Untrained consumer assessment of the eating quality of European beef: 1. A single composite score can predict beef quality grades

S. P. F. Bonny[1,2,a], J.-F. Hocquette[2,3] D. W. Pethick[1], I. Legrand[4], J. Wierzbicki[5] P. Allen[6], L. J. Farmer[7], R. J. Polkinghorne[8] and G. E. Gardner[1]

[1] School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA 6150, Australia;

[2] INRA, UMR1213, Recherches sur les Herbivores, F-63122 Saint Genès Champanelle, France;

[3] Clermont Université, VetAgro Sup, UMR1213, Recherches sur les Herbivores, F-63122 Saint Genès Champanelle, France

[4] Institut de l'Elevage, Service Qualite' des Viandes, MRAL, 87060 Limoges Cedex 2, France;

[5] Polish Beef Association Ul. Kruczkowskiego 3, 00-380 Warszawa, Poland;

[6] Teagasac Food Research Centre, Ashtown, Dublin 15, Ireland;

[7] Agri-Food and Biosciences Institute, Newforge Lane, Belfast BT9 5PX, UK;

[8] 431 Timor Road, Murrurundi, NSW 2338, Australia;

[a] *Present address: School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA 6150, Australia;*

Corresponding author: Sarah Bonny. Email: S.Bonny@Murdoch.edu.au

Short title: Quality grades for beef can be predicted with a single score

Abstract

Quantifying consumer responses to beef across a broad range of demographics, nationalities and cooking methods is vitally important for any system evaluating beef eating quality. Based on previous work, it was expected that consumer scores would be highly accurate in determining quality grades for beef, thereby providing evidence that such a technique could be used to form the basis of and eating quality grading system for beef. Following Australian MSA (Meat Standards Australia) testing protocols, over 19 000 consumers from Northern Ireland, Poland, Ireland, France and Australia tasted cooked beef samples, then allocated them to a quality grade; unsatisfactory, good-every-day, better-than-every-day and premium. The consumers also scored beef samples for tenderness, juiciness, flavour-liking and overall-liking. The beef was sourced from all countries involved in the study and cooked by four different cooking methods and to three different degrees of doneness, with each experimental group in the study consisting of a single cooking doneness within a cooking method for each country. For each experimental group, and for the dataset as a whole, a linear discriminant function was calculated, using the four sensory scores which were used to predict the quality grade. This process was repeated using two conglomerate scores which are derived from weighting and combining the consumer sensory scores for tenderness, juiciness, flavour-liking and overall-liking, the original meat quality 4 score (oMQ4) (0.4, 0.1, 0.2, 0.3) and current meat quality 4 score (cMQ4) (0.3, 0.1, 0.3, 0.3). From the results of these analyses, the optimal weightings of the sensory scores to generate an ‘ideal meat quality 4 score (MQ4)’ for each country were calculated, and the MQ4 values that reflected the boundaries between the four quality grades were determined. The oMQ4 weightings were far more accurate in categorising European meat samples than the cMQ4 weightings,

highlighting that tenderness is more important than flavour to the consumer when determining quality. The accuracy of the discriminant analysis to predict the consumer scored quality grades was similar across all consumer groups, 68%, and similar to previously reported values. These results demonstrate that this technique, as used in the MSA system, could be used to predict consumer assessment of beef eating quality and therefore to underpin a commercial eating quality guarantee for all European consumers.

Keywords: Consumer testing, Beef, Quality, Grading, Europe

Implications

To combat the variable eating quality of beef, which is a major factor in declining beef consumption, the European beef industry needs an accurate and reliable description of eating quality. These results demonstrate that a single score can accurately describe a consumers eating experience, and therefore could be used to form the basis of a commercial eating quality grading scheme. If this model is pursued, the industry must continue to evaluate consumer assessment of eating quality, as there are indications that the relative importance of different sensory traits will change over time.

Introduction

Quantifying consumer responses to beef across a broad range of demographics, nationalities and cooking methods is vitally important for any system guaranteeing beef eating quality to the consumer. Within Europe, the introduction of an eating-quality guarantee system would be well accepted by consumers and strengthen the

European beef industry (Verbeke *et al.*, 2010). However, there are many diverse cultures with different cooking preferences, and it is unknown if a single system could accurately predict consumer responses to beef across Europe. One system that has the potential to reliably predict the eating quality of beef in Europe is the Australian quality prediction system called MSA (Meat Standards Australia). This system was judged uniquely innovative by French experts (Hocquette *et al.*, 2011) and is 50-70% accurate in allocating beef to one of four categories; unsatisfactory, good-every-day, better-than-every-day or premium using inputs known to impact on palatability (Thompson, 2002, Polkinghorne, 2006). At present, the MSA system is largely used within Australia, however it has also been tested in a number of other countries including Korea, Japan, South Africa and France (Hocquette *et al.*, 2014).

At the same time that consumers allocate meat samples to quality grades, they also score the same sample for tenderness, juiciness, flavour-liking and overall-liking. Australian and Japanese experiments have shown strong correlations between these four sensory scores ranging from 0.70 to 0.99 (Thompson *et al.*, 2008, Polkinghorne *et al.*, 2011). In contrast, the correlations for the Korean and South-African consumers were more variable (0.31-0.76, 0.38-0.96) (Thompson *et al.*, 2008, Thompson *et al.*, 2010). In all tested countries, the correlations were quite consistent between diverse cattle groups and cooking methods. The consistency of these correlations suggests that they will be strongly present where-ever consumers are tested, indicating some common eating behaviours of consumers towards beef in spite of the cultural diversity of these countries.

Initial consumer panels in Australia indicated that tenderness was the most important factor determining the quality grade of a piece of beef (Watson *et al.*, 2008b). However, in recent years more recent Australian consumer testing has described flavour-liking as having equal importance with tenderness (Thompson *et al.*, 2010). Work by Legrand *et al.* (2013) demonstrated French consumers also considered flavour and tenderness to be equally important. However, in both Japan and South Africa, flavour-liking was the most important determinant of consumer satisfaction (Thompson *et al.*, 2010, Polkinghorne *et al.*, 2011). Therefore, we expect that the importance of different sensory characteristics will vary between different countries.

This paper looks at consumers from France, Poland, Ireland, Northern Ireland and Australia, who tasted beef cooked by four different cooking methods. Based on the balance of the testing in other countries and of different cooking methods, we hypothesise that the consumer assessed sensory scores will be able to accurately predict the consumer assessed quality grades. However, we also expect that the importance of each sensory score in determining the quality grades of beef will vary between countries despite being strongly correlated with each other.

Material and methods

Consumer demographics

The consumer demographics are described in Table 1. In brief, there were 19 492 consumers, with each individual consumer scoring 6 beef samples. In addition to scoring beef samples, consumers also answered a demographic questionnaire

(Table 1). The English version of this questionnaire is detailed by Anonymous (2008). In all countries, there were slightly more female than male consumers, except for Ireland. The majority of consumers came from households with 1-3 adults and 0-2 children (Table 1).

Animals and muscle samples

The carcasses used for this experiment are described in detail by Bonny *et al.* (2016a). Briefly, the data set was formed through combining the records of a number of specific, smaller, experiments. As a result, this data set provides a cross-section of European cattle types form France, Poland, Ireland and Northern Ireland. The cattle were slaughtered commercially according to standard practice in each country. There was a range of 5-28 days *post mortem* wet ageing at 1 degree Celsius in vacuum pack bags. A total of 25 different muscles were collected, though not all muscles were collected from each carcass.

Meat preparation and consumer panels

Muscle samples were prepared for consumer testing by four different cooking methods, grill, roast, slow cook and Korean barbeque, and three different levels of cooking doneness, rare, medium and well-done, according to protocols for Meat Standards Australia (MSA) testing by personnel trained in MSA testing procedures (Anonymous, 2008, Watson *et al.*, 2008a, Legrand *et al.*, 2013). A brief description of the cooking methods, including preparation and final cooking doneness, which have been employed in this experiment has also been published by Bonny *et al.* (2016c).

The grill cooking method was performed in all countries and the roast cooking method was performed in all countries except for France. The slow cooking method was only tested in Poland and the Korean barbeque was tested in Ireland. Each consumer scored samples prepared by a single cooking method, cooked to a single cooking doneness. Not all muscles were prepared by each cooking method, the distribution of muscles and cooking method is described in more detail by Bonny *et al.* (2016c). A consumer received and evaluated seven meat portions: in every case the first portion was derived from either a generic striploin or rump muscle and expected to be of average quality. The remaining 6 portions were derived from different muscle samples selected to present each consumer with a diverse quality range, from poor to excellent, which were served in accordance with a 6x6 Latin square to balance potential order effects, as described by Watson *et al.* (2008b).

Consumers scored meat from their country of origin, except for two sessions where meat was tested between Poland and France in a complete factorial design. In all countries, consumers were sourced through both commercial consumer testing organisations and local clubs and charities. They were selected to reflect the general population. Consumers scored samples for tenderness, juiciness, flavour-liking and overall-liking, by making a mark on a 100 mm line scale, with the low end of the scale representing poor acceptability and the high end of the scale representing high acceptability. Consumers were also asked to allocate each sample into one of four categories that best described the sample; unsatisfactory, good-every-day, better-than-every-day or premium quality. For a more detailed description of the testing procedures and the questionnaire see Anonymous (2008).

Statistical analyses

The ability of the four sensory scores (tenderness, juiciness, flavour liking and overall liking) to predict the consumer grades (Unsatisfactory, good-every-day, better-than-every-day and premium) for the entire data set was calculated using a linear discriminant function (PROC DISCRIM SAS v9.2) (SAS, 2002). This was also performed for subsets of the data split into separate muscles, cooking methods, and cooking doneness by country. From these results, we can also effectively compare the quality grades, and in particular locate the point where a consumer is deciding between two adjacent categories. This point is assumed to be where adjacent functions are equal and is calculated by subtracting one adjacent function from the other. The boundary between those two quality grades (a number from 0 to 100) is calculated as the intercept after the resulting equation is divided by the sum of the coefficients for the four sensory scores. This procedure, including the discriminant analysis and boundary determination, was repeated using two different weighted combinations of the four sensory scores. The weightings used were the current weightings, employed by the Australian MSA model cMQ4: 0.3*tenderness, 0.1*juiciness, 0.3*flavour-liking and 0.3*overall-liking (Thompson *et al.*, 2010), and the initial weightings, (oMQ4) used when the MSA model was first developed: 0.4*tenderness, 0.1*juiciness, 0.2*flavour-liking and 0.3*overall-liking (Watson *et al.*, 2008a). Optimal weightings for the sensory scores were calculated averaging the discriminant functions as performed by Watson *et al.* (2008a). The discriminant procedure was performed both with and without overall-liking, and then the two resulting functions averaged to redistribute more of the components of overall-liking

back to the other sensory scores. This was performed for all the data and for each country individually.

Results

Allocation of beef samples to different quality grades by untrained consumers

The proportion of samples allocated by consumers to each of the four quality grade was relatively consistent between the different experimental groups (Table 2). In all cases the good-every-day category contained the highest proportion of samples, ranging from 30-42 %. In contrast, consumers only considered 10-19 % of samples tested to belong to the premium category. The Irish consumers tasting grilled and roasted samples, and the Polish consumers overall considered a higher proportion samples to be unsatisfactory than other consumer groups. Fifteen out of the 22 muscles tested had over 20% of the samples classified as unsatisfactory. And for five of those muscles this increased to over 35% (Table 3). In contrast, consumers classified over 20 % of samples from the *m. iliacus, m. spinalis dorsi* and the *m. psoas major,* as premium quality.

Simple correlations between sensory scores

All of the sensory scores were highly correlated (Table 4). Overall-liking and flavour-liking had the strongest correlation of all the sensory scores for all of the countries, cooking methods, cooking doneness levels, the correlation between overall-liking and flavour-liking were greater than or equal to 0.9 ($P<0.05$) except for the Polish scores,

which were greater than or equal to 0.86 (P<0.05). Similarly, overall-liking and tenderness had the second highest correlations across all the groups. The scores for the Polish grill and roast generally had lower correlations between the sensory scores than the rest of the data. Juiciness and flavour-liking were the least correlated of all the sensory scores, with values greater than or equal to 0.52 (P<0.001). Overall, the different muscles tested followed a similar pattern to the different countries, cooking methods and levels of cooking doneness (data not shown). Lower correlations, down to a minimum of 0.42 between flavour liking and juiciness, were seen in muscles with less data (data not shown).

Relative importance of sensory scores and the accuracy of the discriminant analyses

The importance of the sensory scores in determining the quality grades of beef were inferred from the discriminant analyses using all four sensory scores. In all cases, juiciness was the least important when consumers were determining the quality grade of beef (Table 5). Flavour-liking was the most important factor for consumers from Ireland and Poland. However, tenderness, flavour-liking and overall-liking were all of similar importance for Australian, French and Northern Irish consumers (Table 5).

The accuracies of the three discriminant functions, using the optimally weighted sensory scores and two different composite scores, are shown in Table 6. The optimally weighted to sensory scores and the oMQ4 weightings were similar, and more accurate than the cMQ4 weightings. For all groups, the accuracy was higher at the extreme ends of the quality grades, categories unsatisfactory and premium. For

the models with the optimally weighted sensory scores, and the oMQ4 weightings, an average of 81% and 79% of the samples were correctly allocated into categories unsatisfactory and premium. In contrast, only 57% and 49% of samples were correctly identified as either good-every-day or better-than-every-day (Table 6).

The accuracies for each muscle were highest for the unsatisfactory and premium categories, similar to the accuracy pattern for each country (Table 7). The accuracy of the discriminant analysis was similar for all muscles tested, with slight decreases in accuracy for muscles where there was less data. The average or overall accuracy of prediction for each of the muscles was similar, between 60 and 70% for all but two muscles. Muscle groups with accuracies lower than 66% tended to have less than 1000 consumer responses (Table 3, Table 7). If all the samples categorised as unsatisfactory by the discriminant function were removed, then only 6.98% of the remaining samples would be unsatisfactory according to the consumer classification. This percentage varied by muscle with the *m. psoas major* only having 1.29% of unsatisfactory samples remaining, and the *m. pectoralis profundus* having 17.17%.

Boundaries between the quality grades

Boundaries between the quality grades calculated from the discriminant analysis using the optimally weighted sensory scores, split by country, cooking method and cooking doneness are shown in table 8. There is little difference between the boundaries between the groups, with a maximum difference for the better-than-every-day to premium boundary of 12 points between the Irish Korean barbeque samples and the Northern Irish grill samples. The lower boundary cut-off was about 5

points lower for the French than the Polish, whereas the upper boundary was about 5 points higher for the French than the Polish. This is reflected in the greater range between the different cut-offs demonstrated by the French, with the upper and lower boundaries differing by approximately 44 points. In contrast, the Polish had the narrowest range, with the difference between the upper and lower boundaries only 32-35 points. Within a country the Northern Irish had the most variation in cut-offs, with the medium grill samples having a range between the upper and lower cut-offs 10 points narrower than the well done roast samples.

Discussion

Simple correlations between sensory scores

Supporting our hypothesis the European untrained consumer sensory scores for tenderness, juiciness, flavour-liking and overall-liking all had strong correlations, aligning well with similar studies in Australia and Japan (Thompson *et al.*, 2008, Thompson *et al.*, 2010, Polkinghorne *et al.*, 2011). Correlations were also similar between different cooking methods following the results of Thompson *et al.* (2008). These results validate the protocol of combining the sensory scores into a single MQ4 score to describe the consumers eating experience. Additionally, they potentially indicate that the use of a single score, such as overall liking, may be sufficient for untrained consumers. However, the correlations between the sensory scores are not perfect, indicating some ability of untrained consumers to differentiate between certain sensory characteristics of beef, particularly flavour and juiciness. Further supporting this, the relative importance of flavour-liking in Australia has increased over the last 15 years (Watson *et al.*, 2008b), possibly because the MSA

system resolved a large proportion of the tenderness issues in the domestic market (Thompson *et al.*, 2008).

Weightings of sensory scores to determine a single eating quality value

Aligning with the hypothesis, the weightings of the sensory scores differed between countries. For the Australian consumers, the calculated optimal weightings of the sensory scores for the MQ4 score closely match the formula currently used in the Australian MSA system (Thompson *et al.*, 2010). This trend was similar for the French consumers with the results closely aligning with previously published work on the same consumers (Legrand *et al.*, 2013). Flavour-liking was more important in determining the MQ4 score in Ireland, Northern Ireland and Poland, similar to Japan and South Africa (Thompson *et al.*, 2010, Polkinghorne *et al.*, 2011). This variation in the importance of tenderness and flavour-liking may relate to cultural differences between consumer groups (Lorenzen *et al.*, 1999, Feuz *et al.*, 2004, Reicks *et al.*, 2011). Additionally, standard beef production practices vary between countries and would also impact the quality of beef produced and therefore the relative importance in these markets. More work with balanced crossover experimental designs would be needed to differentiate the effect of the country of origin of the consumer and the beef.

There were also indications within the results that cooking method may have an influence on the relative importance of the sensory scores. Much of the literature indicates that tenderness is the most important factor in determining consumer satisfaction (Huffman *et al.*, 1996, Verbeke *et al.*, 2010). However, in this study tenderness had lower weightings and therefore was less important than flavour liking

in determining the quality grade for both the roast and slow cook, though the exact differences varied greatly. As tenderness and flavour have equal importance in consumer purchasing decisions for both steaks and roasts (Reicks *et al.*, 2011), this may indicate that these cooking methods resulted in a more consistent tenderness across the samples, thereby increasing the importance of flavour liking in discriminating between the quality grades. A relative increase in the importance of flavour liking was also seen for the samples prepared by the Korean barbeque method for this experiment, perhaps reflecting the thin slices reducing the importance of tenderness. This is similar to the low importance of tenderness found for the shabu shabu cooking method when compared with grill cooking methods in Japan (Polkinghorne *et al.*, 2014). However, the authors note that these differences have not been tested for significance.

The boundaries between the quality grades

The boundaries between the quality grades were similar between all groups, however the higher (better-than-every-day to premium) and the lower (unsatisfactory to good-every-day) boundaries differed between countries. These small differences may have resulted from differences in how consumers of different countries use the scale to score meat. For example, the French consumers had a lower MQ4 score boundary between unsatisfactory and good-every-day than all of the other countries studied here and previously (Hocquette *et al.*, 2014). Yet they also had one of the highest upper boundaries between better-than-every-day and premium, similar to those reported for the Japanese (83.1) and Australians (Hocquette *et al.*, 2014). As the experiment is designed to present all consumers with meat of a range of

qualities, these boundaries indicate that the French consumers have a greater propensity to use the whole scale for scoring meat. In contrast, Polish consumers tended to avoid the extreme ends of the scale and the consumer scores are more likely to cluster around the middle. However, the authors can't discount the possibility that these are simply random differences between consumer groups, particularly as the discriminant analysis retained high accuracies across the data as a whole. Furthermore, the high accuracies of the discriminant analysis over all the data also indicates that, even if genuine, minor variation in the use of the scale between demographic groups when scoring beef is unlikely to be an important factor in the prediction of eating quality. Cooking method alone had no discernible influence on the boundaries between the quality grades.

The accuracy of using composite values to allocate samples to quality grades

For all the different cuts tested, the discriminant analysis was most accurate for the unsatisfactory and premium categories. This pattern was also seen for the South African and the Japanese consumers and for a larger group of Australian consumers (Watson *et al.*, 2008b, Thompson *et al.*, 2010, Polkinghorne *et al.*, 2011). These results indicate that the four sensory traits used in this study appropriately describes consumer scoring for the higher and lower grades, but further work would be needed to investigate if other factors are influencing the consumer determination of the intermediate quality grades. Alternatively, the variations in accuracy may simply reflect how certain consumers are when they grade a sample as either premium or unsatisfactory with the drop in accuracy representing a relative uncertainty when distinguishing between good-every-day and better-than-every-day. However, this

pattern hasn't been consistently reported in previous work (Thompson *et al.*, 2010, Polkinghorne *et al.*, 2011).

The average accuracy of 67% found in this study is very similar to the accuracy achieved in Australia (Watson *et al.*, 2008b) and other reported accuracies such as 55-65 % in South Africa, 68 % in Korea and 60-67 % in Japan (Thompson *et al.*, 2008, Thompson *et al.*, 2010, Polkinghorne *et al.*, 2011). This indicates that the approach is useful in determining quality grades for an eating quality guarantee system facing a diverse set of consumers and cooking methods.

The drop accuracy between the analysis using optimal weightings and the oMQ4 weightings was minimal, about 2%. This was similar to the reported values for the Korean consumers (1-7%) and South African consumers (3%) (Thompson *et al.*, 2008, Thompson *et al.*, 2010). By comparison, the drop accuracy when using the cMQ4 weightings was around 20%. This result shows that the oMQ4 weightings of 0.4*tenderness, 0.1*juiciness, 0.2*flavour-liking and 0.3*overall-liking would likely function well in a European eating quality grading system. However, it is important to note that the Australian model experienced a change in the relative importance of tenderness and flavour over time, we would speculate that a similar shift would occur if an equivalent model were to be established in Europe. This would have to be monitored by continued consumer testing to follow the evolution of the importance of the sensory scores in determining the MQ4 score and the quality grades over time.

Ability of a prediction model to increase satisfaction for the European consumer

Overall, the chance a consumer had an unsatisfactory eating experience in this experiment was around 25%, with some groups as high 40%. The failure rate of beef was relatively consistent over country, cooking method, doneness and muscle type and supports the theories that variable eating quality is one of the biggest factors negatively impacting the beef industry (Morgan *et al.*, 1991, Polkinghorne *et al.*, 2008). Importantly, the carcasses selected for this study did not impact on these results as Farmer *et al.* (2016) demonstrated in a subset of this data that the failure rate of beef did not change when variation in the carcasses was reduced using standard industry quality cut-offs such as animal age, sex, post mortem ageing and ultimate pH.

The discriminant analysis correctly recognised over 80% of the unsatisfactory samples. Identifying these samples in the industry would reduce the risk of a consumer having an unsatisfactory eating experience from 25% to 7%. As the current beef grading system in Europe is inadequate for the prediction of eating quality (Bonny *et al.*, 2016b), other methods of predicting consumer satisfaction have been investigated. However, most have proven to have limited usefulness or have severe barriers to commercialisation. For example, muscle biochemistry can only explain a small part of beef quality (Renand *et al.*, 2001, Chriki *et al.*, 2013, Bonny *et al.*, 2015). Similarly, the large volume of research that has been conducted into genomic and proteomic techniques has yielded only a relatively low efficiency to predict beef quality (Hocquette *et al.*, 2012). In contrast, the commercialised prediction model used by MSA in Australia is based on carcass and animal traits collected at slaughter. It has been shown that such a system can accurately predict

European consumer scores of beef quality, with minor modifications (Bonny *et al.*, 2016a, Bonny *et al.*, 2016c).

Conclusion

These results highlight the urgent need of the European beef industry to address variable beef eating quality. Improving beef eating quality is achievable using an eating quality grading system modelled on the MSA system, as evaluated in this study. Over 80% of unsatisfactory beef was correctly identified by our analysis and the chance that a consumer would have an unsatisfactory eating experience was reduced from 25% to 7%. In a commercialised system, consumer scores would need to be predicted using animal and carcass traits. This information would be used to greatly increase the consistency of the product available to the consumer and the redirection of the vast majority of low quality product from the fresh beef supply to other markets.

Acknowledgements

This research was supported by Meat and Livestock Australia and Murdoch University. Data were obtained through the financial contributions of the European research project ProSafeBeef (Contract No. FOOD-CT-2006-36241), the Polish ProOptiBeef Farm to Fork project funded by the EU Innovative (POIG.01.03.01-00-204/09), the French ‘Direction Générale de l’Alimentation’ and FranceAgriMer, the Irish Department of Agriculture Food and The Marine under the FIRM programme, and the Northern Ireland Department of Agriculture and Rural Development ‘Vision’

programme. Furthermore, this project would not have been possible without the practical support of the Association Institut du Charolais, the Syndicat de Défense et du promotion de la Viande de Boeuf de Charolles and the gourmet restaurants 'Jean Denaud' and representatives of the beef industry across Europe. The international travel required for this project has been funded by 'Egide/Fast' funds from the French and Australian governments respectively (project no. FR090054) and by 'Egide/Polonium' funds from the French and Polish governments respectively. The assistance and participation the Beef CRC and Janine Lau (MLA, Australia), Alan Gee (Cosign, Australia), Ray Watson (Melbourne University, Australia) and John Thompson (UNE) are also gratefully acknowledged.

References

Anonymous 2008. Accessory Publication: MSA sensory testing protocols Australian Journal of Experimental Agriculture 48, 1360-1367.

Bonny SPF, Gardner GE, Pethick DW, Legrand I, Polkinghorne RJ and Hocquette J-F 2015. Biochemical measurements of beef are a good predictor of untrained consumer sensory scores across muscles. Animal 9, 179-190.

Bonny SPF, Hocquette J-F, Pethick DW, Farmer LJ, Legrand I, Wierzbicki J, Allen P, Polkinghorne RJ and Gardner GE 2016a. The variation in the eating quality of beef from different sexes and breed classes cannot be completely explained by carcass measurements. Animal 10, 987-995.

Bonny SPF, Pethick DW, Legrand I, Wierzbicki J, Allen P, Farmer LJ, Polkinghorne RJ, Hocquette J-F and Gardner GE 2016b. European conformation and fat scores have no relationship with eating quality. Animal 10, 996-1006.

Bonny SPF, Pethick DW, Legrand I, Wierzbicki J, Allen P, Farmer LJ, Polkinghorne RJ, Hocquette JF and Gardner GE 2016c. Ossification score is a better indicator of maturity related changes in eating quality than animal age. Animal 10, 718-728.

Chriki S, Renand G, Picard B, Micol D, Journaux L and Hocquette J-F 2013. Meta-analysis of the relationships between beef tenderness and muscle characteristics. Livestock Science 155, 424-434.

Farmer LJ, Bowe R, Troy DJ, Bonny SPF, Birnie J, Dell'Orto V, Polkinghorne RJ, Wierzbicki J, de Roest K, Scollan ND, Henchion M, Morrison SJ, Legrand I, Roehe R, Hocquette JF and Duhem K 2016. Report of the workshop "Sustainable beef quality for Europe – A workshop for industry and scientists". Viandes & Produits Carnés VPC-2016-32-1-6.

Feuz DM, Umberger WJ, Calkins CR and Sitz B 2004. U.S. Consumers' Willingness to Pay for Flavor and Tenderness in Steaks as Determined with an Experimental Auction. Journal of Agricultural and Resource Economics 29, 501-516.

Hocquette J-F, Legrand I, Jurie C, Pethick DW and Micol D 2011. Perception in France of the Australian system for the prediction of beef quality (Meat Standards Australia) with perspectives for the European beef sector. Animal Production Science 51, 30-36.

Hocquette J-F, Botreau R, Picard B, Jacquet A, Pethick DW and Scollan ND 2012. Opportunities for predicting and manipulating beef quality. Meat Science 92, 197-209.

Hocquette J-F, Van Wezemael L, Chriki S, Legrand I, Verbeke W, Farmer L, Scollan ND, Polkinghorne R, Rødbotten R, Allen P and Pethick DW 2014. Modelling of

beef sensory quality for a better prediction of palatability. Meat Science 97, 316-322.

Huffman KL, Miller MF, Hoover LC, Wu CK, Brittin HC and 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.

Legrand I, Hocquette J-F, Polkinghorne RJ and Pethick DW 2013. Prediction of beef eating quality in France using the Meat Standards Australia system. Animal 7, 524-529.

Lorenzen CL, Neely TR, Miller RK, Tatum JD, Wise JW, Taylor JF, Buyck MJ, Reagan JO and Savell JW 1999. Beef customer satisfaction: cooking method and degree of doneness effects on the top loin steak. Journal of Animal Science 77, 637-644.

Morgan JB, Savell JW, Hale DS, Miller RK, Griffin DB, Cross HR and Shackelford SD 1991. National beef tenderness survey. Journal of Animal Science 69, 3274-3283.

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

Polkinghorne RJ 2006. Implementing a palatability assured critical control point (PACCP) approach to satisfy consumer demands. Meat Science 74, 180-187.

Polkinghorne RJ, Nishimura T, Neath KE and Watson R 2011. Japanese consumer categorisation of beef into quality grades, based on Meat Standards Australia methodology. Animal Science Journal 82, 325-333.

518 Polkinghorne RJ, Nishimura T, Neath KE and Watson R 2014. A comparison of
519 Japanese and Australian consumers' sensory perceptions of beef. Animal
520 Science Journal 85, 69-74.
521 Reicks AL, Brooks JC, Garmyn AJ, Thompson LD, Lyford CL and Miller MF 2011.
522 Demographics and beef preferences affect consumer motivation for purchasing
523 fresh beef steaks and roasts. Meat Science 87, 403-411.
524 Renand G, Picard B, Touraille C, Berge P and Lepetit J 2001. Relationships between
525 muscle characteristics and meat quality traits of young Charolais bulls. Meat
526 Science 59, 49-60.
527 SAS 2002. Applied statistics and the SAS programming language. In SAS institute,
528 Cary, USA.
529 Thompson J, Polkinghorne R, Gee A, Motiang D, Strydom P, Mashau M, Ng'ambi J,
530 deKock R. and Burrow H 2010. Beef palatability in the Republic of South Africa:
531 implications for niche-marketing strategies. In ACIAR Technical Reports, pp. 1-
532 56. Australian Centre for International Agricultural Research ACIAR, Canberra,
533 ACT, Australia.
534 Thompson JM 2002. Managing meat tenderness. Meat Science 62, 295-308.
535 Thompson JM, Polkinghorne R, Hwang IH, Gee AM, Cho SH, Park BY and Lee JM
536 2008. Beef quality grades as determined by Korean and Australian consumers.
537 Australian Journal of Experimental Agriculture 48, 1380-1386.
538 Verbeke W, Van Wezemael L, de Barcellos MD, Kügler JO, Hocquette J-F, Ueland Ø
539 and Grunert KG 2010. European beef consumers' interest in a beef eating-
540 quality guarantee: Insights from a qualitative study in four EU countries.
541 Appetite 54, 289-296.

542 Watson R, Polkinghorne R and Thompson JM 2008a. Development of the Meat
543 Standards Australia (MSA) prediction model for beef palatability. Australian
544 Journal of Experimental Agriculture 48, 1368-1379.
545 Watson R, Gee A, Polkinghorne R and Porter M 2008b. Consumer assessment of
546 eating quality – development of protocols for Meat Standards Australia (MSA)
547 testing. Australian Journal of Experimental Agriculture 48, 1360-1367.
548
549

550 ***Table 1*** Number of consumers within each of the demographic categories for each country who scored beef samples

Demographic categories	Distribution of traits					
Gender	Male	Female	Unreported			
Australia	148	191	02			
France	672	826	2			
Ireland	921	755	64			
Northern Ireland	3 938	4 994	60			
Poland	3 217	4 030	13			
Age (years)	<20	20-30	31-45	46-50	>50	Unreported
Australia	35		37	267		0
France	50	413	431	137	468	1
Ireland	0	603	509	213	393	22
Northern Ireland	3935		2404	2611		42
Poland	502	4123	1804		822	9
Income	*a*	*b*	*c*	*d*	*e*	Unreported
Australia	-	-	-	-	-	339
France	128	446	493	302	129	2
Ireland	184	727	773	-	-	56
Northern Ireland	2 760	4 814	1 224	-	-	194
Poland	817	1 033	2 184	2 234	940	52
Occupation	Trade	Professional	Admin[1]	Technical	Service	Labourer
Australia	43	95	52	31	38	10
France	39	232	542	130	0	100
Ireland	147	519	213	282	110	28
Northern Ireland	924	2 093	1 125	627	709	633
Poland	297	517	1526	479	815	834
	Unemployed	Student	Retired	Homemaker	Other	Unreported
Australia	4	56	6	4	0	0
France	82	82	257	26	8	2
Ireland	34	170	0	203	0	34

Northern Ireland	617	944	0	1 177	0	143	
Poland	219	1190	0	103	1 257	23	
Adults in the home	0	1	2	3	4	5+	Unreported
Australia	0	29	207	64	29	10	0
France	4	311	877	189	83	33	3
Ireland	0	138	707	379	311	181	24
Northern Ireland	505	1 136	3 844	1 678	1 178	576	75
Poland	45	1 001	2 457	1 816	1 304	631	6
Children in the home	0	1	2	3	4	5+	Unreported
Australia	118	56	116	39	9	1	0
France	959	240	222	62	13	2	2
Ireland	531	517	247	212	85	46	102
Northern Ireland	2	5 890	1 099	1 130	421	163	287
Poland	5 080	1 349	537	125	35	8	126

[1] Admin=Administration

[2] Income categories are different for each country. France (€/month): *a*=<1 000, *b*=1 000-2 000, *c*=2 000-3 000, *d*=3 000-4 000, *e*=>4 000; Ireland (€/year): *a*=<20 000, *b*=20 000-50 000, *c*=>50 000; Northern Ireland (£/year): *a*=<20 000, *b*=20 000-50 000, *c*=>50 000; Poland (zl/month): *a*=≤1 000, *b*=1 001-1 400, *c*=1 401-2 200, *d*=2 201-4 000, *e*=>4 000.

555 ***Table 2*** Frequency distribution (and number of samples) of eating quality grades allocated
556 to beef samples by untrained consumers across countries, cook methods and degrees
557 of doneness.

Experimental group	Unsatisfactory	Good[1]	Better[2]	Premium
All	0.26(29 927)	0.38(44 153)	0.24(28 466)	0.12(14 010)
Australia, grill, medium	0.25(115)	0.39(177)	0.22(100)	0.14(66)
France, grill medium	0.25(1 810)	0.30(2 133)	0.26(1 896)	0.19(1 349)
France, grill, rare	0.25(583)	0.31(723)	0.26(602)	0.18(423)
Ireland, grill, medium	0.29(1 862)	0.38(2 440)	0.21(1 356)	0.12(739)
Ireland, roast, medium	0.31(528)	0.41(701)	0.20(349)	0.09(148)
Ireland, BBQ[3] medium	0.19(407)	0.39(825)	0.28(581)	0.14(300)
NI[4], grill, medium	0.22(1 156)	0.43(2 293)	0.25(1 360)	0.11(567)
NI[4], grill well done	0.23(6 724)	0.40(11 525)	0.26(7 464)	0.12(3 454)
NI[4], roast, medium	0.23(582)	0.37(938)	0.27(671)	0.12(311)
NI[4], roast, well done	0.22(3 565)	0.42(6 835)	0.26(4 280)	0.10(1 628)
Poland, grill, medium	0.28(9 465)	0.36(12 021)	0.23(7 813)	0.12(4 055)
Poland, roast, medium	0.32(1 594)	0.37(18 32)	0.21(1 048)	0.10(523)
Poland, slow cook, medium	0.33(1 536)	0.37(17 10)	0.20(946)	0.96(447)

558 [1] Good=Good-every-day
559 [2] Better=Better-than-every-day
560 [3] BBQ= Korean barbeque
561 [4] NI=Northern Ireland
562 Cooking methods degrees of doneness are described by Bonny *et al.* (2016c) and
563 Anonymous (2008)

564 ***Table 3*** Frequency distribution (and number of samples) of eating quality grades allocated

565 to beef samples by untrained consumers across the different cuts tested.

Muscle	Unsatisfactory	Good[1]	Better[2]	Premium
m. psoas major	0.07 (526)	0.22 (1 715)	0.34 (2 690)	0.37 (2 935)
m. spinalis dorsi	0.13 (144)	0.27 (288)	0.33 (351)	0.27 (287)
m. iliacus	0.22 (13)	0.20 (12)	0.35 (21)	0.23 (14)
m. gracilis	0.15 (41)	0.30 (82)	0.36 (99)	0.19 (53)
m. infraspinatus	0.22 (499)	0.31 (711)	0.28 (655)	0.19 (442)
m. semispinalis capitis	0.18 (27)	0.31 (45)	0.33 (48)	0.18 (27)
m. gluteus medius[4]	0.15 (857)	0.37 (2 201)	0.32 (1912)	0.16 (932)
m. biceps femoris[5]	0.19 (775)	0.37 (1518)	0.29 (1 196)	0.15 (605)
m. longissimus[6]	0.21 (423)	0.37 (748)	0.29 (594)	0.14 (284)
m. longissimus[7]	0.19 (5 962)	0.39 (12 262)	0.28 (8 732)	0.13 (4 140)
m. serratus ventralis cervicis	0.26 (275)	0.38 (400)	0.23 (243)	0.12 (130)
m. gluteus medius[8]	0.22 (2 754)	0.40 (5 074)	0.27 (3 459)	0.11 (1 394)
m. rectus femoris	0.22 (1 121)	0.40 (2 015)	0.27 (1 374)	0.11 (545)
m. triceps brachii caput longum	0.27 (637)	0.44 (1 048)	0.20 (487)	0.09 (216)
m. tensor fasciae latae	0.31 (256)	0.38 (316)	0.24 (197)	0.07 (57)
m. adductor femoris	0.29 (834)	0.44 (1 281)	0.20 (593)	0.06 (187)
m. supraspinatus	0.36 (129)	0.37 (132)	0.21 (74)	0.06 (23)
m. pectoralis profundus	0.46 (82)	0.39 (70)	0.09 (17)	0.06 (11)
m. semitendinosus	0.32 (1 057)	0.43 (1 412)	0.19 (628)	0.06 (192)
m. vastus lateralis	0.41 (1 254)	0.38 (1 181)	0.16 (495)	0.05 (159)
m. semimembranosus	0.38 (7 124)	0.41 (7 756)	0.17 (3 141)	0.05 (950)
m. biceps femoris[9]	0.47 (5 137)	0.36 (3 886)	0.13 (1 460)	0.04 (427)
Total	0.26 (29 927)	0.38 (44 153)	0.24 (28 466)	0.12 (14 010)

566 [1] Good=Good-every-day

567 [2] Better=Better-than-every-day

568 [3.]Percentage of samples that were scored as unsatisfactory but predicted to be 3 star or

569 above

570 [4] Eye of rump

571 [5] Rump cap

572 [6] *m. longissimus thoracis et lumborum,* Cube roll

573 [7] *m. longissimus thoracis et lumborum,* Striploin

574 [8] Eye rump centre

575 [9] Silverside

576 ***Table 4*** Simple correlation coefficients (*P*<0.001) between untrained consumer sensory
577 scores for beef samples across all data or within sub-groups defined by their country, cook
578 method, and doneness.
579

Experimental group	Tn[1]/Ju[2]	Tn/Fl[3]	Tn/Ov[4]	Ju/Fl	Ju/Ov	Fl/Ov
All	0.65	0.66	0.76	0.65	0.69	0.89
Australia, grill, medium	0.73	0.78	0.84	0.80	0.83	0.92
France, grill medium	0.61	0.77	0.87	0.63	0.67	0.92
France, grill, rare	0.63	0.75	0.87	0.64	0.68	0.91
Ireland, grill, medium	0.67	0.71	0.79	0.7	0.73	0.90
Ireland, roast, medium	0.66	0.72	0.75	0.67	0.68	0.91
Ireland, BBQ[6] medium	0.70	0.70	0.77	0.75	0.76	0.90
NI[5], grill, medium	0.72	0.70	0.79	0.72	0.76	0.90
NI[5], grill well done	0.75	0.75	0.82	0.76	0.80	0.92
NI[5], roast, medium	0.67	0.68	0.75	0.65	0.69	0.91
NI[5], roast, well done	0.74	0.75	0.8	0.77	0.80	0.94
Poland, grill, medium	0.59	0.57	0.68	0.52	0.56	0.87
Poland, roast, medium	0.69	0.57	0.65	0.53	0.58	0.88
Poland, slow cook, medium	0.75	0.62	0.73	0.67	0.75	0.86

580 [1] Tn=Tenderness
581 [2] Ju=Juiciness
582 [3] Fl=Flavour-liking
583 [4] Ov=Overall-liking
584 [5] NI=Northern Ireland
585 [6]barbeque
586 Cooking methods degrees of doneness are described by Bonny *et al.* (2016c) and
587 Anonymous (2008)

588 ***Table 5*** Optimal weightings of sensory scores overall, and for each country from a

589 discriminant analysis using the sensory scores to allocate beef to quality grades

Experimental group	Tenderness	juiciness	flavour-liking	overall-liking
All	0.28	0.04	0.37	0.31
Australia, grill, medium	0.34	0.05	0.27	0.33
France, grill medium	0.33	0.02	0.32	0.33
France, grill, rare	0.31	-0.01	0.34	0.36
Ireland, grill, medium	0.31	0.10	0.35	0.24
Ireland, roast, medium	0.21	0.16	0.27	0.36
Ireland, BBQ[3] medium	0.24	0.07	0.42	0.27
NI[4], grill, medium	0.29	0.08	0.34	0.28
NI[4], grill well done	0.29	0.08	0.33	0.30
NI[4], roast, medium	0.21	0.10	0.39	0.30
NI[4], roast, well done	0.19	0.12	0.35	0.34
Poland, grill, medium	0.23	0.04	0.42	0.32
Poland, roast, medium	0.18	0.06	0.43	0.33
Poland, slow cook, medium	0.25	0.13	0.35	0.28

590 [1.]BBQ=Korean Barbeque

591 [2] NI=Northern Ireland

592 ***Table 6*** Percentage accuracy of the discriminate functions to allocate beef samples to the
593 quality grade assigned by the consumer (expressed as a percentage of the number
594 allocated to the correct grade relative to the number predicted for a particular grade)
595

Experimental group	Unsatisfactory	Good[1]	Better[2]	Premium	Overall
Optimally weighted Sensory scores					
All	80.7	57.2	51.9	79.0	67.2
Australia, grill, medium	81.7	62.2	58.0	86.4	72.1
France, grill medium	81.4	68.3	62.2	84.5	74.1
France, grill, rare	81.1	75.1	66.8	86.1	77.3
Ireland, grill, medium	80.9	50.3	42.7	83.1	64.3
Ireland, roast, medium	76.1	37.1	34.4	41.0	47.2
Ireland, BBQ[3] medium	84.8	52.4	55.1	76.7	67.2
NI[4], grill, medium	74.3	40.6	34.0	71.8	55.2
NI[4], grill well done	84.8	59.3	54.5	78.9	69.4
NI[4], roast, medium	84.0	63.7	54.8	80.1	70.6
NI[4], roast, well done	84.4	58.5	54.8	79.6	69.3
Poland, grill, medium	81.8	61.6	52.7	78.5	68.6
Poland, roast, medium	80.9	58.4	52.2	77.8	67.3
Poland, slow cook, medium	80.4	59.2	48.6	75.0	65.8
cMQ4[5] weightings					
All	64.1	22.4	20.5	62.2	42.3
Australia, grill, medium	61.7	40.9	24.2	74.2	50.3
France, grill medium	66.8	26.7	28.5	72.4	48.6
France, grill, rare	63.3	30.8	38.8	69.0	50.5
Ireland, grill, medium	64.9	19.7	18.2	61.0	41.0
Ireland, roast, medium	61.3	12.7	13.1	60.8	37.0
Ireland, BBQ[3] medium	64.3	17.6	17.3	61.5	40.2
NI[4], grill, medium	58.1	18.2	16.2	57.9	37.6
NI[4], grill well done	64.8	23.3	17.5	47.6	38.3
NI[4], roast, medium	58.2	21.7	16.1	60.3	39.1
NI[4], roast, well done	65.4	22.0	20.8	62.3	42.6
Poland, grill, medium	65.5	20.7	16.9	63.7	41.7
Poland, roast, medium	63.8	23.9	17.1	60.8	41.4
Poland, slow cook, medium	66.6	21.8	19.8	61.9	42.5
oMQ4[6] weightings					
All	79.4	53.0	48.6	78.7	64.9
Australia, grill, medium	81.7	58.2	54.0	58.0	63.0

France, grill medium	80.6	63.3	61.7	85.7	72.8
France, grill, rare	81.0	69.6	66.6	85.3	75.6
Ireland, grill, medium	78.3	48.1	40.8	83.4	62.6
Ireland, roast, medium	72.4	33.1	31.0	81.1	54.4
Ireland, BBQ[3] medium	79.4	50.3	52.8	75.7	64.5
NI[4], grill, medium	73.4	36.2	30.7	72.0	53.1
NI[4], grill well done	84.8	56.9	51.7	79.2	68.2
NI[4], roast, medium	79.7	60.6	52.0	82.0	68.6
NI[4], roast, well done	81.9	53.5	53.0	79.7	67.0
Poland, grill, medium	79.1	52.8	45.4	77.6	63.7
Poland, roast, medium	78.6	48.5	43.4	77.1	61.9
Poland, slow cook, medium	80.5	52.6	45.1	77.2	63.8

[1] Good=Good-every-day

[2] Better=Better-than-every-day

[3] BBQ= Korean barbeque

[4] NI=Northern Ireland

[5] cMQ4 (current MQ4) = a weighted combination (0.3, 0.1, 0.3, 0.3) of four sensory scores, tenderness, juiciness, flavour-liking and overall-liking;

[6] oMQ4 (initial MQ4) = a weighted combination (0.4, 0.1, 0.2, 0.3) of four sensory scores, tenderness, juiciness, flavour-liking and overall-liking;

Table 7 Percentage accuracy of the discriminate functions, using all four sensory scores, to allocate beef samples to the quality grade assigned by the consumer for each muscle (expressed as a percentage of the number allocated to the correct grade relative to the number predicted for a particular grade)

Muscle	Unsatisfactory	Good[1]	Better[2]	Premium	Overall	Pass/Fail[3]
m. psoas major	82.5	57.3	52.6	78.8	67.8	1.29
m. spinalis dorsi	81.3	70.5	55.0	75.6	70.6	2.92
m. iliacus	92.3	58.3	52.4	85.7	72.2	2.13
m. gracilis	82.9	48.8	50.5	79.3	65.4	3.20
m. infraspinatus	81.4	64.7	51.5	77.4	68.7	5.17
m. semispinalis capitis	88.9	62.2	43.8	29.6	56.1	2.56
m. gluteus medius[4]	81.0	57.7	53.5	77.3	67.4	3.40
m. biceps femoris[5]	81.8	58.7	54.5	82.8	69.5	4.42
m. longissimus[6]	82.5	60.3	51.5	77.8	68.0	4.68
m. longissimus[7]	80.2	53.4	50.6	78.5	65.7	5.03
m. serratus ventralis cervicis	85.8	60.5	48.2	72.3	66.7	5.18
m. gluteus medius[8]	81.8	57.7	51.7	79.6	67.7	5.31
m. rectus femoris	81.2	60.7	54.2	78.2	68.6	5.57
m. triceps brachii caput longum	82.6	59.0	45.6	76.9	66.0	6.65
m. tensor fasciae latae	80.5	57.6	48.2	80.7	66.8	8.88
m. adductor femoris	79.4	61.1	55.8	79.1	68.9	8.60
m. supraspinatus	78.3	59.9	48.7	69.6	64.1	11.97
m. pectoralis profundus	79.3	47.1	29.4	90.9	61.7	17.17
m. semitendinosus	75.7	49.1	43.3	82.3	62.6	12.11
m. vastus lateralis	79.0	61.1	53.1	78.0	67.8	13.94
m. semimembranosus	80.8	57.8	48.2	77.8	66.2	11.75
m. biceps femoris[9]	80.6	59.4	51.9	76.4	67.1	16.35
Total	80.7	57.0	51.9	79.0	67.2	6.98

[1] Good=Good-every-day

[2] Better=Better-than-every-day

[3] Percentage risk that a consumer would have an unsatisfactory eating experience if purchasing beef classified as good-every-day, better, or premium by the discriminant analysis

[4] Eye of rump

[5] Rump cap

[6] *m. longissimus thoracis et lumborum,* Cube roll

[7] *m. longissimus thoracis et lumborum,* Striploin

[8] Eye rump centre

[9] Silverside

620 ***Table 8*** Grade boundaries calculated from the discriminate functions for the optimally

621 weighted sensory scores (from Table 5)

	Grade boundaries		
Experimental group	2-3	3-4	4-5
All	39.22	59.80	75.75
Australia, grill, medium	38.72	62.86	81.03
France, grill medium	35.98	61.04	80.14
France, grill, rare	36.98	61.66	81.47
Ireland, grill, medium	39.66	60.42	76.94
Ireland, roast, medium	41.56	60.27	72.36
Ireland, BBQ[1] medium	42.33	65.54	81.71
NI[2], grill, medium	40.56	57.51	69.79
NI[2], grill well done	36.23	58.75	75.86
NI[2], roast, medium	37.63	60.35	76.26
NI[2], roast, well done	35.38	57.57	74.60
Poland, grill, medium	42.51	60.77	74.56
Poland, roast, medium	41.88	62.18	77.38
Poland, slow cook, medium	41.65	60.83	75.15

622 2-3=Grade boundaries between unsatisfactory and good-every-day scored beef;

623 3-4=Grade boundaries between good-every-day and better-than-every-day scored beef;

624 4-5=Grade boundaries between better-than-every-day and premium scored beef;

625 [1] BBQ= Korean barbeque

626 [2] NI=Northern Ireland

627