Journal American Pomological Society 56(4):223-229 2002

Development of a Model for Prediction of Consumer Liking from Visual Attributes of New and Established Apple Cultivars

MARJORIE C. KING[1] AND MARGARET A. CLIFF[1]

Abstract

Consumer liking of the appearance of new and established apple cultivars was determined on an hedonic scale. Descriptive analysis, with a panel of 18 judges, was used to characterize the appearance of 10 red apple cultivars which varied in overall size, taper, ribbing, symmetry, uniformity-of-shape, length/width ratio, degree of striping, percent overcolor, brightness of red over-color, groundcolor and uniformity-of-color. Stepwise multiple regression was used to develop models to predict liking from the visual attributes. A one-variable model with extent of ribbing predicted liking, for an experienced panel while a one-variable model with degree of apple taper predicted consumer liking. The effectiveness of the consumer model was validated with a second consumer evaluation, on collections of red or yellow apples with varying taper values. While the developed model was not able to predict the actual liking score, it was able to approximate the rank liking of the yellow apples.

Introduction

Fruit appearance is one factor in the consumers' purchase decisions (6, 8, 16). While it is generally acknowledged that fruit defects are not acceptable, less is known about the visual characteristics that are attractive to consumers. One study which used digital images to examine the preferences of Canadian and New Zealand consumers found that red apples were preferred to green or yellow apples and round or conical shapes were preferred to oblong apples (4). Hampson and Quamme (8) found the ideal size of an apple to be approximately 7.5 cm in diameter, with smaller and larger apples being less acceptable. This study found no significant differences in consumer preferences of shape for 26 different cultivars, however the authors noted a range of hedonic scores for genotypes of similar shape and suggested that factors other than shape may have influenced some scores. Lawless (10) has suggested that hedonic acceptance testing 'tends to be a very global integrated perception' with consumers responding to the product as a whole rather than considering individual product attributes. A study of regional preferences in Canada (3) showed a tendency for consumers to have a greater liking for cultivars that were familiar. In the same study, red apples were preferred to yellow apples. Although these studies have attempted to define consumer preference in terms of objective attributes, there are still many unanswered questions.

Fruit breeding and selection is similar to the development of a product for the market. The more that is known about the factors that influence liking, the better the potential of selecting for fruit that will be commercially successful. Consumer sensory evaluation is expensive and time consuming, requiring large numbers of consumers who are representative of the target market (15). Ideally, consumer research should take place throughout the development process right from product conception to final production (7), however, time and money constraints often prevent this from happening. In addition in fruit breeding there is often a shortage of product for testing.

Researchers have shown that sensory results from smaller trained panels can sometimes be in agreement with consumer panels (9,13, 14) and be useful for predict-

[1]Agriculture and AgriFood Canada, Pacific AgriFood Research Station, Summerland, British Columbia, Canada VOH 1Z0. Contribution no. 2143

ing consumer liking. Although it is preferable to use consumers for the prediction of liking, multivariate statistics have been used to successfully relate laboratory sensory and analytical data to consumer responses (1, 5, 10). This research was undertaken to:1) document the scope of visual characteristics present in a broad range of apple cultivars, 2) select the visual characteristics that are most relevant to consumers, using multiple regression and 3) validate the proposed model(s) for prediction of consumer liking.

Materials and Methods

Apples

Ten apple selections grown at the Pacific Agri-Food Research Centre (PARC), Summerland, British Columbia (BC) were used for the consumer evaluation and for the in-house visual profiling. They consisted of four commercial cultivars ['Delicious' (Bisbee strain), 'Jonagold', 'Royal Gala', 'Jonagored'], two new cultivars ('Creston', 'Ambrosia'), and four unnamed selections ('8B-14-56', '8S-29-18', '8C-05-62', '8NE-07-72'). 'Creston'and the numbered selections were from the breeding program at Summerland. 'Ambrosia' is a chance seedling discovered by a grower in Cawston, BC. Six defect and disease free apples of each cultivar were selected. These were red cultivars since previous studies had already demonstrated that consumers preferred red to yellow colored apples.

For the validation data set, a collection of six yellow ('Golden Delicious', 'Silken', 'Ginger Gold', 'Peypring Cervecko',8C-06-20) and six red ('Discovery', 'Summerred', 'Karin Schneider', '9P-15-30', '11W-12-11', '11W-12-85') selections were collected from the PARC orchards to represent a broad range of taper values. Taper was defined as the perceived change in width from the widest to the narrowest part of the apple. The measured height/diameter ratios for these cultivars are given in Table 4.

Descriptive analysis

Descriptive analysis was used to profile the visual characteristics of each selection. The apple breeder and two members of the PARC sensory team examined selected visual attributes, commonly used in the literature, which would describe the apples (Table 1). Six representative fruit from each of the 10 selections were polished and placed on white polyethylene trays labelled with three digit random numbers. Three apples were oriented with their stem up and three apples were oriented with their calyx up. Nineteen panelists were selected from the pool of PARC employees who regularly participated in the sensory evaluation of selections in the apple breeding program. All of these panelists had received training in apple profiling and were experienced in scoring attribute intensities. Prior to the evaluation, each panelist reviewed the descriptions of the attributes and were shown examples of the visual anchors. The selections were evaluated in random order under natural light by placing marks on a 10 cm line scale according to the perceived intensities. All evaluations were done in duplicate. Line scales were anchored at 2 and 8 cm with either verbal descriptors or reference samples/objects, as described in Table 1. The panelists' scores were quantified by measuring the distance of their mark from the origin (maximum score =10).

Consumer evaluation

One hundred Consumers ranging in age from 5 to 80 years were recruited from visitors to the Living Landscapes Festival (Kelowna, B.C. October, 1996). All of the consumers liked apples and were interested in evaluating new apple selections. Six fruit of each selection were polished and placed on white plastic trays, as described above. Trays were labelled with three-digit random numbers and placed on tables covered with white paper.

The 10 selections were divided randomly into two sets for the consumer visual evaluation with 'Royal Gala' included in both sets. All apple selections were evaluated according to a completely randomized design by placing a mark on the line scale (10 cm) according to the degree of liking. Scales were labelled at 1, 3, 5 and 7 cm with the terms dislike, neutral, like-moderately and like-very-much, respectively. The evaluation was based on visual attributes only.

Table 1. Definition of the visual attributes evaluated by the 'experienced' profiling panel.

Attribute	Definition
Length/width ratio	The perceived ratio of length to width. Anchored by computer generated diagrams at 2 cm (length 8 cm and width 8.5 cm) and 8 cm (length 10.5 cm and width 8.5 cm).
Taper	The perceived change in width from the widest to the narrowest point of the apple. anchored at 2 cm with 'Cox's Orange Pippin apple which has no taper and at 8 cm with a 'Delicious'.
Ribbing	The appearance of the outline of the apple when viewed from the stem. a non-ribbed apple has a smooth outline while a ribbed apple has promi nences at five points where the carpels meet. The 10 cm line scale was anchored at 2 cm with the verbal descriptor 'non-ribbed' and at 8 cm with the verbal anchor 'ribbed'.
Symmetry	The uniformity of the shape of individual apples. Anchored at 2 cm and 8 cm with the verbal anchors 'unsymmetrical' and 'symmetrical', respectively.
Uniformity-of-shape	The uniformity of shape among the five apples that formed the set that was being evaluated. The verbal anchors were not uniform (2cm) and uniform (8 cm), repsectively.
Size	Size from small to large anchored at 2 cm by a small 'Royal Gala' (weight = 200 g, length = 65 mm, width = 66 mm) and at 8 cm by a large 'Jonagold' (weight = 330 g, length = 97 mm, width = 91 mm.)
Striping	The relative amount of striping on the apple. anchored at 2 cm with a seedling that was 100% striped and at 8 cm with a seedling that was a solid color apple.
Ground-color	The character of the ground color, anchored at 2 cm with a 'green/yellow' Munsell color chip (10Y 8/8) and at 8 cm with a 'darker yellow/gold ' Munsell color chip (5Y 8/10).
Percent over-color	The percent of red color that was laid over the ground color. The verbal anchors, 40% and 90%, were at 2 cm and 8 cm, respectively.
Red over-color	The character of the red color on the apple. the line scale was anchored at 2 cm with a 'bright red/orange' Munsell color chip (7.5R 4/6) and at 8 cm with a 'dark red' Munsell color chip (5R 2/6).
Uniformity-of-color	The uniformity of color among the five apples that formed the set that was being evaluated. the line scale was anchored at 2 cm and 8 cm with the verbal anchors 'not uniform' and 'uniform'.
Visual liking	A measure of how much the panelists liked the appearance of the apples. The scale was anchored at 2 cm and 8 cm with the verbal anchors 'dislike very much' and 'like very much' were at 2 cm and at 8 cm respectively.

Validation of model

Two hundred and sixty-six consumers ranging in age from 5 to 80 years who attended a PARC Open House (Sept 1998) evaluated visual liking for six yellow or six red selections. As outlined above, fruit were assigned a three digit random code, polished and placed on white trays. Consumers rated their overall visual liking of the apples in random order, under natural light, by placing a mark along a 10 cm line scale. The scale was labelled at 1, 3, 5, 7 and 9 cm with dislike very much, dislike moderately, neither like not dislike, like-moderately, and like-very-much, respectively.

Twelve judges who were employees of PARC, evaluated the degree of apple taper by placing a vertical mark on a 10 cm line scale. Sets were evaluated in random order and in duplicate. The mean taper score for each of the selections was substituted into the proposed model (regression equation) to obtain a predicted consumer liking score.

Statistical analysis

One-way analyses of variance (ANOVA) were used to determine differ-

Table 2. Mean physical dimensions (n=6) and experienced panel profiling scores for the sensory attributes (n=38)[2] from the experienced panel.

Apple Cultivar	Length/Width Ratio	Average Weight (g)	Perceived Length/Width Ratio	Taper [1]	Ribbing[1]	Symmetry[1]	Uniformity of-shape[1]	Size[1]	Striping[1]	Percent-over-color[1]	Red over-color[1]	Ground color[1]	Uniformity of-color[1]
Jonagold	0.905	304.6	5.5	4.1	4.2	4.7	5.6	6.9	6.0	5.5	4.2	5.0	5.5
Royal Gala	0.919	182.1	4.7	5.1	3.6	6.7	6.9	3.3	7.1	6.5	3.9	6.0	6.7
Creston	0.969	241.2	6.3	5.4	4.2	5.3	5.0	5.3	7.4	3.6	3.6	5.1	5.1
Ambrosia	0.886	190.9	4.1	4.9	4.3	4.8	5.6	3.7	4.7	4.6	4.1	6.0	4.9
8B-14-56	0.939	271.3	4.9	4.6	4.0	4.8	4.8	6.6	2.3	7.8	5.3	4.7	7.0
8S-29-18	0.875	201.8	3.6	4.2	5.0	4.7	5.3	4.5	3.4	7.2	5.6	5.3	5.9
Delicious, Bisbee	0.996	243.6	7.3	7.1	7.5	4.2	5.7	6.2	2.2	8.4	8.1	3.6	7.6
8C-05-62	0.876	216.6	3.5	3.6	2.6	5.9	5.5	5.1	3.0	6.3	4.4	5.0	4.8
8NE-07-72	0.909	210.7	4.5	4.1	5.1	3.5	4.5	5.0	2.8	6.3	4.4	6.2	5.4
Jonared	0.912	290.1	4.2	3.7	3.7	5.2	5.7	6.6	6.1	6.7	6.2	5.8	5.7
significance			0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
min	0.875	182.1	3.5	3.6	2.6	3.5	4.5	3.3	2.2	3.5	3.6	3.6	3.8
max	0.996	304.6	7.3	7.1	7.5	6.7	6.9	6.9	7.4	8.4	8.1	6.2	7.6
standard deviation (std)	0.04	42.27	1.2	1.03	1.29	0.88	0.65	1.25	2.01	1.45	1.37	0.78	0.94
least significant difference (LSD)			0.48	0.55	0.47	0.51	0.56	0.38	0.49	0.46	0.56	0.64	0.51

[1]maximum intensisty score = 10
[2]19 judges x 2 replications

ences, for the visual attributes and liking scores. Fisher's least significant difference (LSD) was used to compare mean scores for the apple selections. Stepwise multiple regression with forward selection, was used to develop models to predict experienced panel and consumer liking scores. All statistical analyses were conducted using SAS (SAS Institute Inc, Cary, NC).

Once the models were developed, the predicted consumer liking score was calculated using the degree of taper and compared to the actual consumer liking. For each of the validation data sets (yellow apples, red apples), the predicted and actual consumer liking scores were compared using the square root of the prediction error $[\Sigma(\text{predicted-observed})^2/n]$.

Results and Discussion

Descriptive profiling was used to describe the visual characteristics of the apple selections (Table 2). All attributes

Table 3. Mean liking scores for both consumers and an 'experienced' panel.

Apple Cultivar	Consumer Liking[1] (n=100)	'Experienced panel' Liking[2] (n^3=38)
Delicious	4.6	4.1
Creston	5.4	5.4
Jonagored	5.7	5.1
8NE-07-72	5.7	5.3
Ambrosia	5.7	5.5
8S-29-18	6.0	5.9
8C-05-62	6.1	6.3
8B-14-56	6.0	6.1
Jonagold	6.3	5.3
least significant difference (LSD) (p=0.05)	0.12	0.55

[1,2] maximum score = 10
[3] 19 judges x 2 replicates

were significant in describing differences among the cultivars (Table 2). The relatively small LSD values were attributed to the experience of the panel as well as the relative ease of visual scaling compared to scaling for taste and smell. The terms size, ribbing and percent over color had the smallest LSD values and were the most discriminating of the attributes; whereas, the terms ground color, uniformity-of-shape, and percent over-color were the least discriminating of the attributes.

With the exception of 'Jonagold', there was general agreement between the experienced and consumer panels when the selections were ranked according to the liking score (Table 3). Varietal recognition may have influenced the ratings for 'Jonagold' and 'Delicious' by the experienced panel since product knowledge can influence hedonic scoring (15). At the same time, the results were consistent with other work which suggested that if product knowledge is limited, smaller 'in house' panels may provide a preliminary indication of a more broad based hedonic consumer evaluation (11).

Multiple regression analysis of the mean scores for each attribute was used to select the attributes most important for predicting the liking score of the experienced panel. Although all visual attributes were significant in characterizing differences among the cultivars (Table 2) only the term ribbing was useful in the prediction (n=38). The prediction equation was:

$$\text{Experienced panel liking} = -0.421\ \text{Ribbing} + 7.40 \qquad (1)$$

This model describes a negative linear relationship between liking and ribbing, where the greater the ribbing the less the visual appearance of the apple was liked. The equation had a correlation coefficient of 0.77 and a coefficient of determination of 0.59, indicating that 59 percent of the variation in the experienced liking could be explained by ribbing. 'Delicious' had the lowest liking score (Table 3) and had the highest degree of ribbing (Table 2). In contrast, the selection 8C-05-62, with the lowest ribbing score (Table 2), was the most preferred (Table 3).

Multiple regression of the attribute mean scores from the experienced panel against the consumer liking scores was used to develop an additional prediction model. Taper, was the only significant variable for predicting consumer liking. The prediction equation was:

$$\text{Consumer liking} = -0.444\ \text{Taper} + 7.70 \qquad (2)$$

This equation describes a negative linear relationship between liking and taper, where the greater the taper the less the consumer liking. The equation had a correlation coefficient of 0.85 and a coefficient of determination of 0.72. The similarity of equations (1) and (2) was due to the interrelationship of the terms ribbing and taper. These attributes were the most highly correlated ($R=0.52$, $p< 0.001$) of all the visual characteristics (Table 1).

The influence of taper on visual liking was examined in an additional consumer evaluation. Consumers (n=266) indicated their degree of liking for a collection of red or yellow apples that varied in their amount of taper from round to oblong. For the yellow cultivars, the rounder apples

Table 4. Actual versus predicted rank order of consumer liking of yellow and red apples with different taper values.

	Length/Width Ratio	Mean Taper Scores[2] (n^3=24)	Observed Consumer Liking[1] (n=266)	Observed Consumer Rank Order	Predicted Consumer Liking[1] (n=266)	Predicted Consumer Rank Order
YELLOW APPLES						
Golden Delicious	1.0	6.1	4.2	1	5.0	3
Silken	.9	5.2	3.8	2	5.4	1
8S-06-20	1.0	5.8	3.8	3	5.2	2
Ginger Gold	.9	7.2	3.7	4	4.6	4
Peypring Cervecko	1.0	8.6	3.2	5	4.0	5
Seedling	1.2	9.3	2.6	6	3.7	6
RED APPLES						
11W-12-11	1.0	6.9	4.2	1	4.7	5
9P-15-30	.9	4.6	3.6	2	5.8	3
Discovery	.8	3.2	3.4	3	6.4	1
11W-12-85	1.0	5.5	3.2	4	5.4	4
Summered	1.0	8.7	3.1	5	3.9	6
Karin Schneider	.7	3.2	2.5	6	6.4	2

[1]Maximum consumer liking score = 10
[2]Maximum taper score = 10
[3]12 judges x 2 replications

('Golden Delicious', 'Silken', 8S-06-20) were most liked, while the oblong apples ('Ginger Gold', 'Peypring Cervecko', seedling) were least liked (Table 4). This was consistent with other research that showed that, independent of apple color, Canadian and New Zealand consumers preferred round and conical apples to oblong apples (4). The taper scores were highly correlated with the measured length /width ratio. The correlation coefficients were 0.91 for the yellow apples and 0.99 for the red apples. Although the models were developed on visual estimations of taper, the high correlation between this rating and the measured ratio should allow plant breeders to use objective measurements in screening their selections.

For the collection of yellow apples the equation was somewhat successful in predicting the order of consumer preference with a square root of the prediction error of 1.49. For the red apples, the equation was less successful as indicated by the large variation in the ranked order between the actual and predicted liking (Table 4) and a higher square root of the prediction error (2.38). The least liked red cultivar ('Karin Schneider', Table 4) was predicted to be the most liked on the basis of its round shape. However, this cultivar had heavy russeting, an extremely dark red color and conspicuous lenticels - all factors which were unaccounted for by the model. Stem bowl russet and large lenticles have been shown in other studies (8) to have a negative influence in consumer acceptance. This result indicates attributes other than shape may be more important in determining liking.

Prediction models must be interpreted within the context that they were developed. The model developed by this study was applicable to red apples in good condition with a size range comparable to that normally found in the marketplace. In addition all apples were free of russet and lacked conspicuous lenticles. The consumers in this work were residents of Southern British Columbia representing a wide range of ages and socio-economic status. The work showed that within these parameters consumers were less likely to like tapered apples. This is consistent with work by Hampson and Quamme (8) where length, taper and calyx-end-lobing all had a negative influence on shape preferences.

Conclusions

This study showed that apple shape has an influence on how much the consumer likes the appearance of an apple. It showed that in the context of red apples in good

condition consumers were less likely to like apples that were highly tapered.. A one-variable regression model, with the variable taper was moderately successful in predicting consumer liking of new apple cultivars. To this end, this work will serve apple breeders, market researcher and scientists who wish to estimate consumer liking prior to release of a new cultivar into the marketplace.

Acknowledgments

The authors would like to thank Kimberly Dever, Paul Ferguson and Dogan Yuksel for their assistance.

Literature Cited

1. Carr B.T., D. Craig-Petsinger and S. Hadlich. 2001. A case study in relating sensory descriptive data to product concept fit and consumer vocabulary. Food Quality and Preference 12:407-412.
2. Cliff, M.A., M.C. King and C. Hampson.1998. Comparison of mean scores and R-indices for consumer preferences of apple cultivars. HortScience 33(7):1239-1240.
3. Cliff, M.A., K. Sanford and E. Johnston. 1999. Evaluation of hedonic scores and R-indices for visual, flavour and texture preferences of apple cultivars by British Columbian and Nova Scotian consumers. Can. J. Plant Sci. 79:395-399.
4. Cliff, M, K. Sanford, W. Wismer and C. Hampson. 2001. Use of digital images for evaluation of factors responsible for visual preference of apples by consumers. HortScience XX: in press.
5. Daillant-Spinnler, B., H.J.H. MacFie, P.K. Beyts and D. Hedderley. 1996. Relationship between perceived sensory properties and major preference directions of 12 varieties of apples from the southern hemisphere. Food Quality and Preference 7(2):113-126.
6. Kappel, F., R. Fisher-Fleming and E.J. Hogue. 1995. Ideal pear sensory attributes and fruit characteristics. HortScience 30 (5):988-993.
7. Goldman, A. 1994. Gaining the competitive edge: Sensory science in the market place. Cereal Foods World 39 (11/12):822-826.
8. Hampson, C.R. and H.A. Quamme. 2000. Use of preference testing to identify tolerance limits for fruit visual attributes in apple breeding. HortScience 35(5):921-924.
9. Hampson, C.R, Quamme, H.A., Hall, J.W., MacDonald, R.A., King, M.C. and Cliff,M.A. 2000. Sensory evaluation as a selection tool in apple breeding. Euphytica 111: 79-90.
10. Lawless, H.T. 1994. Getting results you can trust from sensory evaluation. Cereal Foods World 39 (11/12) 809-814.
11. Munoz, A.M., E. Chambers and S. Hummer. 1996. A multifaceted category research study: How to understand a product category and its consumer responses. J Sensory Studies 11:261-294.
12. O'Mahony M. 1995. Sensory measurement in food science: fitting methods to goals. Food Technol. 49:73-82.
13. Roberts, A.K. and Z.M. Vickers. 1994. A comparison of trained and untrained judges' evaluation of sensory attributes and liking of cheddar cheeses. J Sensory Stud. 9:1-20.
14. Shepherd, R., C.A. Farleigh, D.G. Land and J.G. Franklin. 1984 Validity of a relative-to-deal rating procedure compared with hedonic rating. *In*: Progress in flavour Research ed. J. Adda. Elsevier Science Publishers. Amsterdam. pp. 103-110.
15. Sidel, J. L., H. Stone, and H.A. Thomas. 1994. Hitting the target: sensory and product optimization. Cereal Foods World 39 (11/12):826-830.
16. von Alvenseleben, R. And T. Meier. 1990. The influence of origin and variety on consumer perception. Acta-Horticulturae. 239:151-161.

Orchard Soil Compaction and Earthworms

Up to 200 passages of farm machinery per year considerably reduced both earthworm density and biomass in orchards. Juvenile worms were more affected by soil compaction than adult ones mostly because juveniles are usually found in upper soil layers and adults in deeper layers. Authors found more juveniles per adults in untrafficked than in traficked plots. From Pizl 1992. Soil Biol Biochem 24(12):1573-1575.