Comparison of Rate-All-That-Apply (RATA) and Descriptive sensory Analysis (DA) of model double emulsions with subtle perceptual differences

A.K.L. Oppermann\textsuperscript{a,b}, C. de Graaf\textsuperscript{a}, E. Scholten\textsuperscript{b}, M. Stieger\textsuperscript{a}, B. Piqueras-Fiszman\textsuperscript{c,*}

\textsuperscript{a} Division of Human Nutrition, Wageningen University, P.O. Box 8129, 6700 EV Wageningen, The Netherlands
\textsuperscript{b} Physics and Physical Chemistry of Foods, Wageningen University, P.O. Box 17, 6700 AA Wageningen, The Netherlands
\textsuperscript{c} Marketing and Consumer Behaviour Group, Wageningen University, P.O. Box 8130, 6706 KN Wageningen, The Netherlands

\textbf{A B S T R A C T}

The Rate-All-That-Apply (RATA) method, an intensity-based Check-All-That-Apply (CATA) variant, has recently been developed for sensory characterization involving untrained panellists. The aim of this study was to investigate the sensory profiles of ten model (double) emulsions with subtle perceptual differences obtained from the Rate-All-That-Apply (RATA) method with untrained panelists (n = 80). For this purpose two different analysis approaches were followed (treating the data as frequencies and as intensities) and then compared to results obtained from Descriptive Analysis (DA) with trained panellists (n = 11). The RATA method was adapted by including a short familiarization session to acquaint participants with the RATA methodology, the use of the scale, the sensory terms, and product differences. The comparison involved discriminative ability and configuration similarity by means of Multiple Factor Analysis (MFA) and RV coefficients.

The results in our study show that the RATA intensity approach resulted in higher discriminative ability compared to the RATA frequency approach. Both RATA frequency and RATA intensity resulted in similar overall configurations compared to DA. However, important differences between the use of RATA and DA scales suggest that these overall similarities should be interpreted with caution and warrant a deeper investigation on how RATA scales are understood and used by consumers.

\* Corresponding author.

\textbf{1. Introduction}

Generic Descriptive Analysis (DA) with trained panels has been widely used to profile sensory properties of foods and beverages since it provides detailed, consistent, and reliable results (Stone & Sidel, 2004). However, the economic and time consuming aspect of training a sensory panel can be an issue for academic organizations and food industry (Varela & Ares, 2014). Therefore, several consumer-based sensory profiling methodologies have been developed in sensory testing as more rapid and flexible alternatives to DA (Varela & Ares, 2012). Reduced time investment, costs, and training requirements in combination with consumers describing the sensory properties of products instead of analytically trained subjects are key advantages. Consumer-based sensory profiling methodologies can be based on the evaluation of global differences (e.g., sorting and Napping\textsuperscript{*}), the comparison with product references (polarized sensory positioning), global description of individual products (open-ended questions), and on the evaluation of individual terms (e.g., free choice profiling, flash profiling, and Check-All-That-Apply (CATA; Varela & Ares, 2014). For sensory product profiling, also attribute-based methods like intensity scaling, Just-About-Right tasks and Ideal Profile Method have been explored with untrained panellists to indirectly or directly provide sensory profiles (Varela & Ares, 2014).

With the CATA method, consumers are provided with a checklist of predefined terms and asked to select all those terms that apply to describe a given product (Adams, Williams, Lancaster, & Foley, 2007). Previous studies have shown that CATA provides reliable product descriptions comparable to those generated by trained panellists (Ares, Deliza, Barreiro, Giménez, & Gámbaro, 2010; Jaeger et al., 2013). However, the binary response of CATA does not allow a direct measurement of the intensity of the evaluated sensory terms. As elaborated by Ares et al. (2014), this could hamper detailed descriptions and discrimination between products with similar sensory properties. Therefore, Ares et al. (2015)...

\* E-mail address: betina.piquerasfiszman@wur.nl (B. Piqueras-Fiszman).
recommended against the use of CATA questions with untrained consumers for sensory profiling of products with small sensory differences. In this case, CATA questions used by untrained consumers may not provide equivalent information as DA with trained panelists. Recently, Rate-All-That-Apply questions (RATA) have evolved from CATA questions by including intensity ratings of the terms that have been selected. Across a number of studies, Ares et al. (2014) showed that RATA questions, in comparison to CATA, led to an increase in the total number of selected terms. However, even though the total number of selected terms increased, the percentage of selected terms that significantly discriminated the products only increased slightly. Reinbach, Giacalone, Ribeiro, Bredie, and Frost (2014) even found a decrease in significant differences when RATA was used, which could be due to the lack of training in rating the intensity of terms and therefore inconsistent use of scales, leading to variability in consumers’ ratings. Giacalone and Hedelund (2016) recently investigated the reproducibility at assessor-, attribute- and panel-level of the RATA methodology with semi-trained assessors. Assessors were trained during four sessions and subsequently evaluated chocolate samples in quadruplicate. They found that within-assessor reproducibility was moderate and that product maps obtained from individual replicates at panel-level showed high configurational agreement.

It is also worth noting that reported RATA studies so far have involved generally familiar foods such as beer, bread, gummy lollies, peanuts, and apples (Ares et al., 2014; Meyners, Jaeger, & Ares, 2016; Reinbach et al., 2014). These products are relatively easy for consumers to profile. However, it is unclear whether RATA results with untrained panellists would be similar to DA results with trained panellists, or how results from RATA as frequencies of selection only (CATA) would compare to results from RATA as intensities. Overall, we concluded that the differences between stimuli were subtle.

The motivation of the current research is to investigate whether untrained panelists could also discriminate such a set of non-commercial samples with subtle perceptual differences of a generally unfamiliar product category.

The main aim of this study was therefore to compare different approaches to analyse RATA data of samples with subtle differences and to investigate whether data from untrained subjects performing RATA can lead to the same overall general conclusions as data obtained by a trained panel. In contrast to Giacalone and Hedelund (2016), who trained a panel of employees with considerable product expertise during four sessions, we used one familiarization session only to acquaint naïve subjects with the RATA methodology, the use of the scale, and the samples. Intensity rating was implemented by asking participants to rate the intensities for applicable terms on a 9-box scale.

In order to fully investigate the data and compare results to similar studies, we performed RATA and compared RATA data to results from DA using RATA by untrained panelists. In contrast to Giacalone and Hedelund (2016), who trained a panel of employees with considerable product expertise during four sessions, we used one familiarization session only to acquaint naïve subjects with the RATA methodology, the use of the scale, and the samples. Intensity rating was implemented by asking participants to rate the intensities for applicable terms on a 9-box scale.

2. Materials and methods

2.1. Samples

Two sets of emulsions with either 30 or 50% dispersed phase were evaluated. Each set comprised of five model food emulsions differing in the degree of fat reduction. The dispersed phase was either oil droplets or oil droplets filled with varying amounts of (gelled) water droplets, hence various levels of fat reduction (see Fig. 1). Double emulsions were prepared in a two-step process and the process is described in more detail elsewhere (Oppermann et al., 2016). First, primary (w1/o) emulsions were prepared by mixing 30, 40 or 50 wt% inner water phase containing 0.4 wt% NaCl with sunflower oil containing a lipophilic emulsifier (polyglyceryl polyricinoleate, PGPR) in a high shear blender (Waring blender 8011 ES, Stamford, CT). In case of gelled water droplets, gelatin (10 wt%) was added to the inner water phase, and gelation was induced by subsequent cooling a heated emulsion. Double emulsions were prepared by dispersing 30 wt% of the primary (w1/o) emulsions in 70 wt% of the outer water phase containing 1 wt% whey protein isolate as hydrophilic emulsifier and 0.2 wt% NaCl using a high shear blender (Ultra Turrax T25 with the dispersing tool S25-N 18G, IKA, Staufen, Germany). Mixing speeds during preparation were adapted to obtain emulsions with similar oil droplet sizes. Fig. 1 and Table 1 provide an overview of the emulsion characteristics. As explained previously in Oppermann et al. (2016), sample names indicate the emulsion type (“OW” for single emulsions, “WOW” for double emulsions), the amount of inner...