

International perspective: characterisation of United States Department of Agriculture and Meat Standards Australia systems for assessing beef quality

G. C. Smith^{A,B}, J. D. Tatum^A and K. E. Belk^A

^AMeat Science Program, Department of Animal Sciences, Colorado State University, Fort Collins, CO 80523-1171, USA.

^BCorresponding author. Email: gary.smith@colostate.edu

Abstract. The intent, in this manuscript, is to characterise the United States Department of Agriculture (USDA) and Meat Standards Australia (MSA) systems for assessing beef quality and to describe the research evidence that supports the principles involved in grade application. USDA beef quality grading standards rely on carcass-trait-only assessments of approximate age of the animal at harvest and amount of intramuscular fat (as marbling) inside the muscles. USDA beef quality grading started 82 years ago. Then, as now, because no traceability system was in place, each animal's history (exact age, feeding regimen, management practices, etc.) was incomplete; those who assigned quality grades used indicators of age (physiological maturity) and plane of nutrition (amount of marbling), and they do so still. Since 1926, research studies have identified a multitude of palatability-determining live-animal factors (e.g. genetics, use of hormonal growth promotants, high-energy diet finishing) and carcass-treatment factors (e.g. electrical stimulation, tenderstretch carcass suspension, postmortem aging) that cannot be incorporated into a carcass-trait-only quality assessment system. The USA beef industry has depended on development of more than 100 beef brands – some using palatability assurance critical control point plans, total quality management (TQM) philosophies, USDA certification and process verification programs, or combinations of live-animal factors, carcass-treatment factors and carcass-trait constraints – to further differentiate fresh beef products. The MSA grading system is a TQM grading approach that incorporates animal-specific traits (e.g. genetics, sex, age), control of certain pre-harvest and post-harvest processes in the beef chain, cut-specific quality differences and consumer preferences, into a beef pricing system. A unique aspect of the MSA grading system is that the grades are assigned to cuts or muscles, not carcasses; cuts or muscles from the same carcass are assigned individual (and in many cases, different) grades that reflect differences in expected eating quality performance among the various cuts of beef further adjusted to reflect the influence of cut or muscle aging and alternative cooking methods. The MSA grading system is still being modified and refined (using results of an extensive, ongoing consumer testing program), but it represents the best existing example of a TQM grading approach for improving beef quality and palatability. Research studies have shown that the accuracy of palatability-level prediction by use of the two systems – USDA quality grades for US customers and consumers and MSA grades for Australian customers and consumers – is sufficient to justify their continued use for beef quality assessment.

Introduction

The quality of a fresh (raw, uncooked) beef steak or roast, as discerned by the 'customer' (the person who purchases it), is determined by appearance characteristics (e.g. ratios of muscle, fat and bone; amount of marbling; colour of muscle and fat; freedom from defects). The quality of a cooked beef steak or roast, as perceived by the 'consumer' (the person who eats it), is decided by palatability characteristics (e.g. flavour, juiciness, tenderness). Research with US consumers who were characterised as 'frequent beef users' revealed that their overall perceptions of the 'taste' of beef are associated with those of the three primary sensory attributes – flavour, juiciness and tenderness (Neely *et al.* 1998).

The palatability of cooked beef is determined by the aggregated effects of differences in flavour, juiciness and tenderness experienced by human subjects when they eat it. Flavour desirability differences are related to the proteins in muscle, the types and amounts of fat (marbling) in the muscle and the types and amounts of protein and fat degradation products

generated during postmortem aging. The juiciness of cooked beef is determined by the amounts of intramuscular moisture and fat (marbling) that remain in the muscle after cooking. The tenderness of cooked beef muscle is determined by the amounts of connective tissue left unsolubilised after cooking, the amounts of intramuscular moisture and fat (marbling) remaining after cooking and the structural integrity of the sarcomeres, myofibrils and muscle fibres at the time of consumption.

Delivering a quality eating experience is essential to the continued success of the beef industry's efforts to build consumer demand for beef products (NCBA 2001). Tatum (2006) said 'Consumer survey results suggest that eating quality (defined by most consumers simply as 'taste') is a primary driver of food purchase decisions, across a variety of product categories'. From 1983 through 2002, US supermarket shoppers identified 'taste' as the most important factor in food selection (Food Marketing Institute 2002). Quinn (1999) said:

'If you can get away from the straitjacket of regarding meat as a commodity, you will concentrate on how you can best satisfy the needs of consumers. The end product you sell is not meat. . . it is taste. Consumers won't pay more for food that satisfies their nutritional requirements or fits their food safety requirements. People will pay more for greater satisfaction. . . and, taste is their measure of satisfaction in food'.

Shook *et al.* (2008) reported that domestic merchandisers of US beef (purveyors, retailers, restaurateurs) identified inadequate flavour and inadequate tenderness among the 'top five quality challenges' for the beef industry. The intent, in the present manuscript, is to characterise the United States Department of Agriculture (USDA) and Meat Standards Australia (MSA) systems for assessing beef quality and to describe the research evidence that supports the principles involved in grade application.

USDA beef quality grading system

Official United States Standards For Grades Of Carcass Beef (USDA 1997) define 'quality grade' as 'the palatability-indicating characteristics of the lean', and state that:

'for steer, heifer and cow beef, quality of the lean is evaluated by considering its marbling and firmness as observed in a cut surface in relation to carcass evidences of maturity. . . Maturity is determined by evaluating the size, shape and ossification of the bones and cartilages – especially the split chine bones – and the colour and texture of the lean flesh'.

There are eight quality grades: prime, choice, select, standard, commercial, utility, cutter and canner (USDA 1997). Across the entire spectrum (prime to canner), USDA quality grades are intended to predict flavour, juiciness and tenderness of all major muscle cuts from the carcass. For cuts from youthful, grain-finished steers and heifers, the top four USDA quality grades (prime to standard) predict well the flavour and juiciness of all major muscle cuts, but best predict tenderness of the middle meats (rib and loin) because cuts from these primals are usually cooked with dry heat (e.g. broiling, grilling, roasting). The top four USDA quality grades are recognised as less useful predictors of tenderness of the end meats (chuck and round) because cuts from those primals are most often cooked with moist heat (e.g. braising) to soften the connective tissue.

Smith *et al.* (2005) said that beef carcass maturity is determined by evaluating the: (i) size, shape and ossification of the bones and cartilages, especially the split chine bones; and (ii) colour and texture of the lean flesh. Determination of the 'maturity group' among beef carcasses (A, B, C, D and E, which are, nominally, from animals harvested at 9–30, 30–42, 42–72, 72–96 and >96 months of age, respectively) and of the 'position within a maturity group' (e.g. A⁰⁰, A¹⁰, A²⁰ . . . A¹⁰⁰) is determined by evaluating the: (i) split chine bones of the vertebral column – usually, ossification (conversion of cartilage to bone) occurs earliest in the posterior vertebrae (sacral), later in the middle vertebrae (lumbar) and latest in the anterior vertebrae (thoracic); (ii) size and shape of the rib bones – usually, rib bones grow wider

and flatter, and have less blood in their surfaces in more mature carcasses; (iii) colour of muscle – usually, colour becomes progressively darker red, in progressively more mature carcasses; and (iv) texture of muscle – usually, texture becomes progressively coarser, in progressively more mature carcasses (Smith *et al.* 2005).

Smith *et al.* (2005) described the nine 'degrees' (sometimes called 'scores') of marbling (the intermingling of fat deposits in muscle) in beef quality grading standards as (from highest to lowest): abundant (AB), moderately abundant (MA), slightly abundant (SA), moderate (MD), modest (MT), small (SM), slight (SL), traces (TR) and practically devoid (PD). Marbling is assigned to 'degrees' based on the: (i) amount of intramuscular fat – the percentage of intramuscular fat is 1.77 for the lowest marbling degree (PD), increases 1.24 per degree, and is 11.69 for the highest marbling degree (AB) (Savell *et al.* 1986; Lunt *et al.* 1989); (ii) size of individual deposits – the average deposit is tiny in the lowest marbling degree (PD), increases in size as degree increases, and is large (but not excessive or in streaks) in the highest marbling degree (AB); and (iii) dispersion and distribution of deposits – the more perfectly the deposits are dispersed throughout the entire surface of the cut muscle, the higher the degree (when amount and size are held constant). Marbling dispersion and distribution is considered 'most desirable' when some of it occurs in every bite-sized portion of the cut-muscle surface; if intramuscular fat occurs in poorly distributed chunks or streaks, with large void areas in many parts of the surface, the 'amount' may be chemically high but the 'degree' could be intermediate to low (Smith *et al.* 2005).

Relationships of marbling and maturity in determining the quality grade (USDA 1997) are as follows:

- (1) Among 'A' maturity carcasses, those with AB, MA or SA are graded prime, those with MD, MT or SM are graded choice, those with SL are graded select, and those with TR or PD are graded standard.
- (2) Among 'B' maturity carcasses, those with AB or MA are graded prime, those with SA are graded prime or choice, those with MD or MT are graded choice, and those with SM, SL, TR or PD are graded standard.
- (3) Among 'C', 'D' and 'E' maturity carcasses, those with AB, MA or SA marbling are graded commercial, those with MD, MT or SM marbling are graded commercial or utility, those with SL or TR marbling are graded utility or cutter, and those with PD marbling are graded utility, cutter or canner.

Use of *longissimus dorsi* as an indicator muscle

Relative to whether or not it is reasonable to predict the palatability of other major muscles by use of traits that can be assessed by viewing the *longissimus dorsi* (LD) at the 12th–13th rib interface (i.e. colour, texture, marbling) or by instrument or sensory panel evaluations, we offer the following supporting evidence:

- (1) Smith *et al.* (1980) reported that relationships between overall palatability and USDA quality grade were such that for steaks from carcasses grading prime, upper two-thirds choice, lower one-third choice, select and standard, the percentages of undesirable eating experiences for broiled loin

- steaks were 4, 5, 15, 19 and 52%, respectively, and the percentages of undesirable eating experiences for broiled top round steaks were 33, 46, 56, 61 and 64%, respectively.
- (2) Smith *et al.* (1978) reported that measures of tenderness of the broiled LD-rib are significantly ($P < 0.01$) related to the tenderness (shear force values and sensory panel tenderness ratings) of muscles from broiled or roasted cuts from every primal cut region (chuck, rib, loin, round) of the beef carcass. The harmonic mean correlation for LD-rib overall tenderness rating and shear force values for 14 muscles was -0.48 .
 - (3) Correlations among LD sensory panel tenderness ratings or shear force values and those of other muscles in the beef carcass have been reported to be: (i) 0.20 – 0.32 (Joseph and Connolly 1979) based on sensory panel tenderness ratings, and (ii) 0.40 (Slanger *et al.* 1985) and 0.26 – 0.43 (Shackelford *et al.* 1995) based on shear force values.
 - (4) Researchers at the US Meat Animal Research Center (Clay Center, NE, USA) originally reported (Shackelford *et al.* 1995) that: (i) shear force values did not accurately reflect differences among muscles in overall tenderness, and (ii) shear force value of the LD was not highly related to the shear force values of other muscles in the carcass. However, once they decided that USDA quality grades should be replaced by use of a slice shear force value of the LD shortly after carcass chilling, Wheeler *et al.* (2000) reported correlation coefficients of -0.31 to -0.58 between early postmortem, day 2, LD slice shear force values and day 14 sensory panel tenderness ratings of four major muscles and concluded that early postmortem LD slice shear force value could be used to classify top sirloins, top rounds and bottom rounds for tenderness. Then, Rhee *et al.* (2004) reported statistically significant correlations ($P < 0.05$ or lower) between sensory panel tenderness ratings of the LD with 7 of 10 ($r = 0.38$ to 0.73) other major muscles and between shear force values of the LD with 7 of 10 ($r = 0.38$ to 0.73) other muscles, and that if carcasses are sorted, based on shear force value of the LD, into 'tough' v. 'tender' groups, chuck muscles differed in average sensory panel tenderness rating by 0.4 units, rib and loin muscles differed in average sensory panel tenderness rating by 1.0 unit, and round muscles differed in average sensory panel tenderness rating by 0.6 units.
 - (5) Belew *et al.* (2003) reported that of 800 possible combinations of comparisons of shear force value among 40 muscles: (i) 166 produced correlations ranging from $r = 0.70$ to $r = 0.99$; (ii) 314 produced correlations ranging from $r = 0.50$ to $r = 0.69$; (iii) 173 of the correlations ranged from $r = 0.30$ to $r = 0.49$; and (iv) tougher muscles (such as *flexor digitalis superficialis*), in many cases, had negative correlations (there were 26 of those) with other muscles.

The results of these studies support the contention that LD tenderness is an imperfect but very useful predictor of the tenderness of most of the major muscles in the beef carcass. So, to the extent that we can predict the tenderness of the LD by assessing carcass and LD surface characteristics, we can indirectly predict the tenderness of most of the other muscles. And, inasmuch as marbling deposition is correlated with flavour

and juiciness of beef, plus the fact that the amount of marbling in the LD is correlated with the amount of marbling in the other major muscles of the carcass, the amount of marbling in the LD is an imperfect but very useful predictor of flavour and juiciness in most of the major muscles in the beef carcass.

Palatability prediction using USDA quality grades

The philosophy used in the grading of any agricultural commodity involves sorting of the products into groups – usually in some hierarchical fashion – that differ in utility, desirability and value. Grading means ranking, classifying or categorising. USDA quality grades were never intended to provide point estimates for expected beef palatability. Neither quality grades nor palatability ratings are perfectly assigned because both are subjective estimates. When statistically analysing data from studies of quality grades \times palatability assessments, correlation and regression analyses are appropriate only if interest is in absolute ranking on an individual carcass or cut basis and only if the experimental design assures that there are reasonable numbers of carcasses across the entire quality grade spectrum.

To expect near-perfect linearity when correlating data in which quality grade is assigned to the nearest 10% of a grade (e.g. prime⁴⁰, choice²⁰) and palatability ratings are assigned to the nearest one-hundredth of a rating (e.g. 7.28, 4.39) is irrational. Alternatively, the authors of this manuscript believe that it is reasonable (and appropriate) to non-parametrically evaluate the effectiveness of the USDA beef carcass quality grading system by determining its ability to successfully categorise carcasses according to the relative desirability of the beef in that category (i.e. within a quality grade, or portion of a grade), as compared with the desirability of the beef from carcasses in other quality grades or portions of a grade.

In some studies, for example that of Smith *et al.* (1987), the experimental design is such that measures of both ranking ability and categorisation are appropriate. In that study, three experts assigned quality grades, 1005 carcasses (with quality grades of prime¹⁰⁰ to canner¹⁰) were selected from eight packing plants in six plants, cuts were stored for 10–14 days at $1 \pm 1^\circ\text{C}$ postmortem, and cooked cuts were evaluated sensorily by 40 highly trained sensory panelists (10 members from each of three universities plus 10 members from the USDA Meat Science Research Laboratory). Correlation and regression analyses revealed that USDA quality grades (across the entire eight-grade range) accounted for 40–47% of the observed variation in overall palatability of dry heat-cooked loin and top round steaks, and 25–33% of the variation in shear force for loin, top round, bottom round and eye-of-round steaks. Mean overall palatability ratings for loin steaks were 6.02 (prime), 5.71 (choice), 5.33 (select), 4.63 (standard), 4.93 (commercial), 3.99 (utility), 3.39 (cutter) and 2.84 (canner) with $P < 0.05$ significance of prime > choice > select > commercial > standard > canner = cutter. Non-parametric analyses revealed that the 'percentage incidence of loin steaks rated very desirable' values in a composite of all sensory panel ratings and shear force values were 63.6 (prime), 49.4 (choice), 35.3 (select), 20.3 (standard), 30.3 (commercial), 11.1 (utility), 3.2 (cutter) and 0.0 (canner).

In the USA, the USDA beef quality grades actually used in commerce are those for carcasses from cattle less than 42 months

of age (nominally) – prime, choice and select (named ‘good’ before 1987), with standard essentially never sold as such at retail because of the negative connotation of the term. Since 1978, when Certified Angus Beef carved out a niche for beef from A-maturity carcasses with MT and MD marbling, such carcasses (actually, in most ‘premium choice’ programs, those of both A and B maturity with MT and MD marbling) are recognised by the commercial industry as representing a quality grade even though such beef is almost always sold under a brand designation. For research purposes, quality grade descriptors used can be full-width grades (prime, choice, select, standard) or segments within grades (high, average or low prime; high, average or low choice; upper two-thirds choice or lower one-third choice; high, average or low select; upper half select or lower half select; upper half standard or lower half standard).

Briskey and Bray (1964) reviewed the then-available research results and concluded that those studies suggested that mean palatability values between USDA quality grades are not greatly different, but that the risk of having an ‘odd ball’ in the higher quality grades is markedly reduced in comparison with the lower quality grades. Jeremiah *et al.* (1970) reviewed the then-existent literature and concluded that previous studies indicated that USDA quality grades are relatively inaccurate indicators of flavour, juiciness and tenderness of cooked beef. Smith and Carpenter (1974) reviewed the then-available research evidence and concluded that existing data suggested that USDA quality grades have a low relationship to cooked beef flavour, and low to moderate relationships to juiciness and tenderness of cooked beef. Campion *et al.* (1975) reported correlation coefficients between quality grade and sensory panel ratings for flavour, juiciness, tenderness and overall palatability of 0.20, 0.27, 0.21 and 0.25, respectively (all statistically significant at the 0.01 level), but considered their value minimal for predicting eating characteristics of young carcass beef.

Smith *et al.* (1982) concluded that carcasses of A maturity produced ‘very desirable’ loin steaks 1.2, 1.5 and 8.0 times, ‘acceptable’ loin steaks 1.0, 1.1 and 1.6 times, ‘very desirable’ round steaks 3.0, 3.3 and 5.7 times, and ‘acceptable’ round steaks 1.4, 1.5 and 3.1 times as often as did carcasses of B, C or E maturity, respectively. Overall, in comparison to carcasses of B, C or E maturity, carcasses of A maturity produced broiled steaks that: (i) had higher ($P < 0.05$) palatability ratings in 62–86% of comparisons; (ii) were decidedly less variable; (iii) were more likely to be assigned high (≥ 6.00) and less likely to be assigned low (≤ 2.99) sensory panel ratings; and (iv) were more likely to have low (≤ 3.63 kg) and less likely to have high (≥ 6.35 kg) shear force values. Position within the A or A + B maturity groups explained 0–4% (loin steaks) and 10–18% (round steaks) of the observed variation in overall palatability ratings and shear force values.

Smith *et al.* (1984) concluded that:

- (1) Differences in marbling explained ~33% (loin steaks) and 7% (top round steaks) of the variation in overall palatability ratings for cuts from A, B, C or A + B maturity carcasses.
- (2) Among marbling groups for carcasses of A + B maturity, the percentage of steaks with a composite of sensory panel ratings of ≥ 6.00 and a shear force value of ≤ 3.63 kg was 66, 59, 56, 48, 41, 33, 21 and 15% for loin steaks and 18, 19, 5, 13, 8, 12, 5 and 8% for round steaks from carcasses with

MA, SA, MD, MT, SM, SL, TR and PD marbling, respectively.

- (3) As marbling increased from PD to MA, loin steaks were more palatable about two-thirds of the time ($P < 0.05$), round steaks were more palatable about one-eighth of the time ($P < 0.05$), and loin steaks were more likely to be assigned high (≥ 6.00) panel ratings and to have low (≤ 3.63 kg) shear values.
- (4) Coefficients of determination for USDA marbling score (by scores and percentages with scores) for flavour, juiciness, tenderness and overall palatability in A + B carcasses were 27, 20, 26 and 33%, respectively, for loin steaks and 4, 20, 7 and 7% for top round steaks, respectively.

Smith *et al.* (1983) said:

- (1) USDA quality grade is related to flavour of beef because quality grade measures the extent to which flavour and aroma compounds are likely to be present in high v. low concentrations in the meat.
- (2) Carcasses from older animals, leaner animals and animals not fed large amounts of grain – animals for which there is high likelihood that they would produce meat that is less desirable in flavour – are assigned low USDA quality grades.
- (3) Carcasses from young animals, fatter animals and animals fed large quantities of grain – animals for which there is high likelihood that they would produce meat that is ‘beefy’ and more desirable in flavour – are assigned high USDA quality grades.

Smith *et al.* (1987) reported mean sensory panel ratings for: (i) flavour, that ranked loin steaks as prime > choice > select > standard and top round steaks as prime > choice = select = standard ($P < 0.05$); (ii) juiciness, that ranked loin steaks as prime > choice > select = standard and top round steaks as prime > choice = select = standard ($P < 0.05$); (iii) tenderness, that ranked loin steaks as prime > choice > select > standard and top round steaks as prime > choice = select = standard ($P < 0.05$); and (iv) overall palatability, that ranked loin steaks as prime > choice > select > standard and top round steaks as prime > choice = select = standard ($P < 0.05$). Overall, among prime through standard carcasses, quality grade predicted flavour, tenderness and overall palatability of loin steaks with 30–38% accuracy, but could explain no more than 8% of the variation in panel ratings and shear force values of top round steaks.

The National Consumer Retail Beef Study (NCRBS) was an industry-wide endeavour with support from government, producer, feeder, packer and retailer sectors of the US beef industry designed to: (i) determine the role of USDA beef quality grades and taste appeal (first in College Station, TX and Houston, TX, then in Philadelphia, PA, Kansas City, MO and San Francisco, CA); and (ii) identify the interplay among taste, price and leanness (plate waste) in determining consumer acceptability (in Philadelphia, PA and San Francisco, CA) of retail cuts from the four major primal cuts of beef (Branson *et al.* 1984).

Branson *et al.* (1986) reported that in the College Station, TX and Houston, TX portion of the NCRBS: (i) 10 expert laboratory taste panellists provided 2700 product ratings; (ii) 200 consumer laboratory panellists made 4000 observations; and (iii) 180

households in Houston, TX provided 2800 product ratings of top loin steaks from the striploins of 300 beef carcasses from cattle differing in sex (bullocks – young intact males – full-width of select), cattle diet (steers, short-fed, full-width of select) and quality grade (steers or heifers, long-fed, of low prime, high choice, average choice, low choice, upper half select, lower half select and upper half standard). Branson *et al.* (1984, 1986) reported overall palatability ratings from:

- (1) The expert laboratory panel in College Station, TX, which ranked loin steaks as low prime > high choice > average choice > low choice > upper half select = upper half standard > lower half select = bullocks full-width of select = steers full width of select ($P < 0.05$).
- (2) The consumer laboratory panel in College Station, TX, which ranked loin steaks as low prime > high choice > average choice = low choice > upper half select = upper half standard > lower half select = bullocks full-width of select = steers full-width of select ($P < 0.05$).
- (3) The household panel in Houston, TX, which ranked loin steaks as low prime = high choice = average choice > low choice = upper half select = lower half select = upper half standard > bullocks full width of select = steers full width of select ($P < 0.05$).

Savell *et al.* (1987) reported that:

- (1) Eight expert laboratory taste panellists provided 16 800 sensory ratings (SR) and shear force values were obtained for 700 top loin steaks; analyses of those results revealed that low prime was superior to low choice in flavour, juiciness and tenderness (both SR and shear force value) and that low choice was superior to lower half select in flavour, juiciness and tenderness (shear force value but not SR), but that upper half select, lower half select and upper half standard did not differ in flavour, juiciness or tenderness (both SR and shear force value).
- (2) Composited overall desirability ratings from households in PA, MO and CA (6408 responses) revealed that the percentages of ratings that were intermediate and lower for top loin steaks of low prime, high choice, average choice, low choice, upper half select, lower half select and upper half standard were 18, 21, 24, 24, 26, 32 and 32%, respectively. Branson *et al.* (1986) combined the 8018 responses from household panellists in all four states represented in the NCRBS and reported that overall desirability ratings were low prime > high choice > average choice = low choice > upper half select = upper half standard > lower half select ($P < 0.05$).

Consumers in San Francisco, CA and Philadelphia, PA were asked to purchase, in simulated retail markets, beef retail cuts of different grades (choice or select) or with different amounts of external fat (regular trim = 13 mm, extra trim = 8 mm or super trim = 0 mm), all priced at parity or premium (parity plus 10%) prices (Savell *et al.* 1989). Consumers in Philadelphia purchased significantly more extra trim and super trim steaks and roasts than regular trim. At the time of purchase, consumers in both cities could not detect the visual differences in choice *v.* select, but upon eating them found that choice cuts were better tasting – but also fatter – and that select cuts were leaner – but had problems with

taste and texture. Savell *et al.* (1989) concluded from this study that both choice and select were rated high for consumer acceptance but for different reasons – taste for choice, leanness for select. Branson *et al.* (1986) reported results of a retail store pilot test (four supermarkets; 12 weeks/store) that supported the concept that a segmented market existed for beef, and that name-brand lean beef (i.e. branded beef, of select quality) should be introduced in retail food chains.

Relationships between overall palatability and USDA quality grade in the Smith *et al.* (1980) study were such that for steaks from carcasses grading prime, upper two-thirds choice, lower one-third choice, select and standard, respectively: (i) the percentages of undesirable eating experiences for broiled loin steaks were 4, 5, 15, 19 and 52%; and (ii) the percentages of undesirable eating experiences for broiled top round steaks were 33, 46, 56, 61 and 64%. Jones and Tatum (1991a, 1991b, 1994) determined that the percentages of steaks that were 'tender' were 52, 41, 45, 34 and 19% and those that were 'tough' were 3, 6, 10, 9 and 23% for carcasses of \geq average choice, low choice, high select, average select and \leq low select, respectively. Huffhines *et al.* (1992a, 1992b) reported that average choice > low choice > upper half select = lower half select > upper half standard ($P < 0.05$) in tenderness (measured by shear force), overall palatability and percentage 'desirable or higher' in overall eating satisfaction; percentages 'desirable or higher' for average choice, low choice, upper half select, lower half select and upper half standard were 58, 40, 35, 32 and 12%, respectively. George *et al.* (1999) studied the tenderness of beef available at supermarkets throughout the USA and reported that: (i) the odds of having a tough steak from carcasses of prime, upper two-thirds choice, lower one-third choice and select ($P < 0.05$) were 0 (none), 1 in 19 (5.3%), 1 in 9 (11.2%) and 1 in 6 (17.8%) for top loin steaks, respectively and 0 (none), 1 in 6 (18.0%), 1 in 5 (20.2%) and 1 in 4 (28.3%) for top sirloin steaks, respectively; and (ii) as marbling score and USDA quality grade increased, sensory panel ratings for flavour, juiciness, freedom from connective tissue, myofibrillar tenderness, overall tenderness and overall palatability increased ($P < 0.05$).

Smith (2005) composited data from the studies of Smith *et al.* (1980, 1983, 1984, 1987), Branson *et al.* (1984, 1986), Savell *et al.* (1987), Jones and Tatum (1991a, 1991b, 1994), Huffhines *et al.* (1992a, 1992b, 1993) and George *et al.* (1999) and concluded that the odds of having an unpleasant eating experience are 1 in 33 (3%) if a middle-meat steak comes from a prime carcass, as compared with 1 in 10 (10%), 1 in 6 (16%), 1 in 4 (27%) or 1 in 2 (50%) if a middle-meat steak comes from a carcass of upper two-thirds choice, lower one-third choice, select or standard grades, respectively. Platter *et al.* (2003b) reported that: (i) marbling score displayed a significant relationship to acceptance of steaks by consumers; and (ii) the shape of the predicted probability curve for steak acceptance was approximately linear over the entire range of marbling scores (TR^{67} to SA^{97}), suggesting that the likelihood of consumer acceptance of steaks increases ~10% for each full marbling score increase between SL and SA. Platter *et al.* (2005) used an experimental auction technique to determine consumer purchasing behaviour and willingness to pay for beef strip loin steaks and determined that: (i) prime steaks received a US\$2.47/kg premium and upper two-thirds choice received a

US\$0.89/kg premium over the mean bid price for select steaks; and (ii) mean bid prices for steaks decreased by US\$1.02/kg for each 1 kg increase in shear force value, with 'very tender' steaks receiving bid premiums of US\$0.83/kg, US\$2.09/kg and US\$2.55/kg compared with 'slightly tender,' slightly tough' and 'very tough' steaks.

Are there means other than USDA quality grading for beef palatability prediction used in the USA?

The USDA beef quality grading system has served the USA well in both domestic and international commerce but most assuredly has limitations for a US beef industry that is in transition. Sweeping structural changes are transforming the US beef industry – from a commodity-oriented industry, dominated by small, independent producers, to consumer-driven production systems in which firms and producers can manage product attributes, from farm to table, to generate value-added kinds of beef (Tatum 2006). The system currently used to ensure the quality of US beef involves mass inspection (i.e. USDA quality grading) at the end of the production process. USDA quality grading generally categorises beef according to expected palatability, but product value is lost due to the imprecision of grading methodology and because some inferior beef has been produced and now must be sold at a discount (Tatum 2006). Deming (1986) said 'Cease dependence on inspection to achieve quality; eliminate the need for mass inspection by building quality into the product'.

Tatum (2006) reported that in the early 1990s, US beef producers began to embrace the principles of total quality management (TQM) and process control developed by 'quality guru' W. Edwards Deming, who is credited with transforming post-World War II Japan into a leader in international business and industry and was viewed by many as the father of the modern quality revolution that began reshaping US industry in the 1980s. According to Tatum (2006), attention was given to: (i) reducing costs throughout the beef chain (Lambert 1991); (ii) identifying product defects and quality shortfalls (Smith *et al.* 1992); (iii) learning more about the preferences, needs and expectations of beef consumers (NLSMB 1995); (iv) linking segments of the beef chain to facilitate application of TQM principles and implementation of process control (NCA 1993); and (v) improving demand by identifying that as the beef industry's single most important goal and by making quality an industry-wide priority (Industry-Wide Long Range Plan Task Force 1993).

Tatum (2006) described beef 'alliances' and 'supply chains' as exhibiting a variety of distinctive features:

- (1) Nearly all focus on improving quality and adding value to cattle and beef products.
- (2) Most feature value-based marketing agreements to provide economic incentives for production of cattle and beef carcasses that meet program specifications.
- (3) Most are at least partially integrated (or vertically coordinated) with producers retaining some share of ownership through much or all of the beef value chain.
- (4) Two essential features of these coordinated business structures are that of providing cattle producers with an opportunity to capture a share of the product value that is added by the processing and marketing sectors and that of

enabling producer participants to receive market signals directly from consumers.

- (5) Many include breed specifications (based on genotype or phenotype) for program cattle in an effort to improve consistency of genetic inputs into the system.
- (6) Many include information systems that facilitate data acquisition, information sharing among program participants and measurement of system performance.
- (7) Some feature branded beef products designed to target consumer preferences for specific product attributes.
- (8) Many involve source verification and process verification, and some utilise third-party verification to instil consumer confidence in product quality, consistency or safety.

Morgan (1992) was the first person to propose the use of a TQM approach for improving beef palatability; at the Strategy Workshop for the National Beef Quality Audit in 1991 (Smith *et al.* 1992), Morgan coined the term 'palatability assurance critical control points' (PACCP). Shortly thereafter, the implementation of PACCP systems to improve beef tenderness was advocated as a key action point in the National Beef Tenderness Plan (NCA 1994). Tatum (2006) believes that rather than continued, singular focus on measurement and categorisation of beef quality differences at the end of the production process, an alternative and more comprehensive approach – consistent with TQM philosophy – is to focus on understanding the causes of product variability and then work to improve the production process by measuring and monitoring critical variables known to affect variability in finished products. Colorado State University scientists used PACCP decision trees in studies by Sherbeck *et al.* (1995, 1996) and Tatum *et al.* (1997, 1998, 1999, 2000). ConAgra Beef, working with Colorado State University scientists (Anon. 2000), used PACCP principles to develop and implement the 'chain of tenderness'.

The PACCP concept has been used in Australia by Polkinghorne (1996, 1998, 2003, 2006), Polkinghorne *et al.* (1999, 2008a, 2008b, 2008c), R. Polkinghorne, J. Thompson and R. Watson (unpubl. data) and Polkinghorne and George (1998a, 1998b) for that country's 'eating quality assurance scheme' and for the grading system used presently by MSA (Ferguson *et al.* 1999; Watson 2000; Watson *et al.* 2008a, 2008b). Thompson *et al.* (1999) and Thompson (2002a, 2002b) have described the TQM approach taken by MSA for managing beef tenderness using critical control points (CCP) from the production, pre-slaughter, processing and value-adding sectors of the beef supply chain; among CCP in MSA are breed, growth paths, pH and temperature window, alternative carcass suspension, aging, and – for the cut-based grading system – method of cooking. Valin (2000) reported that PACCP-like systems for improving beef palatability were being used in the UK ('blue print' system) and in France ('label rouge').

Colorado State University studied a PACCP model (Tatum *et al.* 1997, 1998, 1999) for improving beef tenderness and reported that the two interventions that were most effective were: (i) selecting the top 25% of sires based on progeny group means for 14-day top-loin steak shear force values; and (ii) high-voltage electrical stimulation followed by a postmortem aging period of 14 or 21 days. Use of these two intervention strategies reduced the expected rate of non-conformance from 54% (worst-case scenario) or 29% (normal-case scenario) to

5% (1 in 20) for top sirloin steaks and from 64% (worst-case scenario) or 28% (normal-case scenario) to 1% (1 in 100) for top-loin steaks. Smith (2002) described use of a PACCP-like approach to assure tenderness of beef from Brahman-cross cattle in the Nolan Ryan Tender Aged Beef (NRTAB) program; constraints in that program include: (i) assuring that cattle are no more than 50% *Bos indicus* genetics; (ii) electrically stimulating (with high voltage) carcasses; (iii) using gentle (36–48 h) carcass chilling; (iv) assuring that cattle are no more mature than 30 months (in 'A' maturity) at harvest; (v) requiring 'slight' or higher marbling; (vi) requiring that carcasses meet muscle colour constraints using BeefCAM (Research Management Corporation, Fort Collins, CO); and (vii) aging all beef primal cuts for at least 14 days. Bradbury (2003) reported that 'At the point US\$4 million of NRTAB had been sold, the company (Beefmaster Cattlemen's, LP) had refunded US\$1100 to consumers asking for their money back (16% due to product toughness)'. Savell (2003) and Smith (2003a) attribute success of guaranteeing tenderness in the NRTAB program to the combination of use of electrical stimulation, aging constraints and BeefCAM technology.

Dolezal (2005) described elements of Rancher's Registry (branded beef programs of Cargill Meat Solutions) as follows:

- (1) Rancher's Registry beef is sold under five supermarket brand names in 2101 US and Canadian supermarkets.
- (2) The target is '90% tender middle meats', so a suspension-stimulation process (snip and shock), vision grading (computer vision system) and automated tenderness sampling (shear force values; every lot, every shift, every day, in four testing laboratories in North America) are used to help assure tenderness compliance.
- (3) Cattle are sourced from five company-owned feedlots, three beef supply chain alliances and from purchases by 45 cattle buyers in the USA and Canada.
- (4) Because '90% tender middle meats' is the target, Cargill Meat Solutions disqualifies 17% of the pens of cattle offered by feedlots and all of the cattle from 8.8% of feedlots because buyers believe, or tenderness sampling proves, they will not hit the target.

Tatum (2006) cited experimental market research that establishes a direct link between the eating qualities (flavour and tenderness) of beef and actual purchase behaviour of beef consumers; included were:

- (1) Boleman *et al.* (1997) reported that when consumers were aware of tenderness differences ('tender', 'intermediate' and 'tough', offered at US\$4.35/lb, US\$3.85/lb and US\$3.35/lb, respectively) nearly 95% of the steaks purchased were from the 'tender' category.
- (2) Lusk *et al.* (1999) determined that when consumers were offered 'guaranteed tender' v. 'probably tough' steaks, 84% of participants preferred the 'guaranteed tender' steak, and 51% were willing to pay an average of US\$1.84/lb more to obtain the 'guaranteed tender' steak.
- (3) Umberger *et al.* (2000) reported that – when tenderness was held constant – consumers preferred and were willing to pay higher prices per pound for steaks with high v. low marbling scores and that were 'US corn-fed' rather than 'Argentine grass-fed' in cattle production system or origin.

- (4) Platter *et al.* (2005) concluded that the prices consumers were willing to pay to purchase steaks increased as marbling increased, and increased as Warner–Bratzler shear force values decreased; consumers in that study were likely to purchase steaks if they had marbling scores of MT⁵⁰ or higher, or Warner–Bratzler shear force values of 3.9 kg or lower.

MSA grading systems

Rationale for development of the MSA grading system, as described by Polkinghorne *et al.* (2008b), is as follows:

- (1) Beef consumption in Australia declined – significantly and continuously from the 1970s to 2000 – in part because consumers found beef inconsistent in eating quality and confusing to purchase, shortcomings that were exacerbated by a decline in consumer knowledge and cooking skills in combination with dietary concerns and a perceived lack of convenience.
- (2) Research was initiated in 1994 to evaluate beef by consumer testing to answer questions regarding whether consumers agree on beef quality and, if they did, could industry grading systems accurately predict the eating quality of beef cuts as sold.
- (3) The MSA research program established that consumers did have a reasonable consensus view of beef eating quality and identified a scoring system that utilised a weighted combination of sensory scores for palatability attributes that eventually evolved into the commercial MSA grading system.
- (4) The support for development of the MSA beef grading system came from the 1996 meat industry strategic plan, where three of six objectives involved the need for better description of product and marketing systems that would deliver a more consistent beef eating experience to the consumer.

The MSA grading system aims to predict the eating quality of individual cuts when aged for a defined number of days and cooked by a specified method (R. J. Polkinghorne, pers. comm.). The prediction is made by a computerised model that calculates the interaction of a range of inputs to produce an MQ4 (meat quality, four variables) score expressed in points between 0 and 100. Grades are assigned to each cut on the basis of estimated MQ4 points with those <46 deemed unsatisfactory, 47–63 graded three star, 64–76 graded four star and >76 graded five star. In effect, 137 grade results, each a cut-by-cook combination, are produced for each carcass. There is no carcass grade as such (R. J. Polkinghorne, pers. comm.).

According to R. J. Polkinghorne (pers. comm.) cattle presented for MSA grading must be supplied from registered MSA producers and be accompanied by a statutory declaration declaring the maximum percentage *B. indicus* content in the consignment, whether the cattle are classified as milk-fed veal (calves weaned immediately before sale) and whether the cattle have ever been implanted with HGP together with a time of departure from the farm and noting if the consignment has been sold through an MSA-accredited saleyard. The abattoir must also be MSA licensed and meet minimum conditions of slaughtering MSA cattle within 24 h of despatch from the supplying farm, not mixing groups in lairage and operating slaughter floor

equipment (including electrical stimulation) according to procedures that have been monitored to control the relationship of carcass pH and temperature decline within a defined window where the loin temperature at pH 6 is below 35°C and above 12°C. The chilled carcass is ribbed and an MSA grader assesses the carcass and enters data into a hand-held data-capture unit (R. J. Polkinghorne, pers. comm.). MacPherson (2004) described MSA graders as assigning grades to specific carcasses by measuring the pH and temperature of the ribeye muscle (using probes connected to a hand-held data-capture unit), entry of the identifying scan barcode, collection of a DNA sample, entry into the data-capture unit of measurements of marbling, ossification, rib-fat depth and hump height, and computation of a carcass grade. A listing of the measurements taken at Australian packing plants in association with MSA quality and yield grading (MacPherson 2004) includes: (i) DNA; (ii) breed; (iii) hump height; (iv) sex; (v) HGP; (vi) milk-fed veal; (vii) stockyard or saleyard; (viii) rinse or flush; (ix) hotscale carcass weight; (x) hang; (xi) maturity or ossification; (xii) marbling; (xiii) rib fat; (xiv) P8 measure (a carcass fat thickness measurement); (xv) pH_u (ultimate pH value of the *longissimus* muscle); (xvi) fat distribution; (xvii) meat colour; and (xviii) temperature.

According to Meat and Livestock Australia (2005), carcasses that fail to comply with MSA specifications are subsequently ungraded to non-MSA products; the factors that downgrade carcasses are: (i) rib fat (less than 3 mm), because a minimum of 3 mm of rib fat will reduce temperature variation through the carcass during chilling, which will counteract the onset of cold shortening; (ii) ossification maturity (300 score or more); (iii) fat distribution (uneven distribution over the loin, butt and forequarter); (iv) pH (5.71 and above); (v) meat colour (4 and above); (vi) miscellaneous (bruising, ecchymosis, etc.); (vii) temperature (must be below 12°C); (viii) hide puller damage (excess damage to the carcass over the primal cuts); and (ix) company specification (at the discretion of the establishment where carcasses are presented for grading). The ossification maximum of 300 has since been removed following release of a new version of the prediction model (R. J. Polkinghorne, pers. comm.). Failure to comply with qualifying conditions such as time from despatch to slaughter will also result in cattle being ineligible for grading (Meat and Livestock Australia 2005).

According to R. J. Polkinghorne (pers. comm.):

- (1) MSA model data inputs from the supplier declaration are percentage *B. indicus*, HGP implant status, milk-fed veal 'yes' or 'no', and if from a saleyard.
- (2) Inputs from the slaughter floor are carcass suspension method, whether a vascular infusion treatment has been applied, carcass weight (in kg) and sex.
- (3) MSA model inputs entered in the chiller from the quartered carcass are also defined under AUS-MEAT chiller assessment language and are identical to the AUS-MEAT language classification described above for rib fat, meat colour and fat colour.
- (4) Of these, the model algorithm only uses rib fat for computation, meat and fat colour being censoring variables applied commercially to meet trace appearance standards.
- (5) Additional MSA specific inputs are: (i) an MSA marbling standard, developed from USDA and Japanese Meat Grading Authority (JMGA) standards, with scores ranging from MSA 100 through MSA 1100 for fat deposited between individual muscle fibres of the LD muscle; (ii) maturity (or ossification) with scores from 100 to 590 – assessment of physiological age of a bovine animal using ossification in the spinous processes of vertebrae and the shape and colour of rib bones, developed from USDA scores; (iii) hump height – measured in gradients of 5 mm, used primarily to verify the tropical breed content indicated on the MSA vendor declaration; and (iv) ultimate pH – as a measure of lactic acid within the muscle, the speed at which pH declines from the live state (~pH=7.0) to the ultimate pH (pH 5.3 to 5.7 is optimal) affects eating quality.

Polkinghorne (1996, 1998, 2003) described the beef Eating Quality Assessment (EQA)–PACCP pathway scheme as including:

- (1) On farm sector: (i) genetics; and (ii) growth, development and handling.
- (2) Transport: (i) industry code of practice for transport and lairage handling; and (ii) monitor handling and consider animal temperament.
- (3) Processing sector: (i) slaughter (prevent stress); (ii) 23 h post-slaughter (pH 6.0 before loin temperature of 12°C, ultimate pH of 5.3–5.7, deep butt temperature <30°C within 10 h, deep butt temperature <16°C within 20 h, and effective electrical stimulation); (iii) chiller assessment (USDA maximum maturity, AUS-MEAT marbling, estimated minimum fat percentage, meat colour, fat colour, texture, and rib fat depth); and (iv) aging (7-, 14- or 21-day minimums).

According to R. J. Polkinghorne (pers. comm.):

- (1) These PACCP criteria were initially combined to describe production pathways, similar in concept to the British Livestock & Meat Commission (LMC) blue print and to various US branded beef programs.
- (2) Specified cuts (striploin, ribeye and tenderloin) from carcasses that met all PACCP criteria in a pathway received a MSA grade.
- (3) Several pathways incorporating alternative PACCP criteria were developed allowing some flexibility in production; for example, a 50% *B. indicus* content might be graded equal to a 0%, due to 28 days (v. 7 days) aging, tenderstretch suspension or through additional marbling.
- (4) The pathways provided a good result to the consumer in meeting maximum eating quality failure criteria but often at the expense of removing considerable product that, while failing one pathway PACCP criteria, still had sufficient eating quality due to other PACCP criteria being beyond threshold levels.
- (5) The pathways system was supplanted by the MSA prediction model, which provided muscle-specific interactive computation of inputs including many PACCP criteria.

- (6) Other PACCP criteria were retained as screening variables whereas some were replaced by alternative inputs or interactions that proved more robust.
- (7) Introduction of the prediction model allowed a higher proportion of consumer-acceptable beef to be graded while maintaining a conservative rejection rate for product predicted to be unsatisfactory.
- (8) Successive versions of the prediction model have increased the number of cuts graded, the number of cooking methods, the range of eligible cattle and modified the prediction approach from new data to improve accuracy and incorporate additional inputs such as HGP use.

A key feature of the TQM grading approach developed by Meat and Livestock Australia is that it incorporates several important elements – animal-specific traits (e.g. genetics, sex, age), control of processes in several sectors of the beef chain (including both pre-harvest and post-harvest processes), cut-specific quality differences, consumer preferences – into the beef pricing system (Tatum 2006). As a result of the latter approach, a much clearer economic signal can be transmitted through the entire beef chain, which provides producers and processors with economic incentives to become more quality conscious, and facilitates consumer-driven improvement in product performance (Tatum 2006).

Polkinghorne (1998), reporting the development and use of the production pathways system, stated that MSA had, through January 1998:

- (1) Completed consumer testing (in Sydney, Brisbane and Melbourne) of 32 000 steaks and 12 000 roasts.
- (2) Found that agreement in relative palatability, across cuts (steaks and roasts) is 'very good' (all are cooked to 'medium' degree of doneness).
- (3) Generated a meat quality score based on $MQ4 = 0.4$ (tenderness) + 0.3 (overall palatability) + 0.2 (flavour) + 0.1 (juiciness).
- (4) Identified three levels of quality: (i) absolute premium: five-star MSA grade, 80 minimum MQ4 score, failure rate = 0%, price per lb at retail AU\$9 at foodservice and AU\$6 at supermarkets, comment 'cannot meet demand for foodservice, supermarket and export trades'; (ii) premium: four-star MSA grade, 64 minimum MQ4 score, failure rate = 7.5%, price per lb at retail AU\$5 at foodservice and AU\$4 at supermarkets, comment 'targeted for Australian restaurants and export to Japan'; and (iii) good, everyday beef: three-star MSA grade, 48 minimum MQ4 score, failure rate = 20%, price per lb at retail AU\$3 at supermarkets, comment '50% will be "tenderstretched" by Summer 1998 for sale at Coles and Woolworths supermarkets'.

Polkinghorne (1998) evaluated consumer test MQ4 scores by grade in relation to mean values for pathway PACCP criteria; the four criteria that were most useful (in determining the meat quality score) were: (i) percentage *B. indicus*: levels of 8, 23, 31 and $\geq 39\%$ corresponded to five-star, four-star, three-star and failure designations, respectively; (ii) USDA marbling score: levels of SM^{87} , SL^{83} , SL^{63} and $\leq SL^{55}$ corresponded to five-star, four-star, three-star and failure designations, respectively; (iii) calculated growth: computed from carcass weight in relation to ossification

score, values of 0.8, 0.7, 0.6 and 0.6 corresponded to five-star, four-star, three-star and failure designations, respectively; and (iv) fat thickness 12th rib: levels of 8.4, 6.4, 6.1 and ≤ 6.0 mm corresponded to five-star, four-star, three-star and failure designations, respectively (R. J. Polkinghorne, pers. comm.). According to R. J. Polkinghorne (pers. comm.):

- (1) Dentition proved to be counterintuitive, levels of 1.5, 0.7, 0.6 and ≤ 0.5 corresponded to five-star, four-star, three-star and failure, respectively; higher quality being associated with greater age. This was influenced by a high percentage of older but long-term grain-fed cattle in the dataset at that point.
- (2) The principal PACCP components proven to strongly relate to eating quality by consumer testing were adopted as input variables to the prediction model with their impact converted to continuous scales rather than as set PACCP-style criteria.
- (3) Dentition was not used as an input variable, with carcass weight, ossification and sex interactions proving more robust as eating quality predictors (Polkinghorne 1998). Thompson (2002a) identified the primary CCP in the meat quality prediction model as: (i) percentage *B. indicus*; (ii) sex of the animal; (iii) the animal's growth path; (iv) milk-fed veal classification; (v) carcass hanging method; (vi) marbling score; (vii) ultimate muscle pH; (viii) length of aging period; and (ix) cooking method; with (x) use of HGP then under study as a potential CCP for use in the MSA grading system.

Relationships of MQ4 score, USDA quality grade and MSA grade are such that (Polkinghorne 2004):

- (1) A typical carcass that would grade US prime (using USDA beef quality grade standards) would produce a tenderloin grading five star, a cube roll and a striploin that would grade four star, a rump and a knuckle that would grade three star, and a brisket that would grade no star.
- (2) A typical carcass that would grade US choice would produce a tenderloin grading four star, a cube roll, a striploin and a rump that would grade three star, and a knuckle and a brisket that would grade no star.
- (3) A typical carcass that would grade US select would produce a tenderloin, a cube roll and a striploin that would grade three star, plus a rump, a knuckle and a brisket that would grade no star.

Tatum (2006) said: (i) a unique aspect of the MSA grading system is that the grades are assigned to cuts, not carcasses; (ii) cuts from the same carcass are assigned individual (and in many cases, different) grades that reflect differences in expected eating quality performance among the various cuts of beef; and (iii) eligibility of beef cuts for a specific MSA grade requires adherence to specific beef production and processing methods, as well as conformance to several live-animal and carcass specifications.

Research results have been integrated into the commercial MSA grading model. To the question 'Do MSA grades (allocated by the pathways system) sort beef according to expected palatability?', Polkinghorne and George (1998a, 1998b) reported that the percentage of unacceptable eating experiences expected from the consumption of five star-graded cuts is 0% (zero); comparable percentages of unacceptable eating

experiences expected from consumption of four star-, three star- and no star-graded cuts are 6% (1 in 17), 9% (1 in 11) and 50% (1 in 2), respectively. Polkinghorne (2006) demonstrated the efficacy of using the MSA prediction model for palatability of individual muscles by conducting a 5-year commercial trial of a retail-to-farm trading model.

Watson *et al.* (2008a, 2008b) described the evolution of the development of the MSA consumer sensory protocol as follows:

- (1) Beef eating quality needed to be routinely measured in order to systematically benchmark existing retail product and to establish and verify the effect and interaction of all product and processing factors.
- (2) It was decided to use consumer taste panels because of the need to have a reliable, transparent system of testing samples that would engender confidence within both the beef industry and consumer sectors.
- (3) It was determined that a weighted average of four sensory scores (tenderness, juiciness, flavour and overall liking) best characterised meat quality of cooked beef samples in terms of a star rating.
- (4) The final recommendation was to calculate an MQ4 score by weighting the scale results for each consumer as follows: $MQ4 = 0.4 \text{ (tenderness)} + 0.1 \text{ (juiciness)} + 0.2 \text{ (flavor)} + 0.3 \text{ (overall liking)}$.
- (5) Consumer MQ4 scores were then used to assign grades, from highest to lowest, of five star (premium), four star (better than everyday), three star (good everyday) and no star (unsatisfactory).

Watson *et al.* (2008a, 2008c) reported that:

- (1) Estimated percentage *B. indicus* and ossification groups are the most consistent animal-based indicators of meat quality. Beef from animals with lower percentages of *B. indicus* genetics and from more youthful animals 'tend to eat better'.
- (2) The variables 'grass-fed v. grain-fed' and 'number of days in the feedlot' were studied but eventually omitted from the model because the observed variation in palatability was 'well explained by weight and marbling variables already in the model'.
- (3) US marbling score appears consistently in the prediction model for all cuts. Rib fat (fat thickness over the ribeye) has a censoring role (animals with less than 3 mm are rejected), as do ultimate pH and muscle colour (animals with extreme pH or colour are rejected).
- (4) The use of HGP results in a penalty of the order of three to six meat quality points on meat palatability, depending on the cut.

Watson *et al.* (2008c) reported that: (i) significant changes in carcass weight, skeletal maturity and marbling were associated with HGP use in both steers and heifers, with the effects being greater in steers; and (ii) in both heifers and steers, HGP use was associated with significant reductions in flavour, juiciness, tenderness and marbling. Thompson *et al.* (2008b, 2008c) concluded that the effects of HGP use in reducing flavour, tenderness and overall palatability are much greater in beef from 100% *B. indicus* cattle than in beef from 50% *B. indicus* cattle, but that all implant strategies used caused a reduction in

meat quality. To assist in the interpretation of the myriad published results, Watson (2008) undertook a meta-analysis of published recent studies in the USA and Australia. He presented evidence that suggests strongly that HGP has a negative effect on eating quality (both sensory and objective laboratory measurements) of beef, especially that of the LD muscle in the striploin.

Thompson *et al.* (2008a) reported that: (i) for samples assessed by both Australian and Korean consumers, Koreans graded a higher proportion of unsatisfactory and a lower proportion of premium-grade product than Australians; and (ii) the MSA grading model correctly predicted the four eating quality grades for 59 and 53% of the samples for Koreans and Australians, respectively. Moreover, Thompson *et al.* (2008a) said that, based on several MSA investigations using different consumer groups (e.g. Australians from urban v. rural backgrounds; Japanese recently arrived in Australia v. long-term residents of Australia of European descent), and if such results were confirmed for consumers in overseas markets, the MSA palatability score may have value as an international descriptor of beef quality. In the USA, Neely *et al.* (1999) found city differences (among four cities) in consumer overall-liking scores for in-home evaluations of beef palatability. However, in Korea Hwang *et al.* (2008) and Park *et al.* (2008), comparing consumers from diverse demographic backgrounds (Korea v. Australia), reported similar sensory responses toward beef quality.

In summary, the MSA beef grading system: (i) uses a TQM approach to predict beef palatability (Polkinghorne *et al.* 1999); (ii) depends on use of consumer taste panels to identify and quantify the CCP to include in a beef grading scheme to predict palatability (Thompson *et al.* 2008a); (iii) depends on a grading prediction model, progressively developed by use of consumer responses to >38 000 muscle samples that were sourced from a variety of production, processing, value-adding and cooking treatments (Watson *et al.* 2008a); (iv) quantifies both the direct effects and interactions of the CCP on the palatability of individual muscles prepared using a variety of cooking methods (Thompson *et al.* 2008a); (v) uses consumer grades for palatability (assigned during taste panels), and depends on a discriminate analysis to form a composite meat quality score (MQ4) to maximise allocation of samples to the correct palatability grade by optimising the MQ4 boundaries between the grades (Watson *et al.* 2008a); and (vi) allocates between 50 and 70% of the samples to the correct consumer grade (Thompson 2002b).

Tatum (2006) concluded that: (i) Meat and Livestock Australia incorporated TQM principles into their MSA beef grading system; (ii) the MSA grading system identifies CCP in various sectors of the beef chain (from cattle production to meal preparation) that influence consumer acceptance of beef products; (iii) an extensive, ongoing consumer testing program identified CCP (throughout the beef chain) that are associated with consumers' likes and dislikes; (iv) a statistical grading model predicts palatability, using the identified CCP, computing a meat quality score (a combined index of tenderness, flavour, juiciness and overall acceptability), which is then used to assign each beef cut to a specific grade based on predicted consumer acceptability; and (v) the MSA grading system is still being modified and refined, but it represents the best-existing example of a TQM

grading approach for improving beef quality and palatability (Tatum 2006).

Advantages of the TQM approach to assessing beef quality

Examples of palatability-determining factors that could be included in a PACCP approach to grading (like that of the MSA beef grading system) have been identified in reviews of literature on the subject. For example, in a comprehensive review of the scientific research literature, Smith (2005) identified relationships of flavour, juiciness and tenderness of cooked beef to: (i) breed or biological type, physiological age or maturity, sex, fatness, production management history and temperament and handling of the animals from which steaks and roasts are derived; (ii) conditions of harvesting, suspending and chilling of carcasses plus conditions of storage of carcasses and primal and subprimal cuts; and (iii) chemical, physical, structural and histological characteristics of carcasses or muscles. Many of those traits, characteristics and circumstances are not easily incorporated into carcass-trait-only beef quality grading systems (like the USDA beef quality grading system), but can be used in branded-beef programs. For example, Smith *et al.* (2000) identified seven ways that the palatability performance of beef from carcasses of low choice or select could be improved for use in beef-branding endeavours: (i) high-voltage electrical stimulation of carcasses; (ii) controlled aging (i.e. refrigerated storage for prolonged times); (iii) enhancement via marination or injection of salt-phosphate solutions; (iv) sorting of carcasses using physical traits (e.g. conformation, hump height); (v) sorting of carcasses using instruments (e.g. Computer Vision System, BeefCAM); (vi) controlling genetics, with and within breeds and crosses; and (vii) use of PACCP programs (Smith *et al.* 2000).

Smith (2003b) identified the following PACCP-like systems that are used to improve the palatability of branded beef:

- (1) Cattlemen's Collection: time-on-feed, mild implant regimen, electrical stimulation, 14 days of postmortem product aging, ≤ 2 inch hump height and muscle colour by use of instrumentation.
- (2) Harris Teeter Rancher's: time-on-feed, mild implant regimen, electrical stimulation, tender-cut suspension, postmortem product aging, hump height and muscle colour by use of instrumentation.
- (3) Nolan Ryan Tender-Aged Beef: time-on-feed, mild implant regimen, electrical stimulation, postmortem product aging and BeefCAM (for muscle colour by use of instrumentation).
- (4) Harris Ranch Beef: source-verified genetics, no implants, hump height, time-on-feed and electrical stimulation.
- (5) Swift's Chain Of Tenderness: source-verified genetics, time-on-feed, electrical stimulation, high-temperature carcass conditioning, postmortem product aging, hump height and muscle colour by use of instrumentation.
- (6) Safeway Rancher's Reserve Angus: only from Red Angus or Black Angus cattle, time-on-feed, mild implant regimen, electrical stimulation, postmortem product aging, hump height, tender-cut suspension and routine Warner-Bratzler shear force testing.

Of factors considered of primary importance in determining tenderness and overall palatability of cooked beef in the MSA

grading system, there is ample US research-study support for use of marbling (Smith *et al.* 1969, 1984, 2007, 2008; Savell *et al.* 1987, 1989; George *et al.* 1999; Wyle 2000; Platter *et al.* 2003b, 2005; Gruber *et al.* 2006), maturity (Berry *et al.* 1974a, 1974b; Smith *et al.* 1982, 1988, 2008; Hilton *et al.* 1998), amount of *B. indicus* genetics (McKeith *et al.* 1985a, 1985b; Sherbeck *et al.* 1995, 1996; O'Connor *et al.* 1997), sex (Choat *et al.* 2006; Tatum *et al.* 2007), tenderstretch carcass suspension (Smith *et al.* 1971, 1979, 2007, 2008; Orts *et al.* 1971; Hostetler *et al.* 1975), ultimate pH (Smulders *et al.* 1990; Jones and Tatum 1991a, 1994; Eilers *et al.* 1996; Wulf *et al.* 1997), meat colour (Jeremiah *et al.* 1972; Wulf *et al.* 1997; Cannell *et al.* 2000; Wyle *et al.* 2003; Vote *et al.* 2003), fat colour (Hilton *et al.* 1998; Wyle *et al.* 1998, 2003; Wyle 2000; Vote *et al.* 2003) and subcutaneous fat thickness (Dolezal *et al.* 1982; Tatum *et al.* 1982; Jones and Tatum 1991b, 1994; Smith *et al.* 2007, 2008).

Another palatability-determining factor that cannot be incorporated into a carcass-trait-only beef quality grading system involves the use of growth-promoting implants. The effects of such implants on beef palatability have been studied extensively in both the USA and Australia (Thompson *et al.* 2008b, 2008c; Watson 2008; Watson *et al.* 2008c). The NCBA (1996) Beef Palatability Task Force determined that:

- (1) Use of mild implant regimens during the finishing of steers and heifers lowered the percentage of US choice (by 5%), lowered marbling score (by 5%), had a negligible effect on USDA carcass maturity, increased shear force value (by 0.5 lb) and increased the percentage of tough steaks (by 5%).
- (2) Use of strong combination implant regimens during the finishing of steers and heifers lowered the percentage of US choice (by 25%), lowered marbling score (by 25%), increased USDA carcass maturity (by 12%), increased shear force value (by 1.5 lb) and increased the percentage of tough steaks (by 25%).

Across all data presented by Duckett *et al.* (1997), representing 14 217 cattle in 77 trials, and Morgan (1997), representing 19 616 steers and heifers in 107 trials, numerical advantages for control (given no implants) v. treated (single or double implants) occurred in 28 of 30, 8 of 9, 15 of 17, 28 of 30, 9 of 9, and 18 of 21 comparisons of marbling score, skeletal maturity score, USDA quality grade, percentage choice, percentage dark cutters and shear force value, respectively. Roeber *et al.* (2000) reported that use of a simple implant decreased the incidence of carcasses grading prime or choice by 1.2 (least) to 19.4 (most) percentage points and increased the occurrence of carcasses producing tough steaks by none (least) to 21.4 (most) percentage points.

Platter *et al.* (2003a) reported that implanting steers: (i) at branding or weaning did not affect steak marbling, shear force tenderness or overall eating quality; (ii) at backgrounding increased steak shear force but did not affect marbling, tenderness or overall eating quality; and (iii) two, three, four or five times (in their lifetime) resulted in lower marbling scores, higher shear force values, lower tenderness ratings and less desirable overall eating quality. Tatum (2006) reported that:

- (1) In 16 studies, involving direct comparisons of shear force values for steer beef, 36% indicated that implants increased toughness, 64% showed no change in tenderness or toughness, and none (0%) showed that implants increased tenderness.
- (2) In four studies that have depended on consumer panels, all four concluded that beef from non-implanted cattle was less desirable in tenderness than some beef from implanted cattle, but in three of the four studies there was at least one implant scheme that did not toughen beef.
- (3) Implant programs featuring a maximum of two or three lifetime implants, with use of no more than one high-potency combination implant, administered 100 days or more before the anticipated harvest date, seem to be associated with the fewest detrimental effects on carcass quality characteristics and the lowest frequency of unsatisfactory eating experiences among beef consumers.

Gerken (2005) reported that the USDA had officially approved 52 certified, two process-verified and two brand name-validated branded-beef programs. Of those, 50 branded-beef programs had a maximum hump height constraint of ≤ 2 inches to disqualify *B. indicus*-influenced cattle and 42 of these programs used a minimum requirement of 'moderately thick muscling' to disqualify dairy-type cattle.

Polkinghorne *et al.* (2008a) concluded that, whereas industry adoption of the MSA grade technology has been variable (industry application ranges from a very basic overlay of MSA output on conventional beef production and marketing to intensive application in which conventional practice is largely supplanted), the Australian industry has become more focussed on eating quality and made substantial changes in response to MSA findings.

Summary

The USDA beef quality grading system relies solely on after-the-fact sorting (based only on differences in quality-indicating carcass traits) rather than using a quality control mentality (allowing mid-course correction in product of harvest animals) and without incorporating production or processing factors that are now well documented to affect beef palatability. Branded-beef programs in the USA have, in fact, used PACCP plans, TQM philosophies, USDA certification and process verification programs plus combinations of live-animal factors, carcass-treatment factors and carcass-trait constraints to differentiate fresh beef products.

The MSA grading system allows cattle from different production systems to achieve a common grade, and different cuts from the same carcass to be assigned different grades. Whether a cut achieves a four-star rating due to being a tenderloin from a poorer quality carcass or a blade from an excellent carcass is not an issue for the consumer – the MSA grade represents a common eating quality and this may result either from the same cut derived from similar carcasses or different cuts sourced from dissimilar carcasses.

Correlation and regression analyses revealed that USDA quality grades (across the entire eight-grade range) accounted for 40–47% of the observed variation in overall palatability of dry heat-cooked loin and top round steaks, and 25–33% of the

variation in shear force for loin, top round, bottom round and eye-of-round steaks. The percentage incidence of loin steaks rated 'very desirable' in a composite of all sensory panel ratings and shear force values were 63.6 (prime), 49.4 (choice), 35.5 (select), 20.3 (standard), 30.3 (commercial), 11.1 (utility), 3.2 (cutter) and 0.0 (canner). Research results for cuts from youthful, grain-finished steers and heifers suggest that the odds of having an unpleasant eating experience are 1 in 33 (3%) if a middle-meat steak comes from a prime carcass, as compared with 1 in 10 (10%), 1 in 6 (16%), 1 in 4 (27%) or 1 in 2 (50%) if a middle-meat steak comes from a carcass of upper two-thirds choice, low choice, select or standard grades, respectively, using the USDA beef quality grading system.

Percentages of unacceptable eating experiences expected from consumption of beef from a five-star cut is 0% (zero). Comparable percentages of unacceptable eating experiences expected from consumption of beef of four star, three star and no star are 6% (1 in 17), 9% (1 in 11) and 50% (1 in 2), respectively, using the MSA grading system. The MSA beef grading system allocates between 50 and 70% of the beef muscles and cuts to the correct consumer grade.

Research studies have shown that the accuracy of palatability-level prediction by use of the two systems – USDA quality grades for US customers and consumers and MSA grades for Australian customers and consumers – is sufficient to justify their continued use for beef quality assessments.

References

- Anon. (2000) ConAgra beefs up 'eatability'. *National Provisioner* **214**, 26.
- Belew JB, Brooks JC, McKenna DR, Savell JW (2003) Warner–Bratzler shear evaluations of 40 bovine muscles. *Meat Science* **64**, 507–512. doi: 10.1016/S0309-1740(02)00242-5
- Berry BW, Smith GC, Carpenter ZL (1974a) Beef carcass maturity indicators and palatability attributes. *Journal of Animal Science* **38**, 507–514.
- Berry BW, Smith GC, Carpenter ZL (1974b) Relationships of certain muscle, cartilage and bone traits to tenderness of the beef *longissimus*. *Journal of Food Science* **39**, 819–824. doi: 10.1111/j.1365-2621.1974.tb17986.x
- Boleman SJ, Boleman SL, Miller RK, Taylor JF, Cross HR, *et al.* (1997) Consumer evaluation of beef of known categories of tenderness. *Journal of Animal Science* **75**, 1521–1524.
- Bradbury C (2003) Money-back guarantee for Nolan Ryan Guaranteed Tender Beef. Presented at the January meeting of the Board of Directors of Beefmaster Cattlemen's, Limited Partnership, 21 January 2003. Available from Beefmaster Cattlemen's, Limited Partnership, Huntsville, TX.
- Branson RE, Smith GC, Cross HR, Savell JW, Edwards RA (1984) National consumer beef study: Executive summary. Final report from the Department of Animal Science and the Department of Agricultural Economics of Texas A&M University (College Station, TX) to the Agricultural Marketing Service, USDA (Washington, DC) and the National Cattlemen's Research Foundation (Denver, CO).
- Branson RE, Cross HR, Savell JW, Smith GC, Edwards RA (1986) Marketing implications from the National Consumer Beef Study. *Western Journal of Agricultural Economics* **11**, 82–91.
- Briskey EJ, Bray RW (1964) A special study of the beef grade standards. Final Report from University of Wisconsin to the American National Cattlemen's Association, Denver, CO.
- Campion DR, Crouse JD, Dikeman ME (1975) Predictive value of USDA beef quality grade factors for cooked meat palatability. *Journal of Food Science* **40**, 1225–1228. doi: 10.1111/j.1365-2621.1975.tb01057.x

- Cannell RC, Belk KE, Tatum JD, Smith GC (2000) A comparison of expert USDA quality grades, the RMS-Canadian Vision System and the Hunter/CSU BeefCAM for sorting beef carcasses into expected-palatability groups. Mimeograph report. Meat Science Program, Department of Animal Sciences, Colorado State University, Fort Collins, CO.
- Choat WT, Paterson JA, Rainey BM, King MC, Smith GC, Belk KE, Lipsey RJ (2006) The effects of cattle sex on carcass characteristics and *longissimus* muscle palatability. *Journal of Animal Science* **84**, 1820–1826. doi: 10.2527/jas.2004-418
- Deming WE (1986) 'Out of the crisis.' (Massachusetts Institute of Technology, Center for Advanced Educational Services: Cambridge, MA)
- Dolezal HG (2005) New technologies in beef programs of Cargill Meat Solutions. Proceedings of the Arkansas Centennial Beef Conference, Fayetteville, AR. pp. 1–12. Available from Department of Animal Science, University of Arkansas, Fayetteville, AR.
- Dolezal HG, Smith GC, Savell JW, Carpenter ZL (1982) Comparison of subcutaneous fat thickness, marbling and quality grade for predicting palatability of beef. *Journal of Food Science* **47**, 397–401. doi: 10.1111/j.1365-2621.1982.tb10089.x
- Duckett SK, Owens FN, Andre JG (1997) Effects of implants on performance and carcass traits of feedlot steers and heifers. In 'Impact of implants on performance and carcass value of beef cattle'. pp. 63–82. Available from Department of Animal Science, Oklahoma State University, Stillwater, OK.
- Eilers JD, Tatum JD, Morgan JB, Smith GC (1996) Modification of early-postmortem muscle pH and use of postmortem aging to improve beef tenderness. *Journal of Animal Science* **74**, 790–798.
- Ferguson D, Thompson J, Polkinghorne R (1999) A PACCP based beef grading scheme for consumers. *Proceedings of the International Congress of Meat Science and Technology* **45**(2), 16–18.
- Food Marketing Institute (2002) 'Factors "very important" to supermarket shoppers in food selections.' (Food Marketing Institute: Washington, DC)
- George MH, Tatum JD, Belk KE, Smith GC (1999) An audit of retail beef loin steak tenderness conducted in eight US cities. *Journal of Animal Science* **77**, 1735–1741.
- Gerken C (2005) USDA-certified, process-verified and brandname-validated beef programs. pp. 1–17. LSD-AMS. Available from Livestock & Seed Division, Agricultural Marketing Service, United States Department of Agriculture, Washington, DC.
- Gruber SL, Tatum JD, Scanga JA, Chapman PL, Smith GC, Belk KE (2006) Effects of postmortem aging and USDA quality grade on Warner-Bratzler shear force values of seventeen individual beef muscles. *Journal of Animal Science* **84**, 3387–3396. doi: 10.2527/jas.2006-194
- Hilton GG, Tatum JD, Williams SE, Belk KE, Williams FL, Wise JW, Smith GC (1998) An evaluation of current and alternative systems for quality grading carcasses of mature slaughter cows. *Journal of Animal Science* **76**, 2094–2103.
- Hostetler RL, Carpenter ZL, Smith GC, Dutson TR (1975) Comparison of postmortem treatments for improving tenderness of beef. *Journal of Food Science* **40**, 223–226. doi: 10.1111/j.1365-2621.1975.tb02167.x
- Huffhines CP, Ledall GC, Field TG, Head MA, Tatum JD, Wise JW, Smith GC (1992a) Effect of time-on-feed on feedlot performance and carcass characteristics of Hereford long-yearling cattle. In 'CSU beef program report 1992'. pp. 159–166. (Colorado State University: Fort Collins, CO)
- Huffhines CP, Ledall GC, McKissack KE, Field TG, Head MA, Tatum JD, Wise JW, Smith GC (1992b) The effect of time-on-feed, electrical stimulation, marbling and subcutaneous fat thickness on the palatability of Hereford beef steaks. In 'CSU beef program report 1992'. pp. 167–173. (Colorado State University: Fort Collins, CO)
- Huffhines CP, Ledall GC, Cannon JE, Tatum JD, Field TG, Head MA, Morgan JB, Smith GC (1993) Carcass characteristics and cooked-steak palatability of straightbred Hereford steers and heifers and Hereford crossbred steers. In 'CSU beef program report 1993'. pp. 145–152. (Colorado State University: Fort Collins, CO)
- 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.
- Industry-Wide Long Range Plan Task Force (1993) Beef Industry Long Range Plan Task Force report. Beef Industry Council, Cattlemen's Beef Board, US Meat Export Federation, and National Cattlemen's Association, Englewood, CO.
- Jeremiah LE, Carpenter ZL, Smith GC (1972) Beef color as related to consumer acceptance and palatability. *Journal of Food Science* **37**, 476–479. doi: 10.1111/j.1365-2621.1972.tb02667.x
- Jeremiah LE, Carpenter ZL, Smith GC, Butler OD (1970) Beef quality. 1. Marbling as an indicator of palatability. Texas Agricultural Experiment Station, departmental technical report no. 22. Department of Animal Science, Texas A&M University: College Station, TX.
- Jones BK, Tatum JD (1991a) Use of a muscle pH measurement to identify beef carcasses with superior tenderness. In 'CSU beef program report 1991'. pp. 181–188. (Colorado State University: Fort Collins, CO)
- Jones BK, Tatum JD (1991b) Use of a minimum fat thickness constraint to ensure acceptable tenderness of beef. In 'CSU beef program report 1991'. pp. 189–194. (Colorado State University: Fort Collins, CO)
- Jones BK, Tatum JD (1994) Predictors of beef tenderness among carcasses produced under commercial conditions. *Journal of Animal Science* **72**, 1492–1501.
- Joseph RL, Connolly J (1979) Measurement and prediction of tenderness in six beef muscles. *Meat Science* **3**, 21–29. doi: 10.1016/0309-1740(79)90021-4
- Lambert CD (1991). Lost opportunities in beef production. Mimeograph report, National Cattlemen's Association, Englewood, CO.
- Lunt DK, Smith GC, Savell JW, Cross HR, Smith SB (1989) Ether-extractable fat in *longissimus* muscles from beef carcasses with the same marbling scores but different Yield Grades. *Journal of Food Science* **54**, 1364–1366. doi: 10.1111/j.1365-2621.1989.tb05994.x
- Lusk J, Fox J, Schroeder T, Mintert J, Koohmaraie M (1999) 'Will consumers pay for guaranteed tender steak.' Research Bulletin 3-99. Research Institute on Livestock Pricing. Available at www.naiber.org/publications.htm [Verified 18 September 2008]
- MacPherson D (2004) Tender, tasty beef – every time. Nuffield Farming Scholarships Trust Report, Blaston Lodge Farm, Blaston, Market Harborough, Leicestershire, UK.
- McKeith FK, Savell JW, Smith GC, Dutson TR, Carpenter ZL (1985a) Tenderness of major muscles from three breed-types of cattle at different times-on-feed. *Meat Science* **13**, 151–166. doi: 10.1016/0309-1740(85)90055-5
- McKeith FK, Savell JW, Smith GC, Dutson TR, Carpenter ZL (1985b) Physical, chemical, histological and palatability characteristics of muscles from three breed-types of cattle at different times-on-feed. *Meat Science* **15**, 37–50. doi: 10.1016/0309-1740(85)90073-7
- Meat and Livestock Australia (2005) 'Meat Standards Australia carcass grading system.' (Meat and Livestock Australia: Sydney)
- Morgan JB (1992) Tenderness problems and potential solutions. In 'Improving the consistency and competitiveness of beef. Final report of the National Beef Quality Audit 1991'. (Eds GC Smith *et al.*). pp. 180–187. (National Cattlemen's Association: Englewood, CO)
- Morgan JB (1997) Implant program effects on USDA beef carcass quality grade traits and meat tenderness. In 'Impact of implants on performance and carcass value of beef cattle'. pp. 147–154. Available from Department of Animal Science, Oklahoma State University, Stillwater, OK.
- NCA (1993) 'Strategic alliances field study: Executive summary.' (National Cattlemen's Association: Englewood, CO)
- NCA (1994) 'National beef tenderness plan: Executive summary.' (National Cattlemen's Association: Englewood, CO)

- NCBA (1996) 'Growth promotants: Carcass and cut effects.' (National Cattlemen's Beef Association: Denver, CO)
- NCBA (2001) 'Beef industry: Long range planning, 2001–2004.' (National Cattlemen's Beef Association, Centennial, CO)
- Neely TR, Lorenzen CL, Miller RK, Tatum JD, Wise JW, Taylor JF, Buyck MJ, Reagan JO, Savell JW (1998) Beef customer satisfaction: Role of cut, USDA quality grade, and city on in-home consumer ratings. *Journal of Animal Science* **76**, 1027–1033.
- Neely TR, Lorenzen CL, Miller RK, Tatum JD, Wise JW, Taylor JF, Buyck MJ, Reagan JO, Savell JW (1999) Beef tenderness satisfaction, cooking method and degree of doneness effects on top round steaks. *Journal of Animal Science* **77**, 653–660.
- NLSMB (1995). 'Beef customer satisfaction: A comprehensive in-home product test among frequent beef users.' (National Live Stock and Meat Board: Chicago)
- O'Connor SF, Tatum JD, Wulf DM, Green RD, Smith GC (1997) Genetic effects on beef tenderness in *Bos indicus* composite and *Bos taurus* cattle. *Journal of Animal Science* **75**, 1822–1830.
- Orts FA, Smith GC, Hostetler RL (1971) Texas A&M Tenderstretch. Bulletin L-1003, 1–8. Texas Agricultural Extension Service, Texas A&M University, College Station, TX.
- Park BY, Hwang IH, Cho SH, Yoo YM, Kim JH, Lee JM, Polkinghorne R, Thompson JM (2008) Effect of carcass suspension and cooking method on the palatability of three beef muscles as assessed by Korean and Australian consumers. *Australian Journal of Experimental Agriculture* **48**, 1396–1404.
- Platter WJ, Tatum JD, Belk KE, Scanga JA, Smith GC (2003a) Effects of repetitive use of hormonal implants on beef carcass quality, tenderness, and consumer ratings of beef palatability. *Journal of Animal Science* **81**, 984–996.
- Platter WJ, Tatum JD, Belk KE, Chapman PL, Scanga JA, Smith GC (2003b) Relationships of consumer sensory ratings, marbling score, and shear force value to consumer acceptance of beef strip loin steaks. *Journal of Animal Science* **81**, 2741–2750.
- Platter WJ, Tatum JD, Belk KE, Koontz SR, Chapman PL, Smith GC (2005) Effects of marbling and shear force on consumers' willingness-to-pay for beef strip loin steaks. *Journal of Animal Science* **83**, 890–899.
- Polkinghorne R (1996) Beef eating quality assurance. PACCP pathway schematic. II. Part A: Proposed EQA pathways. Meat Research Corporation, Sydney.
- Polkinghorne R (1998) Meat Standards Australia: key elements of the scheme. Mimeograph report. Meat Standards Australia, Meat and Livestock Australia, Sydney.
- Polkinghorne R (2003) Meat Standards Australia: key elements of the scheme. Mimeograph report. Meat Standards Australia, Meat and Livestock Australia, Sydney.
- Polkinghorne R (2004) Comparison of USDA and MSA grading systems. Mimeograph report. Meat Standards Australia, Meat and Livestock Australia, Sydney.
- Polkinghorne RJ (2006) Implementing a palatability assured critical control point (PACCP) approach to satisfy consumer demands. *Meat Science* **74**, 180–187. doi: 10.1016/j.meatsci.2006.05.001
- Polkinghorne R, George MH (1998a) Accuracy and precision of the MSA PACCP grading scheme for predicting consumer responses to beef. Mimeograph report. Meat and Livestock Australia, Sydney.
- Polkinghorne R, George MH (1998b) Supply, distribution and marketing of a better, more consistent, beef eating experience: Meat Standards Australia. Mimeograph report. pp. 1–15. Meat and Livestock Australia, Sydney.
- Polkinghorne R, Watson R, Porter M, Gee A, Scott J, Thompson J (1999) A 'PACCP'-based beef grading scheme for consumers. 1. The use of consumer scores to set grade standards. *Proceedings of the International Congress of Meat Science and Technology* **45**, 14–15.
- Polkinghorne R, Watson R, Thompson JM, Pethick DW (2008a) Current usage and future development of the Meat Standards Australia (MSA) grading system. *Australian Journal of Experimental Agriculture* **48**, 1459–1464.
- Polkinghorne R, Philpott J, Gee A, Doljanin A, Innes J (2008b) Development of a commercial system to apply the Meat Standards Australia grading model to optimise the return on eating quality in a beef supply chain. *Australian Journal of Experimental Agriculture* **48**, 1451–1458.
- Polkinghorne R, Thompson JM, Watson R, Gee A, Porter M (2008c) Evolution of the Meat Standards Australia (MSA) beef grading system. *Australian Journal of Experimental Agriculture* **48**, 1351–1359.
- Quinn F (1999) Tracing the future of meat. *Proceedings of the World Meat Congress* **12**(2), 1–10.
- Rhee MS, Wheeler TL, Shackelford SD, Koohmaraie M (2004) Variation in palatability and biochemical traits within and among eleven beef muscles. *Journal of Animal Science* **82**, 534–550.
- Roeber DL, Cannell RC, Belk KE, Miller RK, Tatum JD, Dean K, Smith GC (2000) Implant strategies during feeding: Impact on carcass grades and consumer acceptability. *Journal of Animal Science* **78**, 1867–1874.
- Savell JW (2003) Electrical stimulation, aging grid, consumer evaluations. Available from Beefmaster Cattlemen's, Limited Partnership, Huntsville, TX.
- Savell JW, Cross HR, Smith GC (1986) Percentage of ether-extractable fat and moisture content of beef *longissimus* muscle as related to USDA marbling score. *Journal of Food Science* **51**, 838–840. doi: 10.1111/j.1365-2621.1986.tb13946.x
- Savell JW, Branson RE, Cross HR, Stiffler DM, Wise JW, Griffin DB, Smith GC (1987) National consumer retail beef study: Palatability evaluations of beef loin steaks that differed in marbling. *Journal of Food Science* **52**, 517–519.
- Savell JW, Cross HR, Francis JJ, Wise JW, Hale DS, Wilkes DL, Smith GC (1989) National consumer retail beef study: Interaction of trim level, price and grade on consumer acceptance of beef steaks and roasts. *Journal of Food Quality* **12**, 251–274. doi: 10.1111/j.1745-4557.1989.tb00328.x
- Shackelford SD, Wheeler TL, Koohmaraie M (1995) Relationship between shear force and trained sensory panel tenderness ratings of 10 major muscles from *Bos indicus* and *Bos taurus* cattle. *Journal of Animal Science* **73**, 3333–3340.
- Sherbeck JA, Tatum JD, Field TG, Morgan JB, Smith GC (1995) Feedlot performance, carcass traits and palatability traits of Hereford and Hereford × Brahman steers. *Journal of Animal Science* **73**, 3613–3620.
- Sherbeck JA, Tatum JD, Field TG, Morgan JB, Smith GC (1996) Effect of phenotypic expression of Brahman breeding on marbling and tenderness traits. *Journal of Animal Science* **74**, 304–309.
- Shook JN, VanOverbeke DL, Scanga JA, Belk KE, Savell JW, Lawrence TE, Morgan JB, Griffin DB, Hale DS, Smith GC (2008) The National Beef Quality Audit 2005. Phase I. Views of producers, packers and merchandisers on current quality characteristics of the beef industry. *The Professional Animal Scientist* **24**, 189–197.
- Slinger WD, Marchello MJ, Danielson RB, Haugse CN, Johnson VK, Vidal AS, Dinusson WE, Berg PT (1985) Muscle tenderness, other carcass traits and the effect of cross-breeding on these traits in beef cattle. *Journal of Animal Science* **61**, 1402–1410.
- Smith GC (2002) Beef palatability. Presented at The Beef Palatability Enhancement Workshop. pp. 1–26. Available from Canadian Beef Council, Calgary, Alberta, Canada.
- Smith GC (2003a) Quality, safety and consistency of beef. Beefmaster Cattlemen's, Limited Partnership Think Tank on 9 January 2003. Available from Beefmaster Cattlemen's, Limited Partnership, Huntsville, TX.
- Smith GC (2003b) PACCP systems for improving palatability of US select beef. Mimeograph report. HEB Supermarkets/Swift and Company Palatability Seminar, Greeley, CO.

- Smith GC (2005) The taste of beef. Mimeograph report. Meat Science Program, Department of Animal Sciences, Colorado State University, Fort Collins, CO.
- Smith GC, Carpenter ZL (1974) Eating quality of animal products and their fat content. In 'Proceedings of the symposium on changing the fat content and composition of animal products'. pp. 1–63. (National Research Council, National Academy of Sciences: Washington, DC)
- Smith GC, Carpenter ZL, King GT (1969) Considerations for beef tenderness evaluations. *Journal of Food Science* **34**, 612–618. doi: 10.1111/j.1365-2621.1969.tb12103.x
- Smith GC, Arango TC, Carpenter ZL (1971) The effects of physical and mechanical treatments on the tenderness of the beef *longissimus*. *Journal of Food Science* **36**, 445–449. doi: 10.1111/j.1365-2621.1971.tb06384.x
- Smith GC, Culp GR, Carpenter ZL (1978) Postmortem aging of beef carcasses. *Journal of Food Science* **43**, 823–826. doi: 10.1111/j.1365-2621.1978.tb02430.x
- Smith GC, Jambers TG, Carpenter ZL, Dutson TR, Hostetler RL, Oliver WM (1979) Increasing the tenderness of forage-fed beef. *Journal of Animal Science* **49**, 1207–1218.
- Smith GC, Murphey CE, Carpenter ZL, Savell JW, Cross HR (1980) An evaluation of the USDA beef carcass quality grade standards. Final report to Agricultural Marketing Service, US Department of Agriculture, Washington, DC, from the Meats and Muscle Biology Section, Department of Animal Science, Texas A&M University, College Station, TX.
- Smith GC, Cross HR, Carpenter ZL, Murphey CE, Savell JW, Abraham HC, Davis GW (1982) Relationship of USDA maturity groups to palatability of cooked beef. *Journal of Food Science* **47**, 1100–1107. doi: 10.1111/j.1365-2621.1982.tb07627.x
- Smith GC, Savell JW, Cross HR, Carpenter ZL (1983) The relationship of USDA quality grade to beef flavor. *Food Technology* **37**, 233–238.
- Smith GC, Carpenter ZL, Cross HR, Murphey CE, Abraham HC, Savell JW, Davis GW, Berry BW, Parrish FC Jr (1984) Relationship of USDA marbling scores to palatability of cooked beef. *Journal of Food Quality* **7**, 289–308. doi: 10.1111/j.1745-4557.1985.tb01061.x
- Smith GC, Savell JW, Cross HR, Carpenter ZL, Murphey CE, Davis GW, Abraham HC, Parrish FC Jr, Berry BW (1987) Relationship of USDA quality grades to palatability of cooked beef. *Journal of Food Quality* **10**, 269–286. doi: 10.1111/j.1745-4557.1987.tb00819.x
- Smith GC, Berry BW, Savell JW, Cross HR (1988) USDA maturity indices and palatability of beef rib steaks. *Journal of Food Quality* **11**, 1–13. doi: 10.1111/j.1745-4557.1988.tb00860.x
- Smith GC, Savell JW, Clayton RP, Field TG, Griffin DB, *et al.* (1992) Improving the consistency and competitiveness of beef. Final report of the National Beef Quality Audit 1991, National Cattlemen's Association, Englewood, CO.
- Smith GC, Belk KE, Scanga JA, Tatum JD (2000) Quality of beef and its certification for the public in the future. *Proceedings of the World Buiatrics Congress* **XXI**, 147–174.
- Smith GC, Tatum JD, Belk KE, Scanga JA (2005) Beef carcass quality and yield grading in the USA, Canada and Australia. pp. 1–8. 2^a Jornada Técnica de la Fundación ANETIF, Mexico City, 13–14 June 2005.
- Smith GC, Tatum JD, Belk KE, Scanga JA (2007) Post-harvest practices for enhancing beef tenderness. Final report to the National Cattlemen's Beef Association by the Meat Science Program. Department of Animal Sciences, Colorado State University, Fort Collins, CO.
- Smith GC, Tatum JD, Belk KE, Scanga JA (2008) 'Post-harvest practices for enhancing beef tenderness.' (National Cattlemen's Beef Association: Centennial, CO)
- Smulders FJM, Marsh BB, Swartz DR, Russell RL, Hoenecke ME (1990) Beef tenderness and sarcomere length. *Meat Science* **28**, 349–362. doi: 10.1016/0309-1740(90)90048-B
- Tatum JD (2006) 'Pre-harvest cattle management practices for enhancing beef tenderness.' (National Cattlemen's Beef Association: Centennial, CO)
- Tatum JD, Smith GC, Carpenter ZL (1982) Interrelationships between marbling, subcutaneous fat thickness and cooked beef palatability. *Journal of Animal Science* **54**, 777–784.
- Tatum JD, George MH, Belk KE, Smith GC (1997) Development of a palatability assurance critical control points (PACCP) model to reduce the incidence of beef palatability problems. Final report to the National Cattlemen's Beef Association. Department of Animal Sciences, Colorado State University, Fort Collins, CO.
- Tatum JD, George MH, Belk KE, Smith GC (1998) An overview of a TQM approach for improving beef tenderness. Mimeograph report. Meat Science Program, Department of Animal Sciences, Colorado State University, Fort Collins, CO.
- Tatum JD, Belk KE, George MH, Smith GC (1999) Identification of quality management practices to reduce the incidence of retail beef tenderness problems: Development and evaluation of a prototype quality system to produce tender beef. *Journal of Animal Science* **77**, 2112–2118.
- Tatum JD, Smith GC, Belk KE (2000) New approaches for improving tenderness, quality and consistency of beef. *Journal of Animal Science*. Available at <http://www.asas.org/jas/symposia/proceedings> [Verified 18 September 2008]
- Tatum JD, Gruber SL, Schneider BA (2007) 'Pre-harvest factors affecting beef tenderness in heifers.' (National Cattlemen's Beef Association: Centennial, CO)
- Thompson J (2002a) Managing meat tenderness. *Proceedings of the International Congress of Meat Science and Technology* **48**, 17–27.
- Thompson J (2002b) Managing meat tenderness. *Meat Science* **62**, 295–308. doi: 10.1016/S0309-1740(02)00126-2
- Thompson J, Polkinghorne R, Watson R, Gee A, Murison B (1999) A 'PACCP' based beef grading scheme for consumers. 4. A cut based grading scheme to predict eating quality by cooking method. *Proceedings of the International Congress of Meat Science and Technology* **45**, 20–21.
- Thompson JM, Polkinghorne R, Hwang IH, Gee AM, Cho SH, Park BY, Lee JM (2008a) Beef quality grades as determined by Korean and Australian consumers. *Australian Journal of Experimental Agriculture* **48**, 1380–1386.
- Thompson JM, Polkinghorne R, Porter M, Burrow HM, Hunter RA, McGrabb GJ, Watson R (2008b) Effect of repeated implants of oestradiol-17 β on beef palatability in Brahman and Brahman cross steers finished to different market end points. *Australian Journal of Experimental Agriculture* **48**, 1434–1441.
- Thompson JM, McIntyre BM, Tudor GD, Pethick DW, Polkinghorne R, Watson R (2008c) Effects of hormonal growth promotants (HGP) on growth, carcass characteristics, the palatability of different muscles in the beef carcass and their interaction with aging. *Australian Journal of Experimental Agriculture* **48**, 1405–1414.
- Umberger WJ, Feuz DM, Calkins CR, Killinger KM (2000) Consumer preference and willingness to pay for flavor in beef steaks. Paper presented at the 2000 IAMA Agribusiness Forum, Chicago, IL. Available from the International Food and Agribusiness Management Association, College Station, TX.
- USDA (1997) 'United States standards for grades of carcass beef.' (United States Department of Agriculture, Agricultural Marketing Service: Washington, DC)
- Valin C (2000) Research objectives and requirements in meat science and technology. *Proceedings of the International Congress of Meat Science and Technology* **46**, 24–29.
- Vote DJ, Belk KE, Tatum JD, Scanga JA, Smith GC (2003) Online prediction of beef tenderness using a computer vision system equipped with a BeefCam module. *Journal of Animal Science* **81**, 457–465.
- Watson R (2000) Predicting meat quality. Report to the Meat Science Program, Department of Animal Sciences, Colorado State University, Fort Collins, CO.

- Watson R (2008) Meta-analysis of the published effects of HGP use on beef palatability in steers as measured by objective and sensory testing. *Australian Journal of Experimental Agriculture* **48**, 1425–1433.
- Watson R, Polkinghorne R, Thompson JM (2008a) Development of the Meat Standards Australia (MSA) prediction model for beef palatability. *Australian Journal of Experimental Agriculture* **48**, 1368–1379.
- Watson R, Gee A, Polkinghorne R, Porter M (2008b) Consumer assessment of eating quality – development of protocols for Meat Standards Australia (MSA) testing. *Australian Journal of Experimental Agriculture* **48**, 1360–1367.
- Watson R, Polkinghorne R, Gee A, Porter M, Thompson JM, Ferguson D, Pethick D, McIntyre B (2008c) Effect of hormonal growth promotants on palatability and carcass traits of various muscles from steer and heifer carcasses from a *Bos indicus*–*Bos taurus* composite cross. *Australian Journal of Experimental Agriculture* **48**, 1415–1424.
- Wheeler TL, Shackelford SD, Koohmaraie M (2000) Relationships of beef *longissimus* tenderness classes to tenderness of *gluteus medius*, *semimembranosus* and *biceps femoris*. *Journal of Animal Science* **78**, 2856–2861.
- Wulf DM, O'Connor SF, Tatum JD, Smith GC (1997) Using objective measures of muscle color to predict beef *longissimus* tenderness. *Journal of Animal Science* **75**, 684–692.
- Wyle AM (2000) An evaluation of the palatability attributes of six beef product lines and the effectiveness of using the HunterLab BeefCAM System to predict beef palatability. MS Thesis. Department of Animal Sciences, Colorado State University, Fort Collins, CO.
- Wyle AM, Cannell RC, Belk KE, Goldberg M, Riffle R, Smith GC (1998) An evaluation of the portable HunterLab Video Imaging System (BeefCAM) as a tool to predict tenderness of steaks from beef carcasses using objective measures of lean and fat color. Final report to the National Cattlemen's Beef Association. Department of Animal Sciences, Colorado State University, Fort Collins, CO.
- Wyle AM, Vote DJ, Roeber DL, Cannell RC, Belk KE, Scanga JA, Goldberg M, Tatum JD, Smith GC (2003) Effectiveness of SmartMV prototype BeefCam System to sort beef carcasses into expected palatability groups. *Journal of Animal Science* **81**, 441–448.

Manuscript received 16 July 2008, accepted 12 September 2008