1-P1

MEAT STANDARDS AUSTRALIA, A 'PACCP' BASED BEEF GRADING SCHEME FOR CONSUMERS.
1) THE USE OF CONSUMER SCORES TO SET GRADE STANDARDS

Rod Polkinghorne[1], Ray Watson[2], Mary Porter[3], Alan Gee[4], Justin Scott[5] and John Thompson[5]

[1] Marrinya, Agricultural Enterprises, Bairnsdale, 3875, Victoria, Australia,
[2] Department of Mathematics and Statistics, The University of Melbourne, Parkville, 3052, Victoria, Australia,
[3] 2 Oliver Street, Ashburton, 3147, Victoria, Australia,
[4] Cosign Pty Ltd, 20 Eleventh Avenue, Sawtell, 2452, NSW, Australia,
[5] Co-operative Research Centre for the Cattle and Beef Industries, University of New England, 2351, NSW, Australia.

Background

Meat and Livestock Australia is currently developing a meat grading scheme called Meat Standards Australia (MSA) for use on the Australian domestic market initially, but with the intention to extend its use into Australia's export markets. The MSA approach differs in two important aspects from previous meat grading schemes. Firstly, MSA has focused on providing a guarantee of eating quality to the consumer. To this end, it has implemented a large consumer-testing program, the results of which have been used to set the grading standards. Secondly, MSA has taken a total systems approach to grading meat, in that it aims to control those important factors that impact on meat quality from the production, processing and value-adding sectors of the meat production chain, rather than relying solely on carcass assessment. The MSA grading scheme is based on the principles of Palatability Assurance at Critical Control Points (PACCP). This is a concept taken from the food safety sector. The objective of PACCP is to identify and carefully control those production and processing factors, which have the largest effect on palatability so that it is possible to accurately predict the quality of the final product. From a study of the literature and an ongoing research program, MSA has identified those Critical Control Points (CCPs) that impact on eating quality in Australian beef production and processing systems and has combined these into a workable grading system.

This series of papers describes the development of the MSA grading scheme. This paper outlines the consumer testing system and the development of the grading standards. It was essential to develop a testing protocol that described consumer satisfaction in an accurate and repeatable manner. The grading system was initially based on placing a grade on the whole carcass. More recently it has endeavored to grade individual cuts. Subsequent papers detail the database and the CCPs from the production and processing sectors which MSA have identified as having a significant effect on beef eating quality. The last paper describes the integration of these results into a model to predict eating quality of beef, as a basis for the implementation of a cuts-based grading scheme in the Australian domestic market.

Developing the Tasting Protocol

Eating quality is generally measured using one, or a combination of objective and sensory measurements. The former includes shear force, compression and adhesion measurements, which all measure specific attributes of texture. Sensory evaluation is performed by either a trained, or consumer (i.e. untrained) taste panel. A trained taste panel is skilled in scoring specific attributes of eating quality, independently of other sensory dimensions. It will have a smaller variance, but will often be biased. A consumer panel is unbiased by definition, but with a larger variance. Although the variance of the consumer panel can be reduced by averaging, the bias of a trained panel cannot be removed simply by averaging and needs to be regularly checked against consumer panels. A untrained consumer taste panel, at least in the first instance, was considered the most appropriate method to score beef palatability characteristics as it had the ability to integrate all sensory dimensions. This was predicated on the importance of having a reliable transparent system of testing that would engender consumer confidence in the outcome. The consumer sensory testing protocol used by MSA was based on existing protocols in use by Australian sensory groups and the American Meat Science Association protocols (Anon 1995).

Briefly, the important steps of the protocol are as follows. Consumer groups are recruited from a broad range of socio-economic backgrounds by an independent recruitment company. Generally a donation is given to community groups who provide the required number of consumers to attend a tasting session. These consumers are screened to include only individuals who prefer their steak cooked to medium doneness and eat beef a minimum of once per week. Socio-economic data is recorded for each consumer. For the testing session, venues are selected which are convenient to the consumer groups and the tests are conducted in an informal atmosphere. Consumers are only used once. Four 100 mm lines are shown on the score sheet, which are anchored by the words very tender/very tough for tenderness, very juicy/very dry for juiciness and extremely like/extremely dislike for both flavour and overall acceptability. To allow the sensory scores to be allocated to grades the consumer is asked to tick one of four grades that they consider best describes the quality of the sample: 'unsatisfactory' (no grade), 'good everyday' (3 star), 'better than everyday' (4 star), or 'premium quality' (5 star).

Design: For the grilling sessions, steaks are cooked on a Silex grill and each consumer presented with a total of seven warm steaks served over a 35 minute session. The first sample is a link product designed to allow equilibration of the consumer scores. The data from this sample is subsequently discarded in the analysis. . Following this, a further six samples are presented to each consumer. The tasting design is a latin square, where five samples from each cut are presented in different presentational positions in a minimum of three different sessions to be each tasted by two consumers. This design caters for six products to be tested in any one week. All samples are coded by a unique four digit alpha-numeric code. Once a steak from a specific treatment cell is selected from the inventory a computerised system generates the design number, cooking instruction for time and order and numbered labels for the plates and consumer questionnaires. Each code is independently checked at cooking and consumers asked to cross-check the code on

each plate against their score sheet. Completed score sheets are double entered into a computer and cross-checked prior to forwarding for entry onto the database. Separate protocols have been developed for different cooking methods ie. grilling, roasting, stir frying and slow cooking. Ultimately, as MSA extends into the export markets, separate protocols for cooking procedures specific to those markets will be developed. The protocols have been published (Gee et al. 1998) and are also available in electronic form (web address http://msa.une.edu.au).

Setting the grade standards: Initial studies in which consumers ate up to 32 samples from the same striploin indicated that ten consumer responses per sample provided a sufficiently low standard error to detect a difference of 8 sensory units on a 1 to 100 scale. Whilst tenderness and overall acceptability were highly correlated, juiciness and flavour had lower correlations with the other sensory dimensions. To combine the four sensory dimensions into a single Meat Quality Score (MQ4), weightings were formulated from a discriminant analysis. The weightings were 0.4, 0.1, 0.2 and 0.3 for tenderness, juiciness, flavour and overall acceptability, respectively. If only the first three sensory dimensions were used, the high correlation between tenderness and overall acceptability placed greater weighting on the tenderness element. In practice, the incorporation of the 'overall acceptability' dimension was considered preferable as it added some stability, possibly smoothing out any erratic movement in the other scores and resulted in a slight improvement in the discriminatory efficiency of the vector. The MQ4 scores were then used to calculate the optimum boundaries for the grades assigned by the consumers. A linear discriminant analysis was used to produce a maximal separation of the grade means to the within-grade variance. The first 50 weeks of consumer testing produced MQ4 boundaries of 41.8 between 'unsatisfactory' and '3 star', 64.5 between '3 star' and '4 star' and 81.3 between '4 star' and '5 star'. These boundaries were remarkably consistent between weeks, given that they were determined by different consumer groups. As the 42 boundary represented the pass/fail criteria, this boundary was adjusted up to 48 in order to ensure a high level (95%) of confidence that any product that was passed, was actually a satisfactory product. Outliers among the consumer samples were relatively prevalent, and so it was found necessary to devise a method for minimising their effect. This had to be an automatic and uniformly applied procedure. It was found that for samples of up to ten consumer assessments, the smallest standard error was achieved by using a sample of ten MQ4 assessments, with the top two and the bottom two trimmed (CMQ4).

Sensory design and demographic effects on sensory scores: Considerable attention was paid to balancing the order, spread of sessions and other test details. This was done to minimise the influence of any session, taster, group, socio-economic, or presentational order effect on any one sample. It was considered preferable to use the CMQ4 scores in the unadjusted form. Analysis to support this was undertaken on MQ4 scores from the first 19 weeks of testing. The analysis was conducted within weeks and examined design (session, taster/session, order and animal) and socio-economic factors (income, postcode, age, gender, occupation, frequency of eating meat, number of adults and children in the household, beef appreciation and preferred degree of doneness) on approximately 24,000 MQ4 scores. The results showed that animal and taster within session, were significant ($P<0.0001$) in all 19 weeks. Even though taster was a highly significant source of variation, there were still highly significant differences between animals. Session effects were only significant in a small proportion of the testing weeks. Order of presentation was significant ($P<0.05$) in less than half of the weeks tested. The order effect appeared to be associated with deviations in the scores given to the link steak that was presented first, which supported the use of a link steak in the tasting protocol. Generally, the frequency with which socio-economic factors were significant was low (<4/19 weeks) and when significant, the direction and magnitude of the effects variable. This suggested that the sensory scores were largely unaffected by the above socio-economic factors, with no consistent source of bias found.

The MSA Database: Australia has a wide diversity of production systems over a range of environments that supply both the domestic and export markets. To underpin the development of a meat grading scheme which is relevant to this wide diversity of production systems, a database which collates information on production, processing systems and value-adding sectors of the meat production chain has been developed. Meat samples have been collected from a variety of commercial and experimental cattle and the samples tasted using the above consumer testing protocols. More details of the categorisation of samples in the database is given by Thompson *et al.* (1999). In March 1999 the database comprised a total of over 12,700 records, which had been summarised from 127,000 consumer responses. These samples comprised approximately 6,000 sensory records for the striploin, with the balance distributed between ten other cuts. Five different cooking techniques including, grilling, roasting, stir-fry, slow cook and corning have been tested to date.

Conclusion

As part of the development of a consumer focused meat grading scheme MSA has developed a detailed protocol for consumer testing of meat samples and used this data to set the grade standards. The procedure was initially set up for grilling but has subsequently been extended to include a variety of cooking techniques, including roasting, stir frying, slow cooking and corning. The sensory data have been combined with relevant production, processing and value-adding data to provide a unique database that has underpinned the development of the MSA scheme.

References

Anon. (1995). Amercian Meat Science Association. Chicago, Illinois.

Gee, A., Coffey, D., Porter, M., and Polkinghorne, R. (1998). Design and protocol for steak grilling trials. MSA, Sydney.

Thompson, J., Polkinghorne, R., Hearnshaw, H. and Ferguson, D. (1999). Proc. 45th ICoMST, Yokohama, Japan.