

Consumer evaluation of scab-resistant apple cultivars in Sweden

Åsa Jönsson and Hilde Nybom

Balsgård – Department of Plant Breeding and Biotechnology, Swedish University of Agricultural Sciences, Fjälkestadvägen 459, SE-29194 Kristianstad, Sweden, e-mail: asa.jonsson@vv.slu.se

The feasibility of using untrained consumers for sensory evaluation was investigated on Swedish-grown scab-resistant apple cultivars in 2002 (11 cultivars) and 2003 (15 cultivars). The consumers evaluated five cultivars each, including the control cultivar 'Aroma'. Each cultivar was scored for appearance, texture, juiciness, taste and overall impression by a minimum of 100 consumers in 2002 and by 200 in 2003. All quality attributes were positively correlated when calculations were performed across all consumer scores, with overall impression and taste having the strongest association. 'Rubinola' and K:1160 emerged as the best liked cultivars overall. Cultivar means were analysed for relationships with some previously scored or measured variables and with consumer profile data. Positive correlations with overall impression were obtained for amount of red colour in 2002, and for sugar levels in 2003. Among consumer profile variables, age had the largest influence on consumer scores, with some indications that older consumers preferred (or at least tolerated) soft-fleshed cultivars better than did younger consumers. Minimum number of consumers needed to obtain reliable results was determined by computer simulations indicating that the error rate became stabilized already with 50 consumers.

Key-words: *Malus*, scab resistant, apple, fruit attributes, fruit characteristics, sensory evaluation, sugar content

Introduction

Most of the commonly grown apple cultivars are susceptible to detrimental fungal diseases like apple scab and powdery mildew and to various insects, thereby requiring frequent use of fungicides and pesticides. Recently, there has been an increas-

ing interest in disease-resistant apples that can be grown with a more restricted use of chemicals (Janse 1993, Kellerhals et al. 1997). All over the world, apple breeding programs now have disease resistance as one of the main goals, and more than 100 new disease-resistant apple cultivars have been released since 1970 (Brown and Maloney 2003).

Certainly only a few of the new apple cultivars will become commercially important. The comparatively short and cool growing season in the Nordic countries makes it impossible to grow many of the new cultivars that have been developed elsewhere. The success of a new cultivar is, however, ultimately determined by consumer preferences, which are affected by many different parameters (review in Harker et al. 2003). Although appearance was the most important factor when choosing fruit and vegetables in the shop (von Alvensleben and Meier 1990), consumers rated both texture and taste as more important when given the possibility to taste the fruit prior to making a purchase (Daillant-Spinner et al. 1996). Moreover, consumer preferences can vary considerably over geographical distances as well as between groups within the population (Hampson and Quamme 2000, Richards 2000). A visual consumer test performed in Canada and New Zealand thus showed that preferences for the appearance of apples differed both between the two countries as well as between different cities in the same country (Cliff et al. 2002).

Sensory evaluations are often performed with trained panels (Watada et al. 1980, Daillant-Spinner et al. 1996, Hampson et al. 2000, Weibel et al. 2003). These panels generally consist of at least 20–30 people, although Hampson et al. (2000) found 11 trained panellists to be sufficient for routine screening of breeding selections. The use of a trained panel can be rather costly and suitable panellists are not always easily available. Therefore, some other kind of consumer testing may be needed before deciding on e.g. commercial release of a new cultivar or the planting of many trees of a relatively unknown cultivar (Hampson et al. 2000).

The objective with our study was to investigate the feasibility of using relatively inexpensive and easily available mass testing by untrained consumers. This was therefore investigated during two years in Sweden. A set of 11 scab-resistant apple cultivars together with the control 'Aroma' was evaluated in connection with an apple fair in the year 2002. The evaluation was repeated in the following year with 15 scab-resistant cultivars. Con-

sumer scores were subsequently analysed for relationships with some previously measured variables (shape, background colour, relative area covered by red colour, firmness and sugar content) and with consumer data (age, gender, access to homegrown apples, average apple consumption, and interest in buying organically grown apples).

Material and methods

A large apple fair with apple-related entertainment for the general public is arranged each year in Kivik, situated in the apple district Österlen in SE Sweden. In connection with this fair, a sensory evaluation of 11 (28–29 September 2002) and 15 (27–28 September 2003) scab-resistant apple cultivars was arranged at a small museum called *Äpplets hus* (translated as the Apple house) in Kivik. The well-known and appreciated but not scab-resistant Swedish cultivar 'Aroma' was included to provide a reference. Cultivars analysed in both years were Co-op 12, 'Dayton', 'Scarlet O'Hara' and 'William's Pride' from the US, 'Rubinola' and 'Vanda' from the Czech Republic, and 'Frida', 'Fredrik', K:1016, K:1160 and K:1210 from the Swedish apple breeding program at Balsgård. In addition to these, 'Waleria' from Poland, 'Redfree' from the US and 'Reglindis' from Germany were analysed in 2003 (Table 1). In this year, the evaluated K:1016 fruit was obtained from two different orchards and therefore treated as belonging to two different cultivars; K:1016 B (Balsgård) and K:1016 ÖS (Östra Sönnarslöv). Most of the analysed cultivars have the scab resistance gene V_f from *Malus floribunda*. 'Reglindis' has instead inherited polygenic scab resistance (V_A) from 'Antonovka'.

The cultivars were divided into three groups with five cultivars in 2002 and four groups with five cultivars in 2003. 'Aroma' was included in each group. Fruit was kept at room temperature in 2002 and outdoors in 2003 for at least 12 hours before tasting. Each consumer tasted all five cultivars in one of these groups. The consumers were

Table 1. Presence of scab resistance genes according to the respective breeders, and parameters measured within this study: shape, visual scores of background colour, covering colour and percentage covering colour, sugar measured with a refractometer and firmness measured with a penetrometer (fruitpressure tester).

Cultivar	Scab resistance	Shape	Background colour	Covering colour	%	Sugar % SS	Firmness N
2002							
Aroma	None	Globose to short-conical	Yellow	Red	60	11.2	60.8
Co-op 12	Vf	Globose to conical	Yellow	Purple-red	50	12.3	47.1
Dayton	Vf	Slightly oblate to globose or short-conical	Green	Red	40	13.3	72.6
Fredrik	Vf	Globose to conical	Green	Red	50	14.9	57.9
Frida	Vf	Globose to conical	Yellow	Red	70	13.4	80.4
K:1016	Vf	Oblate to globose	Dark green	Red	40	13.8	17.7
K:1160	Vf	Conical	Green	Red	70	14.2	54.9
K:1210	Vf	Truncate conical	Dark green	Dark red	40	12.6	62.8
Rubinola	Vf	Slightly oblate to short-conical	Green	Red	50	14.2	78.5
Scarlet O'Hara	Vf	Globose to short-conical	Green	Red	50	13.2	101.0
Vanda	Vf	Oblate	Green	Red	50	14.3	59.8
William's Pride	Vf	Globose to conical	Yellow	Purple-red	80	13.1	58.9
2003							
Aroma	None	Globose to short-conical	Green	Red	70	13.9	64.7
Co-op 12	Vf	Globose to conical	Yellow	Purple-red	80	14.2	47.1
Dayton	Vf	Slightly oblate to globose or short-conical	Green-yellow	Red	60	13.0	77.5
Fredrik	Vf	Globose to conical	Green	Red	60	13.3	65.7
Frida	Vf	Globose to conical	Green	Red	64	12.4	96.1
K:1016 B	Vf	Oblate to globose	Green	Red	50	13.5	62.8
K:1016 ÖS	Vf	Oblate to globose	Green	Red	60	12.7	73.6
K:1160	Vf	Conical	Green-yellow	Red	70	13.8	37.3
K:1210	Vf	Truncate conical	Green-yellow	Dark red	76	12.3	77.5
Redfree	Vf	Oblate to globose	Yellow	Red	80	12.8	59.8
Reglindis	VA	Globose to short-conical	Yellow	Red	50	15.5	90.3
Rubinola	Vf	Slightly oblate to short-conical	Yellow	Red	68	15.5	105.9
Scarlet O'Hara	Vf	Globose to short-conical	Green	Red	60	13.2	112.8
Vanda	Vf	Oblate	Green-yellow	Red	56	12.9	68.7
Waleria	Vf	Short-conical	Green	Red	80	14.0	63.8
William's Pride	Vf	Globose to conical	Yellow	Purple-red	80	11.9	68.7

Vf = scab resistance from *Malus floribunda*, VA = scab resistance from 'Antonovka', SS = soluble solids

asked to score each cultivar for appearance, texture, juiciness, taste and overall impression on a 100-mm hedonic line scale on a form sheet. These variables were not defined by us in any way, thus the scores are more likely to reflect the consumer's

own interpretation. Five fruits of each cultivar were placed on a plate for judging the appearance. Apples were cut in approximately ten slices for judgement of the other attributes. For appearance, taste and overall impression, the scales were la-

belled “very poor” and “very good” at the left and right anchor points. For texture, the scale was labelled “very soft” and “very firm”. For juiciness, the scale was labelled “very dry” and “very juicy”.

The consumers were also asked to fill in their age (10–15, >15–20, >20–30, >30–40, >40–50, >50–60 and above 60, although for calculations only three classes were used, 10–30, >30–50 and >50 years old), gender, whether they had access to home-grown apples, their average apple consumption (one or more apples a day, 2–5 apples a week, one apple a week, one apple a month or a maximum of five apples a year), if they prefer to buy organically grown apples and whether they would still prefer organic apples if the price was five Swedish crowns (0.6 Euro) higher per kilo than the price for conventionally grown apples.

On the day before the evaluation weekend, five apples of each cultivar were taken randomly for physical and chemical analysis. Each cultivar was classified into one of 9 different shape categories (Table 1). Background colour was scored visually on a 4-level scale: dark green, medium green, greenish yellow and yellow. Amount of covering red colour was estimated visually for each fruit to the nearest 10% of the whole surface. After removal of the skin, firmness was investigated with a fruit pressure tester, Facchini, mod. FT 327. Each fruit was penetrated twice, on opposite sides, to a depth of 8 mm with a probe with the diameter 11.3 mm. Sugar content (soluble solids) was measured with an Atago digital refractometer PR-1 and expressed as % SS.

Statistical analysis

Consumer scores for the quality attributes (appearance, texture, juiciness, taste and overall impression) were normally distributed and parametric statistics were therefore used. In 2003, scores for the reference cultivar ‘Aroma’ differed significantly between the four test groups, possibly affected by the sequence in which the cultivars were evaluated and by the perceived relative quality of the other cultivars in the group. In order to compare all

cultivars, mean scores for ‘Aroma’ were subtracted from the mean scores for each of the other cultivars in the same test group. Subsequently, the overall mean for ‘Aroma’ (across all test groups) was added to each of the cultivar means.

Pairwise associations between consumer scores for the five attributes were analysed across the whole data set with Pearson correlation coefficients. The five attributes as well as measurements of the amount of covering colour, firmness and sugar content, were also compared using non-parametric Spearman rank correlation based on cultivar means. In addition, one-way ANOVA together with Tukey HSD were used to detect significant differences among the cultivars for each of the five attributes.

Possible influence of the consumer profile variables on the evaluation scores was analysed for each cultivar separately, using one-way ANOVA together with Tukey HSD (age and consumption) or Mann-Whitney U-test (gender, access to home-grown apples and interest in buying organic apples).

The 352 consumer scores for overall impression of ‘Aroma’ in 2002, were used to determine the minimum amount of consumers required for a reliable consumer test. Means were calculated for 100 random subsets consisting of 10, 20, 30, 40, 50, 60, 80, 90 and 100 evaluation scores, respectively. Standard deviation was calculated for each group of subsets and then plotted as a function of the number of consumer scores in the group.

All calculations were carried out using SPSS 10 Data Analysis Package (Norusis 1990).

Results

Associations between quality attributes

All five quality attributes were significantly correlated in 2002 when analyses were performed on all the individual test scores. Possibly there was a tendency for consumers to either like or dislike a certain cultivar in all respects. Nevertheless, it is in-

teresting to note that the highest correlation was obtained between taste and overall impression ($r = 0.84$, $P = 0.00$). Relatively high correlations with overall impression were noted also for appearance ($r = 0.50$, $P = 0.00$) and juiciness ($r = 0.41$, $P = 0.00$). Fruit texture showed surprisingly low correlations with the other attributes, possibly because a firm texture could be perceived either as desirable (crispy) or undesirable (unripe and too hard).

In 2003, overall impression was again correlated with all other quality attributes, with taste being highest also this year, ($r = 0.80$, $P = 0.00$). Obviously, appearance ($r = 0.40$, $P = 0.00$), juiciness ($r = 0.38$, $P = 0.00$) and fruit texture ($r = 0.15$, $P = 0.00$) had less importance for overall impression.

Appearance

Cultivar means for appearance varied from 5.0 to 7.8 (a range of 2.8 units) in 2002 (Table 2) and from 6.3 to 8.7 (2.4 units) in 2003 (Table 3). K:1160 had the highest mean for appearance in 2002, followed by 'Aroma', while the two bottom cultivars were 'Vanda' and K:1210. In 2003, 'Rubinola' had a significantly higher mean than all

others, whereas 'William's Pride' and 'Waleria' had the two lowest means within a group of eight cultivars that were not significantly separated.

Appearance was correlated with overall impression ($r = 0.74$, $P = 0.01$) and taste ($r = 0.68$, $P = 0.02$) when cultivar means for 2002 were compared (Table 4). In 2003, appearance was not associated with any of the other attributes (Table 5). No associations were detected between appearance on the one hand and shape or background colour on the other hand, and only a very weak association with the amount of red colour in 2002 ($r = 0.49$, $P = 0.10$) and none in 2003.

Texture

For texture, cultivar means ranged widely, from 3.5 to 7.2 (3.7 units) in 2002 (Table 2) and from 2.6 to 6.9 (4.3 units) in 2003 (Table 3). In 2002, 'Scarlet O'Hara' was scored as significantly harder than the other cultivars, and Co-op 12 as softer. In 2003, Co-op 12 again had the lowest score while 'Frida' had the highest.

Cultivar means for texture were not associated with any of the other attributes in either of the two

Table 2. The mean scores given by consumers for overall impression, taste, appearance, texture and juiciness of 12 apple cultivars in 2002. (n = 97–352).

Cultivar	Overall impression		Taste		Appearance		Texture		Juiciness	
	stdev	stdev	stdev	stdev	stdev	stdev	stdev	stdev	stdev	
Aroma	6.8	2.1	6.3	2.5	7.8	1.8	5.8	1.6	6.1	1.7
Co-op12	5.4	2.2	5.1	2.7	6.8	1.9	3.5	1.8	4.9	2.0
Dayton	5.1	2.0	4.3	2.2	6.6	2.0	6.0	1.6	5.0	1.8
Fredrik	6.3	2.1	6.0	2.5	6.4	2.1	5.5	1.7	5.7	1.8
Frida	6.3	2.1	5.8	2.4	7.6	1.7	6.4	1.7	5.0	1.9
K:1016	4.8	2.3	4.3	2.7	5.9	2.4	6.1	1.7	5.0	2.1
K:1160	7.0	2.1	6.6	2.5	8.3	1.5	4.7	1.7	5.6	1.7
K:1210	5.0	2.0	4.6	2.3	5.0	2.1	6.1	1.7	4.8	1.9
Rubinola	7.5	1.7	7.4	2.0	7.5	1.8	5.7	1.5	5.7	1.5
Scarlet O'Hara	6.1	2.2	5.8	2.5	6.9	2.2	7.2	1.6	4.5	2.1
Vanda	5.8	2.0	5.3	2.5	5.2	2.2	6.0	1.7	6.0	1.5
William's Pride	5.5	2.0	5.6	2.3	5.4	2.6	5.2	1.8	4.9	1.9

Table 3. The mean consumer scores for overall impression, taste, appearance, texture and juiciness for the year 2003. Number of consumer scores varied between 197 and 215 for all cultivars except for Aroma (n = 829–839) and Fredrik (n = 418–431).

Cultivar	Overall impression		Taste		Appearance		Texture		Juiciness	
	stdev	stdev	stdev	stdev	stdev	stdev	stdev	stdev	stdev	stdev
Aroma	7.0		6.8		7.0		6.0		5.8	
Co-op 12	6.4	3.3	6.7	4.2	7.2	3.3	2.6	2.6	5.1	3.1
Dayton	6.7	2.6	6.4	3.1	7.2	2.8	6.4	1.9	5.7	1.8
Fredrik	6.4	2.4	6.0	2.9	6.9	2.5	5.9	2.0	5.7	1.9
Frida	5.8	2.2	5.0	2.9	7.2	2.0	6.9	1.9	4.9	1.9
K:1016 B	6.8	2.2	6.1	2.8	6.9	2.1	6.1	1.8	5.6	1.6
K:1016 ÖS	6.5	2.3	5.7	2.7	7.4	2.2	6.4	1.7	5.7	1.6
K:1160	6.6	3.1	6.4	3.6	7.6	2.7	4.6	2.1	5.6	2.4
K:1210	6.1	2.7	5.5	3.4	7.4	2.7	5.9	2.3	5.5	2.3
Redfree	5.5	2.7	5.2	3.4	7.6	2.4	4.0	2.5	4.5	2.5
Reglindis	7.0	2.8	7.0	3.5	7.0	2.7	4.6	2.6	5.8	2.2
Rubinola	8.4	2.4	8.7	3.1	8.7	2.5	6.1	1.9	5.8	2.1
Scarlet O’Hara	6.7	2.4	6.9	3.2	6.6	2.4	6.7	1.9	5.0	2.3
Vanda	5.8	2.3	4.6	3.0	6.6	2.4	6.7	1.8	5.5	1.9
Waleria	5.3	3.0	4.8	3.7	6.3	2.8	5.1	2.4	4.5	2.3
William’s Pride	6.0	2.7	5.7	3.0	6.5	2.7	5.5	2.1	5.3	2.0

Table 4. Correlation between cultivar means scored by consumers (appearance, texture, juiciness, taste and overall impression) and visually scoring (covering colour), measurements of sugar content with a refractometer and firmness with a penetrometer in 2002 conducted by the authors.

	Appearance		Texture		Juiciness		Taste		Overall impression		Covering colour, %		Sugar % SS	
	r	P	r	P	r	P	r	P	r	P	r	P	r	P
Texture	-0.16	0.62												
Juiciness	0.29	0.36	-0.27	0.40										
Taste	0.68	0.015	-0.30	0.34	0.48	0.12								
Overall impression	0.74	0.006	-0.27	0.40	0.56	0.06	0.98	<0.001						
Covering colour	0.49	0.10	-0.02	0.94	0.18	0.56	0.63	0.03	0.62	0.03				
Sugar	-0.05	0.87	-0.02	0.94	0.46	0.14	0.26	0.42	0.31	0.33	-0.04	0.90		
Firmness	0.10	0.76	0.86	<0.001	-0.20	0.52	-0.04	0.90	-0.01	0.98	-0.25	0.43	0.00	1.00

SS = soluble solids

years (Tables 4 and 5), indicating that the consumers did not have any distinct preference for high, medium or low scoring of texture. The cultivar means for texture were, however, positively correlated with the penetrometer-derived values for firmness both in 2002 ($r = 0.61$, $P = 0.04$) and in

2003 ($r = 0.64$, $P = 0.01$), showing that physical measurements and consumer scoring yield similar results. In 2003, texture means were negatively correlated with the amount of red colour ($r = -0.56$, $P = 0.02$), possibly because several of the firmest apples had not yet ripened fully.

Table 5. Correlation between cultivar means scored by consumers (appearance, texture, juiciness, taste and overall impression) and visually scoring (covering colour), measurements of sugar content with a refractometer and firmness with a penetrometer in 2003 conducted by the authors.

	Appearance		Texture		Juiciness		Taste		Overall impression		Covering colour, %		Sugar % SS	
	r	P	r	P	r	P	r	P	r	P	r	P	r	P
Texture	-0.12	0.66												
Juiciness	0.01	0.96	0.11	0.69										
Taste	0.21	0.43	-0.13	0.64	0.61	0.01								
Overall impression	0.22	0.41	0.12	0.64	0.81	<0.001	0.90	<0.001						
Cover colour	0.19	0.47	-0.56	0.02	-0.50	0.05	-0.22	0.42	-0.48	0.06				
Sugar	0.02	0.95	-0.36	0.17	0.40	0.12	0.63	0.008	0.54	0.03	-0.11	0.67		
Firmness	-0.003	0.99	0.60	0.02	0.12	0.67	0.16	0.57	0.20	0.47	-0.33	0.21	-0.09	0.74

SS = soluble solids

Juiciness

Cultivar means for juiciness ranged only between 4.5 and 6.1 (1.6 units) in 2002 (Table 2) and between 4.5 and 5.8 (1.3 units) in 2003 (Table 3). ‘Aroma’ had the highest value in both years but there were few significant differences between the cultivars.

When based on cultivar means, juiciness was barely significantly correlated with overall impression in 2002 ($r = 0.56$, $P = 0.06$) (Table 4), but in 2003, the correlation was much stronger ($r = 0.81$, $P = 0.00$) (Table 5). A positive correlation between juiciness and taste ($r = 0.61$, $P < 0.01$) was noted in 2003.

Taste

Cultivar means for taste ranged rather widely, from 4.3 to 7.4 (3.1 units) in 2002 (Table 2) and from 4.6 to 8.7 (4.1 units) in 2003 (Table 3). Significant differences mostly involved the high-rating cultivars. ‘Rubinola’ had the highest mean in 2002, followed by K:1160 and K:1016 while ‘Dayton’ had the lowest mean. Taste scores were very similar to the scores for overall impression ($r = 0.98$, $P = 0.00$); ‘William’s Pride’ and K:1210 moved down one placement compared to overall impression whereas ‘Vanda’ and ‘Dayton’ moved up one

placement. ‘Rubinola’ had the highest taste score also in 2003 but the following placements differed somewhat from those obtained in the preceding year; ‘Reglindis’, ‘Scarlet O’Hara’, K:1160, ‘Fredrik’, ‘William’s Pride’, ‘Redfree’ and ‘Waleria’ moved up 1-3 placements compared to overall impression. Correlation between the mean scores for overall impression and taste was again very high ($r = 0.90$, $P = 0.00$) (Table 5).

The taste means showed a significant, positive correlation with the measured sugar content ($r = 0.63$, $P = 0.01$) in 2003 (Table 5) whereas no association was found in 2002 (Table 4). Taste and amount of red colour were positively correlated in 2002 ($r = 0.63$, $P = 0.03$). The background colour of all the cultivars in the evaluation is greenish yellow or yellow when fully ripe. However, only four cultivars in 2002 and five cultivars in 2003 were rated as yellow (Table 1). Especially ‘Dayton’, ‘Scarlet O’Hara’, ‘Vanda’, K:1016 and K:1210 may have been perceived as not fully ripe by the majority of consumers in 2002 and would probably have gained higher scores if evaluated later in the season. In 2003, K:1016 B, K:1016 ÖS, ‘Vanda’ and ‘Fredrik’ did not appear to have reached optimum eating quality. Some consumers expressed a preference for these apples during the evaluation but none reached the top values. Fully ripe were ‘William’s Pride’, K:1210, ‘Dayton’, ‘Aroma’, ‘Rubinola’, Co-op 12, ‘Redfree’ and ‘Waleria’.

Overall impression

Overall impression of the different cultivars ranged from 4.8 to 7.5 (2.7 units) in 2002 (Table 2) and from 5.3 to 8.4 (3.1 units) in 2003 (Table 3). ‘Rubinola’ obtained the highest mean value in 2002, followed by K:1160 while K:1210 and K:1016 had the lowest values. ‘Rubinola’ obtained the highest value for overall impression also in 2003, followed by ‘Aroma’ while ‘Redfree’ and ‘Waleria’ had the lowest values.

Judging from the strength of the correlations with other scored attributes, taste is the most important factor for overall impression, followed by juiciness and appearance whereas texture had no influence at all. Overall impression was also correlated with sugar content ($r = 0.54$, $P = 0.03$) in 2003 (Table 5), probably due to the strong association between taste and overall impression, and with the amount of red colour ($r = 0.62$, $P = 0.03$) in 2002.

Consumer profile data

A total of 352 and 850 consumers, respectively, participated in the 2002 and 2003 testings. Some values were missing since the consumers did not always fill in the personal information on their score sheets. In 2002, there were 193 females and 148 males, and in 2003 485 females and 350 males. In 2002, 73, 121 and 150 consumers were 10–30, >30–50 and >50 years old, and in 2003, the corresponding values were 201, 278 and 362. More than half of the consumers, i.e. 205 in 2002 and 468 in 2003, had access to home-grown apples. Of the consumers from 2002, 104 ate at least one apple a day, 144 ate 2–5 apples a week, 63 ate one apple a week, 25 ate one apple a month and only 2 consumers ate less than 6 apples in a year. In 2003, the corresponding values were 317, 350, 106, 35 and 8. Thus more than 90% of the participating consumers ate at least one apple a week, and can therefore be regarded as potential future consumers of the tested cultivars.

The only parameter for which the two years differed appreciably, was interest in organically

grown apples. In 2002, 64% of the female consumers and 61% of the male consumers preferred organic apples. Forty-two % of the females and 43% of the males were also prepared to pay a higher price. In 2003, 75% of the female consumers preferred organic apples and 64% of the male consumers. A total of 51% and 46%, respectively, were also prepared to pay the higher price. A significant increase was noted from 2002 to 2003 in both the number of consumers preferring organic apples ($P = 0.00$) and the number prepared to pay a higher price ($P = 0.00$). Preference for organic apples was not related to age (69, 69 and 70% in the three age groups preferred organic apples) nor to access to home-grown apples (69% had home-grown apples and 66% had not).

For each cultivar, possible effects of various consumer profile parameters on the evaluation scores were calculated. Altogether, 28 cultivar-year combinations (12 cultivars in 2002, 16 in 2003) were evaluated in the two years. Consequently, one or two cultivar-year combinations should be expected to show a significant result by chance only (i.e. type 1 error: rejection of a true null hypothesis) when using a probability level of $P < 0.05$. For the variables gender and access to home-grown apples, only one or two cultivar-year combinations appeared to have been significantly affected by consumer parameters. These had P -values between 0.01 and 0.05, therefore not ruling out a type 1 error.

For the variables concerning amount of apple consumption and interest in buying organically produced fruit, 1–4 cultivar-year combinations showed significant results. Tables with the results are available from the authors upon request. The most highly significant results involving the apple consumption ($P = 0.00$) suggested that consumers who ate few apples appreciated the taste of ‘Rubinola’ and the overall impression of ‘Vanda’ more than did consumers who ate more apples. The most highly significant results involving the interest in organic apples ($P = 0.00$) suggest that consumers with a high interest tended to like ‘Redfree’ and to dislike K:1210 compared to consumers with no interest in buying organic apples. However, when these results were compared to the other

data, it was impossible to find any variables that could explain why some cultivars were better liked than others.

Consumer age was the variable that showed the highest impact on the scoring (Table 6). In 2002, older consumers preferred the appearance of 'Aroma', 'Frida' and 'Co-op12', results that could not be explained in terms of differences in shape or colour. For texture the medium-firm 'Aroma' was better liked among the older consumers. For both

taste and overall impression, 'Fredrik' was most liked among the younger consumers whereas K:1160 and 'Co-op12' (the two softest cultivars!) were most liked among the oldest consumers.

Consumer age had some impact on scoring also in 2003. In texture, 'Rubinola' and 'Reglindis' were better liked by the older consumers. According to the penetrometer values, 'Reglindis' and 'Rubinola' were the two firmest cultivars but 'Reglindis' was perceived as one of the softest accord-

Table 6. Comparison of the mean consumer scores for the age groups 10–30, >30–50, >50 years, for appearance, texture, juiciness, taste and overall impression of the cultivars in 2002 and 2003. Only cultivars with significant differences between the different age groups are shown. Number of consumer scores varied between 12 and 150 in 2002 and 32 and 154 in 2003.

Attribute Cultivar	Year	Age			P-value
		10–30	>30–50	>50	
Appearance					
Aroma	2002	7.6		8.3	0.02
Aroma	2002		7.3	8.3	<0.01
Co-op 12	2002	5.9		7.6	<0.01
Frida	2002		7.3	8.1	0.05
Texture					
Aroma	2002		5.6	6.1	0.03
Reglindis	2003	3.7		4.9	0.02
Rubinola	2003	5.4		6.2	0.05
Juiciness					
Reglindis	2003	5.3	6.5		0.01
Reglindis	2003		6.5	5.5	0.02
Waleria	2003		4.1	5.0	0.03
Taste					
Co-op 12	2002	4.3		5.8	0.02
Fredrik	2002	7.0	5.1		0.01
K:1160	2002	5.3		7.3	0.05
Overall impression					
Co-op 12	2002	4.8		6.1	0.01
Co-op 12	2002		5.0	6.1	0.03
Fredrik	2002	7.0	5.5		0.01
Fredrik	2002		5.5	6.7	0.02
K:1160	2002	5.9		7.6	0.03
K:1160	2002		6.4	7.6	0.03
Fredrik	2003	6.0		6.7	0.02
K:1016 B	2003	7.6		6.5	0.03

ing to consumer scores. For juiciness, ‘Reglindis’ obtained higher scores from medium-old consumers than from the other consumer age groups. Mean scores for overall impression for the three consumer age categories can be compared in Table 7. Ten of the 15 test cultivars had the medium-age category score in the middle, indicating the possible existence of a linear relationship between age and some other variable but we were not able to identify any measured or scored variable which in itself could explain why some cultivars were better liked than others.

Number of consumers required for evaluation

The standard deviation for 100 calculations of the mean score for overall impression of ‘Aroma’ was plotted against number of scores analysed (Fig. 1). From this graph, it appears that raising the number of consumers above 50 will not lower the variance

appreciably and may therefore be considered as a sub-optimal use of resources, provided that the sampled consumers are sufficiently homogenous and also representative of the target population.

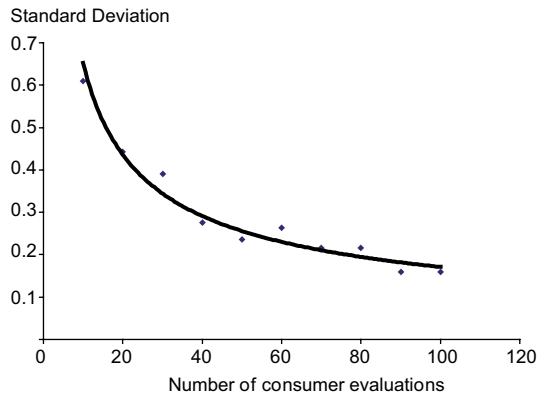


Fig. 1. From the 352 scores for overall impression of ‘Aroma’ in 2002, subsets of 10, 20, 30, 40, 50, 60, 80, 90 and 100 evaluation scores were taken at random. Means of 100 subsets for each size class were calculated, and the standard deviation plotted as a function of the number of consumer scores in the group.

Table 7. Mean evaluation scores provided by consumers belonging to three age groups (10–30, >30–50 and above 50 years old) for their overall impression (OI) of 16 tested apple cultivars in 2003 (n = 32–198 for 10–30, 63–275 for >30–50 and 72–352 for >50 years old).

Cultivar	10–30		30–50		>50	
	OI	stdev	OI	stdev	OI	stdev
Aroma	7.02		7.02		7.02	
Co-op 12	5.60	3.8	6.66	3.1	6.48	3.2
Dayton	7.10	2.9	6.67	2.6	6.53	2.4
Fredrik	5.98	2.9	6.46	2.0	6.74	2.3
Frida	5.62	3.0	5.77	2.1	5.88	1.9
K:1016 B	7.64	2.2	6.85	2.0	6.55	2.2
K:1016 ÖS	6.29	2.2	6.18	2.1	6.89	2.4
K:1160	6.12	3.5	6.61	2.5	7.13	3.1
K:1210	6.04	2.6	6.10	2.8	6.28	2.6
Redfree	4.96	2.6	5.50	2.4	5.87	3.0
Reglindis	6.36	3.4	7.47	2.6	6.96	2.6
Rubinola	8.18	2.7	8.74	2.3	8.35	2.4
Scarlet O’Hara	6.85	2.4	6.98	2.3	6.46	2.5
Vanda	6.07	2.2	5.80	2.4	5.64	2.2
Waleria	4.84	2.9	4.97	2.8	5.87	3.1
William’s Pride	6.55	3.7	5.96	2.4	5.88	2.5

Discussion

Cultivar preference

For overall impression, 'Rubinola' turned out to be the best liked cultivar both in 2002 and 2003, followed by K:1160 and 'Aroma' in 2002, and by 'Aroma' and 'Reglindis' in 2003. In a Danish study of 22 scab-resistant apple cultivars, 'Rubinola' also emerged as the one having the highest eating quality (Kühn and Thybo 2001a).

Associations between consumer-scored attributes

Previous reports have already shown that new apple cultivars must have a desirable appearance to catch the interest of the consumers (von Alvensleben and Meier 1990) but once the apple is bought, taste is more important and probably decides if this particular cultivar will be bought again, provided that the consumer knows and remembers the cultivar name or can recognize the cultivar through its distinctive appearance (Harker et al. 2003).

Taste was the most important factor contributing to overall impression, which agrees with results previously reported by Dailliant-Spinner et al. (1996) and Jaeger et al. (1998). In our trial, both appearance and juiciness turned out to be of lower importance than taste but still very highly correlated with overall impression in one of the years. For appearance, none of the cultivars had a mean score below 5.0 in 2002, or below 6.0 in 2003, which suggests that all of them were acceptable in this respect. By contrast, the rating of texture had no importance for overall impression.

Colour and shape

For a given cultivar, yellow background colour is usually preferred since it is perceived to indicate riper, softer and sweeter fruits (Christensen 1983). All of the apple cultivars in our trial have a more or

less yellow background colour when ripe. Therefore, cultivars having a green background colour at the time of the sensory evaluations may not have been quite ripe. We did, however, not find any association between background colour and any of the scored attributes. In an evaluation performed in Great Britain, no association was found between the colour of the skin (yellow or green background or red streaks) and consumer preference for different sensory properties of apples (Dailliant-Spinnler et al. 1996).

The existence of a red blush or stripes over a large part of the fruit surface is usually perceived as desirable among apple breeders. Although considerable geographic differences were noted in preferences for green or yellow apples, about half of the consumers at each location (Canada, British Columbia & Nova Scotia, and New Zealand) preferred red apples when given a choice between self-coloured red, yellow and green apples (Cliff et al. 2002). We expected that the amount of red colour on the fruit would affect the evaluation of appearance but this was not the case in our trial. A large amount of red colour did however have a positive influence on the scoring of taste as well as overall impression in 2002, possibly because the more reddish apples had reached a level of ripeness that appealed to a majority of the consumers.

Very small effects of fruit shape have been noted for consumer preferences (Table 1), with most shapes being equal and somewhat lower scores being obtained only for oblong, ovate, cylindrical and cordate apples (Hampson and Quamme 2000). In another study, round and conical apples were equally liked whereas oblong apples were less liked (Cliff et al. 2002). Apples in our trial varied only from ovate to globose and conical, which may explain the fact that we did not find any effects of shape on the scoring of the different attributes.

Texture and firmness

Although not synonymous, a crispy texture as well as a hard texture have been shown to be pos-

itively linked, and to contribute (especially crispiness) to the liking of apple cultivars (Hampson et al. 2000, King et al. 2001). Daillant-Spinnler et al. (1996) found an association with texture and taste attributes in a consumer evaluation of apples. Apart from scoring texture in sensory evaluations, attempts have also been made to use instrumental measurements. Penetrometer-derived values were thus correlated with sensory descriptor values for hardness and crispiness (King et al. 2001). Harker et al. (2002) report, however, that fruit lacking significant differences in instrumental firmness may still vary significantly when scored for texture properties by a sensory evaluation panel.

For simplicity, consumers were asked to evaluate texture using a single scale in our trial. Although we found an association between the texture scores and the penetrometer-derived firmness in both 2002 and 2003, there was no significant association between these data and the overall impression or any other scored or measured variable. Possibly, there was a slight tendency for older consumers to like (or at least tolerate) softer cultivars but more data is needed to ascertain this. Analyses of penetrometer-derived measures for firmness indicate that North-American consumers like apples around 60 N whereas apples below 45 N may be regarded as too soft (DeEll et al. 2001). In our trial, cultivar means ranged between 47.3 and 101.0 N in 2002 and between 37.7 and 112.8 in 2003, thus almost all cultivars may have been perceived as acceptable.

Juiciness may be regarded as a texture component but it is also a factor in the perception of taste, probably by acting as a carrier for flavour components. The strong correlation sometimes reported between juiciness and crispiness may be explained by the strength of inter-cellular bonds in the middle lamella of adjacent cells; fracturing of strong bonds is perceived as crispiness and simultaneously causes the release of juice (King et al. 2001). In our trial, juiciness was more closely associated with taste than with texture, possibly because texture might have been considered as hardness and not only (or mainly) crispiness.

Taste

Apple 'flavour' is to a major extent determined by volatile molecules present in the fruit flesh (Fellman et al. 2000) and by the sugar/acid balance. In consumer evaluation tests, properties like sweetness, sourness and aromatics are sometimes scored individually. All of these are associated with flavour liking in apples and appear to be more useful in predicting consumer preferences than e.g. % SS and titratable acid (Hampson et al. 2000). Kühn and Thybo (2001b) however report that high % SS and consumer perception of sweetness was positively related to taste preferences for children between 9 and 13 years old.

Our evaluation form included only one simple taste scoring in order to keep the testing simple and quick. This taste score proved to be the most important factor in determining overall impression. Interestingly, three of the four cultivars that had the highest scores for overall impression in 2002 also had the highest sugar content. The only high-sugar cultivar that did not obtain a high score for overall impression was 'Vanda' (on the 8th place) and the only low-sugar cultivar that received a high overall impression score was 'Aroma', which is well-known and appreciated for its very aromatic flavour. A strong aroma appears to be at least as desirable as sweetness, or even more so. The four cultivars in 2003 with highest overall impression scores included the two sweetest cultivars and two above-average sweet cultivars. Cultivars that differed in sugar content between the two years always placed better for taste and overall impression in the year of the highest sugar content. This may be partly due to a simultaneous development of both sugar and volatile compounds as the ripening process proceeds.

Consumer profile data

Hampson et al. (2000) and Cliff et al. (1998) did not find any gender differences in their consumer evaluation studies, whereas women have been found to prefer smaller fruits in other studies (Redalen 1988, Hampson and Quamme 2000). We did

not investigate fruit size, and the few significant gender-related differences for other variables in our study may have been due mainly to type 1 error.

According to our data, levels of apple consumption did not influence the consumer's opinion of the evaluated apples in any particular direction, and neither did access to home-grown apples.

Number of consumers who prefer to buy organic apples and are prepared to pay a higher price rose significantly from 2002 to 2003, and showed a strong gender difference; women were more interested in organic apples than men. In a household survey in the US, number of consumers willing to buy eco-labelled apples decreased as the price premium increased, but 40% would still buy eco-labelled apples with a 0.5 € higher price (Blend and Ravenswaay 1999). The amount of consumers willing to pay the price premium was even higher in our evaluation, 49% in 2002 and 58% in 2003. Consumers interested in buying organic apples did, however, not differ from other consumers in the way they scored the tested apple cultivars.

The only consumer characteristic that appears to have had a major impact on consumer evaluation scores is age. Still, it is difficult to find a single variable that explains why some cultivars are more liked by younger consumers and others by older. A certain tendency for older consumers to like (or at least tolerate) more soft-fleshed cultivars was discerned but more data is needed. The only age-related differences we have found in the literature concerns fruit size; older consumers have been reported to prefer smaller fruits (Hampson and Quamme 2000) whereas shape and colour showed no association with age-related preferences (Cliff et al. 2002).

Number of consumers required for evaluation

Trained taste panels are often regarded as important and useful tools for evaluating new cultivars and promising selections. Using a trained panel is, however, generally more expensive compared to a

volunteer consumer evaluation. As a complement, mass-testing by volunteers may therefore provide valuable data to a lower cost. Sometimes there is also need for a survey of a larger number of consumers, especially if there are any indications of heterogeneity in the future target population.

The exact number of panellists or consumers needed of course depends on the type of study. Already 11 panellists were considered sufficient for discrimination at one point on a 0-9 scale (Hampson et al. 2000). Cliff et al. (1998) report that results from 50 consumers do not provide a reliable estimate of consumer preference. We however found that raising the number of consumers above 50 reduces the variance only to a minor extent, probably due to our consumers being relatively homogenous.

Acknowledgements. Gustav Redalen provided valuable advice on the design and evaluation of the study. Paul Ilg helped with harvesting and storing of the fruit. We would also like to thank Ingemar Jönsson, Larisa Gustavsson, Barbro Johnsson, Helena Mattisson, Helena Persson-Hovmalm, Larry Tvingmark, and Gun Werlemark for all help during the practical evaluation arrangements at Äpplets hus.

References

- Blend, J. & Ravenswaay, E. 1999. Measuring consumer demand for ecolabeled apples. *American Journal of Agricultural Economics* 81: 1072–1077.
- Brown, S.K. & Maloney, K.E. 2003. Genetic improvement in apple: breeding, markers, mapping and biotechnology. In: Ferree, D.C. & Warrington, I.J. (eds.). *Apples – botany, production and uses*. CABI Publishing. 672 p.
- Christensen, C. 1983. Effects of color on aroma, flavour and texture judgments of foods. *Journal of Food Science* 48: 787–790.
- Cliff, M., King, M. & Hampson, C. 1998. Comparison of mean scores and R-indices for consumer preferences of apple cultivars. *HortScience* 33: 1239–1240.
- Cliff, M., Sanford, K., Wismer, W. & Hampson, C. 2002. Use of digital images for evaluation of factors responsible for visual preference of apples by consumers. *HortScience* 37: 1127–1131.
- DeEll, J., Khanizadeh, S., Saad, F. & Ferree, D. 2001. Factors affecting apple fruit firmness – a review. *Journal of the American Pomological Society* 55: 8–27.
- Dailant-Spinner, B., MacFie, H.J.H., Beyts, P.K. & Hedderley, D. 1996. Relationships between perceived sensory

- properties and major preference directions of 12 varieties of apples from the southern hemisphere. *Food Quality and Preference* 7: 113–126.
- Fellman, J., Miller, T., Mattinson, D. & Mattheis, J. 2000. Factors that influence biosynthesis of volatile flavor compounds in apple fruits. *HortScience* 35: 1026–1032.
- Hampson, C. & Quamme, H. 2000. Use of preference testing to identify tolerance limits for fruit visual attributes in apple breeding. *HortScience* 35: 921–924.
- Hampson, C., Quamme, H., Hall, J., MacDonald, R., King, M. & Cliff, M. 2000. Sensory evaluation as a selection tool in apple breeding. *Euphytica* 111: 79–90.
- Harker, F., Gunson, F. & Jaeger, S. 2003. The case for fruit quality: an interpretive review of consumer attitudes, and preferences for apples. *Postharvest Biology and Technology* 28: 333–347.
- Harker, F., Maindonald, J., Murray, S., Gunson, F., Hallett, I. & Walker, S. 2002. Sensory interpretation of instrumental measurements 1: texture of apple fruit. *Postharvest Biology and Technology* 24: 225–239.
- Jaeger, S., Andani, Z., Wakeling, I. & MacFie, H. 1998. Consumer preferences for fresh and aged apples: a cross-cultural comparison. *Food Quality and Preference* 9: 355–366.
- Janse, J. 1993. Breeding of resistant apple varieties in the Netherlands. *Acta Horticulturae* 347: 143–148.
- Kellerhals, M., Mouron, P., Goerre, M. & Gessler, C. 1997. Breeding apples for disease resistance – an integrated approach towards sustainable and economic orchard ecosystems. *Proceedings of the 4th workshop on integrated control of pome fruit diseases*, Croydon. p. 134–140.
- King, G., Lynn, J., Dover, C., Evans, K. & Seymour, G. 2001. Resolution of quantitative trait loci for mechanical measures accounting for genetic variation in fruit texture of apple (*Malus pumila* Mill.). *Theoretical and Applied Genetics* 102: 1227–1235.
- Kühn, B. & Thybo, A. 2001a. Sensory quality of scab-resistant apple cultivars. *Postharvest Biology and Technology* 23: 41–50.
- Kühn, B. & Thybo, A. 2001b. The influence of sensory and physiochemical quality on Danish children's preference for apples. *Food Quality and Preference* 12: 543–550.
- Norusis, M. 1990. *SPSS® Base systems user's guide*. SPSS inc, Chicago. 507 p.
- Redalen, G. 1988. Quality assessment of apple cultivars and selections. *Acta Horticulturae* 224: 441–447.
- Richards, T. 2000. A discrete/continuous model of fruit promotion, advertising, and response segmentation. *Agribusiness* 16: 179–196.
- von Alvensleben, R. & Meier, T. 1990. The influence of origin and variety on consumer perception. *Acta Horticulturae* 259: 151–161.
- Watada, A., Abbott, J. & Hardenburg, R. 1980. Sensory characteristics of apple fruit. *Journal of the American Society for Horticultural Science* 105: 371–375.
- Weibel, F., Schmid, A. & Häseli, A. 2003. Efficient multi-location testing of scab resistant cultivars for organic apple production in Switzerland. *Acta Horticulturae* 622: 335–342.