POLKINGHORNES
Beef at its Best

The Secretary,
Senate Standing Committee on Rural and Regional Affairs and Transport
Parliament House
Canberra ACT 2600

20th March 2009

SUBMISSION REGARDING BEEF LABELLING & GRADING

Rod. Polkinghorne OAM, Director Polkinghornes Pty Ltd.

In my opinion the one critical issue in regard to beef labelling is that it accurately reflects a consumer sensory experience. Existing description based solely on cut names cannot deliver this outcome. Some form of grading system is required to accurately predict cooked product performance.

A grading system becomes relevant and valuable when it conveys a clear simple cooked result to a consumer. This result should not require additional consumer knowledge or information. If this is delivered beef can compete in the food sector as a contemporary consumer product with the many layers of traditional mystique and confusion removed.

Consumers understand the relative value offer in purchasing petrol – unleaded, premium unleaded, Ultimate etc or airline tickets – economy, business, first, but beef is mostly a lottery. The description system used reflects anatomy; rump, chuck, tenderloin. The consumer is assumed to have a base knowledge of cuts and cooking that can somehow deliver a desired result. This is clearly impossible; I may know I have a rump but off which animal? How old was it? How fast has it grown? Has it had a hormone implant? What breed was it? What sex? How was it killed? How long has the rump been aged? Which of the five main muscles within the rump are we talking about? These matters are mostly not known by the butcher, let alone the consumer and they interact in a complex manner.

A consumer wants a decent meal and is delivered an unsolvable mental challenge. A grading system must deliver a simple good, extremely good or fantastic guaranteed meal at a price point. This allows quality to be related to occasion with the best value likely to be different for different occasions.

In truth traditional beef grading systems have never attempted to do this and fail badly. The USDA system for example was designed in 1926 to classify carcasses into groups as a trade description. The notion that these might reflect eating quality came much later and again is nonsense. Other systems such as the EUROP system used in Europe have always focused on yield with no regard to eating quality. Later

15 Speedwell St, Somerville Vic 3912
PH: 03 5977 8940 FAX: 03 5977 7988
ACN: 122 589 726
EMAIL: warehouse@polkinghornes.com.au

systems such as the Japanese JMGA, Canadian and Korean systems have attempted to convey an eating quality guide but work poorly. All grade the carcass as a single unit.

The original AUSMEAT language was an attempt to provide a base which could be used to provide a flexible specification between buyers and sellers. The inclusion of marbling, meat colour, fat colour, sex and dentition was thought to provide a means of defining grades, hopefully related to eating quality. This was the basis we first used in trying to establish an Australian grading system for ALFA (Australian Lot Feeders Association) and others. Unfortunately the effort failed, with the failure providing the stimulus to embark on a consumer testing program to unravel the true connections between sensory experience and potential grading inputs.

There are a number of critical issues that preclude carcass grading ever being effective:

1. Carcasses that look the same don't eat the same. Grouping on the basis of sex, dentition, weight range, fat and meat colour provides a pleasant looking group of carcasses with an unacceptable range of eating quality.

2. A carcass is not a uniform item. It is composed of hundreds of different muscles with different structure, function and eating quality. There are over 40 major muscles used to create conventional beef cuts. A tenderloin (fillet) will not eat the same as a brisket, shin or chuck despite coming from the same carcass.

3. The individual muscles within a carcass do not have a constant relationship. This appears to be the common assumption in grading carcasses. It is acknowledged that the tenderloin and shin will eat differently but assumed that their relative performance is constant. If it were then all carcass muscles could be estimated from a single indicator muscle. This is generally assumed to be the striploin (porterhouse) as major grading systems assess the face of this cut to assign marbling, meat and fat colour scores.

4. Consumers do not purchase or eat carcasses. They typically consume a 200 gm portion of beef in a meal. The description they require is the sensory result they can rely on for a specific 200gm portion. The other 200kg or so of product from the source carcass is irrelevant.

The problem with muscle relationships is illustrated by tables 1 and 2. (Polkinghorne, 2005). Table 1 presents the sensory results by cooking method for a range of major muscles from a single carcass. The sensory scores have been converted to a ratio of the grilled score of the striploin (*m. longissimus lumborum*). Consequently another cut with a ratio of 80 has a sensory result 80% of the striploin whereas a 120 ratio indicates a sensory result 20% above the striploin. For the carcass shown the striploin itself would eat 4% better as a stir fry than as a steak whereas the GMD (large portion of the rump) would eat 10%

below the striploin if both were grilled but 5% better than the grilled striploin if roasted.

It is immediately evident that sensory response is directly affected by cooking method. The ratio of each muscle changes by cooking method so that their individual sensory performance and their relativity to other muscles is dependent on the cooking method selected for each. This demands that any description system needs to be implemented within a cooking method framework to provide a reliable consumer guide or guarantee.

Table 1. Ratio of cut by cook MQ4 scores for selected muscles (Base carcass)

MUSCLE	Code	Grill	Roast	Stir Fry	Thin Slice	Slow Cook
m.longissimus lumborum	**LD**	100	101	104	105	na
m.spinalis dorsi	**SD**	136	118	135	128	na
m.psoas major	**PM**	136	134	140	130	na
m.infraspinatus	**IF**	113	109	118	123	na
m.triceps brachii caput longum	**TB**	96	102	104	106	107
m.gluteus medius ("D")	**GMD**	90	105	99	109	95
m.gluteus medius (eye)	**GME**	95	109	108	107	na
m.biceps femoris (cap)	**BFC**	104	na	118	119	na
m.biceps femoris (distal)	**BFD**	na	71	75	98	103
m.rectus femoris	**RF**	83	106	97	103	84
m.vastus lateralis	**VL**	65	85	79	91	93
m.semitendinosus	**ST**	77	84	80	85	88
m.adductor femoris	**AF**	71	na	91	95	89
m.semimembranosus	**SM**	62	77	77	100	93
m.serratus ventralis cervicis	**SV**	95	97	98	104	118

Table 2 adopts a similar approach of presenting muscles as a ratio of a grilled striploin score, in this case displaying only the grilled result for muscles from eight carcasses. The carcasses represented have different characteristics; sex, weight, ossification, breed, carcass hanging method, days aged etc. Given these differences we would expect the carcasses to have a range of eating quality. However if the relationship between muscles were constant the ratios should also be constant within each carcass as the striploin score has been set to 100 in each case.

Again the ratios are anything but constant due to individual muscles responding differently to the contributing factors. Increased *bos-indicus* (Brahman) content has a greater effect on striploins than most other muscles as does use of hormonal growth promotants. Tenderstretch carcass hanging has a big effect on some muscles and none on others. Ageing effects differ between muscles as does the influence of marbling and ossification (a measure of maturity).

This does not mean that each muscle needs to be individually viewed or assessed in a grading process but rather that a single set of carcass based inputs – carcass weight, sex, marbling etc need to be applied with differential weightings for each muscle. Fortunately this is not difficult within a computerised system .

Table 2. The ratio of predicted MQ4 scores (Grill) for selected muscles from a range of carcasses.
** AT=achilles tendon, TX=obturator foramen

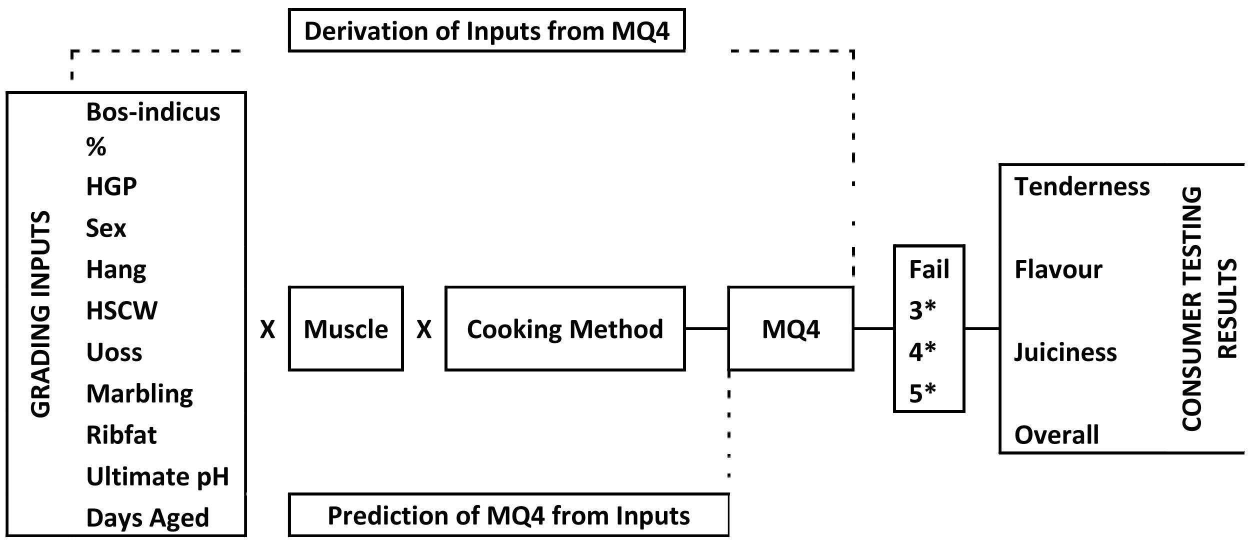
MODEL INPUTS		CARCASS INPUTS							
		Base	**A**	**B**	**C**	**D**	**E**	**F**	**G**
% bos indicus		0	100	0	0	60	35	0	60
Sex		M	F	M	M	M	F	M	M
HGP implant		No	No	Yes	No	No	No	No	Yes
Carcass Wt (Kg)		250	250	380	280	290	250	380	290
HANG **		AT	AT	AT	TX	TX	AT	AT	AT
Ossification		150	120	170	120	170	500	190	190
Marbling		250	200	330	350	330	200	500	300
Rib fat (mm)		5	5	5	5	5	5	15	5
pHU		5.5	5.5	5.5	5.5	5.5	5.5	5.5	5.5
Loin temp°C		3	3	3	3	3	3	3	3
Days aged		7	7	14	21	14	28	14	21
MUSCLES	**CODE**	**RATIO TO LD MQ4**							
m.longissimus lumborum	**LD**	100	100	100	100	100	100	100	100
m.spinalis dorsi	**SD**	136	171	134	112	124	147	121	143
m.psoas major	**PM**	136	162	134	114	125	144	122	142
m.infraspinatus	**IF**	113	142	110	91	99	120	100	114
m.triceps brachii caput longum	**TB**	96	114	95	84	85	87	87	95
m.gluteus medius ("D")	**GMD**	90	121	89	90	95	92	82	95
m.gluteus medius (eye)	**GME**	95	128	94	94	100	98	87	101
m.biceps femoris (cap)	**BFC**	104	139	101	96	102	106	94	109
m.rectus femoris	**RF**	83	104	83	79	84	77	77	85
m.vastus lateralis	**VL**	65	80	66	65	67	56	61	66
m.semitendinosus	**ST**	77	98	79	72	71	63	72	80
m.adductor femoris	**AF**	71	88	70	72	78	70	68	74
m.semimembranosus	**SM**	62	77	62	65	70	60	60	65
m.serratus ventralis cervicis	**SV**	95	117	96	83	86	97	89	97

The inescapable conclusion is that a single carcass grade cannot be remotely effective as a consumer tool. To be effective a consumer grading system must in effect grade individual muscles within a cooking method framework and describe or label a beef portion according to its' cooked performance rather than anatomical origin. A further conclusion is that retail description by muscle can also be more misleading than helpful due to the same muscle, with the same visual appearance, having radical eating quality differences. These conclusions were not evident when we first attempted to develop an Australian grading system in the early 1990's but have evolved as we addressed the problem from the consumer down rather than from the animal up.

Our initial efforts between 1992 and 1995 utilised base AUSMEAT language and compared this to using USDA grading. Neither system worked when tested by consumers, even on a single cut and cook combination – grilled striploins. We assumed this was due to pre slaughter factors which influenced eating quality but could not be seen at the point of grading. This lead us to develop and trial what was dubbed a Pathways system under which a set of rigid parameters relating to animal history _were combined with defined abattoir procedures for electrical stimulation and other matters and conventional grading inputs. This approach mirrored the MLC Blueprint approach used in the UK and proposed PACCP systems from the USA. Again these systems were applied to an entire carcass.

A level of success was achieved under the Pathways approach when a single grilled striploin result was the measure. A sufficiently tight set of Pathway parameters could be used to deliver an acceptable level of consumer satisfaction. A major problem however was that as the Pathway criteria were strengthened to achieve acceptable levels of consumer guarantee many rejected cuts actually performed well whereas if criteria were relaxed to achieve an acceptable level of inclusion the failure rate became unacceptable. The core reason was that individual inputs inter-reacted so that a minimal failure in marbling level might be offset by lower ossification or longer cut ageing for example. To address these issues multiple alternate Pathways were devised and tested, each delivering a common designated quality result (for a grilled striploin). This became difficult to manage however and, as individual cut and cooking method testing commenced, unworkable due to the plethora of possible combinations and their different outcomes at the cut level.

The trial work had however established a sizeable data base which was analysed from a different consumer down perspective. The principal is represented in the following diagram.



Under this approach consumer test results were analysed to firstly establish a consumer benchmark or score that could be used as both a measure and ultimately as a grade standard or target. The MQ4 (meat quality, 4 variable) score was developed statistically to represent weightings between tenderness, flavour, juiciness and overall satisfaction. Having combined the four parameters into a single score on a 1 to 100 scale further statistical analysis was used to determine cut-off scores which best _allocated a grade on the basis of score (unsatisfactory <47, premium >77 etc).

This was and remains a vital principal; the grade standard, by definition, is the sensory response from untrained consumers independent of the combination of physical inputs which interact to create the result. Further, the result applies to an individual consumer meal sized beef portion cooked by a designated method. The standards can and should be adjusted over time or country to best reflect consumer response. This contrasts to all other grading system specifications where grades are defined by carcass characteristics – marbling, maturity, colour etc – and applied to the entire carcass.

The continuing consumer based trial work became the foundation of the Meat Standards Australia (MSA) program developed under an MLA structure following the disbanding of the Meat Research Corporation and AMLC structure in the mid 1990's.

Having established this consumer standard all available criteria from live animal to final meal were evaluated against their ability to predict the observed MQ4. This resulted in creation of an interactive computer model, the technical base of MSA grading. A diagrammatic representation of the model inputs and output is attached. The 3, 4 and 5 star results shown in the right hand table are allocated from the individual MQ4 scores between 1 and 100, assigned as 3, 4 or 5* according to the established cut-offs. In all there are currently 146 cut by cook combinations produced for every carcass.

The inputs shown in the left hand table are those that have been found to assist in prediction. Many interact and most have differential effects on a cut by cut basis. The individual cut by cook grade results are created by differential use of the input criteria.

The model has been revised over time since its' initial release in 2000 with a new version currently being developed for scientific approval in April. The current data includes over 45,000 cuts each evaluated by 10 consumers including testing in the USA, Korea, Japan, Northern Ireland and Ireland. Over 85,000 consumers have participated in testing, each scoring 7 cooked beef samples. This is the largest consumer testing program ever conducted in beef and is ongoing. It has established that people are remarkably consistent in their sensory response to beef which means that a grading system based on this science can be expected to meet consumer needs and create value.

Further analysis of willingness to pay data collected in conjunction with testing in USA, Japan, the Irish Republic and Australia confirms that consumers place substantial value differences on beef placed within alternate grade levels. Table 3 presents this data with the nominated price for each grade presented as a percentage of the 3* "good everyday" value.

Table 3. Nominated price – relationships by grade level.

	n	Fail	3*	4*	5*
Aust (2006)	2280	44.9%	0.0%	145.5%	197.4%
Aust (2008/9)	2400	47.5%	0.0%	145.9%	199.3%
Japan	1620	52.4%	0.0%	167.2%	293.6%
USA	1440	47.2%	0.0%	153.0%	214.5%
Irish	1380	48.1%	0.0%	131.0%	164.0%

The net result of the research work is that Australia now has a unique grading system which is acknowledged as the global benchmark by a substantial margin. No other system attempts to predict consumer satisfaction by meal portion. Logical application of the outcomes challenges traditional description systems at all levels of the trading chain from farm to retail counter. Traditional retail purchasing relies on a mix of retailer reputation, cut name, pricing and raw meat appearance. Many consumers find these parameters challenging and confusing allied with insecurities regarding their cut or cooking knowledge. Experience has shown that they achieve a variable outcome from identically described product. This is not their fault; the product does vary and the information provided cannot separate eating quality levels.

We have the tools to supplant traditional description with a system that is simple and effective. Here is a 3* roast or 4* steak with no further knowledge needed. The quality grade can be established by the grading process. This can be used directly at retail or indirectly as a silent brand support where a Retailer X premium brand might be only sourced by product grading 4* or to a specification such as over 70 MQ4 points. This offers exciting opportunity to transform the beef category by making beef purchasing simple and effective. If retail value points are accurately transmitted via payment systems to the processor and farmer the entire chain will adopt a consumer focus.

While the base science has been accepted within the industry detailed application has been variable and least visible at consumer level. In my opinion the quality of Australian beef has improved due to understanding and partial adoption of the science although we are yet to utilise the full potential.

Efforts to improve retail description are commendable and sensible for the consumer and the entire industry. In my view however they are only worthwhile to the extent that they provide a genuine indication of cooked meal performance. Categorisation based on simplified carcass description measures, however well intentioned, cannot work and may add to confusion. We have the worlds' best knowledge in this area and should build any proposed new retail description framework on its' detailed application. A simple retail description accurately conveying the cooked meal result empowers the consumer, provides an accurate value proposition and is able to be implemented with existing technology. The responsibility for an eating quality outcome moves from the customer to the supply chain which is how it should be.

Representation of MSA model inputs and outputs.

Description	Format	Name	Input
Estimated % Bos Indicus	% or X if doubt	**EPBI**	**25**
Animal Sex Type	M/F	**Sex**	**M**
Hormone Growth Promotent	Y or ? / N	**HGP**	**N**
MilkFedVealer	Y/N	**MFV**	**N**
SaleYard	Y/N	**SlYrd**	**N**
Rinse/Flush	Y/N	**RnFl**	**N**
Hot Std Carcase Weight	Weight in Kg	**HSCW**	**280**
HangMethod	AT/TS/TL/TC/TX	**Hang**	**TS**
Hump Height	mm	**Hump**	**5**
Ossification USDA	USDA measure	**uoss**	**120**
Marbling USDA	USDA measure	**umb**	**350**
RibFat	mm	**RbFt**	**5**
Ulitimate pH	Metered pH	**UpH**	**5.5**
Loin Temp at Grade	Metered Temp C	**Utmp**	**3**
Days of Ageing from Kill	Days Aged	**Age**	**21**

cut	muscle	GRL	RST	SFR	TSL	SCT	CRN
spinalis	SPN081	5	4	5	4		
tenderloin	TDR034	5		4			
tenderloin	TDR062	5	4	5	4		
tenderloin	TDR063	4					
cube roll	CUB045	4	4	4	4		
striploin	STA045	4	4	4	4		
striploin	STP045	4	4	4	4		
oyster blade	OYS036	3	3	4	4		
blade	BLD095						
blade	BLD096	3	3	3	3	4	
chucktender	CTR085		3	3	3	4	
rump	RMP131	4	4	4	4	4	
rump	RMP231	4	4	4	4		
rump	RMP005	4		4	5		
rump	RMP032			4	5		
rump	RMP087		3	4	4	4	
knuckle	KNU066	3	4	4	4	3	
knuckle	KNU098			4	4	4	
knuckle	KNU099	3	3	3	3	3	
knuckle	KNU100			4	4	4	
outside flat	OUT005		3	3	4	4	3
outside flat	OUT029			3	4	3	
eye round	EYE075	3	3	3	3	3	3
topside	TOP001	3		3	4	3	
topside	TOP033	3		3	4	4	
topside	TOP073	3	3	3	4	4	
chuck	CHK068			3	3	4	
chuck	CHK074	4	3	3	4	4	
chuck	CHK078	3	3	3	3	4	
chuck	CHK081			3	4		
chuck	CHK082			3	3		
thin-flank	TFL051			3		3	
thin-flank	TFL052			4	4		
thin-flank	TFL064			4	3	4	
rib-blade	RIB041			3			
brisket	BRI056				3	3	
brisket	BRI057				3	4	
shin	FQshin					4	
shin	HQshin					4	
intercostal	INT037			3			

References:

Polkinghorne R (2005) Does variation between muscles in sensory traits preclude carcass grading as a useful tool for consumers? In *Proceedings of 51st international congress of meat science and technology, 7 – 12 August 2005, Baltimore, USA.*

–