Animal (2017), **11:8**, pp 1399–1411 © The Animal Consortium 2017
doi:10.1017/S1751731117000076

animal

Untrained consumer assessment of the eating quality of European beef: 2. Demographic factors have only minor effects on consumer scores and willingness to pay

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

[1]*School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA 6150, Australia;* [2]*INRA, UMR1213, Recherches sur les Herbivores, F-63122 Saint Genès Champanelle, France;* [3]*VetAgro Sup, UMR1213, Recherches sur les Herbivores, Clermont Université, F-63122 Saint Genès Champanelle, France;* [4]*Teagasc Food Research Centre, Ashtown, Dublin 15, Ireland;* [5]*Institut de l'Elevage, Service Qualité' des Viandes, MRAL, 87060 Limoges Cedex 2, France;* [6]*Polish Beef Association Ul. Kruczkowskiego 3, 00-380 Warszawa, Poland;* [7]*Agri-Food and Biosciences Institute, Newforge Lane, Belfast BT9 5PX, UK;* [8]*431 Timor Road, Murrurundi, NSW 2338, Australia*

(Received 13 May 2016; Accepted 13 December 2016; First published online 13 February 2017)

The beef industry must become more responsive to the changing market place and consumer demands. An essential part of this is quantifying a consumer's perception of the eating quality of beef and their willingness to pay for that quality, across a broad range of demographics. Over 19 000 consumers from Northern Ireland, Poland, Ireland and France each tasted seven beef samples and scored them for tenderness, juiciness, flavour liking and overall liking. These scores were weighted and combined to create a fifth score, termed the Meat Quality 4 score (MQ4) (0.3 × tenderness, 0.1 × juiciness, 0.3 × flavour liking and 0.3 × overall liking). They also allocated the beef samples into one of four quality grades that best described the sample; unsatisfactory, good-every-day, better-than-every-day or premium. After the completion of the tasting panel, consumers were then asked to detail, in their own currency, their willingness to pay for these four categories which was subsequently converted to a proportion relative to the good-every-day category (P-WTP). Consumers also answered a short demographic questionnaire. The four sensory scores, the MQ4 score and the P-WTP were analysed separately, as dependant variables in linear mixed effects models. The answers from the demographic questionnaire were included in the model as fixed effects. Overall, there were only small differences in consumer scores and P-WTP between demographic groups. Consumers who preferred their beef cooked medium or well-done scored beef higher, except in Poland, where the opposite trend was found. This may be because Polish consumers were more likely to prefer their beef cooked well-done, but samples were cooked medium for this group. There was a small positive relationship with the importance of beef in the diet, increasing sensory scores by about 4% in Poland and Northern Ireland. Men also scored beef about 2% higher than women for most sensory scores in most countries. In most countries, consumers were willing to pay between 150 and 200% more for premium beef, and there was a 50% penalty in value for unsatisfactory beef. After quality grade, by far the greatest influence on P-WTP was country of origin. Consumer age also had a small negative relationship with P-WTP. The results indicate that a single quality score could reliably describe the eating quality experienced by all consumers. In addition, if reliable quality information is delivered to consumers they will pay more for better quality beef, which would add value to the beef industry and encourage improvements in quality.

Keywords: consumer testing, beef, quality, demographics, Europe

Implications

A single quality descriptor of beef eating quality will likely be applicable to the entire European market due to the small impact of demographics on consumer scores. This descriptor could form the basis of an eating quality-based grading system for beef which would allow consumers to select beef of a desired quality when purchasing beef. As European consumers are also willing to pay more for better quality beef, such a system would provide a price signal in the market, creating a financial incentive for producers to include eating quality in their management strategies.

[a] Present address: School of Veterinary and Life Sciences, Murdoch University, Murdoch, WA 6150, Australia.
[†] E-mail: S.Bonny@Murdoch.edu.au

Introduction

There is interest in developing an eating quality-based grading system for the European beef industry to reduce the variability in eating quality of European beef (Verbeke *et al*., 2010). It has already been shown that such a system would be well accepted (Hocquette *et al*., 2011) and could be based upon the Meat Standards Australia (MSA) model (Watson *et al*., 2008a) which uses predictors such as carcase weight, ossification, rib fat and intramuscular fat to predict consumer eating quality. Indeed, Bonny *et al*. (2016a and 2016b) have previously shown that this model functions well when using European beef, with only minor adjustments. However, it is unclear if a single quality descriptor would be applicable to all European consumers. Demographics are well established as factors that influence the beef quality scores and purchasing decisions of consumers (Berry and Hasty, 1982; Thompson *et al*., 2005). Therefore, these factors must be investigated in order to properly design taste panel experiments (Thompson *et al*., 2005) and to validate the use of a single quality descriptor for all consumers, as the basis of an eating quality beef grading system. Furthermore, willingness to pay information would allow the beef industry to take full advantage of a beef grading system based on eating quality with realistic price differentials, and highlighting groups of consumers who place a greater value on quality.

Previous work on Australian and Korean consumers identified only very minor demographic effects on sensory scores of beef and lamb (Thompson *et al*., 2005; Hwang *et al*., 2008). The main response was that consumers who considered beef to be a more important part of their diet scored lamb more favourably (Thompson *et al*., 2005). Thompson *et al*. (2005) also found a small difference between the genders, with men scoring beef around 2 points out of 100 lower than women. However, this trend is not consistent in the literature, Huffman *et al*. (1996) found no differences between the sexes when scoring beef, and Kubberød *et al*. (2002) found that men scored beef more favourably than women.

A consumer's preferred level of cooking doneness also has a small effect on consumer scores (Thompson *et al*., 2005) with consumers who preferred beef cooked medium-well or well-done scoring beef prepared medium about 2 points higher than consumers who preferred their beef cooked medium or rare. This appears to contradict the results of Cox *et al*. (1997) who found that consumers tasting beef cooked to their preferred level of cooking doneness rated beef higher. However, the Cox study was performed in restaurants where consumers ordered and paid for their steaks, potentially altering the result.

Previous studies have shown that Australian, American, Japanese, Irish and South African consumers are willing to pay at least twice as much for better quality beef (Lyford *et al*., 2010; Thompson *et al*., 2010), although there were small variations between different demographic groups. Consumer age was found to have a negative relationship with willingness to pay in three studies (Lusk *et al*., 2001; Lyford *et al*., 2010; Thompson *et al*., 2010). In contrast, Reicks *et al*. (2011) found that age had no effect on the importance of price when consumers were purchasing beef, though this survey was more directed to factors influencing purchasing decisions, rather than willingness to pay for different quality levels.

There are several other demographic factors (sex, occupation, number of children in the household or cooking method) that have been investigated and found to have no effect on willingness to pay for quality beef (Cox *et al*., 1997; Lyford *et al*., 2010; Thompson *et al*., 2010; Reicks *et al*., 2011). In addition, both Feuz *et al*. (2004) and Lusk *et al*. (2001) found no effect of income bracket on willingness to pay for American consumers. In addition, Feuz *et al*. (2004) and Reicks *et al*. (2011) found that these patterns were similar across different regions in the United States.

In this study, we explore the demographic effects on consumer scores from four different European countries, and willingness to pay from these same countries and from Australia. Based on the results of consumer testing in other countries, we hypothesise that there will be only small demographic effects on sensory scores. These effects will be limited to a positive relationship with the importance of beef in a consumer's diet and a small increase in sensory scores when consumers prefer their beef prepared medium-well and well-done. We also hypothesise that consumers will be willing to pay approximately double for beef of a premium quality compared to good-every-day product, and that this will not vary with demographic factors outside of a small negative relationship with consumer age.

Material and methods

Animals and muscle samples

The carcasses used for this experiment were described in detail by Bonny *et al*. (2016a) and Legrand *et al*. (2013). Briefly, the data set was formed through combining the records of a number of specific, smaller, experiments. As a result, this data set provides a cross-section of commonly used European cattle types from France, Poland, Ireland and Northern Ireland. The cattle were slaughtered commercially according to standard practice in each country. There was a range of 5 to 28 days *postmortem* ageing for the samples, and all samples were wet aged. A total of 25 different muscles were collected, which reflected a wide range of different eating qualities, though not all muscles were collected from each carcass.

Meat preparation and consumer panels

Meat preparation and consumer assessment of eating quality for the four cooking methods were performed according to the protocols for MSA testing by personnel trained in MSA testing procedures (Anonymous, 2008; Watson *et al*., 2008a) in each of the participating countries, France, Ireland, Northern Ireland and Poland. Each consumer only tasted beef cooked by a single cooking method and degree of cooking doneness. The slow cooking method was only used in Poland and the Korean BBQ (barbeque) was tested only in Ireland. The grill cooking

method was performed in all countries and the roast cooking method was performed in all countries except for France. Grill samples were prepared to either a rare or a medium doneness in France. Legrand *et al.* (2013) has previously demonstrated that there were no significant differences in consumer responses between the two levels of cooking doneness used in this experiment. In Northern Ireland, the roast and grill samples were prepared to either a medium or a well-done cooking doneness (Anonymous, 2008; Bonny *et al.*, 2016b). All other samples were prepared to a medium cooking doneness with cooking doneness determined by a combination of consistent sample size, cooking temperature and time (Anonymous, 2008; Bonny *et al.*, 2016b).

For each cooking method, consumers received seven portions: the first portion (a 'link' sample) was derived from either a generic striploin or rump muscle and expected to be of average quality – the sensory scores for this portion were not part of the final statistical analysis. Each of the remaining six portions was derived from one of the muscle samples collected. These were taken from a variety of different muscles and selected to present each consumer with a diverse quality range from unsatisfactory to premium. These were served in accordance with a 6×6 Latin square to balance potential order effects (Thompson *et al.*, 2005; Hwang *et al.*, 2008). Consumers scored meat from their country of origin except for two sessions where meat was tested between Poland and France in a complete factorial design, and a series of sessions where French consumers tasted Australian beef. In all countries, consumers were sourced through both commercial consumer testing organisations and local clubs and charities. They were selected to reflect the general population. Consumers scored samples for tenderness, juiciness, flavour liking and overall liking, by making a mark on a 100 mm line scale, with the low end of the scale representing a negative response and the high end of the scale representing a positive response. Consumers were also asked to place the beef in one of four categories; unsatisfactory, good-every-day, better-than-every-day or premium. After the completion of the tasting panel, consumers were then asked to detail, in their own currency, their willingness to pay for these four quality categories. All consumers indicated their willingness to pay by marking a point on a line, except for the Irish consumers who were required to tick a box indicating a fixed value. For a more detailed description of the testing procedures and the questionnaire, see Anonymous (2008).

Consumer demographics

In addition to scoring beef samples, consumers answered a short demographic questionnaire in their native language. The English version of this questionnaire is detailed elsewhere (Anonymous, 2008). This included questions about their age, gender, occupation, number of children and adults in the household and total income of the household. Not all taste panel sessions gathered information on willingness to pay; therefore, the numbers of consumers differed between the analyses examining the sensory scores and willingness to pay. In addition, the Australian consumers were not included in the analysis of the sensory scores. The distribution of the demographics for the sensory score analysis is detailed by Bonny *et al.* (2016c). In brief, there were 19 492 consumers and in all countries women outnumbered the men, except for Ireland. The majority of consumers came from households with 1 to 3 adults and 0 to 2 children (Bonny *et al.*, 2016c). This is similar to the distribution of the demographics for the willingness to pay analysis (Table 1). In Australia, there were a greater number of consumers aged over 60 years than any other age group, and information regarding income was not collected (Table 1). Consumers were also asked to rate the importance of beef in their diet, their usual frequency of eating beef and their preferred level of doneness (Table 2). In all countries, most consumers ate beef at least once a week. In most countries, the majority of consumers considered red meat to be at least a regular part of their diet, whereas in Poland there was a more even spread of consumers over the 'importance of beef in the diet' categories than other countries (Table 2).

Statistical analysis of sensory scores

The four sensory scores (tenderness, juiciness, flavour liking, overall liking) were weighted (0.3, 0.1, 0.3, 0.3) and combined to create a fifth score termed the Meat Quality 4 (MQ4) score. These weightings were used in this analysis as they are the weightings currently used in the MSA beef grading system in Australia (Thompson *et al.*, 2010). The important results from this analysis did not differ when the original weightings for the sensory scores (0.4, 0.1, 0.2, 0.3) were used. The effect of demographic factors on the four sensory scores and the MQ4 score was investigated using separate linear mixed effects models with the HPMIXED procedure in SAS (SAS, 2002). The random terms were beef sample identification number, consumer identification number within country and session, and session within country. The experimental design factors of country and sample serve order (2 to 7) were included in the model as fixed effects. This structure accounted for the variance associated with consumer and session, with each consumer providing six responses, and each session consisting of a separate set of 60 consumers. The score of the previous sample was included as a covariate to test for carry-over effects. Demographic effects were included in the model as fixed effects. These were age, income, gender, occupation, number of children in the household, number of adults in the household, importance of red meat in the diet, frequency of eating beef and the preferred level of doneness. Income and age were fitted within country as different scales were used in different countries. All factors in the model were interacted with country, and the score of the previous sample was also interacted with sample serve order. Non-significant terms ($P > 0.05$) were then removed in a step-wise fashion to arrive at the final model. The predicted means for the demographic effects were compared using the LSD, generated using the PDIFF function to perform pairwise *t*-tests in SAS (SAS, 2002).

Table 1 *Number of consumers who scored beef samples (and percentage distribution) within each of the demographic categories for each country for the willingness to pay analysis*

Demographic categories	Total number of individual consumers in each category									
Gender	Men	Women	UR							
Australia	148 (43.7)	191 (56.3)	0 (0)							
France	672 (45.0)	822 (55.0)	1 (0.07)							
Ireland	615 (51.4)	539 (45.0)	43 (3.59)							
Northern Ireland	1643 (45.7)	1941 (54.0)	15 (0.42)							
Poland	2647 (44.0)	3367 (55.9)	13 (0.22)							
Income[1]	*a*	*b*	*c*	*d*	*e*	UR				
Australia	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	339 (100)				
France	128 (8.56)	443 (29.6)	493 (33.0)	300 (20.1)	129 (8.63)	2 (0.13)				
Ireland	129 (10.8)	487 (40.7)	552 (46.1)	0 (0)	0 (0)	29 (2.42)				
Northern Ireland	995 (27.7)	2014 (56.0)	542 (15.1)	0 (0)	0 (0)	48 (1.33)				
Poland	661 (11.0)	851 (14.1)	1866 (31.0)	1858 (30.8)	758 (12.6)	33 (0.55)				
Occupation	Trade	Professional	Admin	Technical	Service	Labourer				
Australia	43 (12.7)	95 (28.0)	52 (15.3)	31 (9.14)	38 (11.2)	10 (2.95)				
France	39 (2.61)	231 (15.5)	540 (36.1)	129 (8.63)	0 (0)	100 (6.69)				
Ireland	92 (7.69)	377 (31.5)	162 (13.5)	181 (15.1)	66 (5.51)	12 (1)				
Northern Ireland	389 (10.8)	937 (26.0)	675 (18.8)	319 (8.86)	240 (6.67)	51 (1.42)				
Poland	240 (3.98)	410 (6.8)	1256 (20.8)	400 (6.64)	689 (11.4)	721 (12.0)				
	Unemployed	Student	Retired	Homemaker	Other	UR				
Australia	4 (1.18)	56 (16.5)	6 (1.77)	4 (1.18)	0 (0)	0 (0)				
France	82 (5.48)	82 (5.48)	256 (17.1)	26 (1.74)	8 (0.54)	2 (0.13)				
Ireland	24 (2.01)	141 (11.8)	0 (0)	126 (10.5)	0 (0)	16 (1.34)				
Northern Ireland	112 (3.11)	494 (13.7)	0 (0)	354 (9.84)	0 (0)	28 (0.78)				
Poland	182 (3.02)	957 (15.9)	0 (0)	89 (1.48)	0 (0)	1083 (18.0)				
Adults in the home	0	1	2	3	4	5+	UR			
Australia	0 (0)	29 (8.55)	207 (61.1)	64 (18.9)	29 (8.55)	10 (2.95)	0 (0)			
France	4 (0.27)	309 (20.7)	875 (58.5)	188 (12.6)	83 (5.55)	33 (2.21)	3 (0.2)			
Ireland	0 (0)	88 (7.35)	511 (42.7)	267 (22.3)	217 (18.1)	109 (9.11)	5 (0.42)			
Northern Ireland	165 (4.58)	450 (12.5)	1497 (41.6)	720 (20.0)	520 (14.5)	242 (6.72)	5 (0.14)			
Poland	43 (0.71)	836 (13.9)	2060 (34.2)	1515 (25.1)	1055 (17.5)	516 (8.56)	2 (0.03)			
Children in the home	0	1	2	3	4	5+	UR			
Australia	118 (34.8)	56 (16.5)	116 (34.2)	39 (11.5)	9 (2.65)	1 (0.29)	0 (0)			
France	954 (63.8)	240 (16.1)	222 (14.9)	62 (4.15)	13 (0.87)	2 (0.13)	2 (0.13)			
Ireland	486 (40.6)	277 (23.1)	179 (15.0)	134 (11.2)	41 (3.43)	22 (1.84)	58 (4.85)			
Northern Ireland	1 (0.03)	2319 (64.4)	477 (13.3)	505 (14.0)	185 (5.14)	80 (2.22)	32 (0.89)			
Poland	4349 (72.2)	1121 (18.6)	430 (7.13)	92 (1.53)	29 (0.48)	5 (0.08)	1 (0.02)			
Age group (years)	15 to 29	20 to 44	45 to 49	⩾60	UR					
Australia	35 (10.3)	37 (10.9)	82 (24.2)	185 (54.6)	35 (10.3)					
Northern Ireland	967 (26.9)	829 (23.0)	1012 (28.1)	3 (0.08)	967 (26.9)					
Age group (years)	20 to 25	26 to 30	31 to 35	36 to 40	41 to 45	46 to 50	51 to 55	56 to 60	⩾61	UR
France	300 (20.1)	163 (10.9)	114 (7.6)	163 (11.1)	0 (0)	287 (19.2)	119 (7.96)	149 (9.97)	196 (13.1)	1 (0.07)
Ireland	274 (22.9)	173 (14.5)	112 (9.4)	109 (9.11)	137 (11.5)	145 (12.1)	110 (9.19)	104 (8.69)	26 (2.17)	7 (0.58)
Age group (years)	<20	20 to 25	26 to 30	31 to 39	40 to 50	⩾51	UR			
Poland	383 (6.4)	2557 (42.4)	844 (14.0)	786 (13.0)	701 (11.6)	748 (12.4)	8 (0.13)			

UR = unreported; Admin = administration.
[1]Income categories are different for each country. In all countries 0 = unreported; France (€/month) – *a*: <1000, *b*: 1000 to 2000, *c*: 2000 to 3000, *d*: 3000 to 4000, *e*: >4000; Ireland (€/year) – *a*: <20 000, *b*: 20 000 to 50 000, *c*: >50 000; Northern Ireland (£/year) – *a*: <20 000, *b*: 20 000 to 50 000, *c*: >50 000; Poland (zł/month) – *a*: ⩽1000, *b*: 1001 to 1400, *c*: 1401 to 2200, *d*: 2201 to 4000, *d*: >4000.

Table 2 *Number of consumers within each of the demographic categories for the willingness to pay (and sensory score[1]) analyses, outlining the role of meat in their diet for each country*

Demographic traits	Total number of individual consumers in each category							
Frequency of eating beef[2]	7	4 to 5	2 to 3	1	0.5	0.25	Never	Unreported
Australia	10	51	180	85	13	0	0	0
France	24 (19)	225 (169)	757 (520)	377 (208)	107 (37)	4 (3)	1 (0)	0 (1)
Ireland	42 (46)	181 (214)	561 (648)	305 (314)	58 (64)	45 (50)	5 (5)	0 (9)
Northern Ireland	192 (470)	576 (1471)	1763 (4026)	734 (1649)	201 (422)	118 (263)	15 (0)	0 (42)
Poland	25 (31)	134 (162)	1216 (1416)	1795 (2134)	1352 (1663)	1420 (1740)	85 (0)	0 (110)
Importance[3]	a	b	c	d	Unreported			
Australia	200	127	12	0	0			
France	411 (315)	871 (561)	211 (82)	2 (2)	0 (0)			
Ireland	462 (534)	485 (551)	210 (243)	27 (34)	13 (18)			
Northern Ireland	1282 (3027)	1522 (3486)	667 (1544)	88 (233)	40 (102)			
Poland	1183 (1393)	1423 (1696)	1882 (2260)	1446 (1795)	93 (116)			
Doneness[4]	Blue	Rare	Med/rare	Medium	Med/well	Well-done	Unreported	
Australia	0	0	128	95	116	0	0	
France	52 (36)	330 (163)	0 (0)	919 (646)	166 (87)	28 (28)	0 (0)	
Ireland	0 (0)	139 (149)	120 (139)	308 (352)	288 (334)	333 (392)	9 (14)	
Northern Ireland	24 (50)	106 (208)	525 (1138)	715 (1570)	797 (1952)	1412 (3415)	20 (59)	
Poland	229 (269)	146 (169)	254 (324)	1654 (2020)	2918 (3495)	798 (950)	28 (33)	

[1]The numbers in brackets indicate the numbers of consumers for each category used in the analysis investigating the effect of demographics on sensory scores.
[2]Number of meals containing red meat eaten by the consumer in an average week.
[3]Importance of beef in the diet.
a = red meat is an important part of my diet; b = red meat is a regular part of my diet; c = red meat is part of my diet but it would not worry me if it was not; d = I rarely/never eat red meat.
[4]Preferred cooking doneness.

Statistical analysis of willingness to pay

As the currencies differed between countries, willingness to pay was expressed as a proportion of good-every-day (P-WTP), for each quality grade; unsatisfactory, good-every-day, better-than-every-day and premium, as previously described by Lyford *et al.* (2010). This was analysed as the dependent variable in a linear mixed effects model, using the MIXED procedure in SAS (SAS, 2002). The random terms were consumer identification number, within country and session. As for the sensory score analysis, the demographic factors were included as fixed effects. The country where the session attended by the consumer and the cooking method, including degree of doneness, used in the session were also included as fixed effects. All fixed effects were interacted with country and quality grade. Non-significant terms ($P > 0.05$) were then removed in a step-wise fashion to arrive at the final model. The predicted means for the demographics effects were compared using the LSD generated using the PDIFF function in SAS (SAS v9.1).

Results and discussion

The effect of a consumer's preferred level of cooking doneness on sensory scores

Confirming our hypothesis, consumers who preferred their beef cooked medium-well or well-done scored beef more favourably than consumers who preferred their beef cooked medium or less ($P < 0.01$), although this effect varied between countries as the interaction between both factors was significant (Table 3). The group of consumers in Northern Ireland who preferred their beef cooked well-done or medium-well scored beef samples ~4 points higher ($P < 0.05$) than consumers who preferred their beef cooked blue/rare, and slightly higher than those who preferred medium (Table 4). This pattern was also present for the Irish consumers for tenderness, overall liking and MQ4, but to a lower extent ($P < 0.05$). In contrast, the Polish consumers who preferred their beef cooked well-done or medium-well scored beef samples ~3 points lower ($P < 0.05$) than consumers who preferred their beef cooked blue, rare or medium (Table 4). There was no effect of preferred cooking doneness for the French consumers.

This result is supported by Hwang *et al.* (2008) and Thompson *et al.* (2005) who found a similar trend in Australian consumers. In contrast, the Polish consumers exhibited the opposite relationship with consumers who preferred beef cooked medium-well or well-done scoring samples less favourably. This may partially be explained by variations in the different degrees of cooking doneness used in this study, as consumers' rate beef cooked to their preferred cooking doneness higher (Cox *et al.*, 1997). More Northern Irish and Polish consumers preferred beef cooked medium-well to well-done than any other category. As the majority of Northern Irish consumers also tasted beef cooked well-done, it would be expected that there would be a positive

Table 3 *The* F*-values for the linear mixed effects model, predicting MQ4[1] and sensory scores for beef samples*

Variables	NDF[2]	Tenderness	*P*	Juiciness	*P*	Flavour liking	*P*	Overall liking	*P*	MQ4[1]	*P*
Country	3	8.01	<0.0001	28.8	<0.0001	9.92	<0.0001	37.8	<0.0001	28.1	<0.0001
Order[3]	5	154	<0.0001	155	<0.0001	150	<0.0001	147	<0.0001	163	<0.0001
Age (country)	27	1.75	0.0094	1.97	0.0019	–	–	–	–	–	–
Gender	2	0.29	0.7497	1.62	0.197	9.35	<0.0001	14.4	<0.0001	9.17	<0.0001
Occupation	9	–	–	–	–	1.99	0.0365	2.28	0.015	–	–
Adults[4]	5	2.61	0.0229	–	–	–	–	2.41	0.0339	–	–
Children[5]	6	–	–	1.58	0.1475	–	–	–	–	–	–
Income (country)	18	–	–	1.78	0.0218	–	–	–	–	–	–
Importance[6]	3	0.82	0.4849	4.30	0.0049	32.1	<0.0001	30.9	<0.0001	8.05	<0.0001
Doneness[7]	3	2.21	0.0842	4.59	0.0032	1.15	0.3263	1.88	0.1309	2.16	0.0899
Carry-over[8]	1	1.93	0.1652	58.9	<0.0001	46.6	<0.0001	80.2	<0.0001	89.0	<0.0001
Carry-over[8]*carry-over[8]	1	13.5	0.0002	20.3	<0.0001	33.2	<0.0001	89.2	<0.0001	104	<0.0001
Order[3]*Country	15	3.15	<0.0001	–	–	3.98	<0.0001	4.49	<0.0001	4.34	<0.0001
Carry-over[8]*country	3	91.6	<0.0001	17.3	<0.0001	19.6	<0.0001	16.1	<0.0001	15.2	<0.0001
Carry-over[8]*order[3]	5	13.3	<0.0001	99.2	<0.0001	95.1	<0.0001	95.1	<0.0001	108.4	<0.0001
Gender*country	6	2.58	0.0169	3.70	0.0012	3.05	0.0055	–	–	–	–
Children[5]*country	18	–	–	1.70	0.0305	–	–	–	–	–	–
Importance[6]*country	9	1.98	0.0369	2.40	0.0093	–	–	–	–	2.13	0.0241
Doneness[7]*country	9	5.70	<0.0001	4.90	<0.0001	9.68	<0.0001	10.53	<0.0001	9.31	<0.0001

NDF = Numerator degrees of freedom; MQ4 = Meat Quality 4.
[1]MQ4 = a weighted combination (0.3, 0.1, 0.3, 0.3) of four sensory scores, tenderness, juiciness, flavour liking and overall liking.
[2]Denominator degrees of freedom is 111 000.
[3]The order in which the product was served to the consumer.
[4]Number of adults in the household.
[5]Number of children in the household.
[6]The importance of beef in their diet.
[7]The preferred degree of cooking doneness of the consumer.
[8]The sensory score of the previously tasted sample.

relationship between preferred cooking doneness and consumer scores. However, while a large proportion of Polish consumers also preferred beef cooked medium-well to well-done, the Polish consumers tasted beef cooked medium, possibly underpinning the negative relationship between cooking doneness and consumer scores seen for this group. However, this theory does not explain the behaviour of the Irish consumers; therefore, other factors, such as preferred cooking method, may also be influencing the results. The lack of response identified for the French consumers may have resulted from the uneven distribution of consumers over the doneness categories, with the overwhelming majority preferring their beef cooked medium. This should not be taken as an indication of the general population however, as consumers who preferred their beef cooked medium were actively recruited for this part of the study.

The effect of the importance of red meat on sensory scores

The hypothesis that consumers would score beef more favourably if they considered red meat to be a more important part of their diet was supported by consumers in Poland, France and Northern Ireland, but not by those tested in Ireland.

For most of the sensory scores, the more important consumers considered red meat in their diet, the more favourably ($P < 0.01$) they scored beef (Table 3). This effect was the most pronounced for the French consumers, with a change in the average sensory scores by over 19 points out of 100 for MQ4, and 4 points for tenderness and juiciness ($P < 0.05$) (Table 5). The responses for both the Northern Irish and Polish consumers were small, with sensory scores changing by 1 to 3.4 points ($P < 0.05$) over the range of importance tested (Table 5). The magnitude of the effect in Poland and Ireland is similar to the findings of Thompson *et al.* (2005), who used the same technique with Australian consumers tasting lamb. The sensory scores for overall liking and flavour changed by 3 points over the range and did not vary by country ($P < 0.05$).

Notably, the effect seen for the French consumers was much larger than for the other groups. This result should be treated with caution due to the poor spread of French consumers over the four possible responses, with only 0.13% in the least important category. In contrast, the Polish data had between 20% and 30% of consumers in each category. Further investigation with a more balanced distribution of consumers is required to fully quantify the effect of the importance of meat in the diet for French consumers on their perception of the eating quality of beef.

In Ireland, there were no relationships detected between sensory scores and the importance of beef in the diet, except for flavour liking and overall liking. Therefore, we reject our hypothesis for this group. This result is supported by work on Korean and Australian consumers by Hwang *et al.* (2008) who also found no relationship. The spread of

Table 4 *Predicted sensory score means (± SE) of beef samples by a consumer's preferred level of doneness*

Countries	Rare/blue	Medium	Medium-well	Well-done
MQ4[1]				
Average	51.5 ±0.85[a]	52.6 ± 0.73[b]	53.0 ± 0.79[b]	53.1 ± 0.91[b]
France	51.8 ± 2.65	51.9 ± 2.29	51.9 ± 2.59	54.4 ± 3.16
Ireland	52.4 ± 1.27[a]	54.3 ± 0.95[ab]	55.7 ± 0.99[b]	55.0 ± 0.94[b]
Northern Ireland	47.7 ± 0.82[a]	50.6 ± 0.54[b]	51.5 ± 0.55[c]	51.9 ± 0.52[c]
Poland	54.0 ± 0.75[ab]	53.8 ± 0.53[a]	53.1 ± 0.51[b]	51.1 ± 0.61[c]
Overall				
Average	53.1 ± 0.78[a]	54.3 ± 0.63[b]	54.5 ± 0.7[b]	54.7 ± 0.86[ab]
France	56.5 ± 1.71	56.3 ± 1.08	56.3 ± 1.65	59.2 ± 2.56
Ireland	52.0 ± 1.27[a]	54.0 ± 0.91[ab]	55.0 ± 0.97[b]	54.4 ± 0.93[ab]
Northern Ireland	48.0 ± 0.93[a]	51.2 ± 0.64[b]	52.2 ± 0.65[c]	52.6 ± 0.63[c]
Poland	55.7 ± 0.87[ab]	55.6 ± 0.66[a]	54.7 ± 0.64[b]	52.3 ± 0.73[c]
Tenderness				
Average	49.4 ± 1.37[a]	50.2 ± 1.27[ab]	51.1 ± 1.32[b]	51.0 ± 1.42[ab]
France	50.3 ± 4.63	49.7 ± 4.34	50.7 ± 4.58	51.6 ± 5.02
Ireland	51.7 ± 1.62[a]	53.4 ± 1.31[a]	55.7 ± 1.37[b]	55.3 ± 1.31[b]
Northern Ireland	45.9 ± 1.20[a]	48.7 ± 0.95[b]	49.3 ± 0.96[bc]	49.8 ± 0.93[c]
Poland	49.9 ± 1.75[a]	49.2 ± 1.64[a]	48.7 ± 1.63[a]	47.4 ± 1.67[b]
Flavour				
Average	54.2 ± 1.00	55.3 ± 0.88	55.2 ± 0.95	55.4 ± 1.06
France	58.7 ± 3.32	58.7 ± 3.04	57.5 ± 3.32	62.0 ± 3.84
Ireland	53.9 ± 1.27	55.8 ± 0.91	56.7 ± 0.97	55.8 ± 0.93
Northern Ireland	48.7 ± 0.99[a]	51.2 ± 0.72[b]	52.0 ± 0.74[c]	52.1 ± 0.71[c]
Poland	55.6 ± 1.47[a]	55.4 ± 1.36[a]	54.4 ± 1.35[b]	51.7 ± 1.39[c]
Juiciness				
Average	51.7 ± 1.49[a]	53.6 ± 1.41[b]	54.3 ± 1.46[b]	53.9 ± 1.54[b]
France	54.1 ± 5.17	54.0 ± 4.95	53.9 ± 5.17	53.1 ± 5.54
Ireland	52.4 ± 1.64	56.0 ± 1.35	57.6 ± 1.39	57.3 ± 1.33
Northern Ireland	43.0 ± 1.71[a]	46.7 ± 1.56[b]	48.0 ± 1.57[c]	48.4 ± 1.55[c]
Poland	57.2 ± 1.84[a]	57.7 ± 1.74[a]	57.8 ± 1.73[b]	56.7 ± 1.76[c]

MQ4 = Meat Quality 4.
[a,b,c]Values within a row with different superscript letters differ significantly at $P < 0.05$.
[1]MQ4 = a weighted combination (0.3, 0.1, 0.3, 0.3) of four sensory scores, tenderness, juiciness, flavour liking and overall liking.

consumers across the four important classes was very similar for the Irish and Northern Irish data, consequently, we expect that the difference in the relationships found are more likely related to actual differences in consumer behaviour. This is supported by the work of Lorenzen *et al.* (1999) and Neely *et al.* (1999) who both found that consumers in different geographical areas scored beef sensory quality differently. However, it is also possible that the analysis was not sensitive enough to pick up such a small effect in the smaller number of consumers tested in Ireland.

The effect of gender on sensory scores

Contrary to our hypothesis, men scored beef samples more favourably than women where a significant difference was found. Men scored beef samples higher than women by about 1 point out of 100 for overall liking and MQ4 ($P < 0.05$) (Table 6). A similar effect was seen for flavour ($P < 0.05$) and juiciness ($P < 0.05$), but only among Irish and Northern Irish consumers where men scored beef samples 1 to 2 points higher. For tenderness, only the Polish showed a difference between genders, with men scoring about 1 point higher ($P < 0.05$).

The difference we found between men and women is supported by Gregory (1997) and Kubberød *et al.* (2002) who also found that men scored meat more favourably than women. However, although this trend was also seen for the other sensory scores, it was not seen consistently across countries. One explanation for the variable responses found between countries could be linked to a consumer's perception of the importance of red meat in their diet. Kubberød *et al.* (2002) in the same study also found that men had a more positive attitude towards red meat. As our study and other previous studies have shown that consumers who consider meat to be important in their diet score beef more favourably, it is possible that the differences in the sexes reported by Kubberød *et al.* (2002) may be confounded by the consumer's attitude towards red meat. In our case, we found that Irish men were more likely to score beef as very important in their diet (data not shown), and Polish women were more likely to never/rarely eat red meat (data not shown). All other categories had fairly even distributions.

Table 5 *Predicted sensory score means (± SE) of beef samples by the importance of red meat in a consumer's diet*

Countries	1	2	3	4
MQ4[1]				
Average	54.7 ±0.53[a]	54.0 ± 0.52[b]	52.8 ± 0.61[c]	48.7 ± 2.19[c]
France	58.7 ± 1.23[a]	57.2 ± 1.13[ac]	54.7 ± 1.60[bc]	39.4 ± 8.33[b]
Ireland	54.4 ± 0.80	53.7 ± 0.81	54.1 ± 0.98	55.1 ± 2.10
Northern Ireland	51.9 ± 0.52[a]	51.5 ± 0.52[a]	49.8 ± 0.56[b]	48.5 ± 0.89[b]
Poland	53.9 ± 0.58[a]	53.5 ± 0.57[a]	52.8 ± 0.55[b]	51.9 ± 0.57[c]
Overall				
Average	55.6 ± 0.64[a]	54.8 ± 0.64[b]	53.5 ± 0.65[c]	52.6 ± 0.69[d]
Tenderness				
Average	52.0 ± 1.14[a]	51.8 ± 1.13[a]	50.8 ± 1.19[b]	47.2 ± 2.68[ab]
France	56.4 ± 3.75[a]	55.3 ± 3.71[ab]	52.5 ± 3.95[b]	38.0 ± 10.11[ab]
Ireland	52.9 ± 1.18	53.1 ± 1.19	54.1 ± 1.34	56.0 ± 2.52
Northern Ireland	49.4 ± 0.94[a]	49.5 ± 0.94[a]	47.9 ± 0.97[b]	47.0 ± 1.25[b]
Poland	49.3 ± 1.66[a]	49.4 ± 1.65[a]	48.6 ± 1.66[ab]	47.9 ± 1.66[b]
Flavour				
Average	56.5 ± 0.90[a]	55.7 ± 0.90[b]	54.4 ± 0.91[c]	53.5 ± 0.94[d]
Juiciness				
Average	55.0 ± 1.30[a]	54.4 ± 1.30[b]	53.5 ± 1.35[b]	50.5 ± 2.66[ab]
France	59.4 ± 4.48[a]	58.0 ± 4.46[ab]	55.4 ± 4.67[b]	42.4 ± 10.07[ab]
Ireland	55.4 ± 1.23	54.7 ± 1.25	55.5 ± 1.38	57.7 ± 2.45
Northern Ireland	47.8 ± 1.55[a]	47.2 ± 1.56[a]	45.8 ± 1.57[b]	45.3 ± 1.74[b]
Poland	57.6 ± 1.75[a]	57.6 ± 1.75[a]	57.5 ± 1.76[a]	56.6 ± 1.75[b]

MQ4 = Meat Quality 4.
1 = red meat is an important part of my diet; 2 = red meat is a regular part of my diet; 3 = red meat is part of my diet but it would not worry me if it was not; 4 = I rarely/never eat red meat.
[a,b,c,d]Values within a row with different superscript letters differ significantly at $P < 0.05$.
Where the effect did not vary by country, only average values were reported.
[1]MQ4 = a weighted combination (0.3, 0.1, 0.3, 0.3) of four sensory scores, tenderness, juiciness, flavour liking and overall liking.

Therefore, the lack of a clear consistent gender effect for all sensory scores and countries may be related to the lack of a consistent gender bias in the importance of red meat in the diet of the consumers in this study.

The effect of the country on willingness to pay and proportional willingness to pay

Supporting our hypothesis, consumers in all countries showed a general trend for the willingness to pay to increase with the quality level (Table 7). This was most evident for the French consumers, increasing from €5 to €23, and least evident for the Irish consumers, where the difference between unsatisfactory and premium was only €3. The consumers from Northern Ireland were willing to pay a similar amount for unsatisfactory and good-every-day as the French consumers, but this trend did not continue, and they were only willing to pay €14.7 for premium beef (Table 7). The Australian consumers were willing to pay almost $30 for premium beef, but only $6.6 for the unsatisfactory category.

The *F*-values for the final model can be seen in Table 8. Country had a significant effect on P-WTP, and this varied by quality grade (Figure 1). All countries except Ireland were willing to pay around half the price for unsatisfactory beef than good-every-day. The French consumers were willing to pay proportionally more for both better-than-every-day (1.78) and premium beef (2.63), closely followed by the Australian consumers (1.56, 2.17) ($P < 0.05$). The results for the Australian consumers line up well with the findings of both Lyford *et al*. (2010) and Thompson *et al*. (2010). The Polish consumers were willing to pay almost double for premium beef (1.89) than good-every-day, and one and a half times more (1.38) for better-than-every-day ($P < 0.05$). Increasing beef quality had a smaller effect on P-WTP in Northern Ireland, with consumers willing to pay only 1.49 times as much for premium beef than good-every-day beef, and 1.25 for better-than-every-day beef ($P < 0.05$).

Consumers from Ireland went against the general trend of the other countries reported in this study. They were willing to pay proportionally more for better quality beef; however, this response was markedly smaller than for the other countries (Figure 1), and the only differences evident were between the unsatisfactory category compared with the better-than-every-day and premium categories ($P < 0.05$).

This is in contrast to the work by Lyford *et al*. (2010) who found that Irish consumers were willing to pay double for premium beef than good-every-day quality. However, for their analysis, Lyford *et al*. (2010) excluded all consumer responses in which the lower quality grades had a higher willingness to pay than the higher quality grades. This assumes that consumers were willing to pay more for better quality beef, and would cloud any other relationship. Such an

assumption was not used in this study so as not to bias the data. Second, the structure of the questionnaire differed between Ireland and the other countries: Irish consumers used a 'tick-box' questionnaire format instead of the 'mark the line' format used in the other countries in this experiment. Lyford *et al.* (2010) found that a 'tick the box' format significantly reduced the range of the responses and therefore blunted the willingness to pay results for Australian consumers. Subsequently, we would treat the result from Ireland with caution, and further work in that area would need to consider the effect of questionnaire format on the responses. Adding further weight to this idea, the Northern Irish consumers, which are geographically and culturally similar to the Irish consumers, had similar responses to the other countries in this experiment, which all used a 'mark the line' format for the questionnaire.

Table 6 *Predicted sensory score means (± SE) of beef samples by the consumer's gender*

Countries	Men	Women
MQ4[1]		
Average	53.1 ±0.60[a]	52.3 ± 0.60[b]
Overall		
Average	54.8 ± 0.50[a]	53.7 ± 0.50[b]
Tenderness		
Average	50.9 ± 0.94	50.7 ± 0.94
France	50.7 ± 3.11	50.9 ± 3.09
Ireland	53.1 ± 1.18	52.8 ± 1.19
Northern Ireland	50.4 ± 0.80	50.5 ± 0.80
Poland	49.5 ± 0.86[a]	48.7 ± 0.85[b]
Flavour		
Average	55.8 ± 0.35[a]	54.5 ± 0.35[b]
France	59.1 ± 1.10	58.0 ± 1.07
Ireland	55.7 ± 0.68[a]	53.3 ± 0.73[b]
Northern Ireland	52.7 ± 0.35[a]	51.4 ± 0.35[b]
Poland	55.6 ± 0.36	55.4 ± 0.35
Juiciness		
Average	53.8 ± 1.14	53.3 ± 1.14
France	53.7 ± 3.99	54.7 ± 3.98
Ireland	55.7 ± 1.26[a]	53.5 ± 1.25[b]
Northern Ireland	48.8 ± 1.48[a]	47.9 ± 1.48[b]
Poland	57.0 ± 1.13	56.9 ± 1.12

MQ4 = Meat Quality 4.
Where the effect did not vary by country, only average values were reported.
[a,b]Values within a row with different superscript letters differ significantly at $P < 0.05$.
[1]MQ4 = a weighted combination (0.3, 0.1, 0.3, 0.3) of four sensory scores, tenderness, juiciness, flavour liking and overall liking.

The effect of consumer age on proportional willingness to pay

Aligning with our hypothesis, consumer age had a negative relationship with P-WTP for four out of the five countries examined in this study (Table 9). For most countries, consumer age only had an effect on P-WTP for better-than every-day and premium quality grades (Table 9). For the Australian consumers, participants between 15 to 29 years had a higher P-WTP for better-than-every-day and premium than older consumers ($P < 0.05$). This is similar to Northern Ireland where consumers between 15 and 29 years had a higher P-WTP than older consumer groups for the premium quality ($P < 0.05$) and older consumers had a lower P-WTP for the better-than-every-day category. Following this same pattern, as consumers got older in Poland the P-WTP for better-than-every-day and premium beef decreased (Table 9). French consumers of different age groups differed in their P-WTP for both the better-than-every-day and premium categories. This aligns well with previous Australian data (Lyford *et al.*, 2010) and similar patterns have also been identified in Japanese, American and South African consumers (Feuz *et al.*, 2004; Lyford *et al.*, 2010; Thompson *et al.*, 2010). Further supporting this relationship, younger consumers' also out-bid older consumers for tender steaks in a mock auction in the United States (Lusk *et al.*, 2001).

Although there were many significant differences for the French consumers, the relationship between age and willingness to pay was less clear. The consumers aged between

Table 7 *Means ± SD of the raw willingness to pay values both in local currency*

	Australia[1]	France[2]	Ireland[2]	Northern Ireland[3]	Poland[4]
Local currency					
Unsatisfactory	6.62 ± 4.26	4.58 ± 3.95	20.0 ± 3.46	5.76 ± 2.69	14.1 ± 12.3
Good[5]	14.3 ± 5.60	11.2 ± 4.66	22.7 ± 4.94	10.3 ± 2.70	26.9 ± 11.7
Better[6]	21.3 ± 8.15	16.7 ± 5.86	22.6 ± 4.91	12.7 ± 2.71	37.0 ± 15.1
Premium	29.4 ± 11.2	23.0 ± 7.92	23.2 ± 4.88	14.7 ± 3.11	49.8 ± 21.0
Ratio					
Unsatisfactory	0.45 ± 0.23	0.40 ± 0.29	0.93 ± 0.29	0.57 ± 0.22	0.52 ± 0.35
Good[5]	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
Better[6]	1.53 ± 0.37	1.61 ± 1.59	1.06 ± 0.35	1.27 ± 0.29	1.43 ± 0.52
Premium	2.14 ± 0.64	2.26 ± 1.72	1.08 ± 0.33	1.49 ± 0.43	1.97 ± 0.93

[1]Australian dollars.
[2]Euros.
[3]Pounds.
[4]Złoty.
[5]Good-every-day.
[6]Better-than-every-day.

Table 8 *The* F*-values for the base model, predicting the ratio of willingness to pay for beef classed as good-every-day compared with unsatisfactory, better-than-every-day and premium*

Variables	NDF[1]	*F*-value	*P*-value
Quality grade[2]	3	455	<0.0001
Doneness[3] (country)	1	5.91	0.0151
Country	4	15.8	<0.0001
Age (country)	30	2.37	<0.0001
Income (country)	18	2.03	0.0059
Frequency[4]	4	1.22	0.3018
Quality grade*age (country)	90	3.21	<0.0001
Quality grade*income (country)	54	2.08	<0.0001
Quality grade*country	12	53.1	<0.0001
Country*frequency[4]	16	2.24	0.003
Quality grade*frequency[4]	12	1.01	0.4329
Quality grade*country*frequency[4]	48	2.05	<0.0001

NDF = nominator degrees of freedom.
[1]Denominator degrees of freedom = 38 000.
[2]Quality grade; unsatisfactory, good-every-day, better-than-every-day and premium.
[3]Final cooking doneness used in the consumer panel before the questionnaire, rare, medium or well-done.
[4]Frequency of eating beef from daily to never.

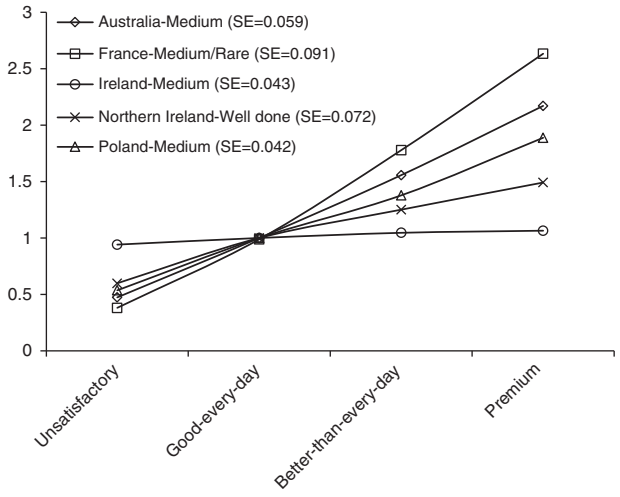

Figure 1 Willingness to pay, expressed as a ratio of the good-every-day grade, for each country (standard error over the continuum), adjusted for demographic and meat consumption preferences.

30 and 40 years tended to have had a higher P-WTP than the consumers aged between 40 and 60. Consumers aged above 60 appeared to have a similar P-WTP to the younger consumers. This may be a result of the greater number of age categories used for the French questionnaire allowing for the more subtle relationships to be elucidated, while concurrently reducing the number of consumers within each category. There is evidence that there is a curvilinear effect in the younger age groups, with P-WTP peaking around 35 years (Lyford *et al*., 2010). Therefore, with an increasing sample size this curvilinear relationship may become clearer for the French consumers. In contrast to other work by Lyford *et al*. (2010), Irish consumers demonstrated no relationship between consumer age and P-WTP. This may be due to the very small variations in P-WTP between the quality grades seen in this study reducing the ability of our statistical analysis to detect such relationships.

The effect of the frequency of beef consumption on proportional willingness to pay

In support of our hypothesis, there was no relationship between beef eating frequency and P-WTP in any of the countries tested except France. This is supported by Lyford *et al*. (2010), who also found no effect in Australia, Japan, the Unite States and Ireland. Consumers from France who ate beef fortnightly or less had a higher P-WTP by ~0.5 for both good-every-day and premium beef than consumers who ate beef more frequently (data not shown) ($P < 0.05$). In addition, French consumers who ate beef two to three times a week had a higher P-WTP for premium beef than consumers who ate beef weekly ($P < 0.05$), though this difference was much smaller (data not shown). The different behaviour of the French consumers may be cultural, with the consumers eating beef less frequently considering beef a premium or luxury product. Uncovering the exact motivations of the French consumers would require further investigation.

The effect of income on proportional willingness to pay

Validating our hypothesis, there was no relationship between income and P-WTP for the Australian, Irish and Northern Irish consumers. This is in alignment with the results of Lyford *et al*. (2010), Feuz *et al*. (2004) and Lusk *et al*. (2001) who also found no relationship between income and P-WTP for Australian, Japanese, American and Irish consumers. However, in contrast to this, we found that income significantly influenced P-WTP for the French and Polish consumers (data not shown), thus we reject our hypothesis for these groups. French consumers exhibited a slight decrease in P-WTP for better-than-every-day quality in the middle income groups (1.76 to 1.56, SE 0.09) ($P < 0.05$, data not shown). Similarly, French consumers with incomes of €1000 to €2000 and >€6000 per month were willing to pay proportionally more for premium beef than consumers in the middle income ranges (2.55 to 2.22, SE 0.09) ($P < 0.05$, data not shown). In contrast, the Polish consumers' P-WTP had a more direct relationship with income. P-WTP for premium beef increasing from 1.87 to 1.98 times good-every-day as income increased from zł 1001–1400 per month to zł 4000 per month and more ($P < 0.05$). This may indicate that the Polish and French consumers differ from the other countries or may be due to the different income brackets used for the countries reflecting different income levels, relative to gross domestic product, between the two countries. These results are supported by Reicks *et al*. (2011) who found that consumers with higher incomes did not consider price as important when purchasing beef. The positive relationship between income and P-WTP in both Poland and France is worth further investigation as it suggests there is a niche for high-quality branded products.

Table 9 *Predicted means for willingness to pay expressed as a ratio of good-every-day for the quality grade for each age group by country*

Quality grades	Age group (years)								
Australia	15 to 29	20 to 44	45 to 49	⩾60					
Unsatisfactory	0.42	0.48	0.46	0.54					
Good-every-day	1.00	1.00	1.00	1.00					
Better-than-every-day	1.78[a]	1.47[b]	1.52[b]	1.46[b]					
Premium	2.47[a]	1.99[b]	2.18[b]	2.04[b]					
SE[1]	0.109	0.106	0.078	0.062					
France	20 to 25	26 to 30	31 to 35	36 to 40	41 to 45	46 to 50	51 to 55	56 to 60	⩾61
Unsatisfactory	0.36	0.39	0.34	0.37	–	0.43	0.38	0.42	0.39
Good-every-day	1.00	1.00	1.00	1.00	–	1.00	1.00	1.00	1.00
Better-than-every-day	1.79[abc]	1.69[b]	1.86[cd]	1.80[abc]	–	1.72[ab]	1.71[ab]	1.70[ab]	1.97[d]
Premium	2.79[ac]	2.59[b]	2.88[c]	2.68[ab]	–	2.53[bd]	2.56[be]	2.46[e]	2.75[acd]
SE[1]	0.070	0.077	0.084	0.078	–	0.071	0.081	0.078	0.075
Ireland	20 to 25	26 to 30	31 to 35	36 to 40	41 to 45	46 to 50	51 to 55	56 to 60	⩾61
Unsatisfactory	0.90	0.97	0.90	0.97	0.95	0.95	0.96	0.98	0.89
Good-every-day	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Better-than-every-day	1.03	1.09	1.03	1.08	1.09	1.08	1.06	1.13	1.00
Premium	1.06	1.11	1.07	1.12	1.09	1.08	1.08	1.10	0.98
SE[1]	0.046	0.055	0.065	0.066	0.061	0.061	0.066	0.067	0.118
Northern Ireland	15 to 29	20 to 44	45 to 49	⩾60					
Unsatisfactory	0.65[a]	0.61[ab]	0.58[b]	0.56[ab]					
Good-every-day	1.00	1.00	1.00	1.00					
Better-than-every-day	1.29[a]	1.27[a]	1.26[ab]	1.21[b]					
Premium	1.54[a]	1.50[ab]	1.47[b]	1.37[c]					
SE[1]	0.028	0.031	0.029	0.031					
Poland	<20	20 to 25	26 to 30	31 to 39	40 to 50	⩾51			
Unsatisfactory	0.51	0.52	0.53	0.53	0.53	0.54			
Good-every-day	1.00	1.00	1.00	1.00	1.00	1.00			
Better-than-every-day	1.41[abc]	1.40[ac]	1.42[c]	1.36[ab]	1.33[bd]	1.29[d]			
Premium	1.94[ab]	1.94[a]	1.99[b]	1.85[c]	1.81[c]	1.73[d]			
SE[1]	0.041	0.033	0.037	0.037	0.038	0.038			

[a,b,c,d,e]Values within a row with different superscript letters differ significantly at $P < 0.05$.
[1]Standard error for each age group by country.

The effect of other demographic factors on sensory scores and proportional willingness to pay

Contrary to our hypothesis, we found small effects of occupation, income and the number of adults and children in the household, on a consumer's evaluation of beef eating quality (data not shown). Tradespeople, professionals and administrators scored beef about 0.5 to 1 point lower than technical personnel, students and unemployed/retired people ($P < 0.05$). The number of adults in the home had a small positive effect on consumer scores for overall liking and tenderness (data not shown). Consumers with two adults in the home scored beef about 0.5 to 1 point lower than consumers with three or four adults in the home ($P < 0.05$). In the case of occupation and adults in the home, the effects were similar to the size of the standard error (data not shown).

Consumer age had a small negative relationship with tenderness in France and Poland, and with juiciness in Ireland, Northern Ireland and Poland (data not shown). The reduction was about 4 points for tenderness and 2 to 3 points for juiciness ($P < 0.05$). In contrast, there was a small positive relationship between consumer age and tenderness scores in Northern Ireland (data not shown). The youngest age group scored beef about 1 point out of 100 lower than the older age groups ($P < 0.05$).

These effects are in contrast to the findings of Hwang *et al.* (2008) who found no relationships between these demographic categories and sensory scores for beef. However, in our study the size of these effects was very small, approaching the standard deviations for the predicted means. Similar, small, effects for some demographic factors were found in the study of Thompson *et al.* (2005). In both cases, the size of the effects indicates that these results may not be repeatable and would be of limited use for any practical applications.

In agreement with our hypothesis, the demographic factors of gender, occupation, the number of adults in the home and the number of children in the home had no effect on P-WTP. This is supported by the results of Lyford *et al.* (2010) who also found no effect of these factors for Australian, Japanese, American and Irish consumers. These results are

further corroborated by Reicks *et al*. (2011) who found that these demographic factors had no effect on the importance of price when consumers were making purchasing decisions.

Design effects on consumer sensory scores

Carry-over effects from the previous sample and the serve order of the sample had strong effects in this study. This was expected on the basis of previous work which used similar experimental protocols with Australian and Korean consumers (Thompson *et al*., 2005; Hwang *et al*., 2008). Our results highlight the importance of using a Latin square design for the presentation order of samples, as opposed to a randomised design. This was first described by Williams (1949) and validated for this style of experiment by Watson *et al*. (2008b) and ensures that the effects of previous samples are equally distributed across the samples. In addition, the carry-over and order effects were by far the most prominent for the second sample (data not shown), demonstrating that the eating quality of the first sample would have a disproportionately large effect on the scores for the second sample. In anticipation of this effect in this experiment, an average quality 'link' sample was the first sample served to consumers. This would then minimise the variation attributable to the halo effects for this group, as was proposed by Watson *et al*. (2008b).

Conclusion

The way consumers score beef eating quality is highly consistent between different demographic groups. Willingness to pay for beef is also transferrable across different demographics. As consumers from different demographic groups have a similar appreciation of beef quality, this provides strong evidence that a single descriptor of eating quality will likely be applicable to the entire European market. If such a descriptor could be predicted from information available at slaughter then our results demonstrate that it could be used as the base of an eating quality-based grading system for beef. The provision of information on eating quality to the consumer would allow consumers to exercise their willingness to pay, realising the 1.5 to 2-fold increase in value for premium beef, and the 50% penalty in value for unsatisfactory beef. If these price differentials were realised, it would send a price signal through the beef supply chain, encouraging producers to include eating quality in their breeding and management strategies.

Acknowledgements

This research was supported by Meat and Livestock Australia and Murdoch University. Data were obtained through the financial contributions of the European research project ProSafeBeef (Contract No. FOOD-CT-2006-36241), the Polish ProOptiBeef Farm to Fork project funded by the EU Innovative (POIG.01.03.01-00-204/09), the French 'Direction Générale de l'Alimentation' and FranceAgriMer, the Irish Department of Agriculture Food and The Marine under the FIRM programme and the Northern Ireland Department of Agriculture and Rural Development 'Vision' programme. Furthermore, this project would not have been possible without the practical support of the Association Institut du Charolais, the Syndicat de Défense et du promotion de la Viande de Boeuf de Charolles and the gourmet restaurants 'Jean Denaud' and representatives of the beef industry across Europe. The international travel required for this project has been funded by 'Egide/Fast' funds from the French and Australian governments, respectively (project no. FR090054) and by 'Egide/Polonium' funds from the French and Polish governments, respectively. The assistance and participation the Beef CRC and Janine Lau (MLA, Australia), Alan Gee (Cosign, Australia), Ray Watson (Melbourne University, Australia) and John Thompson (UNE) are also gratefully acknowledged.

References

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

Berry BW and Hasty RW 1982. Influence of demographic factors on consumer purchasing patterns and preferences for ground beef. Journal of Consumer Studies & Home Economics 6, 351–360.

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

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

Bonny SPF, Hocquette JF, Pethick DW, Legrand I, Wierzbicki J, Allen P, Farmer LJ, Polkinghorne RJ and Gardner GE 2016c. Untrained consumer assessment of the eating quality of beef: 1. A single composite score can predict beef quality grades. Animal 1–10.

Cox RJ, Thompson JM, Cunial CM, Winter S and Gordon AJ 1997. The effect of degree of doneness of beef steaks on consumer acceptability of meals in restaurants. Meat Science 45, 75–85.

Feuz DM, Umberger WJ, Calkins CR and Sitz B 2004. U.S. consumers' willingness to pay for flavor and tenderness in steaks as determined with an experimental auction. Journal of Agricultural and Resource Economics 29, 501–516.

Gregory NG 1997. Meat, meat eating and vegetarianism. A review of the facts. In Proceedings of the 43rd International Congress of Meat Science and Technology (ed. J Bass), pp. 68–85. Elsevier, Aukland, New Zealand.

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

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

Hwang IH, Polkinghorne R, Lee JM and 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.

Kubberød E, Ueland Ø, Rødbotten M, Westad F and Risvik E 2002. Gender specific preferences and attitudes towards meat. Food Quality and Preference 13, 285–294.

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

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

Lusk JL, Fox JA, Schroeder TC, Mintert J and Koohmaraie M 2001. In-store valuation of steak tenderness. American Journal of Agricultural Economics 83, 539–550.

Lyford C, Thompson J, Polkinghorne R, Miller M, Nishimura T, Neath K, Allen P and Belasco E 2010. Is willingness to pay (WTP) for beef quality grades affected by consumer demographics and meat consumption preferences? Australasian Agribusiness Review 18, 1–17.

Neely TR, Lorenzen CL, Miller RK, Tatum JD, Wise JW, Taylor JF, Buyck MJ, Reagan JO and Savell JW 1999. Beef customer satisfaction: cooking method and degree of doneness effects on the top round steak. Journal of Animal Science 77, 653–660.

Reicks AL, Brooks JC, Garmyn AJ, Thompson LD, Lyford CL and Miller MF 2011. Demographics and beef preferences affect consumer motivation for purchasing fresh beef steaks and roasts. Meat Science 87, 403–411.

SAS 2002. Applied statistics and the SAS programming language. SAS institute, Cary, NC, USA.

Thompson J, Polkinghorne R, Gee A, Motiang D, Strydom P, Mashau M, Ng'ambi J, deKock R and Burrow H 2010. Beef palatability in the Republic of South Africa: implications for niche-marketing strategies. In ACIAR Technical Reports, Australian Centre for International Agricultural Research ACIAR, Canberra, ACT, Australia, pp. 1–56.

Thompson JM, Pleasants AB and Pethick DW 2005. The effect of design and demographic factors on consumer sensory scores. Australian Journal of Experimental Agriculture 45, 477–482.

Verbeke W, Van Wezemael L, de Barcellos MD, Kügler JO, Hocquette J-F, Ueland Ø and Grunert KG 2010. European beef consumers' interest in a beef eating-quality guarantee: insights from a qualitative study in four EU countries. Appetite 54, 289–296.

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

Williams EJ 1949. Experimental designs balanced for the estimation of residual effects of treatments. Australian Journal of Chemistry 2, 149–168.