Animal Science Journal (2011) **82,** 325–333 doi: 10.1111/j.1740-0929.2010.00825.x

ORIGINAL ARTICLE

Japanese consumer categorisation of beef into quality grades, based on Meat Standards Australia methodology

Rod J. POLKINGHORNE,[1] Takanori NISHIMURA,[3] Kate E. NEATH[4] and Ray WATSON[2]

[1]*Marrinya, Wuk Wuk, and* [2]*Department of Mathematics and Statistics, University of Melbourne, Vic. Australia; and* [3]*Meat Science laboratory, Graduate School of Agriculture, Hokkaido University, Kita, Sapporo,* [4]*Meat and Livestock Australia, Minato, Tokyo, Japan*

ABSTRACT

The objective was to evaluate the sensory categorisation of beef by Japanese consumers, based on Meat Standards Australia methodology. Various cuts of beef, with a wide range of quality (from Australian and Japanese cattle) and three cooking methods (grill, yakiniku, shabu shabu), were evaluated by 1620 Japanese consumers in Tokyo and Osaka. Consumers rated each sample for four sensory attributes (tenderness, juiciness, flavor and overall satisfaction), then selected one of four grades (unsatisfactory/2-star, good everyday/3-star, better than everyday/4-star, and premium quality/5-star), based on the quality of the beef within each cooking method. Meat quality scores, denoted as MQ4 scores (weighted combination of the four sensory attributes) were calculated from the Japanese consumer test results, to describe the Japanese consumer rating of beef. The distribution of the Japanese consumer MQ4 scores showed a clear distinction between grades, with the majority of scores being included within the boundaries of each grade. The MQ4 score allocated approximately 64% of the samples to their correct consumer grades. The MQ4 score showed potential to be used as a tool in developing and monitoring a consumer-focused grading system that is able to predict Japanese consumer satisfaction of individual beef cuts prepared by different cooking methods.

Key words: *beef, eating quality, Japanese consumer, sensory.*

INTRODUCTION

Sensory attributes including tenderness, juiciness and flavor, each contribute directly and interactively to consumers' eating experience. Many reports have indicated that tenderness is a key component of consumer satisfaction (Boleman *et al*. 1997; Miller *et al*. 2001), whereas others have highlighted the importance of flavor (Huffman *et al*. 1996). In Japan, it has been traditionally thought that the most important sensory attribute that determines meat quality for the Japanese consumer is tenderness due to marbling. Marbling was found to be important for the judging of beef quality by Japanese consumers (Koizumi *et al*. 1986). In addition the sweet and fatty 'Wagyu beef aroma' that is produced in highly marbled Wagyu beef was suggested as highly desirable for Japanese consumers (Okitani 1999; Matsuishi *et al*. 2004). A recent questionnaire-based study by Sasaki and Mitsumoto (2004) found that Japanese consumers are increasingly diverse in their preferences for beef quality, and that safety and price also play an important role in beef selection. However, there are few reports in terms of the sensory response of Japanese consumers to beef, or in regards to more accurately predicting the Japanese consumer preference for beef products. Understanding the relative importance of these sensory attributes for consumer perception is valuable as it can enable establishment of benchmarks for product evaluation. It also has the potential to enable the development of a model to predict the eating quality of individual beef portions for consumer satisfaction.

In Australia, a meat grading system that assures eating quality for the consumer is currently being used, called Meat Standards Australia (MSA). An outline of the evolution of the MSA grading system is described in Polkinghorne *et al*. (2008a). The MSA grading system is

Correspondence: Takanori Nishimura, Meat Science laboratory, Graduate School of Agriculture, Hokkaido University, Kita, Sapporo 060-8589, Japan. (Email: nishi@anim.agr.hokudai.ac.jp)
Received 7 October 2009; accepted for publication 22 April 2010.

unique in that it assigns a grade for each beef cut according to various cooking methods. This is in contrast to the carcass-based grading systems of Japan (JMGA 1988), the United States Department of Agriculture (USDA 1989), Canada (Canadian Beef Grading Agency 1997) and Korea (Kim & Lee 2003). Polkinghorne (2005) demonstrated that a single carcass grade was not capable of accurately predicting eating quality when carcasses were produced from a variety of production systems, as different cuts from any carcass had different eating qualities and the eating quality relationship between cuts varied from carcass to carcass.

The MSA prediction model, statistical analysis and development of protocols, are described in detail in previous publications (Polkinghorne *et al*. 1999; Watson *et al*. 2008a,b). Briefly, extensive consumer test results were utilized to develop a consumer-based composite meat quality score (MQ4). This MQ4 score was derived from weighting and combining individual scores of four sensory traits: tenderness (tn), juiciness (ju), flavor (fl) and overall satisfaction (ov) (MQ4 score = 0.4 tn + 0.1 ju + 0.2 fl + 0.3 ov (Watson *et al*. 2008a)). The MQ4 score derived from consumer data is used to describe the eating quality response of the entire population of consumers.

The MSA prediction model was developed to estimate this MQ4 score, based on animal attributes and carcass characteristics, for a wide range of muscles of various eating qualities when cooked by different methods (Polkinghorne *et al*. 2008a; Watson *et al*. 2008b). The predicted MQ4 score was then used to assign beef to one of four quality grades: 'unsatisfactory' (2 star; no grade), 'good everyday' (3 star), 'better than everyday' (4 star), or 'premium quality' (5 star). The MSA grading system has been utilized in Australia since 1996, and has proved to be a very successful eating quality assurance system.

The current usage and future development of the MSA grading system is discussed in Polkinghorne *et al*. (2008b). Following the success of the MSA grading system in predicting eating quality for Australian consumers, one of the next steps was to test the accuracy of the model in satisfying other consumer groups on an international scale. A comparison of the consumer preference between Australian and Korean consumers was carried out, based on the MSA protocol, and suggested that although Korean consumers were slightly more discriminating, the MSA model can accurately predict eating quality for both Australian and Korean consumers (Thompson *et al*. 2008).

Using MSA methodology, the objective of this study was to evaluate the sensory categorization of beef by the Japanese consumer, across a wide meat quality range and three cooking methods. The results showed that the Japanese consumer categorized beef into four grades according to four sensory attributes (tenderness, juiciness, flavor and overall satisfaction), similar to Australian and Korean consumers. Furthermore, it was possible to combine these four sensory attributes with weightings from the MSA prediction model, to create a single composite meat quality score (MQ4 score). These results suggest the potential to utilize the MQ4 score to develop a consumer-based grading system in Japan, to predict beef eating quality for the Japanese consumer.

MATERIALS AND METHODS

Consumer recruitment

A research company recruited consumers for the study by door-to-door interview following a 2-stage stratified random sampling technique within Tokyo and Osaka. Stage 1 consisted of selecting 1620 respondents who met the relevant screening conditions. Consumers were screened to provide an age range from 20 to 69 years, and to include only those who ate beef at least twice per month at home. The aim of this screening was to recruit consumers who regularly ate beef across a demographic mix representative of the Japanese census. The total of 1620 respondents was divided between Tokyo and Osaka by a ratio of approximately 61% (990) and 39% (630), respectively, to give a reasonably sized base of respondents in each location.

In stage 2, 162 consumer sample locations were randomly selected (99 in Tokyo and 63 in Osaka), with 10 interviews completed at randomly selected dwellings within each sample location.

Consumer descriptive statistics

Consumer descriptive statistics are shown in Tables 1–4. Table 1 shows the consumer age and sex by city.

Table 1 Consumer demographic statistics – age and sex by city

Age (years)	Tokyo				Osaka			
	Male	Female	All	%	Male	Female	All	%
20–25	35	63	98	10%	22	39	61	10%
26–30	33	59	92	9%	19	43	62	10%
31–39	79	154	233	24%	47	92	139	22%
40–50	73	134	207	21%	38	79	117	19%
51–59	54	124	178	18%	42	81	123	19%
60–69	58	124	182	18%	42	86	128	20%
All	332	658	990		210	420	630	
%	34%	66%			33%	67%		

Table 2 Consumer demographic statistics – occupation and sex by city

Occupation	Tokyo				Osaka			
	Male	Female	All	%	Male	Female	All	%
Trade	27	5	32	3%	26	4	30	5%
Professional	13	35	48	5%	7	11	18	3%
Administration	78	106	184	19%	36	63	99	16%
Teacher	43	21	64	7%	30	12	42	7%
Salesperson	90	100	190	19%	54	65	119	19%
Laborer	31	52	83	8%	22	38	60	9%
Homemaker	0	288	288	29%	0	193	193	31%
Unemployed	26	26	52	5%	27	26	53	8%
Student	24	25	49	5%	8	8	16	2%
All	332	658	990		210	420	630	

Table 3 Consumer demographic statistics – frequency of beef consumption (%)

Frequency of eating beef	Tokyo	Osaka	All
Daily	0.1	1.1	0.5
4–5 times/week	1.3	8.6	4.1
2–3 times/week	31.0	46.5	37.0
Once/week	41.9	35.1	39.3
Once each two weeks	25.7	8.7	19.1

Table 4 Consumer demographic statistics – income categories

Income range (million yen)	Tokyo		Osaka	
	No.	%	No.	%
Less than 2.9	63	6	103	16
3.0–3.9	100	10	74	12
4.0–4.9	122	12	88	14
5.0–5.9	136	14	83	13
6.0–6.9	101	10	78	12
7.0–7.9	110	11	54	9
8.0–8.9	82	8	27	4
9.0–9.9	89	9	40	6
10.0–11.9	95	10	37	6
12 and above	63	6	29	5
Missing data	29	3	17	3
All	990		630	

Approximately two-thirds of consumers participating in both Tokyo and Osaka were female. The age distribution of the Tokyo and Osaka sample was similar and distributed across 5-year age bands from 20 to 70 years. Table 2 shows the occupation and sex by city. The occupations of the consumers were diverse, including trades, administrative, professional, technical, sales, laborers, home makers, students and unemployed. Occupational distribution differed by sex, with no males classified as home makers, but being more heavily represented in trade and technical categories. According to the frequency of beef consumption as shown in Table 3, 56% of the Osaka consumers ate beef at least 2–3 times per week, in comparison to 33% of those from Tokyo. Income levels, presented in Table 4, ranged from less than 3 million yen per annum to greater than 12 million yen, with Tokyo consumers tending to have higher incomes.

Sensory testing and sample allocation

The sensory testing was conducted at central venues in Tokyo and Osaka. Consumers attended sessions in groups of 20, and the date and time of each session was arranged in advance. Beef was cooked by grill, yakiniku or shabu shabu methods at each session with a total of 540 consumers allocated to each cooking method.

The MSA sensory testing protocols are explained in detail in the 'accessory publication' of Watson *et al*. (2008a). Briefly, each consumer was served seven samples. The first was a common mid-quality 'starter', with the following six being different quality products presented in balanced order via a latin square design. This ensured all samples were presented an equal number of times in each presentational order, and an equal number of times before and after each other product. The six samples served to every consumer were from a wide range of meat quality and marbling scores. MSA protocols, described in the accessory publication of Watson *et al*. (2008a) were utilized for sample preparation, cooking and serving. Analyses reported by Hwang *et al*. (2008) found that these procedures effectively balanced out effects related to sample order and carry-over.

Meat sample preparation

Four diverse muscles (*m. longissimus dorsi lumborum*, *m. serratis ventralis cervicis*, *m. semispinalis capitus*, and *m. biceps femoris*) were collected from 36 Japanese and 87 Australian carcasses comprising a wide range of breeds and feeding regimes, as summarized in Table 5. This was crucial in order to collect samples with a wide range of eating quality. Over the 123 carcasses, the ultimate pH ranged 5.40–5.76, the rib fat thickness ranged 3–44 mm, and ossification score ranged 140–400.

The four muscles were collected from carcasses within three MSA marbling bands (mb): MSA.mb < 400 (similar to Japan Meat Grading Association Beef Marbling Score, BMS ≤ 2), MSA.mb 400 to <800, and MSA.mb >800 (similar to BMS > 4). Samples for all cooking methods were prepared from each muscle, except for *m. semispinalis capitus*. As *m. semispinalis capitus* is a relatively smaller muscle, at times it was not large enough to prepare all three cooking methods. Therefore, when required, *m. serratis ventralis cervicis* was used instead of *m. semispinalis capitus* in Australian-sourced yakiniku samples. The sample position for each cooking method was rotated.

The grill samples comprised five individual 25 mm thick steaks, whereas the initial yakiniku sample was from an 80 × 60 × 25 mm block with muscle fiber aligned along the

Table 5 Cattle that were sourced for consumer testing to provide wide range of quality and marbling scores for consumer testing

Country of origin	Feeding background	Breed type	No. of head
Australia	Long-fed (>150 days grain-fed)	Wagyu	3
		F1 (Wagyu × Angus)	20
		F1 (Wagyu × Santa Gertrudis)	6
		Angus	18
	Short-fed (<150 days grain-fed)	British Breeds and crosses	11
		Brahman	1
	Grass-fed	British Breeds and crosses	22
		Brahman	6
Japan	Japanese housed system	Japanese Black Cattle	17
		F1 (Japanese Black Cattle × Holstein)	9
		Holstein	10
Total			123

80 mm axis (4 mm strips cut perpendicular to the fiber when prepared from the block). The shabu shabu was an initial 100 × 50 × 20 mm block with muscle fiber aligned with the 20 mm axis (1.5 mm slices cut perpendicular to the fiber when prepared from the block).

Sample allocation was controlled by MSA software, which is explained in detail in the accessory publication of Watson *et al*. (2008a). Allocation to the nominated consumer and serving round was achieved by placing the individual steak or strip on an acetate sheet pre-printed with the sample identification codes in set positions. One acetate sheet was produced for each of the seven serving rounds and vacuum packed to maintain sample position until cooking. Samples were transferred to the cooking apparatus in a set sequence from the sheets.

Each muscle from Japan and Australia was prepared by each cooking method, and was served to 10 consumers. The 10 samples were prepared from the original fabricated blocks for each cooking method by: halving the five steaks after cooking for grilled steak, preparing ten 50 × 75 × 4 mm slices for yakiniku, and ten 100 × 50 × 1.5 mm slices for shabu shabu. The 10 steak halves, or yakiniku and shabu shabu slices, prepared from each sample were served in five different presentational order positions to five consumer pairs. Each consumer pair was selected from five different groups of 12 consumers. This ensured that individual muscles were evaluated across multiple test sessions to minimize potential consumer, session or presentational order bias.

Samples were vacuum packed and stored at −18°C until sensory testing, as outlined in the accessory publication of Watson *et al*. (2008a).

Cooking and serving

Thawing procedures for each cooking method are described in the accessory publication of Watson *et al*. (2008a). The cooked samples were presented to consumers without accompaniments, seasoning or flavoring. Samples were identified by unique four-digit alpha-numeric codes attached to both plates and the questionnaire forms. All samples were prepared and cooked by trained staff.

Steaks were cooked to medium doneness using a double-sided clamshell grill set (Silex S165, Hamburg, Germany) at 220–230°C for 5 min. The steaks were cooked in accordance with MSA protocols as described in the accessory publication of Watson *et al*. (2008a). The grill was switched on 45 min prior to cooking and a set of sacrificed starter steaks used to commence the cooking cycle and stabilize temperature recovery. All cooking operations were conducted with reference to a timing schedule to ensure consistent cooking time for all samples and to control the serving sequence. The grill temperature was stabilized by cooking a set of left-over steaks immediately prior to each test session. Steaks were placed on the grill in the same order as on the acetate sheet to maintain sample identification. After cooking, steaks were transferred to a cutting board in the same order. Steaks were held for 2 min before halving and placing on pre-numbered serving plates. A cross-check was conducted by an independent observer confirming the pre-printed sample identification codes on the plates matched the round sheet identification. A further check was conducted by confirming a pre-printed label identification on each consumer score sheet against the plate sticker at the point of serving.

Yakiniku cooking was carried out with a portable electric hotplate (Ceracoal EC 201, Kyonggi-Do, Korea). Plate temperature was maintained between 250°C and 260°C by adjusting the electrical voltage with a variable voltage auto transformer. The hotplate was warmed up for 15 min at 100 V, and then set at 69.5 V for cooking. Single samples were placed on the hot plate and turned as moisture pooled on the surface. The sample was served to the nominated consumer when the second side pooled (approx. 1 min total cooking time). This visual indicator in combination with temperature control produced a uniform medium degree of doneness in the cooked strip.

For shabu shabu, samples were cooked in boiling water that was poured from a standard electric urn into a 500 mm plastic container. A new container and fresh boiling water was used for each sample to prevent carry-over of flavor. No vegetables, salt or flavorings were added. Samples were gently moved around in the boiling water until the meat color changed from red to light grey. Samples were removed when the meat color turned light grey (approx 30 sec total cooking time).

Samples served in each of the seven presentational rounds were cooked immediately prior to serving that round. Consumers did not view the beef prior to or during cooking, with steaks prepared in a kitchen adjacent to the test area and yakiniku and shabu shabu cooking being done behind a screen in the serving area.

Questionnaire and score sheet

The questionnaire and score sheet are shown in the accessory publication of Watson *et al*. (2008a). The questionnaire comprised a series of demographic and attitudinal questions followed by a single score sheet for each of the seven samples

served. Each score sheet contained four 100 mm line scales and four quality category boxes. The line scales were used to record evaluations for tenderness, juiciness, flavor and overall satisfaction, each anchored by a statement at either end ('not tender' and 'extremely tender' for tenderness, 'not juicy' and 'extremely juicy' for juiciness, and 'dislike extremely' and 'like extremely' for both the flavor and overall satisfaction scales). After eating a beef sample, consumers recorded their scores by placing a mark across each line at the point which best reflected their judgement of that sample. The distance in mm from the left-hand end of the scale to the mark was measured to obtain a score out of 100 for each scale.

The consumer then marked one of four quality category boxes to indicate which description best suited the quality of the sample. The four alternative boxes were labeled unsatisfactory, good everyday quality, better than everyday quality, and premium quality. These category descriptions are also referred to as grades (2, 3, 4 and 5 star) later in the paper. All data was double-keyed and cross-checked to prepare data files for analysis. Software was used for cross-checking the data and confirming that the samples and presentational order for each consumer matched the design criteria.

Statistical analysis

The MQ4 score (meat quality score) was calculated by linear discriminant analyses as described in Watson *et al*. (2008a). The MQ4 score ranges from 0 to 100, and is based on a weighted combination of the consumer scores of the four sensory attributes (tenderness, juiciness, flavor and overall satisfaction). The MQ4 score can be used to describe the consumers' rating of beef.

Weightings are used to describe the coefficients assigned to the four sensory attributes of tenderness, juiciness, flavor and overall satisfaction. The methodology for calculating the weightings is explained in detail in Watson *et al*. (2008a).

The model for predicting the MQ4 score in the MSA grading system has the following fixed weightings (all cooking methods):

MQ4 score = 0.4 tn + 0.1 ju + 0.2 fl + 0.3 ov (tn: tenderness, ju: juiciness, fl: flavor, ov: overall satisfaction; Watson *et al*. 2008a).

The weightings give an indication of the relative importance of the four sensory attributes (tenderness, juiciness, flavor, overall satisfaction) to the final meat quality score. According to the fixed weightings above, when a consumer eats beef their satisfaction is influenced mostly by the tenderness of the beef, followed by overall liking, flavor and then juiciness.

Calculation of MQ4 score using fixed weightings from MSA model

The four consumer sensory scores and quality ratings were analyzed by a linear discriminant function (PROC DISCRIM, SAS 1997), as described in Watson *et al*. (2008a). Briefly, the linear discriminant function was used to assess the accuracy of using the four sensory scores to predict the consumer grades (unsatisfactory, good everyday, better than everyday and premium). The four sensory scores were then combined into a single MQ4 score. To do this, the four sensory scores were multiplied by the fixed weightings from the MSA prediction model (0.4 tn + 0.1 ju + 0.2 fl + 0.3 ov). The reason for using the fixed weightings was to provide direct links with the MSA database. This MQ4 score was used to describe the rating of beef by the Japanese population.

Boundaries between the grades were then estimated, which were the optimal points at which the Japanese consumer distinguished between the grades of 2/3 star, 3/4 star and 4/5 star. Details regarding the calculation of boundaries are explained in Watson *et al*. (2008a). The boundaries were estimated for the fixed linear discriminate function by assuming that the boundaries occurred where the adjacent discriminant functions were equal. The boundaries were the MQ4 scores which would ensure the highest rate of assigning the correct grade (and the lowest rate of assigning the wrong grade) for the Japanese consumer.

Accuracy of assigned grade relative to true category

The accuracy of the MQ4 score in assigning the grade (2, 3, 4, 5 star) was determined, relative to the true category selected by each individual consumer on the consumer score sheet.

Optimal weightings calculated from current data set

In order to investigate the importance of each of the four sensory attributes to the final meat quality score for the Japanese consumer in the current data set, weightings were calculated. Weightings for the four sensory attributes were first calculated for each grade (2, 3, 4 and 5 star) within each cooking method (grill, yakiniku, shabu shabu) and then averaged over the four grades within each cooking method.

RESULTS AND DISCUSSION

Japanese consumer MQ4 score and boundaries between grades

Fig. 1a–d shows the distribution of the Japanese consumer MQ4 scores for each cooking method (grill, yakiniku, shabu shabu and all cooking methods). Within each cooking method, the MQ4 scores are displayed according to the category selected by the individual consumer on the consumer score sheet (unsatisfactory, good everyday, better than everyday, and premium). The boundaries are the points of intersection between the distribution curves, which were the optimal points at which the Japanese consumer distinguished between the grades of 2/3 star, 3/4 star and 4/5 star. The optimal boundaries between 2/3 star, 3/4 star and 4/5 star grades were, respectively: 40.4, 66.8 and 83.1 for grill; 43.4, 68.5 and 83.9 for yakiniku; 43.7, 67.4 and 83.4 for shabu shabu; and 42.5, 67.6 and 83.5 for all cooking methods. Therefore, for example, if the MQ4 score of a grilled sample was 72, it would be assigned to the 4 star grade, as it is above the 3/4 star boundary of 66.8 and below the 4/5 star boundary of 83.1.

A solid pattern was evident across the distribution of MQ4 scores with a relatively clear distinction between grades, and the majority of scores were included within the boundaries of each grade. Therefore, for example, if a consumer's MQ4 score for grilled sample was 72, and they selected the 'Better Than Everyday' (4 star) category on the consumer score sheet, this consumer would be satisfied with the 4 star predicted grade. However, if the same consumer had selected 'Good

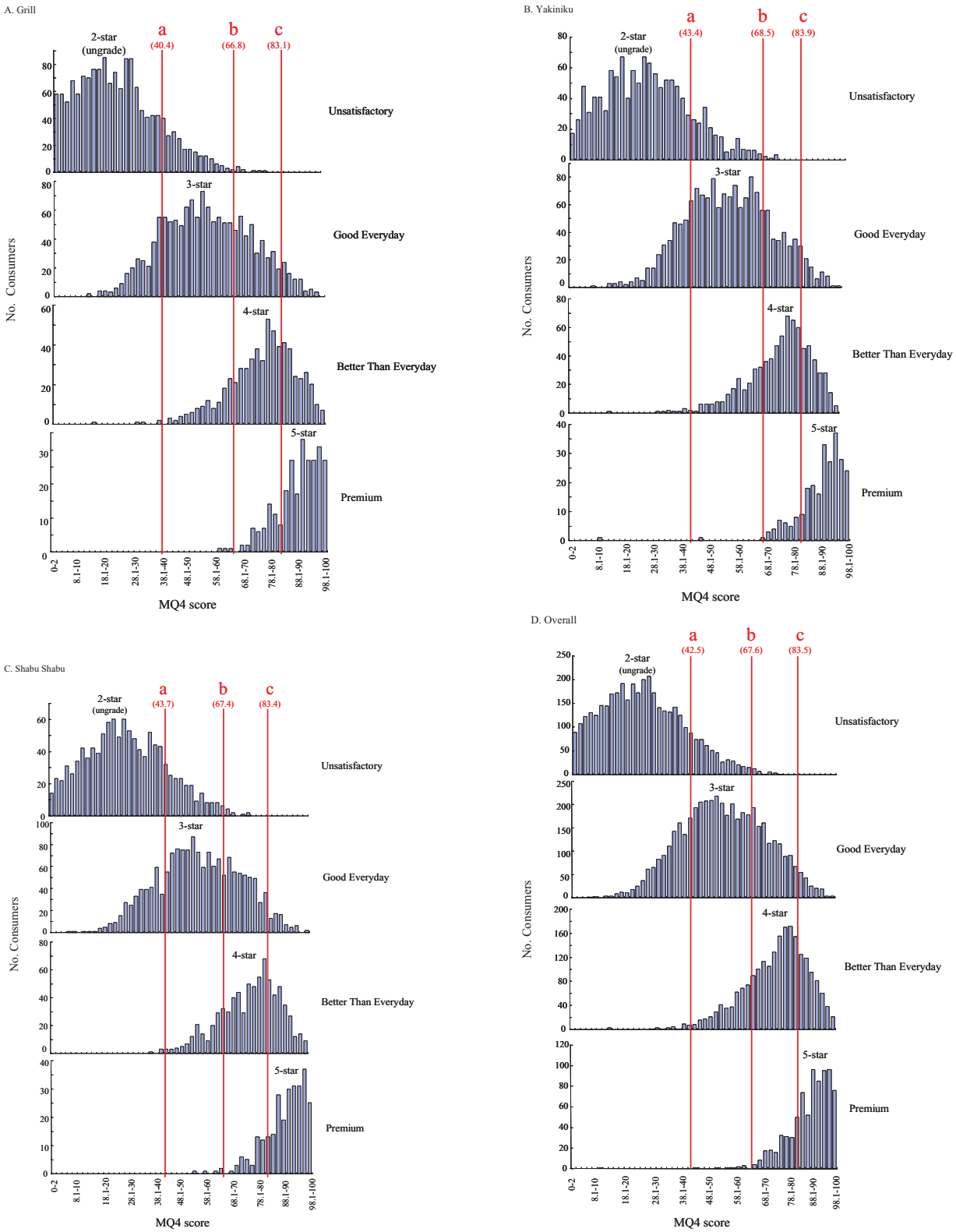

Figure 1 Japanese consumer meat quality (MQ4) scores for each cooking method: (A) grill steak, (B) yakiniku, (C) shabu shabu, and (D) all cooking methods, according to the quality category selected by each individual consumer on the consumer score sheet: unsatisfactory, good everyday, better than everyday, and premium. Boundaries that separate consumer MQ4 scores into grades are shown by vertical lines: (a) boundary between 2 star and 3 star, (b) 3/4 star, and (c) 4/5 star. *Y*-axis (no. of consumers) indicates number of serving rounds (total 540 consumers × 7 samples).

Everyday' (3 star) category, they would not be satisfied with the 4 star rating, as the eating experience would be lower than what they had expected. This accuracy of the assigned grade, relative to the true category as selected by the consumer, is discussed below.

Accuracy of assigned grade relative to true category

Table 6 explains the accuracy of the MQ4 score in assigning the grade, relative to the true category selected by each individual consumer on the consumer score sheet, for each meat sample and cooking method.

The accuracy of assigning the grades was 67.0% for grill, 63.7% for yakiniku, 60.2% for shabu shabu and 63.6% for all cooking methods. This indicates that the MQ4 score allocated approximately 64% of the samples to their correct consumer grades, which is similar to the average accuracy gained for Australian and Korean consumers (64% and 61%, respectively) by Thompson *et al*. (2008) and to the accuracy gained in another study on Australian consumers (68.4%) by Watson *et al*. (2008a). The latter paper suggested that it is not possible to achieve a perfect categorization, and that even though there is overlapping between grades,

Table 6 Accuracy of assigned grade compared to the true grade

A. Grill

Assigned grade (by MQ4)	True grade (by consumer selection)				All grades
	X (2)	3	4	5	
X (2)	**1293**	252	5	0	1550
3	202	**733**	115	3	1053
4	11	307	**290**	51	659
5	0	90	212	**213**	515
Total samples	1506	1382	622	267	3777
Total correct	1293	733	290	213	2529
% correct	85.9%	53.0%	46.6%	79.8%	67.0%

B. Yakiniku

Assigned grade (by MQ4)	True grade (by consumer selection)				All grades
	X (2)	3	4	5	
X (2)	**1012**	285	10	1	1308
3	187	**834**	133	1	1155
4	9	337	**357**	41	744
5	0	95	273	**204**	572
Total Samples	1208	1551	773	247	3779
Total Correct	1012	834	357	204	2407
% Correct	83.8%	53.8%	46.2%	82.6%	63.7%

C. Shabu shabu

Assigned grade (by MQ4)	True grade (by consumer selection)				All grades
	X (2)	3	4	5	
X (2)	**892**	332	1	0	1225
3	205	**808**	143	5	1161
4	11	398	**354**	50	813
5	0	87	269	**221**	577
Total samples	1108	1625	767	276	3776
Total correct	892	808	354	221	2275
% correct	80.5%	49.7%	46.2%	80.1%	60.2%

D. All samples

Assigned grade (by MQ4)	True Grade (by consumer selection)				all grades
	X (2)	3	4	5	
X (2)	**3210**	869	14	1	4094
3	581	**2368**	400	9	3358
4	31	1044	**990**	141	2206
5	0	277	758	**639**	1674
Total samples	3822	4558	2162	790	11332
Total correct	3210	2368	990	639	7207
% correct	84.0%	52.0%	45.8%	80.9%	63.6%

Accuracy of assigned grade (according to the MQ4 boundaries, as shown in Fig. 1) compared to the true grade that was selected by the consumer for each sample on the consumer score sheet.

a solid pattern exists with each grade grouped in a relatively distinct position.

It was evident that the accuracy was higher for assigning 2 star unsatisfactory product and 5 star premium product (84.0% and 80.9%, respectively) compared to that of 3 star and 4 star product (52.0% and 45.8%, respectively) over all samples. These results were similar to that obtained by Watson *et al.* (2008a): 85.2% and 79.6% accuracy for 2 and 5 star, respectively, and 61.7% and 62.6% accuracy for 3 and 4 star. This difference in accuracy can be seen by the overlapping across grades as shown in Figure 1. If the MQ4-assigned grade was higher than the true consumer-selected grade, this would result in consumer dissatisfaction. On the other hand, if the MQ4-assigned grade was lower than the consumer-selected grade, the consumer would have their expectations exceeded. If both the consumers who had their expectations met and exceeded were taken into account, the accuracy of the MQ4-assigned grade would be much higher (approximately 81%). These results indicate that the satisfaction of the majority of the consumers could be achieved.

Weightings calculated from current data set

In order to investigate the importance of the four sensory attributes (tenderness, juiciness, flavor and overall satisfaction) to the Japanese consumer, the weightings of the four sensory attributes were calculated. The data can be displayed as linear functions as follows:

Grill MQ4 score = 0.30 tn + 0.20 ju + 0.24 fl + 0.26 ov

Yakiniku MQ4 score = 0.25 tn + 0.06 ju + 0.41 fl + 0.28 ov

Shabu Shabu MQ4 score = 0.16 tn + 0.20 ju + 0.40 fl + 0.24 ov

Overall MQ4 score = 0.24 tn + 0.15 ju + 0.35 fl + 0.26 ov

For grilled beef, the results showed that consumers rated tenderness as most important (0.30), followed by overall satisfaction (0.26), flavor (0.24) and juiciness (0.20). In contrast, flavor was the dominant attribute for yakiniku (0.41), shabu shabu (0.40) and overall (0.35). Juiciness was more important in grill and shabu shabu, than for yakiniku. The overall weightings suggest that flavor was the most important attribute for rating beef by Japanese consumers, which is in contrast to the MSA prediction model which rates tenderness as the most important attribute (Watson *et al.* 2008a). Further investigation is required in order to clarify the importance of the four sensory attributes for the Japanese consumer.

Consumer ratings by category

The proportion of consumers that rated the samples by the four categories (unsatisfactory, good everyday, better than everyday, and premium) are shown in Table 7. There were differences between cooking methods, with a higher percentage of Japanese consumers rating grilled steaks as unsatisfactory (40%), compared to that for yakiniku (32%) and shabu shabu (30%). The percentage of Japanese consumers that scored the samples as premium (5 star) were similar across the cooking methods. However there was a lower percentage of grill steaks rated as 3 star (37% compared to 41% yakiniku and 40% shabu shabu) and 4 star (16% compared to 20% yakiniku and 21% shabu shabu).

As the samples for each cooking method were prepared from the same cuts (with position rotated), the higher percentage of unsatisfactory grill steak samples suggest that Japanese consumers tended to rate grilled products more harshly than the same beef cooked by yakiniku or shabu shabu methods. Further statistical analysis is required in order to clarify the difference in Japanese consumer attributes among cooking methods.

Correlations between sensory scales

Table 8 shows a simple correlation of tenderness, juiciness, flavor and overall satisfaction for Japanese consumers for all three cooking methods. There was a high degree of correlation between the four sensory attributes, which suggested each consumer tended to rate similarly on each scale; if a product rated highly on one scale then it tended to rate high on the other scales. The degree of correlation between sensory attributes for Japanese consumers was much higher than that for Korean consumers, and very similar to that for Australian consumers (Thompson *et al.* 2008).

Table 7 Percentage of consumers that rated beef by the four grades by cooking method

	Unsatisfactory	Good everyday	Better than everyday	Premium
Grill steak	40	37	16	7
Yakiniku	32	41	20	7
Shabu shabu	30	40	21	9
Overall	34	39	19	8

'Unsatisfactory' (2 star; ungraded), 'good everyday' (3 star), 'better than everyday' (4 star) and 'premium quality' (5 star).

Table 8 Simple correlation coefficients between the four sensory attributes

	Tenderness	Juiciness	Flavor	Overall satisfaction
Tenderness	1.00			
Juiciness	0.79	1.00		
Flavor	0.76	0.80	1.00	
Overall satisfaction	0.81	0.83	0.96	1.00

Data combined from all cooking methods. $n = 11,340$.

Conclusion

Japanese consumers readily identified distinct differences in eating quality of beef samples. This categorization occurred across grill steak, yakiniku, and shabu shabu cooking methods, and was the result of an interaction between sensory traits, including tenderness, flavor, juiciness and overall satisfaction. It was possible to create a workable single composite meat quality score (MQ4 score), by combining the four sensory scales of tenderness, juiciness, flavor and overall satisfaction with the weightings from the MSA prediction model. The MQ4 score could be used to develop a consumer-focused grading system that aims to predict Japanese consumer satisfaction of individual cuts when cooked by different methods. Further studies to increase the range of muscles would be desirable, tested together with further cooking methods.

ACKNOWLEDGMENTS

The contributions of MLA and MSA staff in Australia and Japan and the funding by Australian beef producers through MLA levy funds is acknowledged. Services provided by Cosign Pty Ltd in sample preparation and sensory testing services by Sensory Solutions Pty Ltd in Australia and Synovate in Japan is also acknowledged.

REFERENCES

Boleman SJ, Miller RK, Taylor JF, Cross HR, Wheeler TL, Koohmaraie M, Shackelford SD, Miller RL, West RL, Johnson DD, Savell JW. 1997. Consumer evaluation of beef of known categories of tenderness. *Journal of Animal Science* **75,** 1521–1524.

Canadian Beef Grading Agency. 1997. *Canadian Beef Carcass Grading Regulations*. Canadian Beef Grading Agency, Calgary, AB.

Huffman KL, Miller MF, Hoover LC, Wu CK, Brittin HC, Ramsey CB. 1996. Effect of beef tenderness on consumer satisfaction with steaks consumed in the home and restaurant. *Journal of Animal Science* **74,** 91–97.

Hwang IH, Polkinghorne R, Lee JM, Thompson JM. 2008. Demographic and design effects on beef sensory scores given by Korean and Australian consumers. *Australian Journal of Experimental Agriculture* **48,** 1387–1395.

Japan Meat Grading Association (JMGA). 1988. *New Beef Grading Standards*. Japanese Meat Grading Associations, Tokyo. (In Japanese)

Kim CJ, Lee ES. 2003. Effects of quality grade on the chemical, physical and sensory characteristics of Hanwoo (Korean native cattle) beef. *Meat Science* **63,** 397–405.

Koizumi S, Takase H, Kimura T, Takaku S, Nishino M, Nagano M. 1986. Studies on consumer behaviour for animal products – XI. An investigation on beef. *Bulletin of the College of Agriculture and Veterinary Medicine of Nihon University* **43,** 249–269. (In Japanese)

Matsuishi M, Kume J, Itou Y, Takahashi M, Arai M, Nagatomi H, Watanabe K, Hayase F, Okitani A. 2004. Aroma components of Wagyu beef and imported beef. *Nihon Chikusan Gakkaiho* **75,** 409–415. (In Japanese)

Miller MF, Carr MA, Ramsey CB, Crockett CB, Hoover LC. 2001. Consumer thresholds for establishing the value of beef tenderness. *Journal of Animal Science* **79,** 3062–3068.

Okitani A. 1999. Organoleptic factors responsible for the superior palatability of Japanese Wagyu Beef. In: *Proceedings of 45th International Congress of Meat Science and Technology,* Yokohama, Japan. pp. 354–355.

Polkinghorne R. 2005. Does variation between muscles in sensory traits preclude carcass grading as a useful tool for consumers? In: *Proceedings of the 51st International Congress of Meat Science and Technology*, 7–12 August 2005, Baltimore, USA. p. 4.

Polkinghorne R, Thompson JM, Watson R, Gee A, Rooke M. 2008a. Evolution of the Meat Standards Australia (MSA) beef grading system. *Australian Journal of Experimental Agriculture* **48,** 1351–1359.

Polkinghorne R, Watson R, Porter M, Gee A, Scott J, Thompson J. 1999. Meat Standards Australia. A 'PACCP' based beef grading scheme for consumers. 1) The use of consumer scores to set grade standards. In: *Proceedings of the 45th International Congress of Meat Science and Technology,* 1–6 August 1999, Yokohama, Japan. pp. 14–15.

Polkinghorne R, Watson R, Thompson JM, Pethick DW. 2008b. Current usage and future development of the Meat Standards Australia (MSA) grading system. *Australian Journal of Experimental Agriculture* **48,** 1459–1464.

Sasaki K, Mitsumoto M. 2004. Questionnaire-based study on consumer requirements for beef quality in Japan. *Animal Science Journal* **75,** 369–376.

Statistical Analysis Software (SAS). 1997. *Applied Statistics and the SAS Programming Language*. SAS Institute, Cary, NC.

Thompson JM, Polkinghorne RJ, Hwang IH, Gee AM, Cho SH, Park BY, Lee JM. 2008. Beef Quality grades as determined by Korean and Australian consumers. *Australian Journal of Experimental Agriculture* **48,** 1380–1386.

United States Department of Agriculture (USDA). 1989. *Official United States Standards for Grades of Carcass Beef*. USDA, Washington, DC.

Watson R, Gee A, Polkinghorne R, Porter M. 2008a. Consumer assessment of eating quality – Development of protocols for MSA testing. *Australian Journal of Experimental Agriculture* **48,** 1360–1367. Accessory publication [cited 15 September 2010]. Available from URL: http://www.publish.csiro.au/nid/72/issue/4061.htm

Watson R, Polkinghorne R, Thompson JM. 2008b. Development of the Meat Standards Australia (MSA) prediction model for beef palatability. *Australian Journal of Experimental Agriculture* **48,** 1368–1379.