and pivot profile) and similarity-based methods (free sorting task and projective mapping aka Napping) (Valentin et al. 2012), the paper at hand exclusively focuses on the latter.

Napping, also being referred to as a holistic approach in sensory analysis (Pagès, Cadoret, and Lê, 2010), represents a similarity-based, rapid descriptive sensory method with a wide range of potential applications in marketing and consumer research contexts.

However, while numerous studies focus on applications of Napping in different contexts, methodological issues, such as the optimal sample size for the usage of Napping in consumer research, have yet not been subject to a thorough discussion, neither in sensory, nor marketing literature. While, according to some standards, samples for Napping should comprise nine to 15 untrained or trained persons who, on average, sample twelve stimuli at a time (Schneider-Häder & Derndorfer, 2016), several other publications report rather vague recommendations, such as that “the appropriate sample size will probably depend on the objectives of the study” (Delarue at al., 2014).

A review of extant studies on Napping ($n$ = 74) reveals that, surprisingly, within the field of sensory science, most authors refrain from challenging the recommendation of nine to 15 persons, wherefore only slight discrepancies regarding the authors’ conceptions of appropriate sample sizes can be observed. While several studies consider 15 panellists as optimal for projective mapping methods including napping (see, for instance, Lelievre-Desmas, Valentin, and Chollet, 2017; Louw et al., 2015), there are some downward (e.g. Liu, Grónbeck, Di Monaco, Giacolone, and Bredie (2016), whose sample comprises eight participants) as well as slight upward deviations (Kemp, Pickering, Willwerth, and Inglis, 2018).

To summarize, it can be noted that there is a lack of systematical methodological investigations and statistically validated reference points, wherefore it remains unclear how large a Napping panel should be in order to provide statistically sound and reproducible results. Resting upon these considerations, the paper at hand aims at identifying the optimal sample size for a scientific employment of Napping in marketing contexts.

## 2 Theoretical background on napping as methodolgy

Seen historically, the idea underlying Napping may be traced back to projective tests in clinical psychology. Projective tests, such as the also in marketing contexts well-known Rorschach inkblot test, enabled subjects to "indirectly reveal their personalities […] by projecting themselves through their responses to the stimuli" (Lê, Le, & Cadoret, 2015).

In sensory science literature, the idea of "the idea of asking subjects to reveal themselves through the way they are positioning stimuli (products) on a sheet of paper based on their perceived similarities" (Lê et al. 2015) arose in 1994, encouraged by a publication of Risvik, McEwan, Colwill, Rogers, and Lyon (1994), who come to the conclusion that this approach "…could be a potentially useful technique for linking sensory analysis and consumer research data" (Risvik et al. 1994).

Indeed, despite of initial warnings of sensory scientists, emphasizing that "...as with any untrained panel, beyond the overall acceptance judgment there is no assurance that the responses are reliable or valid" (Stone & Sidel, 1993) or that "...consumers can only tell you what they like or dislike" (Lawless & Heymann, 1999), recent studies provide empirical evidence for the fact that the use of consumers seems to be a good alternative to classical sensory profiling techniques provided by a trained panel (Worch, Lê, & Punter, 2010).

Napping represents a rapid sensory descriptive method, more precisely, a similarity measurement, in sensory sciences, pursuing the goal of obtaining a sensory comparison of several products in terms of their relative similarity to one another. By providing valuable information about products and their sensory properties, as well as consumers' preferences, Napping facilitates comparisons of products with competing products and may provide valuable insights for product development (Schneider-Häder & Derndorfer, 2016).

According to literature, three sub-categories of Napping may be distinguished: In general napping, all samples are served at the same time and arranged by the participants on a sheet of paper, usually with the dimensions 400 x 600 mm, relative to each other. The paper thereby represents a two-dimensional space. If the samples differ, they are placed far away from each other, whereas if they are similar, they are positioned close to each other. Each product can be assigned a position in the coordinate system and be characterized by freely selectable sensory attributes (Derndorfer, 2016).

In the so-called partial napping, samples are analyzed focusing on individual product features, such as appearance, smell, taste or texture (Pfeiffer & Gilbert, 2008)

Sorted Napping extends the positioning of the products by grouping arranged samples with similar sensory properties into product groups. Subsequently, these clusters are described verbally with defined attributes. As a result, in addition to the positioning data of the individual samples, statements about the respective clusters are obtained. Similarities and differences between the individual products can be determined (Kermarrec, 2010).

# 3 Empirical work

## *3.1 Study design*

We conduct a large-scale Napping study with $n$ = 104 participants, hence drastically deviating upwards from current standards, recommending that samples for Napping should comprise nine to 15 persons (Schneider-Häder & Derndorfer, 2016). The rationale for the large sample size is that, after completion of the Napping procedure and coding of the data, several random samples systematically varying in size may be drawn out of this sample. This procedure (i.e., the drawing of random samples out of a large-scale sample) allows a systematical analysis of sets of variables collected on the same observations with Multiple Factor Analysis (MFA) and a subsequent Generalized Procrustes Analysis (GPA), providing an index of consensus (Tucker's congruence coefficient) between the different random samples and the whole sample. Building on these analyses, recommendations pertaining to the optimal sample size for Napping studies may be derived.

For sensory testing, a combination of partial and sorted napping was used, pursuing the aim of not only obtaining the positioning data of the individual samples, but also statements about the respective formed sample clusters. This procedure should allow the drawing of conclusions about the similarities or differences between the different samples and a description of them. The sensory testing was carried out in the sensory laboratory of a Central European University.

## *3.2 Selection of product category and stimuli*

As stimuli, 15 different commercially available strawberry yoghurt samples were selected. Each of the yogurts was provided in transparent plastic cups (filling quantity: 20 g), which were served at room temperature. All samples were purchased in the morning of the day of testing and stored at room temperature until sensory testing. In addition, the samples were coded with four-digit, randomly selected sample numbers.

## *3.3 Procedure*

104 voluntary and untrained students of the University participated in the sensory testing. In advance to their participation, respondents were familiarized with the method of Napping.

During sensory testing, each of the 104 subjects was provided with all 15 samples simultaneously. In addition to the samples, each participant received three sheets of white paper in DIN A3. Respondents were instructed to systematically cluster the samples based on perceived similarities and differences in a way that similar samples were arranged close to each other. In a next step, the positions of each sample were marked on the paper. Moreover, respondents were requested to verbally describe each cluster with sensory attributes, whereby they used their own vocabulary. Overall, three rounds of Napping were performed, as participants were asked to first rate samples based on their appearance, followed by taste and texture.

The marked positions of all samples as indicated by the test persons were transferred into a coordinate system, whereby the lower edge of each sheet of paper formed the x-axis and the left side edge the y-axis. As this coding procedure has to be performed manually, this is a quite time-consuming procedure. Based on the thus evolving coordinate system, all sample positions were recorded in form of coordinates and entered into a table in Microsoft Office Excel for each participant and for all three characteristics (appearance, taste, texture).

The verbal descriptions (sensory attributes) were also collected in the Excel file, regardless of the assessors, for each sample and all three characteristics in order to assess the frequencies of the responses. Very similar attributes (e.g. thick, viscous) were combined into one. Subsequently, these data were analyzed with the help of XLStat® using multiple factor analysis (MFA) and Generalized Procrustes Analysis (GPA).

### *3.4 Analysis*

Due to poor data quality, four respondents had to be excluded from the analysis, resulting in an analyzable sample of n = 100. In a next step, for the evaluation of the napping data, random samples for all of the three characteristics (appearance, taste, texture) were drawn from the entire sample in order to determine the optimal sample size. The samples taken comprised 10, 15, 20, 33 and 50 cohorts and were evaluated using multiple factor analysis (MFA) in the Microsoft Excel add-on XLStat®. Table 1 provides an overview of the sizes and numbers of samples drawn.

| sample size | n = 10 | n = 15 | n = 20 | n = 33 | n = 50 |
|---|---|---|---|---|---|
| **number of samples drawn** | **10** | **6** | **5** | **3** | **2** |

*1: Characteristics of random samples*

Subsequently, the two-dimensional coordinates for each of the 15 stimuli resulting of the various MFAs were subject to Generalized Procrustes Analyses (GPA), following the purpose of analyzing metric multidimensional scaling between each sub-sample and the whole sample by means of scaling, rotation and translation techniques. Tucker's congruence coefficient ($r_c$) serves as an index of consensus between the different random samples and the whole sample.

### *3.5 Results*

Procrustes analyses of the sub-samples with $n$ = 10 and $n$ = 15 respondents reveal Tucker's congruence coefficient ($r_c$) values below 85 %, thus indicating that, in fact, Napping sample sizes below 15 are not recommendable, as they provide only limitedly reliable results. Methodological literature even suggests that congruence coefficients below 0.85 should not be "interpreted as indicative of any factor similarity at all" (Lorenzo-Seva & ten Berge, 2006).

The most important results of the characteristic property texture, which achieved the comparatively lowest consensus indices in the Generalized Procrustes Analyses (GPA), are shown below.

Figure 1 shows the result of the MFA on the characteristic property texture.

The characteristic texture was used as an example, because it is the most demanding and therefore the most difficult to discriminate attribute for untrained probands. The MFA clearly shows overall that the test subjects classify the products based on the two factors of consistency and homogeneity.

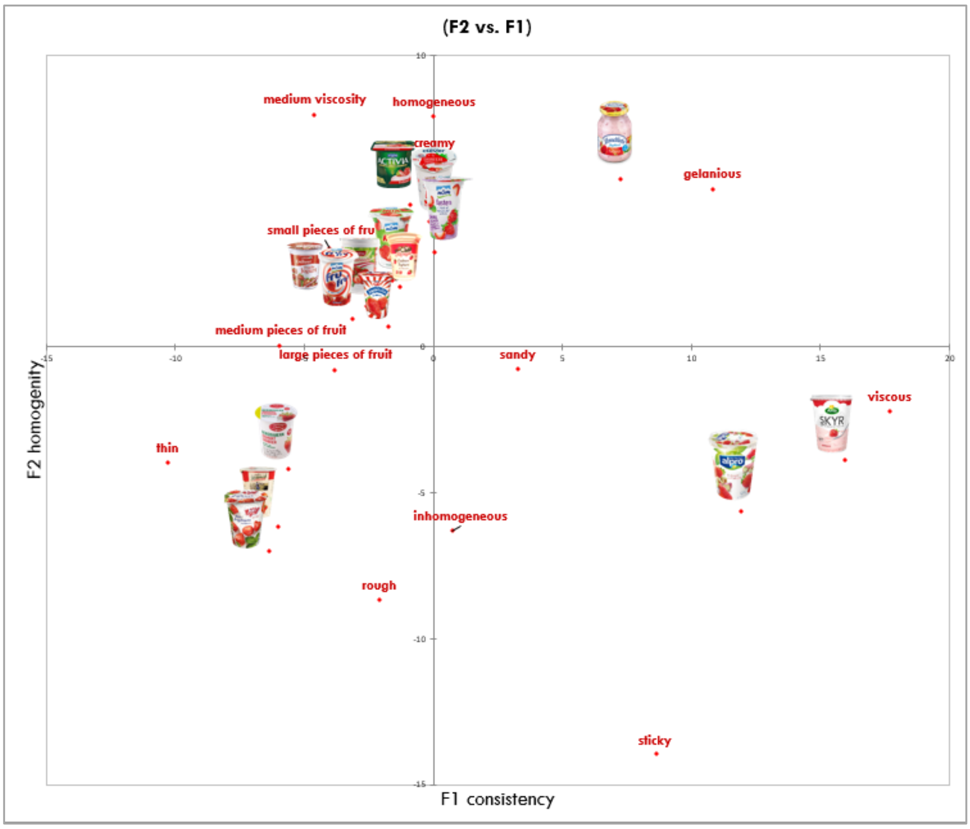


*Figure 1: Results multiple factor analysis, texture, n = 100*

Figures 2 and 3 illustrate the results of Procrustes Analyses for the characteristic property texture for the sub-samples of the sample size n = 20 and n = 33 with the lowest consensus indices of all drawn sample comparisons of all three characteristic properties. Table 2 and 3 show Tucker´s congruence coefficients (rc) for all sub-samples of the sample size n = 20 und n= 30 in comparison.

Figure 2 visualizes the sub-sample with the lowest consensus index of all sample sizes n = 20 of the characteristic property texture. As the Tucker´s congruence coefficient (rc) illustrates, the consensus indices for this sample size are between 0.90 % and 0.95 %. This indicates an appropriate similarity of the individual subsamples to the entire sample (Lorenzo-Seva & ten Berge, 2006).

Table 2 shows Tucker´s congruence coefficients (rc) for all sub-samples of the sample size n = 20 in comparison.

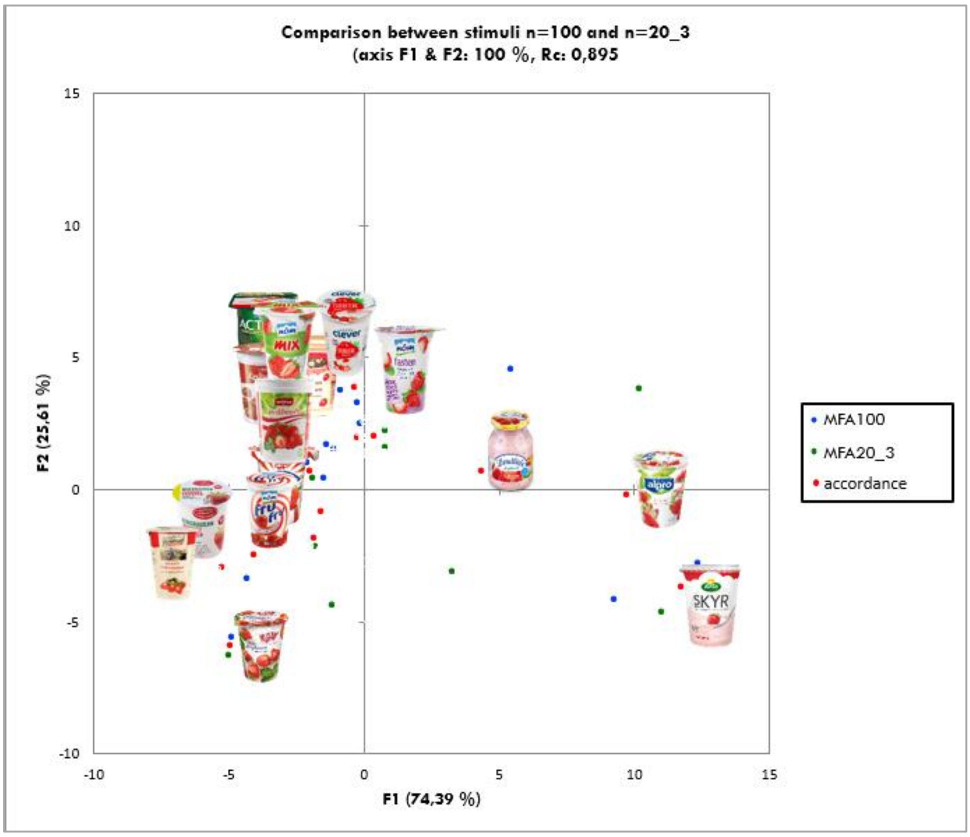


*Figure 2: Comparison between whole sample and sub-sample, texture, n =20_3*

| sub-sample | GPA20_1 | GPA20_2 | GPA20_3 | GPA20_4 | GPA20_5 |
|---|---|---|---|---|---|
| Rc | **0,925** | **0,937** | **0,895** | **0,934** | **0,954** |

*Table 2: Tucker´s congruence coefficient for sub-samples, texture, n = 20, texture*

Figure 3 shows the sub-sample n = 33_3, the sample taken (n = 33), which describes the lowest consensus index of all samples of all characteristics.

Table 3 shows Tucker´s congruence coefficients (rc) for all sub-samples of the sample size n = 30 in comparison.

As can be seen from the Tucker congruence coefficients (rc), all consensus indices for sample sizes of n = 33 are equal to or higher than 97 %, which indicates an extraordinarily high consensus of the individual sub-samples with the entire sample. According to literature, congruence coefficient values over 95 % can be regarded as almost identical factors (Lorenzo-Seva & ten Berge, 2006).

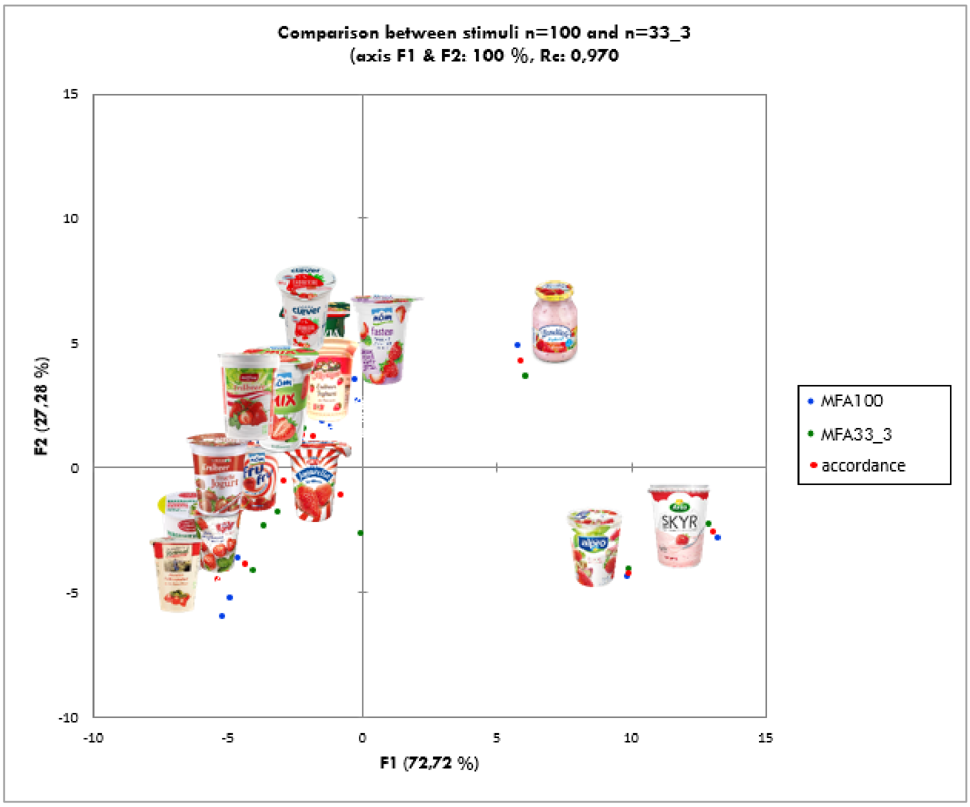


*Figure 3: Comparison between whole sample and sub-sample, texture, n =33_3*

| **sub-sample** | GPA33_1 | GPA33_2 | GPA33_3 |
|---|---|---|---|
| **Rc** | **0,986** | **0,980** | **0,970** |

*Table 3: Tucker´s congruence coefficient for sub-samples, texture, n = 33*

To complete our results, Table 4 shows the results of the characteristic properties taste and appearance based on the worst result of the Tucker´s congruence coefficient.

| **sample size** | n = 10 | n = 15 | n = 20 | n = 33 | n = 50 |
|---|---|---|---|---|---|
| **number of samples drawn** | **10** | **6** | **5** | **3** | **2** |
| **taste $Rc_{min}$** | **0,718** | **0,795** | **0,832** | **0,930** | **0,975** |
| **appearance $Rc_{min}$** | **0,877** | **0,974** | **0,959** | **0,987** | **0,993** |

*Table 4: Tucker´s congruence coefficient for sub-samples, taste and appearance (Rc minimum)*

## 4 Conclusion

Being based on our analyses, current recommendations for Napping sample sizes (nine to 15 respondents) deliver, at best, average results, with all of our sub-samples ($n \leq 15$) exhibiting not even fair similarity to the whole sample.

In conclusion, we consider Napping a promising method for the application not only in sensory science, but also in marketing contexts. However, we advise researchers to build their Napping studies, at least those with untrained respondents (consumers), on at least 20, ideally over 33 respondents in order to get reliable and statistically sound results.

## 5 General discussions, implications and further research

The concept of the study was primarily intended as a decision-making aid for SMEs in order to achieve valid results with the smallest possible number of subjects. However, the methodological strength of the napping method is that it offers the opportunity to combine quantitative mapping results with qualitative, subjective, freely selectable descriptions of the subjects. This study should serve as a food for thought so that this method can also be tested in other marketing-relevant areas apart from sensory marketing.

## 6 References


Delarue, J., Lawlor, B., & Rogeaux, M. (2014). Rapid sensory profiling techniques. Applications in new product development and consumer research. Woodhead Publishing.

Derndorfer, E. (2016). Lebensmittelsensorik. Wien: Facultas Universitätsverlag.

Greenhof, K., & MacFie, H. J. H. (1994). Preference mapping in practice. In MacFee, H.J.H. & Thomson, D.M.H. (eds.), Measurement of Food Preferences. 137-166. London, Blackie Academic and Professional.

Kemp, B., Pickering, G., Willwerth, J., & Inglis, D. (2018). Investigating the use of partial napping with ultra-flashprofiling to identify flavour differences in replicated, experimental wines. Journal of Wine Research, 29(4), 302-309.

Kermarrec, C. (2010): Exploring sorted napping as a tool to be used by Kraft Foods Sensory department, master thesis/internship report for Agrocampus Ouest.

Lawless, H. T., & Heymann, H. (1999). Sensory evaluation of food: Principles and practices. New York: Kluwer Academic/Plenum Publishers.

Lawless, H. T., & Heymann, H. (2010). Sensory evaluation of food: Principles and practice. New York: Springer.

Lê, S., Le, T., & Cadoret, M. (2015). Napping and sorted Napping as a sensory profiling technique. In: Delarue, J., Lawlor, J. B., & Rogeaux, M. (2015). Rapid Sensory Profiling Techniques and Related Methods: Applications in New Product Development and Consumer Research. Woodhead Publishing.

Lelievre-Desmas, M., Valentin, D., & Chollet, S. (2017). Pivot profile method: What is the influence ofthe pivot and product space? Food Quality and Preference, 16, 6-14.

Liu, J., Grónbeck, M. S., Di Monaco, R., Giacolone, D., & Bredie, W. L. P. (2016). Performance of flash profile and napping with and without training for describing small sensory differences in amodel wine. Food Quality and Preference, 48, 41-49.

Lorenzo-Seva, U. & ten Berge, J. (2006). Tucker's Congruence Coefficient as a Meaningful Index of Factor Similarity. Methodology European Journal of Research Methods for The Behavioral and Social Sciences, 2, 57-64.

Louw, L., Oelofse, S., Naes, T., Lambrechts, M., Van Rensburg, P., & Niewoudt, H. (2015). Optimisation of the partial napping approach for the successful capturing of mouthfeel differentiation between brandy products. Food Quality and Preference, 41, 245-253.

Pagès, J., Cadoret, M., & Lê, S. (2010). The sorted napping. A new holistic approach in sensory evaluation. Journal of Sensory Studies, 25, 637-658.

Pfeiffer, J. C., & Gilbert, C. C. (2008). Napping by modality: a happy medium between analytic and holistic approaches. Campden & Chorleywood, Food Research Association, 9th Sensometrics Conference, 20–23, Ontario, Canada.

Risvik, E., McEwan, J. A., Colwill, J. S., Rogers, R., & Lyon, D. H. (1994). Projective mapping: A tool for sensory analysis and consumer research. Food Quality and Preference, 5(4), 263-269.

Schneider-Häder, B., & Derndorfer, E. (2016). Sensorische Analyse: Methodenüberblick und Einsatzbereiche: Klassische beschreibende Prüfungen & neue Schnellmethoden, DLG-Expertenwissen, Deutsche Landwirtschafts-Gesellschaft, 9-11 (in German)

Stone, H., & Sidel, J. L. (1993). Sensory evaluation practices. California: Academic Press.


Complete references available upon request.