Mapping the Preferences of Adult and Child Consumers for California-grown Navel Oranges

Tyler Simons¹, Hanne Sivertsen, and Jean-Xavier Guinard

Department of Food Science and Technology, University of California, One Shields Avenue, Davis, CA 95616

Additional index words. consumer testing, descriptive analysis, preference mapping, navel oranges

Abstract. Citrus consumers from Northern California—235 adults (age 18+) and 106 children (ages 8–12)—rated their overall liking and liking for appearance, flavor, and texture, as well as adequacy of sweetness, sourness, firmness, and juiciness for 10 commercially available, California-grown navel oranges (*Citrus sinensis*). Descriptive analysis measures and sugar/acid values were also collected for the fruit. Four adult and two child preference clusters were identified. Adult positive drivers of liking consisted of sweetness, soluble solids content (SSC), overall flavor, orange flavor, and juiciness. A small adult cluster showed a significant preference for acidic fruit. Both segments of child consumers mirrored most of the adults' preferences. This study shows the homogeneity of multiple cultivars of navel oranges grown in California that are currently available to consumers and confirms past studies that the main drivers of liking consist of sweetness, juiciness, and orange flavor, among others. Just-about-right (JAR) ratings acted to confirm multiple attributes that correlated to liking. The findings presented here have implications for future marketing.

Production of citrus fruit in California has a substantial impact on the economy of the state, bringing in greater than \$2.2 billion in 2016 alone (National Agricultural Statistics Service, 2017). Fresh and processed navel oranges (*C. sinensis*) alone contributed greater than \$706 million in 2016 (National Agricultural Statistics Service, 2017). In the 2015–16 season, navel oranges produced for fresh market consumption represented a large portion of all citrus fruit produced in California, at 42% of total volume (National Agricultural Statistics Service, 2017).

Improving liking for a product requires consumer testing. Data from consumers are often nondescriptive, with consumers providing information solely about their liking. While this information is useful in determining which product each group of consumers may prefer, combining this information with detailed analysis of the sensory or chemical properties of a product allows for characterization of the drivers of liking for each consumer segment that is uncovered by preference mapping (Schlich, 1995). This technique has shown success in characterizing the preferences of consumers for different products in multiple

markets (Greenhoff and MacFie, 1994; Guinard et al., 2001; Helgesen et al., 1997; Naes and Risvik, 1996; Schlich, 1995; Sidel and Stone, 1993; Thybo et al., 2004; Yackinous et al., 1999). Data analysis consisting of partial least squares (PLS) regression to relate instrumental and descriptive analysis measures of sensory attributes to consumer liking has also been successful in the past for identifying key attributes that drive consumer preferences (Helgesen et al., 1997; Tenenhaus et al., 2005; Thybo et al., 2004).

Past work with navel oranges has found that high-quality fruit often has high concentrations of sugar in relation to acid (Ivans et al., 1987; Jordan et al., 2001; Obenland et al., 2009). Both sugar and acid components are measured by growers to determine whether the fruit is market ready by use of the California Standard. The California Standard indeed is a quality measurement, based on soluble solids concentration and acid concentration, that aims to enhance consumer acceptability of the fruit (Ferguson and Grafton-Cardwell, 2014). In other commercial citrus categories such as grapefruits, mandarins, and other oranges, a minimum ratio of soluble solids to acid [SSC/titratable acidity (TA)] is required before the fruit is deemed acceptable for market (Ferguson and Grafton-Cardwell, 2014). This standard was changed in 2012 for navel oranges as it was found that the SSC/TA ratio is not a strong predictor of consumer liking (Jordan et al., 2001; Obenland et al., 2009). The new standard provides a quality rating proportional to the concentrations of both sugar and acid instead of relying on their ratio and is defined as

$$(\text{SSC} - 4 \times \text{TA}) \times 16.5.$$

This study set out to fulfill four objectives: 1) to uncover preference segmentation of adult and child consumers with respect to commercially available navel oranges; 2) to identify the key sensory attributes that drive liking for each preference cluster of consumers; 3) to characterize these segments using demographic, usage, and psychographic information collected through an exit survey; and 4) to test the appropriateness of the current California Standard for navel oranges as it applies to consumers.

Materials and Methods

Fruit samples. Ten types of navel oranges, each consisting of five to six 88-count boxes, were obtained from growers and packers through the California Citrus Research Board during the week of 17 Jan. 2016. The samples consisted of specified and unspecified varieties (Table 1). Unspecified varieties were samples for which no variety information was provided. The fruit were treated according to current industrial processing practices, which include washing, rinsing, waxing, grading, and boxing (Ferguson and Grafton-Cardwell, 2014). On reception, the fruit were stored at 5 °C and 85% humidity. When the fruit were tested, instrumental, sensory, and consumer evaluations were all performed within a 10-d span to limit changes to the fruit. Samples tested were taken out of cold storage ≈12 h before descriptive analysis or consumer testing to equilibrate them to room temperature.

Descriptive analysis. A generic descriptive analysis that combined elements of the Quantitative Descriptive Analysis and Spectrum methods (Lawless and Heymann, 2010) was used with eight judges (six females, two males; ages 21–57 years old) with prior descriptive analysis experience. All panelists completed 10 training sessions before evaluation. The first two sessions involved term generation based on navel oranges found in local retail outlets. The following sessions focused on concept alignment, aided by the use of references, to end a list of descriptive terms (Table 2). Some of these terms

Table 1. Commercial navel oranges tested. All fruit were received during the week of 17 Jan. 2016. The fruit were stored at 40 °F until tested by consumers. Once the fruit were first tested by consumers, the descriptive analysis and instrumental measurements were completed within 10 d.

Products	Dates tested by consumers
Navel 1	
Navel 2	
Navel 3	23 and 30 Jan.
Organic	
Washington 1	
Atwood 1	
Atwood 2	
Fisher 1	13 and 20 Feb.
Thomson Improved	
Washington 2	

Received for publication 26 Jan. 2018. Accepted for publication 19 Mar. 2018.

This research was funded by the California Citrus Research Board (Project #5200-150).

We thank Pauline Lestringant for her help with data analysis and Melissa Pellini for conducting the sugar and acid measurements. We also would like to thank Sunkist and Suntreat for providing us with fruit.

¹Corresponding author. E-mail: tjsimons@ucdavis.edu.

mirrored past descriptive work that has been performed on orange juice (Plotto et al., 2010), such as orange, sweet, sour, astringent, and fruity, non-citrus. During training, judges used a sample ballot, which listed each of the terms with an adjacent 10-cm line

scale anchored at 1 cm indentations to limit scale end-use effects (Lawless and Heymann, 2010). The scale anchors denoted high and low intensity of the respective attribute. An electronic ballot was designed for data collection through FIZZ (Biosystèmes, Couternon,

France) for the actual descriptive analysis. The judges were presented with the fruit in whole form, instructed to peel it, and then evaluate it for the attributes on the ballot. Unsalted crackers (Mondelez, East Hanover, NJ) and water were provided to cleanse their

Table 2. Categories, descriptors, definitions, and references used for descriptive analysis of the navel orange samples. Terms were generated by the panel on a variety of commercially available navel oranges during the weeks before testing.

Category	Descriptor	Definition
Exterior whole fruit	Color hue	Hue of the peel Low = green High = dark orange
	Color intensity	Lightness of the peel Low = dark/black High = light/white
	Speckles	Frequency of black spots on the peel Low = no noticeable dark spots High = many noticeable dark spots
	Shininess	Amount of light reflected by the peel Low = matted High = very shiny
	Skin feel	Roughness of peel Low = smooth peel High = coarse peel
	Fruit firmness	Firmness of the whole fruit Low = very soft fruit High = very firm fruit
	Skin thickness	Thickness of the peel after peeling
	Internal color hue	Hue of the interior segments Low = green High = dark orange
Whole fruit peel	Peelability	Ease of removing the peel Low = peel stuck firmly to fruit High = peel comes off very easily
	Peel elasticity	The ease of bending the peel after it has been taken off of the fruit without tearing it Low = peel breaks easily when bent High = peel is very elastic
Aromas and flavors	Aroma/flavor intensity	The overall intensity of aroma given off by the fruit during peeling (orthonasal) and eating (retronasal)
	Orange intensity	The aroma intensity of orange-like notes. Reference: 0.2 mL Sun Essential Oils orange oil diluted in 40 mL tap water
	Mandarin intensity	The aroma intensity of mandarin-like notes. Reference: 0.2 mL Healing Solutions tangerine essential oil diluted in 40 mL tap water
	Lemon	The aroma intensity of lemon-like notes. Reference: one tablespoon each of lemon zest, lemon membrane, and lemon juice from a fresh Eureka lemon
	Lime	The aroma intensity of lime-like notes. Reference: one tablespoon each of lime zest, lime membrane, and lime juice from a fresh lime
	Grapefruit	The aroma intensity of grapefruit-like notes. Reference: one tablespoon each of grapefruit zest, grapefruit membrane, and grapefruit juice from a fresh grapefruit
	Fruity non-citrus	The intensity of the aroma of fruit other than citrus, such as melon, pineapple, cherry, and berry. Reference: two tablespoons Del Monte fruit cocktail
	Floral	The aroma intensity of flower-like notes, such as orange blossom. Reference: 40 mL Carlo Enterprises orange blossom water
	Pine	The intensity of pine-like notes. Reference: 10 g hand-crushed needles from a white fir tree (<i>Abies concolor</i>)
	Cut grass	The intensity of grass-like notes. Reference: 5 g chopped grass (<i>Poa pratensis</i>)
	Earthy	The aroma intensity of dirt/soil-like notes. Reference: 30 g potting soil mixed with 10 mL water and 5 g chopped white mushroom
	Fermented Cardboard, woody	The intensity of fermented-like notes associated with old or rotten fruit The aroma associated with fresh cut wood and cardboard. Reference: One CedarFresh ball
Taste	Sweet	The intensity of sweet taste, such as canned orange/mandarin. Reference: 30 mL Del Monte canned mandarins in syrup
	Sour	The intensity of sour taste, such as lemon juice. Reference: 30 mL fresh Eureka lemon juice
	Bitter	The intensity of bitter taste, such as fresh grapefruit. Reference: 30 mL fresh grapefruit juice
Texture/mouthfeel	Firmness of membrane	The amount of force required to break the membrane of an orange wedge by the teeth when biting into it
	Firmness of juice vesicles (fibrousness)	The amount of force required to break the juice vesicles inside the membrane
	Juiciness	The amount of juice released when chewing the segment
	Residual fiber	The relative amount of solids left in the mouth after all the juice has been extracted
	Astringency	Sensation of drying in the mouth. Reference: 30 mL of 0.02% alum dissolved in water
	Flavor duration	The amount of relative time the sensation of flavor lasts in the mouth

Table 3. Sugar and acid measurements of the navel oranges. Each value represents the average of eight random sample replicates.

Sample	SSC (%)	pH	TA (%)	SSC/TA	California std
Atwood 1	11.28 b	3.96 bcde	0.64 b	17.78 cd	143.70 cd
Atwood 2	11.03 b	4.03 bcd	0.57 bcd	19.77 bc	144.28 c
Fisher 1	11.00 b	4.07 bc	0.53 cd	21.31 abc	146.54 bc
Navel 1	12.24 a	4.10 ab	0.50 de	25.18 a	168.74 a
Navel 2	9.20 c	4.23 a	0.41 e	23.14 ab	124.48 e
Navel 3	12.18 a	3.93 cde	0.62 bc	21.37 abc	160.03 ab
Organic	11.35 b	3.90 de	0.65 b	18.01 cd	144.59 bc
Thomson Improved	11.56 ab	3.99 bcde	0.65 b	18.09 cd	147.68 bc
Washington 1	10.86 b	3.87 e	0.77 a	14.41 d	128.30 de
Washington 2	12.21 a	4.09 ab	0.51 de	24.59 a	168.05 a

Letters within the same column indicate significant differences ($P \leq 0.05$) according to Fisher's least significant difference.

Table 4. Significantly different attributes among the oranges tested through descriptive analysis. The judges evaluated the samples in triplicate and rated the samples on a 10-point scale. The oranges were tested by the trained judges within 10 d of the consumer evaluations. The judges were presented with a whole fruit, instructed to evaluate the appearance, peel the fruit, and then evaluate it for the remaining attributes.

Attribute	Atwood 1	Atwood 2	Fisher 1	Navel 1	Navel 2	Navel 3	Organic	Thomson Improved	Washington 1	Washington 2
Color hue	7.36 ab	5.90 d	6.63 c	6.94 abc	7.46 ab	7.62 a	6.98 abc	6.79 bc	7.19 abc	7.53 a
Color intensity	7.35 ab	5.61 e	6.48 cd	6.74 bcd	7.91 a	7.45 ab	7.24 abc	6.44 d	7.06 bcd	7.34 ab
Shininess	6.78 b	6.91 b	6.82 b	5.04 c	6.65 b	6.73 b	6.45 b	7.68 a	7.00 ab	5.25 c
Skin feel	4.87 c	6.19 b	6.60 ab	5.93 b	4.81 c	6.26 b	6.14 b	4.78 c	6.78 ab	7.34 a
Fruit firmness	6.20b c	4.91 d	6.43 abc	6.73 abc	7.19 a	6.03 c	6.34 bc	5.96 c	6.83 ab	6.60 abc
Skin thickness	4.89 cd	5.88 ab	6.55 a	6.09 ab	5.59 bc	5.31 bc	6.00 ab	4.34 d	6.07 ab	6.65 a
Peel elasticity	3.87 bc	4.62 ab	4.33 ab	4.03 b	3.57 bc	5.32 a	2.91 c	4.05 b	2.79 c	3.67 bc
Sweet	6.54 cd	7.36 ab	6.78 bc	6.95 abc	5.54 e	6.35 cd	5.90 de	6.48 cd	6.46 cd	7.58 a
Overall flavor	6.69 bcde	7.43 ab	6.73 abcde	6.59 cde	5.42 f	7.26 abc	6.43 de	6.26 e	7.18 abcd	7.47 a
Orange flavor	5.62 bc	6.73 a	6.09 abc	6.42 ab	4.54 d	6.88 a	6.10 abc	5.45 c	6.88 a	6.71 a
Membrane firmness	4.97 cde	5.13 cd	5.46 bc	6.18 a	6.53 a	4.47d e	4.95 cde	5.91 ab	4.38 e	5.94 ab
Fibrousness	4.02 cd	3.80 d	4.34 bcd	4.94 abc	5.75 a	3.93 d	4.18 bcd	4.71 bcd	3.95 cd	5.08 ab
Juiciness	7.72 a	7.27 ab	7.05 abc	5.97 de	5.33 e	6.43 cd	6.84 bc	6.48 cd	6.86 bc	7.17 abc
Residual fiber	4.68 d	5.29 cd	5.76 abc	6.67 a	6.37 ab	4.70 d	4.65 d	5.56 bcd	5.29 cd	6.01 abc

Letters within the same row indicate significant differences ($P \leq 0.05$) according to Fisher's least significant difference test.

palates between samples. The samples were coded with randomized three digit codes and evaluated in triplicate under white light. Presentation order of the samples was randomized using a William's Latin square (Williams, 1949) design provided by the FIZZ system. Performance of the panel (i.e., ability to discriminate among samples, reproducibility, and concept alignment) was checked through evaluation of a two-sample set of oranges during the final training sessions using analysis of variance (ANOVA).

Sugar and acid measurements. Eight oranges of each sample were juiced separately. Fifty milliliters of the juice from each of the eight replicates was then placed into a 60 mL plastic cup, capped, and refrigerated at 3 °C for up to 1 d. Samples of juice were analyzed for SSC using a LTD Refractometer 110 (Bellingham + Stanley, Kent, United Kingdom). The pH was measured with an Orion 5-Star Plus pH meter (Thermo Fisher Scientific, Waltham, MA). Samples of the juice were also used to measure TA by titration of 0.1 M NaOH to end point of pH 8.2 using a Mettler Toledo DL50 titrator (Mettler Toledo, Columbus, OH). The SSC/TA ratio was calculated by dividing the SSC concentration by the TA concentration. The California Standard value was also calculated for each of the eight replicates.

Consumers. Adults and children (ages 8–12 years old) from the local community (Davis, Woodland, and Sacramento, CA) were recruited for the tests. Potential participants were screened for appropriate age, lack

Table 5. Overall liking and liking of specific modalities for adults ($n = 235$) and children ($n = 106$), respectively, for the oranges. The adults rated the products on the 9-point hedonic scale with 1 = "dislike extremely" and 9 = "like extremely." The children rated the products using a modified 7-point hedonic scale with 1 = "I hate it" and 7 = "I love it."

Product	Adults				Children		
	Overall liking	Appearance	Flavor	Texture	Taste	Appearance	Texture
Atwood 1	6.53 ab	7.10 a	6.50 ab	6.67 abc	5.15 bc	5.60 a	5.32 ab
Atwood 2	6.58 ab	6.85 abc	6.47 ab	6.66 abc	5.39 abc	5.12 b	5.25 ab
Fisher 1	6.33 bc	6.54 d	6.22 bc	6.43 c	5.33 abc	5.35 ab	4.99 bc
Navel 1	6.69 a	6.85 abc	6.60 a	6.66 abc	5.73 a	5.40 ab	5.28 ab
Navel 2	5.07 d	6.77 bcd	4.91 d	5.49 e	4.63 d	5.39 ab	4.68 c
Navel 3	6.65 a	7.08 a	6.58 a	6.83 a	5.31b c	5.64 a	5.34 ab
Organic	6.73 a	7.02 ab	6.68 a	6.75 ab	5.56 ab	5.38 ab	5.41 a
Thomson Improved	6.06 c	6.74 cd	6.07 c	6.00 d	5.06 c	5.12 b	4.98 bc
Washington 1	6.59 ab	6.88 abc	6.50 ab	6.78 ab	5.33 abc	5.46 ab	5.11 ab
Washington 2	6.71 a	6.90 abc	6.70 a	6.54b c	5.42 abc	5.50 ab	5.32 ab

Letters within the same column indicate significant differences ($P \leq 0.05$) by Fisher's least significant difference test.

of allergies, and consumption of citrus. The age of the children was selected based on cognitive abilities required to use hedonic scales, intensity scales, and other sensory measures (Guinard, 2000). Children were accompanied by a parent/guardian at all times but seated at their own individual booth.

The tasting sessions took place in the Silverado Vineyards Sensory Theater of the Robert Mondavi Institute for Wine and Food Science at the University of California, Davis, CA. In total, 235 adults and 106 children attended two tasting sessions (23 and 30 Jan. and 13 and 20 Feb.) consisting of five products each for a total of 10 samples.

The consumers were separated by booth dividers. They were given a test ballot with a front and back page for the evaluation of each sample, plain crackers, water, and napkins. The fruit were cut into six wedges and two wedges from the same fruit were served to consumers in clear, plastic cups. The consumers were told that they would be evaluating five samples of navel oranges from California and instructed on how to proceed through the ballot and survey. All consumers consented to participate in the tasting. Consumers received a gift card ranging from \$5–10 based on their age for their participation in the study. This study was approved for the use of human subjects by the

Institutional Review Board of the University of California, Davis, CA.

Exit survey. The adults participating in the study completed an online questionnaire aimed to collect information on their demographics, consumption habits, and psychographics hosted on Survey Monkey (Palo Alto, CA). The questions included single and multiple response types surveying the consumers' marital status, ethnicity, children, income, citrus consumption and knowledge, and food neophobia questions, among others.

Ballots. Adult consumers rated their degree of liking on the 9-point hedonic scale (Peryam and Pilgrim, 1957) for appearance, overall liking, flavor, and texture, as well as the adequacy of sweetness, sourness, firmness, and juiciness on a 5-point JAR scale (Rothman and Parker, 2009). Children consumers rated their liking of the appearance, taste, and texture of the fruit on a smaller, 7-point hedonic scale with cartoon faces (Guinard, 2000), as well as the adequacy of the sweetness, sourness, firmness, and juiciness of the fruit on a 3-point JAR scale from "too low" to "too high," with "just right" as the midpoint.

Data analysis. Data analysis for the descriptive analysis was performed using R (R Core Team, Vienna, Austria). A full-model, three-factor ANOVA was performed on each descriptive attribute. The three factors assessed were judges (J), replications (R), and products (P) along with their two-way interactions, $J \times R$, $J \times P$, and $R \times P$. When a significant $J \times P$ or $P \times R$ was found, a pseudo-mixed model ($MS_{\text{Product}}/MS_{J \times R/P \times R}$) was used to assess product differences. Fisher's least significant difference test followed the ANOVA to compare product means. The level of alpha was set at 0.05 for all statistical parameters.

For the consumer data, univariate statistics were performed on the hedonic questions. Penalty analysis was performed on the JAR scales for sweetness, sourness, firmness, and juiciness in comparison with overall liking (adults) or liking of taste (children). Fisher's least significant difference test was used to determine differences between products. In cases of missing values due to skipped questions, data were imputed using the average score for the product and rounded to the closest integer to stay consistent with the ballot. For preference clustering, values for overall liking and for liking of taste (for adults and children, respectively) were scaled to prevent clustering based on scale usage differences. Following this, a Euclidean distance matrix was computed among the consumers who were then clustered according to Ward's Method (Qannari et al., 1997). A two-way ANOVA was performed with respect to the products and clusters to validate significant differences among products and clusters. Consumer preferences were regressed onto the sensory descriptive and instrumental measurements using a PLS2 regression model (Martens and Martens, 1986) in which the sensory and instrumental values were the

Table 6. Correlations between liking of specific attributes for adult consumers ($n = 235$) of the navel oranges.

Modality	Adult liking correlations			
	Appearance	Overall liking	Flavor	Texture
Appearance	1	0.43*	0.38*	0.46*
Overall liking	0.43*	1	0.90*	0.73*
Flavor	0.38*	0.90*	1	0.67*
Texture	0.46*	0.73*	0.67*	1

*Denotes significance ($P \leq 0.05$) by Pearson correlation.

independent variables and the consumer cluster averages were the dependent variables. χ^2 tests were performed on the consumer segments with respect to the JAR questions and their answers on the exit survey to determine differences in demographics, psychographics, and citrus consumption habits. All data analysis and figures were produced using R (R Core Team) except for the JAR penalty analysis, which was created using XLSTAT-Sensory® 2017 (Addinsoft, New York, NY).

Results

Chemical and descriptive analysis. There were significant differences among the samples in each of the chemical measurements (Table 3). All of the fruit samples had a California Standard value above the minimum requirement of 90, and all samples had a SSC/TA value above 10, the previous minimum standard. The panelists found significant differences in 14 of the 32 sensory attributes measured, consisting of seven appearance and peel attributes, one taste attribute, two flavor attributes, and four textural attributes (Table 4). For the flavor attributes, many showed correlation, with sweeter samples often having high overall and orange flavor ratings. Likewise, some textural attributes were negatively correlated. For example, fibrousness and membrane firmness were rated nearly opposite of juiciness. Some of the samples stood out because of their unique sensory properties. Navel 2 and Thomson Improved, for example, had low ratings for sweet, overall flavor, orange flavor, and juiciness, whereas receiving high ratings for membrane firmness and fibrousness.

Consumer hedonic ratings. The liking ratings of adults and children for all products are shown in Table 5. Overall, both the adult and child consumers liked the samples that were tested. All of the fruit were rated above "neither like nor dislike" for adult overall liking and child liking of taste. One sample, Navel 2, received the lowest ratings from both the adults and children but this fruit was not disliked, receiving a score of 5.07 on a 9-point hedonic scale from the adults and 4.63 on a 7-point scale in "liking of taste" from the children.

The adults' overall liking was significantly correlated ($P < 0.05$) with their liking for the appearance, flavor, and texture of the fruit. The correlations were highest between overall liking and liking of flavor ($r = 0.90$) and between overall liking and texture ($r = 0.73$), with the lowest correlation with liking of appearance ($r = 0.43$), as shown in Table 6.

Table 7. Correlations between liking of attributes for child consumers ($n = 106$) for the navel oranges.

Modality	Child liking correlations		
	Appearance	Taste	Texture
Appearance	1	0.24*	0.34*
Taste	0.24*	1	0.49*
Texture	0.34*	0.49*	1

*Denotes significance by Pearson correlation at $P \leq 0.05$.

There was a weaker correlation among the liking ratings of the children (Table 7). Liking of taste and appearance were the least correlated ($r = 0.24$) compared with the highest correlated likings of taste and texture ($r = 0.49$).

JAR scaling. Consumers also rated the products for adequacy of sweetness, sourness, firmness, and juiciness on a JAR scale. These parameters were chosen for their documented contributions to fruit quality and because they represented both taste and texture dimensions. The JAR scores for each attribute were compared with the overall liking scores for penalty analysis, shown in Fig. 1A for the adults and Fig. 1B for the children. In this analysis, the overall hedonic drop for an attribute that is not at an ideal level is plotted against the percentage of consumers that used the nonideal rating. Too little sweetness had a hedonic penalty of about 1.8 points for both adults (Fig. 1A) and children (Fig. 1B). This indicates that if the fruit did not meet the consumer expectation for sweetness, the hedonic rating suffered substantially. Too little sweetness had the largest penalty on adult overall liking (Fig. 1A) and child liking of taste (Fig. 1B), followed closely by too much sourness. Interestingly, too little sourness was marked nearly as frequently as too little sweetness, giving indication that many consumers felt that the fruit they were tasting were not acidic enough. This could be an indication that the fruit was overly mature.

The textural attributes showed a smaller effect on overall liking and liking of taste for both adults and children. Too little juiciness caused a similar penalty to that of too much sweetness, a rating used very rarely by the consumers in our study. For the adults, "too firm" received a similar penalty to "not juicy enough," a penalty that the children did not enforce. In both groups, too much juiciness was the attribute with the smallest penalty. This is not surprising as past research has found juiciness to be extremely important in consumer acceptance of citrus (Gao et al., 2011; Poole and Baron, 1996).

The distribution of JAR ratings for the adults is shown in Table 8, where the JAR attributes have been condensed to “too little,” “JAR,” and “too much.” None of the oranges received equal distributions across the three JAR levels, meaning that the fruit were skewed. For sweetness and sourness, all of the samples showed a substantial number of “too little” ratings, with Navel 2 standing out in terms of a lack of sweetness and sourness. For the textural attributes, the Navel 2 sample and the Thomson Improved samples show wider distributions in their JAR ratings than the other samples. The children had similar proportions in rating these same fruit (data not shown). The combination of low juiciness, sweetness, and sourness can help to explain why Navel 2 and Thomson Improved received the lowest liking ratings from both adults and children (Table 5).

Preference segmentation. Both the adults and children showed statistically significant

segmentation based on their overall liking and liking of taste as seen in Figs. 2 and 3. The adults were segmented into four distinct clusters and the children into two clusters. Whereas a product may perform well on average (i.e., across all consumers), segmentation is indicative of different groups having different liking patterns. Significant differences in liking were seen among the segments in nine fruit samples for the adults and five for the children, respectively (Figs. 2 and 3). This is more distinct than the liking differences when the consumers are considered together (Table 5).

After the consumers were clustered by their overall liking ratings, their JAR ratings were separated, as shown in Table 9. This table shows the differences between the clusters in their overall use of the JAR ratings across all products. Adult clusters 3 and 4 used the sweetness rating differently from the other two groups. Cluster 3 was stricter with their JAR sweetness ratings, meaning they rated more samples to have “too little” sweetness more

than any other group. The fourth cluster did just the opposite. For sourness, cluster 4 gave more “too little” ratings than the other groups, expressing a desire for tarter fruit. Textural attributes showed much higher acceptability among the consumers. Most of the adults felt the samples were appropriately juicy and firm. One slight difference is that the third adult cluster showed a smaller majority with regards to the juiciness level of the samples, indicating that they may prefer if the fruit they consumed were even juicier.

The children used the JAR scales similarly. Cluster 2 showed slightly more particularity toward sweetness, whereas cluster 1 was more sensitive to textural attributes of juiciness and firmness (data not shown). This trend is reflected in the children’s liking of texture (Fig. 4), where cluster 1 had lower ratings than cluster 2.

Drivers of liking. PLS regression was performed to compare liking of all six of the

Fig. 1. (A) Penalty of suboptimal attribute intensities for adults on overall liking. The just-about-right (JAR) ratings for each fruit were compared with their overall liking scores and an average difference between nonideal and JAR ratings was calculated. Mean drop represents the average decrease in overall liking due to a lower than ideal (“-”) or higher than ideal (“+”) JAR rating across all of the oranges. (B) Penalty of suboptimal attribute intensities for children on liking of taste. The JAR ratings for each fruit were compared with their liking of taste scores and an average difference between nonideal and JAR ratings was calculated. Mean drop represents the average decrease in liking of taste due to a lower than ideal (“-”) or higher than ideal (“+”) JAR rating across all of the oranges.

Table 8. Proportion of adult just-about-right (JAR) ratings for sweetness, sourness, firmness, and juiciness of all orange samples. The judges were given two wedges of the same orange to evaluate all of the JAR attributes on a 5-point scale. Ratings of 1 and 2 and 4 and 5 have been combined into “too little” and “too much,” respectively. χ^2 test was performed on the separated ratings before being converted to proportions and combined.

Sample	Sweetness*			Sourness*			Firmness*			Juiciness*		
	Too little	JAR	Too much	Too little	JAR	Too much	Too little	JAR	Too much	Too little	JAR	Too much
Atwood 1	0.43	0.52	0.05	0.34	0.54	0.11	0.19	0.77	0.03	0.07	0.82	0.11
Atwood 2	0.35	0.58	0.06	0.31	0.52	0.17	0.18	0.76	0.06	0.17	0.77	0.06
Fisher 1	0.49	0.48	0.03	0.41	0.50	0.09	0.16	0.71	0.12	0.23	0.73	0.04
Navel 1	0.37	0.57	0.06	0.40	0.54	0.06	0.09	0.78	0.13	0.19	0.76	0.05
Navel 2	0.75	0.21	0.04	0.63	0.31	0.06	0.10	0.63	0.27	0.46	0.50	0.04
Navel 3	0.40	0.57	0.03	0.30	0.57	0.14	0.16	0.75	0.09	0.17	0.77	0.05
Organic	0.31	0.60	0.09	0.33	0.55	0.12	0.22	0.74	0.04	0.12	0.83	0.05
Thomson Improved	0.49	0.44	0.07	0.37	0.49	0.14	0.14	0.63	0.22	0.38	0.59	0.03
Washington 1	0.45	0.51	0.04	0.28	0.53	0.20	0.13	0.83	0.04	0.08	0.82	0.10
Washington 2	0.29	0.60	0.11	0.37	0.55	0.08	0.10	0.80	0.10	0.17	0.77	0.06
Average	0.43	0.51	0.06	0.37	0.51	0.12	0.15	0.74	0.11	0.21	0.74	0.06

*indicates significance ($P \leq 0.05$) by χ^2 test.

Fig. 2. Adult overall liking by cluster for each navel orange sample. Adults were clustered based on their preferences using Ward's agglomeration method and validated through analysis of variance. Numbers correspond to the 9-point hedonic scale with 1 = "dislike extremely" and 9 = "like extremely." Columns within a product category that share the same letter are not significantly different ($P \leq 0.05$) by Fisher's least significant difference test.

Fig. 3. Child liking of taste by cluster for each navel orange sample. Children were clustered based on their preferences using Ward's agglomeration method and validated through analysis of variance. Numbers correspond to the 7-point hedonic scale with 1 = "I hate it" and 7 = "I love it." Columns within a product category that share the same letter are not significantly different ($P \leq 0.05$) by Fisher's least significant difference test.

consumer clusters with the chemical and descriptive attributes as seen in Fig. 5. This analysis compares the descriptive and chemical analyses to the consumer liking. Objects on the plot that are located closer to each other are more correlated. This analysis shows drivers of liking, or attributes that were strongly correlated to the clustered consumer liking. The drivers of liking for the first two adult clusters were sweetness and SSC, both of which were strongly correlated to the California Standard rating scale. The preferences of the second and third adult clusters were driven by juiciness, orange flavor, and overall flavor. The fourth cluster preferred samples higher in TA.

As for the children, some of them largely mirrored their adult counterparts. The first child cluster was driven strongly by the same attributes as the second and third adult clusters—orange flavor, overall flavor, and SSC. The second cluster of children was not as well described by the model. These children

Table 9. Proportion of just-about-right (JAR) ratings by adult consumer clusters for sweetness, sourness, and juiciness across all of the products. The judges were given two wedges of the same orange to evaluate all of the JAR attributes on a 5-point scale. Ratings of 1 and 2 and 4 and 5 have been combined into "too little" and "too much," respectively. χ^2 test was performed on the separated ratings before being converted to proportions and combined.

Adult cluster	Sweetness*			Sourness*			Juiciness*		
	Too little	JAR	Too much	Too little	JAR	Too much	Too little	JAR	Too much
Cluster 1 ($n = 50$)	0.43	0.51	0.06	0.39	0.48	0.13	0.22	0.73	0.05
Cluster 2 ($n = 88$)	0.43	0.53	0.04	0.34	0.54	0.13	0.16	0.78	0.06
Cluster 3 ($n = 70$)	0.47	0.46	0.07	0.37	0.52	0.11	0.26	0.68	0.06
Cluster 4 ($n = 27$)	0.33	0.56	0.11	0.47	0.44	0.09	0.19	0.74	0.08

*Denotes significance ($P \leq 0.05$) by χ^2 test.

had more homogenous ratings for taste, liking all of the samples. One of the negative drivers of liking for almost all clusters was fibrousness, which makes sense given the preferences of the consumers for juicy fruit.

Exit survey. The adults were asked to answer a few questions about their demographics, consumption of citrus, and

psychographic information. The adults were also asked to answer a few questions about their child's eating habits. These data were collected to help characterize the preference segments of consumers. Adults showed significant differences in what time of the year they ate citrus. Adult cluster 4 ate citrus year around more frequently than the other three

Fig. 4. Child liking of texture by cluster for each navel orange sample. Children were clustered based on their preferences using Ward's agglomeration method and validated through analysis of variance. Numbers correspond to the 7-point hedonic scale with 1 = "I hate it" and 7 = "I love it." Columns within a product category that share the same letter are not significantly different ($P \leq 0.05$) by Fisher's least significant difference test.

Fig. 5. Partial least squares regression of consumer liking clusters for adults (overall liking) and children (liking of taste) on descriptive analysis attributes and sugar/acid measurements. Consumer clusters, chemical and descriptive attributes, and products are shown as triangles, circles, and squares, respectively. Descriptive/chemical attributes and consumer clusters were used as the independent and dependent variables, respectively. Objects located closer together on the plot are more correlated. $R^2Y = 36\%$, 24% ; $R^2X = 34\%$, 13% .

clusters ($P < 0.05$). This cluster also showed liking for the more acidic fruit. They may be more used to citrus that is not quite in season, therefore sourer.

Discussion

Overall, the oranges that were tested were liked by both consumer groups, with all

products but one receiving ratings of "liked slightly" or better (Figs. 2 and 3). If there is room for improvement though, it may be in the intensity of both sourness and sweetness of the fruit. Past work with mandarins has found that highly preferred mandarins had an average total soluble solids and acid content of 13.1% and 1.1%, respectively (Goldenberg et al., 2015). None of the fruit

tested in this study exhibited quantities of either attribute at these levels. In fact, all of the fruit samples were quite low in acidity, with all samples at or below 0.8% TA. Given that the period of testing was late January through early February, it seems surprising that all of the fruit would register so low in acid. Nevertheless, these values are quite low for traditional acid measurements in oranges and do indicate that the fruit was quite mature. Despite the low acid content, however, the consumers still felt that the fruit was not sweet enough. Even though the fruit may have been overly mature, the JAR ratings of "too little" sweetness were still marked at 41% of the time for the adults (Table 8) and 26% for the children, indicating that there are still improvements to be made.

The low acid flavor of the fruit was not the only attribute that was similar among the samples. The trained descriptive panel provided results that confirm the similarities among the fruit between the different sensory attributes (Table 4). Despite the panel having many metrics by which to measure flavor differences, only three flavor attributes showed significant differentiation. Of the 32 attributes in the scorecard, less than half were found to be significant, a small proportion indeed for a descriptive analysis. More variation in the chemical and sensory properties of the fruit would have provided stronger insight into consumer preferences.

Looking to the future, the information presented here may have implications for marketing. One commonly cited issue with using consumers to provide any sort of descriptive data, JAR ratings in this instance, is that consumers often do not agree on the stimulus (Lawless and Heymann, 2010). This issue is remedied in descriptive analysis through the use of reference standards, providing consensus among the panelists. However, in this study, we found certain instances in which consumers used JAR attributes which mirrored their preferences. For the

adults, the fourth cluster rated the samples as “not quite sour/tart enough” more than any other cluster (Table 9) and their liking was strongly correlated to the measured value of TA (Fig. 5). For the children, the first cluster was more sensitive to sweetness, indicating that they felt the samples to be “JAR” less often than the other cluster of children. These children are then shown to be mapped closer to soluble solids than the other child cluster (Fig. 5). Although mapping their preferences to specific attributes provides insight for the growers, evidence presented here shows that consumers may be aware of their attribute preferences. These consumers may respond well to advertisement. For example, adult cluster 3 may be enticed by a label indicating the juiciness of the orange, adult cluster 4 might be inclined to purchase a “tart” fruit, and child cluster 1 might be drawn toward packaging that emphasizes sweetness. This study also confirmed that juiciness, overall flavor, orange flavor, and sweetness are critical attributes in determining consumer liking of navel oranges, which is in agreement with past studies (Gao et al., 2011; Obenland et al., 2009; Poole and Baron, 1996). However, these attributes cannot be appropriately conveyed through the external appearance of the fruit (Tables 6 and 7). In this study, liking of appearance showed the lowest correlation among taste and texture to overall liking of the fruit. To remedy this, packers could provide consumers with additional information, such as a sugar or acid range from the lot of fruit, provided to consumers at their point-of-purchase. Ranges of soluble solids or acids, accompanied by words such as “tart” or “extra sweet,” could provide insight for the consumers. A system like this might reward high-quality producers with higher sales and repeat consumers who appreciate clarity and consistency in their fruit.

These data come together to provide a market assessment of 10 commercially produced navel oranges in Northern California. If the 10 fruit samples we collected for this study were any indication, the fruit offered to consumers appear to be very consistent, of relatively high quality, or one that meets consumer expectations. This sampling of fruit, although liked strongly, may not fully represent the entirety of navel

oranges produced commercially in California. The similarity of soluble solid and acid levels does not represent fruit harvested at varying levels of maturity or ripeness which would be seen across the entire harvest period. The homogeneity observed may be due in part to the legislative change in quality evaluation from the SSC/acid ratio in favor of the California Standard value. For most of the consumers, liking showed a strong correlation to this new standard, a welcome sign that this new tool is working well.

Literature Cited

- Ferguson, L. and E.E. Grafton-Cardwell. 2014. Citrus production manual. UCANR Publications, Richmond, CA.
- Gao, Z., L.O. House, F.G. Gmitter, Jr., M.F. Valim, A. Plotto, and E.A. Baldwin. 2011. Consumer preferences for fresh citrus: Impacts of demographic and behavioral characteristics. *Intl. Food Agribus. Mgt. Rev.* 14:23–40.
- Goldenberg, L., Y. Yaniv, T. Kaplunov, A. Doron-Faigenboim, N. Carmi, and R. Porat. 2015. Diversity in sensory quality and determining factors influencing Mandarin flavor liking. *J. Food Sci.* 80:S418–S425.
- Greenhoff, K. and H.J.H. MacFie. 1994. Preference mapping in practice. p. 137–166. In: *Measurement of food preferences*. Springer, Boston, MA.
- Guinard, J-X. 2000. Sensory and consumer testing with children. *Trends Food Sci. Technol.* 11:273–283.
- Guinard, J-X., B. Uotani, and P. Schlich. 2001. Internal and external mapping of preferences for commercial lager beers: Comparison of hedonic ratings by consumers blind versus with knowledge of brand and price. *Food Qual. Prefer.* 12:243–255.
- Helgesen, H., R. Solheim, and T. Næs. 1997. Consumer preference mapping of dry fermented lamb sausages. *Food Qual. Prefer.* 8:97–109.
- Ivans, E. and M. Ferree. 1987. Early-season navel oranges may be too sour for consumers. *Calif. Agr.* 41:20–21.
- Jordan, R.B., R.J. Seelye, and V.A. McGlone. 2001. A sensory-based alternative to brix/acid ratio. *Food Technol.* 55:36–44.
- Lawless, H.T. and H. Heymann. 2010. *Sensory evaluation of food - principles and practices*. Springer, New York, NY.
- Martens, M. and H. Martens. 1986. Partial least squares regression. *Stat. Proc. Food Res.* p. 293–359.
- Naes, T. and E. Risvik. 1996. *Multivariate analysis of data in sensory science*. Elsevier, Amsterdam, The Netherlands.
- National Agricultural Statistics Service (NASS). 2017. Citrus Fruits 2017 Summary USDA, NASS, Washington D.C. <<http://usda.mannlib.comell.edu/usda/current/CitrFruit/CitrFruit-08-31-2017.pdf>>.
- Obenland, D., S. Collin, B. Mackey, J. Sievert, K. Fjeld, and M.L. Arpaia. 2009. Determinants of flavor acceptability during the maturation of navel oranges. *Postharvest Biol. Technol.* 52:156–163.
- Peryam, D.R. and F.J. Pilgrim. 1957. Hedonic scale method of measuring food preferences. *Food Technol.* 11:9–14.
- Plotto, A., E. Baldwin, G. McCollum, J. Manthey, J. Narciso, and M. Irey. 2010. Effect of liberibacter infection (Huanglongbing or “Greening” disease) of citrus on orange juice flavor quality by sensory evaluation. *J. Food Sci.* 75: S220–S230.
- Poole, N. and L. Baron. 1996. Consumer awareness of citrus fruit attributes. *Br. Food J.* 98:23–28.
- Qannari, E.M., E. Vigneau, P. Luscan, A.C. Lefebvre, and F. Vey. 1997. Clustering of variables, application in consumer and sensory studies. *Food Qual. Prefer.* 8:423–428.
- Rothman, L. and M.J. Parker. 2009. Just-about-right scales: Design, usage, benefits, and risks. *ASTM Manual MNL63*. ASTM Intl., Conshohocken, PA.
- Schlich, P. 1995. Preference mapping: Relating consumer preferences to sensory or instrumental measurements. *Colloq. INRA*, Dijon, France.
- Sidel, J.L. and H. Stone. 1993. The role of sensory evaluation in the food industry. *Food Qual. Prefer.* 4:65–73.
- Tenenhaus, M., J. Pagès, L. Ambroisine, and C. Guinot. 2005. PLS methodology to study relationships between hedonic judgements and product characteristics. *Food Qual. Prefer.* 16:315–325.
- Thybo, A.K., B.F. Kühn, and H. Martens. 2004. Explaining Danish children’s preferences for apples using instrumental, sensory and demographic/behavioural data. *Food Qual. Prefer.* 15:53–63.
- Williams, E.J. 1949. Experimental designs balanced for the estimation of residual effects of treatments. *Austral. J. Chem.* 2:149–168.
- Yackinous, C., C. Wee, and J-X. Guinard. 1999. Internal preference mapping of hedonic ratings for Ranch salad dressings varying in fat and garlic flavor. *Food Qual. Prefer.* 10:401–409.