

CONVENIENCE SAMPLING FOR ACCEPTABILITY AND CATA MEASUREMENTS MAY PROVIDE INACCURATE RESULTS: A CASE STUDY WITH FRUIT-FLAVORED POWDERED BEVERAGES TESTED IN ARGENTINA, SPAIN AND U.S.A.

PAULA CARDINAL¹, MARIA CLARA ZAMORA^{2,3}, EDGAR CHAMBERS IV⁴,
ÁNGEL CARBONELL BARRACHINA⁵ and GUILLERMO HOUGH^{1,6}

¹Comisión de Investigaciones Científicas (CIC) de la Provincia de Buenos Aires, 9 de Julio, Argentina

²Facultad de Ciencias Agrarias, Pontificia Universidad Católica Argentina (UCA), Ciudad de Buenos Aires, Argentina

³Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina

⁴Sensory Analysis Center, Department HN, Kansas State University, Justin Hall

⁵Universidad Miguel Hernández de Elche, Departamento Tecnología Agroalimentaria, Grupo Calidad y Seguridad Alimentaria, Orihuela, Alicante, España

⁶Corresponding author.

TEL/FAX: +54 2317 431309;

EMAIL: guillermo@desa.edu.ar

Accepted for Publication June 9, 2015

doi:10.1111/joss.12158

ABSTRACT

The objective of this study was to measure the sensory acceptability and obtain check-all-that-apply (CATA) responses for fruit-flavored powdered juices, with three different consumer segments: children and women who could be considered target populations, and a convenience sample of food-science-related consumers (FSRC). The study was conducted with a total of 550 consumers in four cities: Alicante (Spain), Buenos Aires (Argentina), Manhattan (U.S.A.) and 9 de Julio (Argentina). The products were reconstituted powdered juices with the following flavors: apple, cherry, grape, grapefruit, orange and pear. Overall, FSRC consumers had the lowest acceptability scores for these products. Regarding CATA results, multiple correspondence analysis showed cherry and grape juices were associated to *artificial-flavor* and *artificial-color*, with the FSRC respondents being mainly responsible for the use of these descriptors. Pear and orange were considered to have *natural-flavor* and *good-color*, mainly by children and women. A generalized linear model was used to analyze the effect of “sample,” “city” and “consumer segment” on the percentage of checks given to each descriptor. The “consumer segment” effect was significant for nearly all descriptors, with FSRC checking samples differently to women and children.

PRACTICAL APPLICATIONS

When convenience samples are assembled for food acceptability or CATA studies, the most convenient consumers to recruit often are food science students, staff or, in the case of companies, food science professionals or technicians. The consequences of using these convenience samples are not always considered. The present research is conclusive in showing the significant differences between FSRC and target populations, both in measuring acceptability and in consumer description using the CATA methodology. Development projects based on acceptability results from convenience samples of food scientists, students or technicians could lead to erroneous development directions.

INTRODUCTION

When it is necessary to measure the acceptability of a food product, it generally is recommended that consumers should be recruited among the target population. To quote two well-known text books:

- “Consumer goods and services try to meet the needs of target populations, select markets, or carefully chosen segments of the populations. Such criteria require that the sensory analyst first determine the population for whom the product (or service) is intended” (Meilgaard *et al.* 2007);
- “Certainly participants should be users of the product category and usually people who also actually like the product. A screening questionnaire will normally include several usage frequency categories, in order to eliminate those consumers that only use the product so rarely that they are really not in the target market” (Lawless and Heymann 2010).

In many research papers, these guidelines are respected. For example, Østli *et al.* (2013) investigated the acceptability of cod with different storage times among Norwegian consumers. A total of 420 consumers were recruited in three different cities through various organizations such as clubs, dancing groups, choirs and soccer teams. Participants had to be regular buyers and consumers of cod. Lawless *et al.* (2013) studied juice products with consumers who had used juice products at least three times per week to qualify and had to be likers of the flavors tested. Similarly, to study acceptance of fish by preschoolers, Donadini *et al.* (2013) used kindergarten students.

However, in other works, convenience consumer samples often are used. This type of sampling may be valid when testing a methodology. For example, Araneda *et al.* (2008) presented the sensory shelf-life methodology whereby each consumer tasted a single sample corresponding to a single storage time. They recruited a convenience sample of consumers among students and staff of the University of Chile. This convenience sample was adequate to test the methodology, but the authors made no pretense that the data obtained from this study represented Chilean consumers.

In some publications, convenience samples are used and conclusions are drawn without considering the limitations of the consumer sample that was used. For example, Devereux *et al.* (2003) recruited 62 consumers among students and staff from of the School of Health Sciences, Deakin University (Melbourne, Australia). They concluded that the incorporation of inulin and oligofructose into standard recipes can be used to produce low fat foods that are acceptable. However, they made no considerations regarding the possible compliance bias their consumers could have had when evaluating samples presented to them by a research group of

their own University. Similarly, Ripoli *et al.* (2013) reported using 106 consumers for a test of difference in beef quality in two breeds of cattle, but indicated that their consumer sample included over 15% consumers who did not eat meat frequently and over 25% that had low or intermediate preference for eating meat. For this apparent convenience sample, no attempt was made to segregate the data from frequent users or those with little preference for meat.

Check-all-that-apply (CATA) questionnaires consist of a list of words or phrases from which respondents select all the words they consider appropriate to describe a product. For example, Dooley *et al.* (2010) compared the sensory map of vanilla ice cream obtained by CATA from 80 regular consumers with the map obtained from external preference mapping. Giacalone *et al.* (2013) used a sensory CATA questionnaire as part of complete demographic, appropriateness and hedonic questionnaire on beers. Perceived attributes of milk were studied using CATA by Vidal *et al.*, (2013). CATA has also been used with concept statements, for example Varela *et al.* (2010) applied CATA to different brands of powdered drinks using eight statements of the type: “it is a good product to go along with meals” or “it is a product for the whole family.” Ng *et al.* (2013) compared the EsSense profile (King and Meiselman 2010) to a consumer developed CATA questionnaire; both methods produced similar emotional spaces and product configurations. In these and other CATA studies (e.g. Ares *et al.* 2014), consumers were recruited based on their consumption of the target product.

When convenience samples are assembled for food acceptability or CATA studies, the most convenient consumers to recruit are usually food science students, staff or, in the case of companies, food science professionals or technicians. This consumer category can be considered to have varying degrees of knowledge that lead them to prefer or reject products in a different manner to the general consumer. For example, ASTM (2014) defines a consumer panel as “a group that is representative of the potential user and that does not have technical knowledge of the products to be tested” where one key aspect is that the participants have no specialized knowledge. Ramírez *et al.* (2001) in their study on sensory shelf life of vegetable oil found that 15% of their consumers actually gave higher scores to more oxidized samples. This result would be highly improbable in a food-science-related convenience sample where scientists often focus on attributes they think are “off.”

The objective of this study was to measure the sensory acceptability and obtain CATA responses for fruit-flavored powdered juices, with three different consumer segments: (1) children and (2) women, who could be considered the target population, and (3) a convenience sample of food-science-related consumers (FSRC). The study was conducted in four cities from three countries: Alicante (Spain), Buenos

Aires (Argentina), Manhattan (U.S.A.) and 9 de Julio (Argentina).

MATERIALS AND METHODS

Samples and Preparation

Six fruit-flavored powdered juices were used. This product was chosen due to the ease of transportation among countries and because generally it is well-accepted worldwide. Four of the juices were bought in a local supermarket in Argentina, the brand was Tang (Kraft Foods Argentina, Villa Mercedes, San Luis, Argentina) and the flavors were: apple, orange, pear and pink-grapefruit. The other two samples were bought in a local supermarket in U.S.A., the brand was Kool-Aid (Kraft Global, Inc., Northfield, IL) and the flavors were: cherry and grape.

The Argentine samples were reconstituted following package instructions: one 35 g packet/L of drinking water. The U.S. samples were also prepared following package instructions: one packet (3.6 g) made 2 L of fruit juice. The U.S. samples had no sweeteners in their formulation and were sweetened with 142 g of sugar/L; a preliminary test was conducted using a trained panel to obtain similar sweetness levels to the Argentine samples. A full equisweet test was not performed as the Argentine samples had different sweetness levels between them and the main objective of the study was the acceptability on the fruit flavors and not on the sweetness levels.

Consumers

Ares *et al.* (2014) in a study on the number of consumers necessary for CATA studies suggested that when working with widely different samples (as it was the case of the present juice study), 60–80 consumers can be regarded as a reasonable compromise to get stable sample and descriptor configurations. For sensory acceptability studies, Hough *et al.* (2006) recommended approximately 100 consumers. In this study, a total of 550 consumers were recruited in four cities:

- Buenos Aires (Argentina): the country's capital city, 3,000,000 inhabitants. The fruit-flavored juices are well-known products; however, grape and cherry flavors do not exist;
- 9 de Julio (Argentina): a small inland city, 40,000 inhabitants. The fruit-flavored juices are well-known products; however, grape and cherry flavors do not exist;
- Alicante (Spain): a Mediterranean port-city, 340,000 inhabitants. Natural fruit juices are widely consumed; the powdered alternatives used in this study are known but not frequently consumed;

TABLE 1. NUMBER OF CONSUMERS PER CITY AND PROFILE

Total	City (Country)	Profile
550	9 de Julio (Argentina): 150	Children: 50 Women: 50 FSRC*: 50
	Buenos Aires (Argentina): 150	Children: 50 Women: 50 FSRC: 50
	Alicante (Spain): 150	Children: 50 Women: 50 FSRC: 50
	Manhattan (U.S.A.): 100	Women: 50 FSRC: 50

* FSRC, food-science-related consumers.

- Manhattan, Kansas (U.S.A.): small inland city, 60,000 inhabitants. The fruit-flavored juices are well-known products; of the Argentine products pear and grapefruit are practically unknown.

The profile and number of corresponding consumers was as follows (Table 1):

- Children: 11–12 years, 150 children in total, 50 in each of the cities of 9 de Julio, Buenos Aires and Alicante. In Buenos Aires and 9 de Julio, they were regular consumers of this type of drink; in Alicante as the product is not well-known, they were nonrejecters and regular users of juices; in Kansas because ingredients in several products (red dye #2 and cyclamate) are banned in the U.S.A., permission was received to test with adults but not children. There were approximately an equal number of girls and boys in the study. Children were recruited in local schools in each one of the cities;
- Women: 18–65 years, 200 in total, 50 in each of the cities. They were recruited from consumer databases in each one of the cities, being regular consumers of fruit-flavored soft drinks in Buenos Aires, 9 de Julio and Kansas; in Alicante they were nonrejecters but regular users of juices;
- FSRC: 18–65 years, 200 in total, 50 in each of the cities. They were recruited among students and teaching staff in university food science departments, being nonrejecters of the product. Recruitment was done by email, classroom invitation and direct contact with colleagues.

Acceptability and CATA Study

Approximately, 25 mL of each sample was presented to consumers refrigerated (8–10°C) in a 70-mL plastic cup, coded with a three-digit number. In each city and each segment, there were 50 consumers; as there were six samples, presentation order was balanced for 48 of the consumers. Samples

were presented monadically, and consumers had water to rinse between samples if they wished.

For each sample, consumers first scored overall acceptability from 1 (dislike) to 10 (like) (Sosa *et al.* 2008). Then, consumers completed the CATA task. Because the objective of this study was to compare responses from different consumer segments, both sensory and nonsensory descriptors were used for the CATA to cover a wider range of responses. The descriptors were developed by consensus by the five authors in collaboration with two other sensory professionals from the 9 de Julio (Argentina) group. The terms were generated based on a nonformal sensory evaluation of the products, followed by discussion. The nonsensory descriptors were generated based on the commercial characteristics and uses of this type of product. The resulting descriptors were:

- Sensory descriptors: artificial-color, artificial-taste, nasty-flavor, natural-flavor, nice-color, not-sweet-enough, pleasant-flavor, sour, strong-flavor, too-sweet and weak-flavor;
- Nonsensory descriptors: cheap, expensive, for-all-the-family, for-kids, for-people-from-other-countries, for-the-kid's-birthday-party, not-familiar-taste, to-drink-any-time, to-drink every-day, to-drink-occasionally, to-drink-with-lunch-and-dinner and when-thirsty.

The order of the descriptors in the CATA questionnaire was randomized between participants

Statistical Analysis

Overall acceptability scores were analyzed by analysis of variance (ANOVA) taking “city,” “consumer segment” and “samples” as fixed effects; and “consumer,” nested within “city” and “consumer segment,” as random effects.

Correspondence analysis (CA, Clausen 1998) has traditionally been used to analyze CATA data. Multiple correspondence analysis, MCA (Le Roux and Rouanet 2010) is an extension, which allows additional categorical variables to be included in the analysis and has been used in other sensory studies (e.g., Varela *et al.* 2013). A separate MCA was performed for each city; thus in each city, the categorical variables were “sample” and “consumer segment.” A single MCA analysis could have been done, but the resulting graphs would have been too busy and complicated to read, thus losing the main objective of CA, which is to summarize results in an overall visual picture. Also to help visualization, separate MCAs were performed for sensory and nonsensory descriptors.

The R_v coefficient was used to compare the sample and descriptor configurations obtained from CA on the CATA data within each country. The R_v coefficient (Escoufier and Robert 1976; Schlich 1996) is a simple way of measuring the similarity between two sets of variables which have been measured on the same samples. The coefficients were calcu-

lated on the first four dimensions of the CA including both the sample and descriptor scores. R_v coefficient has values between 0 and 1 with numbers closer to 1 indicating greater similarity.

Meyners *et al.* (2013) presented existing and new approaches to analyzing CATA data. Cochran's Q test (Siegel and Castellan 1988) is used to determine if a descriptor discriminates between samples. In this study, apart from the samples, there were two additional categorical variables: “city” and “consumer segment.” For each sample, a consumer can either check or not-check a descriptor thus generating a binomial response. This type of data can be analyzed by a generalized linear model, GLM (McConway *et al.* 1999) using the binomial distribution and the logit link function. For each descriptor, the response variable was the number of checks given to a sample by the 50 consumers corresponding to each “city” and “consumer segment”; and the explanatory variables were “sample,” “city” and “consumer segment”; and their two-way interactions. The full three-way interaction model could not be fitted due to insufficient degrees of freedom. To determine significant effects, a stepwise multiple regression was performed (McConway *et al.* 1999).

Once the significant model had been determined, the percent checks were estimated with corresponding 5% least significant differences. These estimated percentages of checks are those presented in the results section to facilitate the comparison with the least significant differences (LSD) error bars.

ANOVA, GLM and MCA were calculated using Genstat 16th Edition (VSN International, Hemel Hempstead, U.K.).

RESULTS AND DISCUSSION

Overall Acceptability

ANOVA analysis showed that the three way “city” \times “segment” \times “sample” interaction was significant ($P < 5\%$). Figure 1 shows the average acceptability for each “city,” “segment” and “sample.” There were no significant differences between the two Argentine cities, 9 de Julio and Buenos Aires; thus, their results were averaged.

As shown in Fig. 1, in each city acceptability varied by “sample” and by “segment.” In Alicante, where these products are little known, there was no uniform tendency of segments across samples. For example, for cherry, FSRC had lower acceptability than children, while for grapefruit, it was the inverse; and for other samples, acceptability was similar. In Buenos Aires-9 de Julio, FSRC had lower acceptability for all samples except grapefruit. In Manhattan, FSRC had slightly lower acceptability than women for most samples. In all cities, obtaining acceptability results from convenience samples of FSRC would be misleading in relation to the target population of these products.

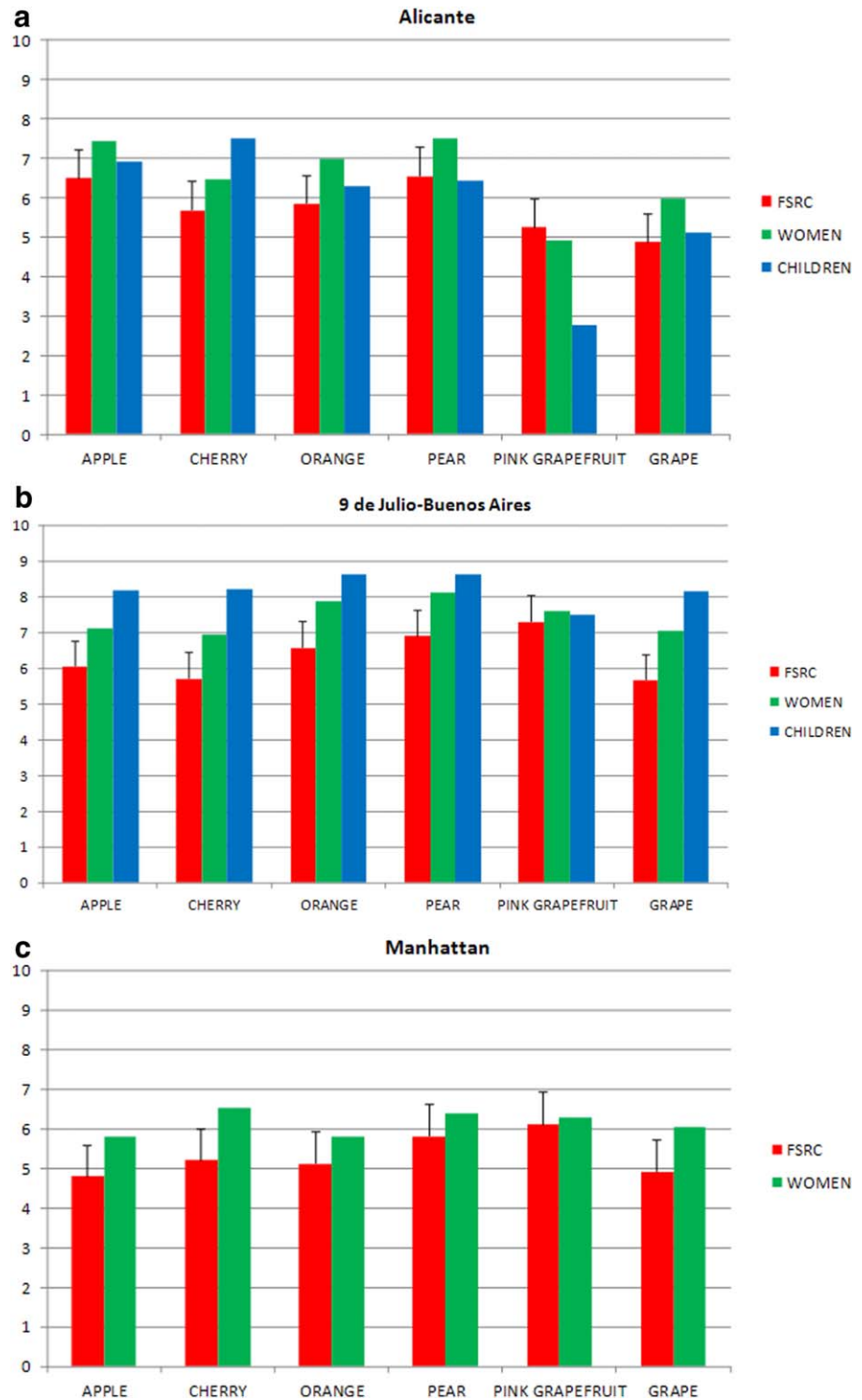


FIG. 1. AVERAGE ACCEPTABILITY FOR ALICANTE (a), 9 DE JULIO-BUENOS AIRES (b) AND MANHATTAN (c) CLASSIFIED BY CONSUMER SEGMENT AND SAMPLE*

Check-All-That-Applies

As explained in the Statistical Analysis section, a separate MCA was performed in each city in order to avoid busy graphs. Results from the MCA of the consumers from 9 de

Julio are in Fig. 2a, b for the sensory and nonsensory descriptors, respectively.

For the sensory descriptors (Fig. 2a), grape and cherry had the highest frequency of responses for *artificial-flavor* and

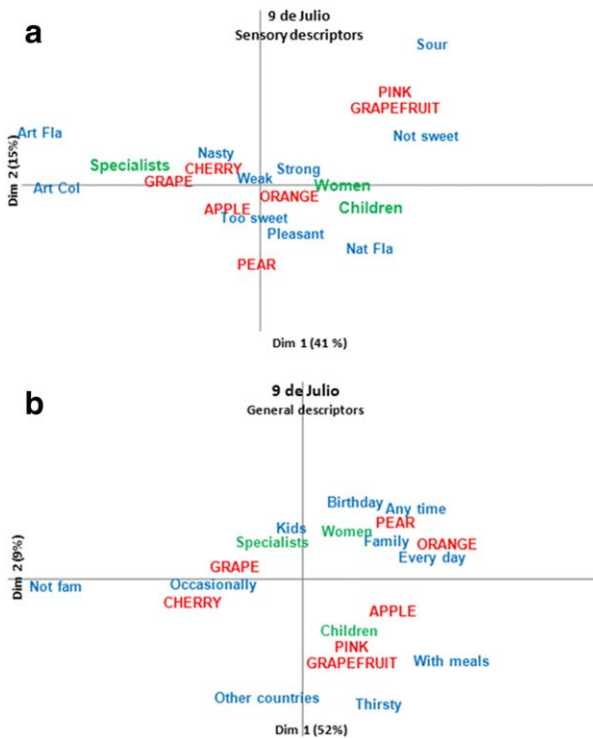


FIG. 2. MULTIPLE CORRESPONDENCE ANALYSIS (MCA) OF THE CATA STUDY FOR 9 DE JULIO CONSUMERS: (a) SENSORY DESCRIPTORS AND (b) NONSENSORY DESCRIPTORS

artificial-color, with the FSRC (specialists) respondents being mainly responsible for the use of these descriptors. Apple, pear and orange were considered to have *natural-flavor and good-color*, mainly by children and women. These respondents also considered grapefruit to be *sour and low-in-sweetness*. MCA maps were similar for other cities (not shown).

For the nonsensory descriptors (Fig. 2b), grape and cherry were on one side of the graph, dominated by *not-familiar-taste*, which is reasonable as these flavors were not on the Argentine market at the time of the study. Conversely, apple, pear and orange, which were on the market, were associated to descriptors such as *for-all-the-family, to-drink-with-lunch-and-dinner* and *to-drink-every-day*. These last descriptors were chosen more by children and women than by FSRC.

The nonsensory descriptor maps for other cities (not shown) were related to whether flavors were available and/or consumed in their respective markets. For example, in Alicante (Spain), *not-familiar-taste* was related to grapefruit while in Manhattan (Kansas, U.S.A.) this descriptor was related to pear. In Manhattan, grape, cherry and orange were considered *cheap, appropriate-for-children* and *for-the-kid's-birthday-party*. The Buenos Aires map was similar to 9 de Julio's, because both cities are in the same country.

The R_v coefficients calculated over the first four CA dimensions are in Table 2. If the objective were to replace

regular consumers of the product (women and children) by a convenience sample of FSRC, the R_v coefficients were not satisfactory. In Buenos Aires, FSRC versus women had an $R_v = 0.85$ for sensory descriptors (Table 2), which could be pointing to similar configurations. However, there is no way of predicting this similarity; for another city and/or product, the R_v could be, as was the case for 9 de Julio, ≤ 0.64 . MCA showed the overall picture including all of the CATA descriptors. GLM analyzed each descriptor independently. Because in one of the cities (Manhattan) children did not do the study, two separate GLM analysis had to be done. One covering three cities (9 de Julio, Alicante and Buenos Aires) where children were considered; and another one covering all four cities where children were not considered. Table 3 shows what effects were significant for each descriptor, considering the analysis with and without children.

The six samples were clearly different in flavor, color and availability and this was reflected in the "sample" effect being significant for all descriptors. The "city" effect was to be expected as these products have different availability and tradition in the cities that were tested.

The main objective of this study was to analyze how the consumer "segment" affected acceptability and CATA results. To show GLM results for each one of the 23 descriptors would take up too much space, and thus, a few descriptors were selected to exemplify the type of results that were obtained. Only four descriptors (*not-sweet-enough, not-familiar-taste, expensive* and *for-people-from-other-countries*) were checked at least once by less than 50% of respondents. Of the remaining descriptors, those that had been checked at least once by more than 50% of respondents, five representative descriptors were chosen: two sensory (*natural-flavor* and *sour*), two nonsensory (*when-thirsty* and *for-all-the-family*) and one that has mainly sensory but some nonsensory implications (*pleasant-flavor*).

- Natural-flavor: the main effects "sample" + "segment" were significant for both models, without Manhattan and without children. Pear had the highest percentage

TABLE 2. R_v COEFFICIENTS BETWEEN SAMPLE + DESCRIPTOR CORRESPONDENCE ANALYSIS CONFIGURATIONS

CITY	COMPARISON	Sensory R_v	Nonsensory R_v
9 de Julio	FSRC versus Women	0.64	0.62
	FSRC versus Children	0.50	0.56
	Women versus Children	0.80	0.68
Buenos Aires	FSRC versus Women	0.85	0.79
	FSRC versus Children	0.54	0.19
	Women versus Children	0.56	0.34
Alicante	FSRC versus Women	0.77	0.59
	FSRC versus Children	0.65	0.47
	Women versus Children	0.82	0.72
Manhattan	FSRC versus Women	0.65	0.79

TABLE 3. MAIN EFFECT AND INTERACTION SIGNIFICANT EFFECTS OF “SAMPLE” (Sam), “CONSUMER SEGMENT” (Seg) and “CITY” (Ci) FOR EACH DESCRIPTOR, CONSIDERING THE ANALYSIS WITH AND WITHOUT CHILDREN

CATA descriptors	Significant terms without Manhattan-with children	Significant terms without children-with Manhattan
Artificial-color	Sam + Ci + Seg + Ci.Seg	Sam + Seg + Ci + Sam.Ci + Seg.Ci
Artificial-taste	Sam + Seg	Sam + Seg + Ci + Seg.Ci
Cheap	Seg + Ci + Seg.Ci	Sam + Seg + Ci
For-all-the-family	Sam + Seg + Ci + Sam.Seg + Ci.Seg	Sam + Seg + Sam.Seg
For-kids	Sam + Seg	Sam + Ci
For-the-kid's-birthday-party	Sam + Seg	Sam + Ci
For-people-from-other-countries	Sam + Seg + Ci + Seg.Ci	Sam + Seg + Ci + Seg.Ci
Nasty-flavor	Sam + Seg + Ci + Sam.Ci + Sam.Seg	Sam + Seg
Natural-taste	Sam + Seg	Sam + Seg
Nice color	Sam + Seg + Ci + Seg.Ci + Sam.Ci	Sam + Seg
Not-familiar-taste	Sam + Ci + Sam.Ci	Sam + Seg + Ci + Sam.Ci
Not- sweet-enough	Sam + Seg + Ci	Sam + Seg + Ci
Pleasant-flavor	Sam + Seg + Ci + Sam.Seg + Sam.Ci + Seg.Ci	Sam + Seg
Sour	Sam + Seg + Ci	Sam + Seg
Strong-flavor	Sam + Seg + Ci	Sam + Seg + Ci
To-drink-any-time	Sam + Seg + Ci	Sam + Seg + Ci
To-drink-every-day	Sam + Seg + Ci + Sam.Seg	Sam + Seg + Ci
To-drink-occasionally	Sam + Seg + Ci	Sam + Ci
To-drink-with-lunch-and-dinner	Sam + Seg + Ci	Sam + Seg + Ci
Too-sweet	Sam + Seg + Ci	Sam + Seg + Ci
Weak-flavor	Sam + City	Sam + Ci
When-thirsty	Sam + Seg + Ci + Sam.Seg	Sam + Ci

of checks and cherry and grape the lowest. Regarding the “segments,” percentage checks were 17, 27 and 38 for FSRC, women and children, respectively. Action on the perceived *natural-flavor* of these products based on FSRC would be very different to action based on the children, who are the primary target population. The without-children model that included Manhattan, showed a similar pattern;

- Sour: In the with-children model, the main effects “city,” “sample” and “segment” were significant. “City” and “segment” differences were small in magnitude. Regarding “samples”, 49% checked grapefruit as *sour*, followed by orange (23%) and apple (11%). Cherry, grape and pear were between 5 and 6%. These values are as expected and show that a CATA sensory descriptor evaluated by consumers can highlight differences in

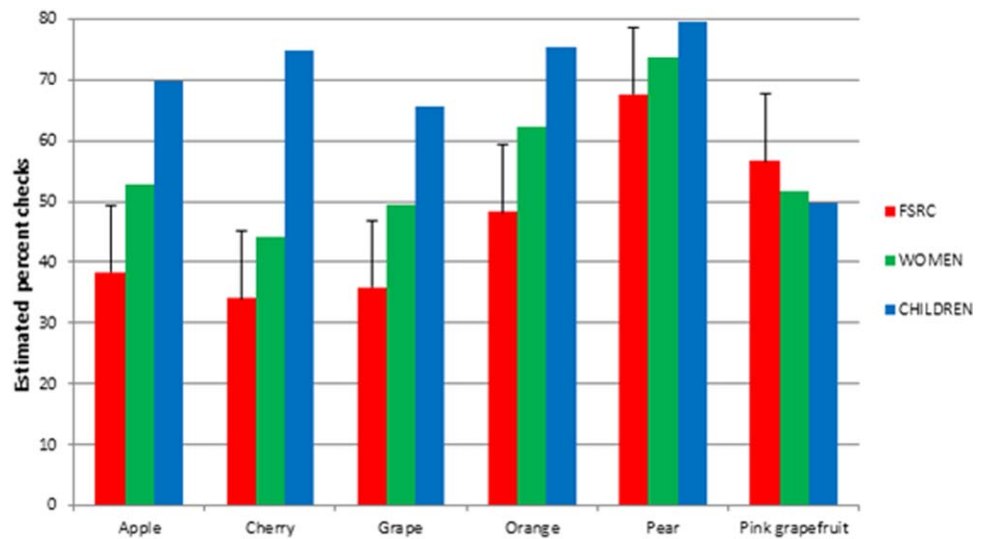


FIG. 3. ESTIMATED PERCENTAGE CHECKS FOR PLEASANT-FLAVOR AVERAGED OVER BOTH ARGENTINE CITIES FOR THE SIX SAMPLES AND DIFFERENT CONSUMER SEGMENTS

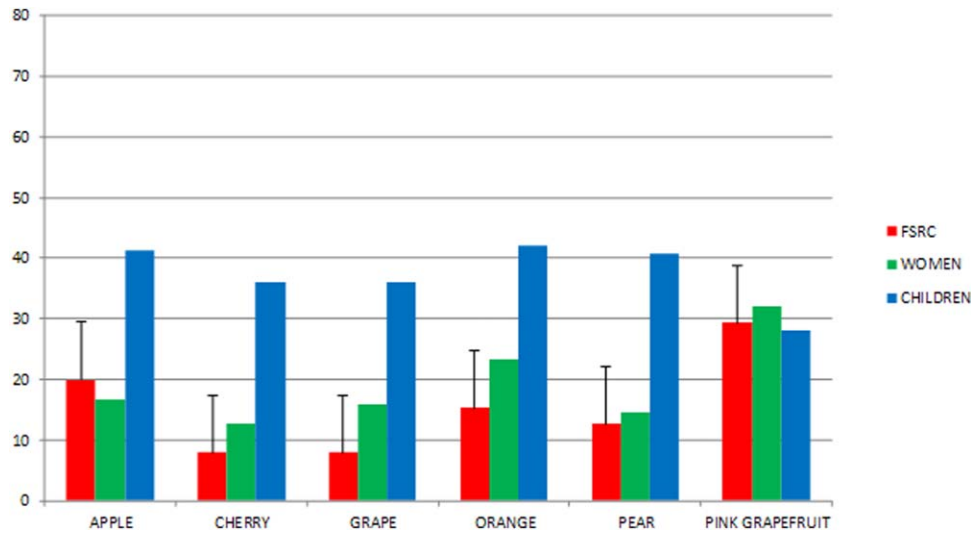


FIG. 4. ESTIMATED PERCENTAGE CHECKS FOR "WHEN-I'M-THIRSTY" AVERAGED OVER 9 DE JULIO, BUENOS AIRES AND ALICANTE FOR THE SIX SAMPLES

a basic taste such as sour. The without-children model was only significant for "sample," with similar values as shown for the other model;

- Pleasant-flavor: Table 3 shows that all two-way interactions were significant for the with-children model. For comparative purposes with overall acceptability (Fig. 1), Fig. 3 is presented with percentage of checks for *pleasant-flavor* averaged over both Argentine cities for the six samples. The overall trend was similar to overall acceptability, thus suggesting that a CATA checking task can be an approximate surrogate to measure acceptability on a scale. It could be argued that checking a box is more in line with what consumers actually do when tasting a drink, they mentally check that it has a *pleasant-flavor* and that might lead them to drink

more or to repeat a purchase. Consumers do not mentally score a product with an 8 on a 1–10 scale as a means of deciding further drinking or future purchase. Regarding the "city" × "sample" and "city" × "segment" interactions, these were influenced by Alicante where these products are not popular and grapefruit was not considered pleasant by any segment. The without-children model was simpler, with "sample" and "segment" being significant as main effects. Pear had the highest percentage of checks (67%) followed by orange (52%) and apple (49%);

- When-thirsty: "City" as a main effect and "sample" × "segment" interaction were significant for the with-children model. Alicante consumers had the lowest percentage of checks (11%) as this product is not popular

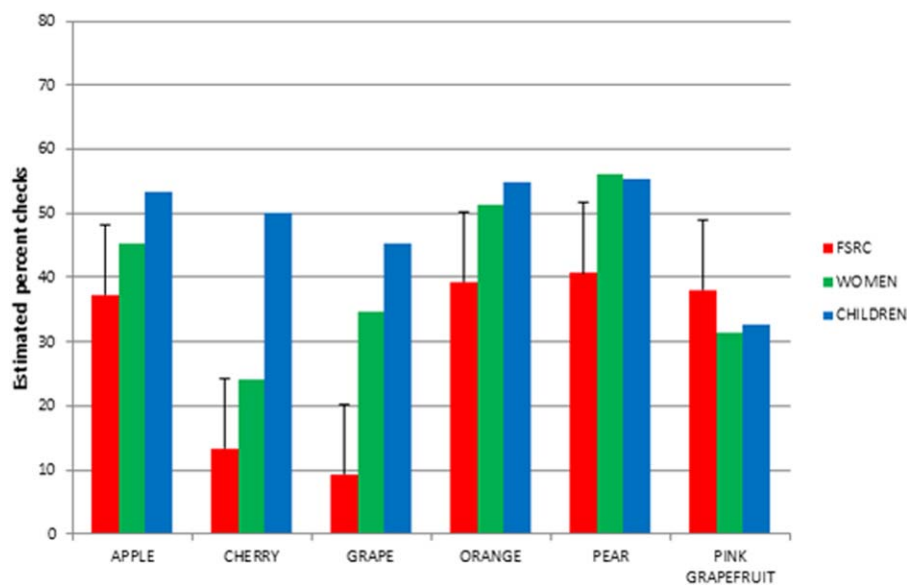


FIG. 5. ESTIMATED PERCENTAGE CHECKS FOR "FOR-ALL-THE-FAMILY" AVERAGED OVER 9 DE JULIO-BUENOS AIRES AND ALICANTE FOR THE SIX SAMPLES

or well-known there; Buenos Aires and 9 de Julio had 28 and 30% checks, respectively. Figure 4 shows how the “consumer segments” differed in how they checked this CATA descriptor for the different “samples.” FSRC and women had similar results, while children considered these products as more appropriate for *when-thirsty*, except grapefruit, most probably due to its lower acceptability. Results from the without-children model showed similar trends, that is, Alicante with the lowest percentage of checks, and, as children were not included, grapefruit had the highest percentage of checks. Overall, regarding “segments,” FSRC and women had similar percentage of checks for this descriptor;

- For-all-the-family: “City” \times “segment” and “sample” \times “segment” interactions were significant for the with-children model. For Buenos Aires and 9 de Julio, FSRC had the lowest percentage of checks (mean 30%) followed by women (mean 41%) and children (mean 58%); in Alicante, the women had the highest percentage of checks (38%) while FSRC and children had lower values (30%). In Manhattan, FSRC and women had similar values. Figure 5 shows the “sample” \times “segment” interaction. Except for the grapefruit flavor, children had high percentage of checks; women had high values for orange and pear and intermediate values for other flavors. FSRC had the lowest values, except for grapefruit. A manufacturer would be misled if he/she decided to guide sales with the “for-all-the-family” concept based on FSRC responses.

From this GLM analysis of CATA descriptors, the “segment” effect was important both as a main effect and in interaction with other factors. The way various consumer segments responded to the CATA questionnaire clearly was different. Varela and Ares (2012), in their review paper on novel methods for product characterization stated that “*the hypothesis that consumers are able to accurately describe products is more and more accepted within the sensory science community.*” Acceptance of this hypothesis must be taken with care as the characterization not only depends on the method but also on the profile of the consumer segment.

CONCLUSIONS

It generally is recommended that acceptability tests should be performed by consumers recruited among the target population. However, there have been few direct comparisons between representative consumer samples and convenience consumer samples that comprised students or professionals with a food science background. In this study, consumers from different segments and cities in different countries were used. Both in sensory acceptability and in CATA results, the “consumer segment” effect was significant and showed dif-

ferences among food-science-related “consumers” and potential target populations. Development projects based on acceptability results from convenience consumer samples comprised of “people with specialized knowledge” could lead to erroneous development directions.

ACKNOWLEDGMENT

Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC).

REFERENCES

- ARANEDA, M., HOUGH, G. and WITTIG DE PENNA, E. 2008. Current-status survival analysis methodology applied to estimating sensory shelf life of ready-to-eat lettuce (*lactuca sativa*). *J. Sensory Studies* 23, 162–170.
- ARES, G., DELIZA, R., BARREIRO, C., GIMÉNEZ, A., and GÁMBARO, A. 2010. Comparison of two sensory profiling techniques based on consumer perception. *Food Qual. Prefer.* 21, 417–426.
- ARES, G., TÁRREGA A., IZQUIERDO, L. and JAEGER, S. 2014. Investigation of the number of consumers necessary to obtain stable sample and descriptor configurations from check-all-that-apply (CATA) questions. *Food Qual. Prefer.* 31, 135–141.
- ASTM INTERNATIONAL. 2014. *Standard terminology relating to sensory evaluation of materials and products, E-253*. Annual Book of Standards, ASTM International, West Conshohoken, PA.
- CLAUSEN, S. 1998. Applied correspondence analysis: An introduction. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-121. Oaks, CA.
- DEVEREUX, H.M., JONES, G.P., MCCORMACK, L. and HUNTER, W.C. 2003. Consumer acceptability of low fat foods containing inulin and oligofructose. *J. Food Sci.* 68, 1850–1854.
- DONADINI, G., FUMI, M.D. and PORRETTA, S. 2013. Hedonic response to fish in preschoolers. *J. Sensory Studies*, 28, 282–296.
- DOOLEY L., LEE YOUNG-SEUNG and MEULLENET J.F. 2010. The application of check-all-that-apply (CATA) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping. *Food Qual. Prefer.* 21, 394–401.
- ESCOUFIER, R. and ROBERT, P. 1976. A unifying tool for linear multivariate statistical methods: The Rv-coefficient. *Appl. Stat.* 25, 257–265.
- GIACALONE D., BREDIE WENDER L.P. and MICHAEL B.F. 2013. “All-In-One Test” (AI1): A rapid and easily applicable approach to consumer product testing. *Food Qual. Prefer.* 27, 108–119.
- HOUGH, G., WAKELING, I., MUCCI, A., CHAMBERS IV, E., MÉNDEZ GALLARDO, I. and RANGEL ALVES, L. 2006. Number of consumers necessary for sensory acceptability tests. *Food Qual. Prefer.* 17, 522–526.

- KING, S.C., and MEISELMAN, H.L. 2010. Development of a method to measure consumer emotions associated with foods. *Food Qual. Prefer.* 21, 168–177.
- LE ROUX B. and ROUANET H. 2010. *Multiple Correspondence Analysis*. Sage, CA.
- LAWLEES H. and HEYMANN H. 2010. In *Sensory evaluation of Food Principles and Practices*, 2nd Ed., pp. 359–360, Springer, NY.
- LAWLESS, L.J.R., THRELFALL, R.T., MEULLENET, J.-F. and HOWARD, L.R. (2013). Applying a mixture design for consumer optimization of black cherry, concord grape and pomegranate juice blends. *J. Sensory Studies* 28, 102–112
- McCONWAY K.J., JONES M.C. and TAYLOR P.C. 1999. Binary regression and What are generalized linear models? In *Statistical Modelling Using GENSTAT^R*, pp. 225–272, The Open University, New York.
- MEILGAARD, M., CIVILLE, G. and CARR, B. 2007. In *Sensory Evaluation Techniques*, 4th Ed., pp. 259–260, CRC Press, Boca Raton, FL.
- MEYNEERS, M., CASTURA J. and CARR, T (2013). Existing and new approaches for the analysis of CATA data. *Food Qual. Prefer.* 30, 309–319.
- NG M., CHAYA C. and HORT J. (2013). Beyond liking: Comparing the measurement of emotional response using EsSense Profile and consumer defined check-all-that-apply methodologies. *Food Qual. Prefer.* 28, 193–205.
- ØSTLI, J., ESAIASSEN, M., GARITTA, L., NØSTVOLD, B. and G. HOUGH. 2013. How fresh is fresh? Perceptions and experience when buying and consuming fresh cod fillets. *Food Qual. Prefer.* 27, 26–34.
- RAMÍREZ, G., HOUGH, G.E. and CONTARINI, A. 2001. Influence of temperature and light exposure on sensory shelf life of a commercial sunflower oil. *J. Food Qual.* 24, 195–204.
- RIPOLL, G., BLANCO, M., ALBERTÍ, P., PANEA, B., JOY, M. and CASASÚS, I. 2014. Effect of two Spanish breeds and diet on beef quality including consumer preferences. *J. Sci. Food Agric.* 94, 983–992.
- SCHLICH, P. 1996. Defining and validating assessor compromises about product distances and attribute correlations. In *Multivariate Analysis of Data in Sensory Science*, Ch. 9 (T. Naes and E. Risvik, eds.) Elsevier, Amsterdam, the Netherlands.
- SIEGEL S. and CASTELLAN N.J. 1988. *Nonparametric Statistics for the Behavioral Sciences*, Ch. 7, 2nd Ed., pp. 170–171, McGraw-Hill, NY.
- SOSA, M., MARTÍNEZ, C., MÁRQUEZ, F. and HOUGH, G. 2008. Location and scale influence on sensory acceptability measurements among low income consumers. *J. Sensory Studies* 23, 707–719.
- VARELA P. and ARES G. 2012. Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Res. Int.* 48, 893–908.
- VARELA P., ARES G., GIMÉNEZ A. and GÁMBARO A. 2010. Influence of brand information on consumers' expectations and liking of powdered drinks in central location tests. *Food Qual. Prefer.* 21, 873–880.
- VARELA, P., ARES, G. and FISZMAN, S. 2013. Texture and semantics: The conceptual structure in consumers' minds. *J. Sensory Studies* 28, 194–204
- VIDAL, L., BARREIRO, C., GÓMEZ, B., ARES, G. and GIMÉNEZ, A. 2013. Influence of information on consumers' evaluations using check-all-that-apply questions and sorting: A case study with milk desserts. *J. Sensory Studies* 28, 125–137.