

Trained Panel Evaluation of Upper 2/3 Prime Beef Strip Loin Steaks

*K.E. Wallace, C.A. Effling, L.M. Frink, M.J. Prester, S.L. Witberler,
G.E. Huber, S.F. Stickley, E.S. Beyer, M.D. Zumbaugh, J.L. Vipham,
and T.G. O'Quinn*

Abstract

The objective of this study was to evaluate trained sensory panel perceptions of beef strip loin steaks representing four marbling categories: Super Prime (moderately abundant and abundant), Low Prime (slightly abundant), Top Choice (modest and moderate), and Low Choice (small). Strip loins ($n = 15/\text{grade}$) were collected from a commercial packing facility, fabricated into approximately 1-in. steaks, and aged 28 days before analysis. Steaks were assigned to one of three degrees of doneness ([DOD]; Rare = 140°F, Medium = 160°F, and Well-done = 170°F) and evaluated by trained panelists for initial juiciness, sustained juiciness, myofibrillar tenderness, connective tissue amount, overall tenderness, and beef flavor intensity using a 100-point line scale. Data were analyzed using a split-plot design with quality grade as a whole-plot factor and DOD as the subplot factor. There were no interactions ($P > 0.05$) between quality grade and DOD for any trained panel-evaluated traits. Quality grade influenced ($P < 0.05$) all sensory attributes. Super Prime steaks received higher ($P < 0.05$) ratings for juiciness, tenderness, and flavor traits than Low Prime, Top Choice, and Low Choice steaks. For DOD, Rare and Medium samples were rated similar ($P > 0.05$) for initial juiciness, myofibrillar tenderness, and overall tenderness. Rare samples scored higher ($P < 0.05$) than Medium for sustained juiciness. Well-done steaks, regardless of grade, received the lowest ($P < 0.05$) scores. However, there was an interaction ($P < 0.05$) between quality grade and DOD for objective measurements, providing additional evidence of the negative effect of increased doneness on palatability and the positive effect of higher marbling scores. These results suggest that both marbling level and DOD are critical to the eating quality of beef strip loin steaks.

Introduction

In the past quarter century, the U.S. beef industry has experienced a steady and measurable improvement in carcass quality, with one of the most notable shifts being the increased prevalence of USDA Prime carcasses. According to the National Beef Quality Audit, only 2.0% of carcasses graded Prime in 2000; however, by 2021, this percentage had increased to nearly 7.8% (Mayer et al., 2025). This improvement reflects industrywide advancements in genetics, nutrition, and management practices. At the same time, heightened consumer demand for high-quality beef has reinforced incentives for producers to pursue carcass traits associated with tenderness, juiciness, and flavor (Gonzalez et al., 2024). Despite these gains, most value-based marketing programs

continue to focus on upper 2/3 USDA Choice carcasses, often treating USDA Prime as a single, undifferentiated category. This suggests that the beef industry may be failing to fully capture the added value potential of carcasses within the upper 2/3 of the USDA Prime grade, which contains substantially more marbling than those at the lower threshold of Prime. As USDA Prime availability grows, there is a unique opportunity to evaluate whether consumers can discern differences in eating quality within the USDA Prime grade. Previous research has clearly established that marbling is a strong predictor of palatability, particularly tenderness, juiciness, and flavor (O'Quinn et al., 2012). However, little work has been done to determine whether the sensory improvements associated with increasing marbling extend beyond the current USDA Prime threshold. The lack of such information has limited the development of branded beef programs that could differentiate and market "Super Prime" products as a premium tier within USDA Prime. The objective of this study was to evaluate palatability characteristics of strip loin steaks from the upper 2/3 USDA Prime (Super Prime), USDA Low Prime, upper 2/3 USDA Choice (Top Choice), and USDA Low Choice grades by trained sensory panelists.

Experimental Procedures

Beef strip loins ($n = 15/\text{grade}$) representing four marbling categories were collected from a commercial packing facility. Marbling categories included: Super Prime (moderately abundant and abundant), Low Prime (slightly abundant), Top Choice (modest and moderate), and Low Choice (small). Subprimals were fabricated into approximately 1-in.-thick steaks, vacuum-packaged, and aged for 28 days. Also, steaks were designated to be one of the three degrees of doneness ([DOD]; Rare = 140°F, Medium = 160°F, or Well-done = 170°F). Following aging, steaks were frozen at -4°F until later analytical testing. Trained sensory panelists ($n = 8$) evaluated samples for initial juiciness, sustained juiciness, myofibrillar tenderness, connective tissue amount, overall tenderness, and beef flavor intensity using a 100-point line scale. Anchors were provided with 0 being extremely dry/tough/bland and 100 being extremely juicy/tender/intense, and 50 being neutral. For off-flavor, there was a "not applicable" option. If there was an off-flavor present, panelists had the option to describe the type of off-flavor. Responses were collected using a digital survey platform on electronic tablets. Steaks designated for trained panels were thawed for 24 hours at 35-39°F before cooking. Samples were cooked on a flat-top electric grill using a weighed steak press to maintain consistent surface contact. Internal temperatures were monitored using a Type K thermocouple. Rare steaks were flipped at 80.6°F and removed from the grill at 127.4°F to achieve a peak temperature of 140°F. Medium steaks were flipped at 86°F and removed at 154.4°F for a peak of 160°F, while Well-done steaks were flipped at 89.6°F and removed at 165.2°F for a peak of 170°F. After cooking, each sample was portioned into two cubes ($0.51 \times 0.51 \times 1$ in.) and served to panelists. Final internal peak temperatures were recorded for all samples. Data were analyzed as a split-plot design using SAS with quality grade as a whole-plot factor and DOD as the subplot factor. Statistical significance was declared at $\alpha = 0.05$.

Results and Discussion

Table 1 shows there was no quality grade $\times$ DOD interaction for all sensory attributes evaluated by trained sensory panelists. Super Prime steaks received the highest ($P < 0.05$) scores for initial and sustained juiciness, myofibrillar tenderness, overall tenderness, and beef flavor intensity. Low Prime samples were rated higher ($P < 0.05$)

for both initial and sustained juiciness when compared to Low Choice. However, for initial and sustained juiciness, there were no differences ($P > 0.05$) between Low Prime and Top Choice, or between Top Choice and Low Choice. There were also no differences ($P > 0.05$) among Low Prime, Top Choice, and Low Choice for myofibrillar tenderness, connective tissue amount, and overall tenderness. Super Prime scored the lowest ($P < 0.05$) for connective tissue amount compared to all other quality grades. Low Prime and Top Choice steaks scored similar ($P > 0.05$) for beef flavor intensity, while Low Choice was rated the lowest ($P < 0.05$) for beef flavor intensity. For DOD, Rare and Medium samples were rated similar ($P > 0.05$) for initial juiciness, myofibrillar tenderness, and overall tenderness. Rare samples scored higher ($P < 0.05$) than Medium for sustained juiciness. Well-done samples were rated the lowest ($P < 0.05$) for initial and sustained juiciness, myofibrillar tenderness, and overall tenderness. There were no differences ($P > 0.05$) among all DOD for connective tissue and beef flavor intensity.

Table 2 shows an interaction ($P < 0.05$) between quality grade and DOD for Warner-Bratzler Shear Force (WBSF), slice shear force (SSF), pressed juice percentage (PJP), and cook loss. In all DOD categories, Super Prime had the lowest ($P < 0.05$) WBSF values. Rare Low Prime and Top Choice were similar ($P < 0.05$) for WBSF, and Low Choice was the highest ($P > 0.05$). Within Medium and Well-done, Top Choice and Low Choice samples had similar ($P > 0.05$) WBSF but were higher ($P < 0.05$) than Low Prime. For SSF, there were no differences ($P > 0.05$) among quality grade treatments cooked to Rare. Within Medium, Low Prime, Top Choice, and Low Choice all were similar ($P > 0.05$) for SSF, but Super Prime and Low Prime were similar ($P > 0.05$). Within the Well-done DOD, Super Prime had the lowest ($P < 0.05$) SSF values. Low Prime and Top Choice had similar ($P > 0.05$) SSF values but were lower ($P < 0.05$) than Low Choice. For PJP, there were no differences ($P > 0.05$) among treatments when cooked to Rare. Within Medium, Super Prime had the highest ($P > 0.05$) PJP values, when compared to all the lower grades, which were similar ($P > 0.05$). Within Well-done samples, Super Prime samples had a higher ($P < 0.05$) PJP value than Top Choice. There were no differences ($P > 0.05$) among the other quality grades. For cook loss in Rare, there were no differences ($P > 0.05$) among the quality grades. In Medium, Super Prime had the lowest ($P < 0.05$) cook loss value, when compared to Low Prime, Top Choice, and Low Choice, which had no differences ($P > 0.05$). Within Well-Done, Top Choice samples had a higher ($P < 0.05$) cook loss compared to Super Prime. There were no differences ($P > 0.05$) among the other quality grades.

Implications

Super Prime steaks consistently rated higher than the other grades in trained sensory evaluations, indicating that marketing them as a premium tier could boost value and ensure a more consistent, high-quality eating experience.

References

- Gonzalez, A.A., E.P. Williams, T.E. Schwartz, A.N. Arnold, D.B. Griffin, R.K. Miller, K.B. Gehring, J.C. Brooks, J.F. Legako, C.C. Carr, G.G. Mafi, C.L. Lorenzen, R.J. Maddock, and J.W. Savell. 2024. National beef tenderness survey—2022: Consumer sensory panel evaluations and Warner-Bratzler shear force of beef steaks from retail and foodservice. *Meat and Muscle Biology* 8(1): 16997, 1-11. doi: <https://doi.org/10.22175/mmb.16997>
- Mayer T.R., S.E. Borders, T.E. Schwartz, K.B. Gehring, D.B. Griffin, C.R. Kerth, K.E. Belk, J.A. Scanga, M.N. Nair, M.M. Pfeiffer, G.G. Mafi, K.M. Harr, T.E. Lawrence, T.C. Tennant, L.W. Lucher, T.G. O'Quinn, E.S. Beyer, P.D. Bass, L.G. Garcia, B.M. Bohrer, J.A. Pempek, A.J. Garmyn, R.J. Maddock, C.C. Carr, T.D. Pringle, T.L. Scheffler, J.M. Scheffler, A.M. Stelzleni, J.M. Gonzalez, K.R. Underwood, B.N. Harsh, C.M. Waters, and J.W. Savell. 2025. National Beef Quality Audit - 2022: Instrument grading assessments of carcass characteristics of fed steers and heifers. *Translational Animal Science* 9. <https://doi.org/10.1093/tas/txaf097>
- O'Quinn, T.G., J.C. Brooks, R.J. Polkinghorne, A.J. Garmyn, B.J. Johnson, J.D. Starkey, R.J. Rathmann, and M.F. Miller. 2012. Consumer assessment of beef strip loin steaks of varying fat levels. *Journal of Animal Science* 90(2):626–634. doi: <https://doi.org/10.2527/jas.2011-4282>

Table 1. Least square means (n = 8/panel) of trained panelist palatability ratings¹ for strip loin steaks across four marbling categories and three degrees of doneness

Treatment	Initial juiciness	Sustained juiciness	Myofibrillar tenderness	Connective tissue amount	Overall tenderness	Beef flavor intensity
Quality Grade ²						
Super Prime	76.0 ^a	70.2 ^a	76.6 ^a	3.4 ^b	74.9 ^a	57.7 ^a
Low Prime	64.9 ^b	58.0 ^b	65.4 ^b	5.0 ^a	62.8 ^b	49.4 ^b
Top Choice	63.5 ^{bc}	56.6 ^{bc}	64.9 ^b	4.5 ^a	62.3 ^b	47.5 ^b
Low Choice	60.4 ^c	52.8 ^c	62.9 ^b	5.3 ^a	59.6 ^b	43.7 ^c
SEM ⁴	1.5	1.6	1.3	0.3	1.5	1.4
P-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Degree of doneness ³						
Rare	71.8 ^a	65.5 ^a	70.8 ^a	4.8	68.3 ^a	51.8
Medium	68.3 ^a	61.8 ^b	69.3 ^a	4.8	66.6 ^a	49.1
Well-done	58.6 ^b	51.0 ^c	62.2 ^b	4.0	59.8 ^b	47.9
SEM ⁴	1.4	1.4	1.0	0.3	1.1	1.3
P-value	< 0.01	< 0.01	< 0.01	0.07	< 0.01	0.07

^{abc} Least square means within the same section of the same column without a common superscript differ ($P < 0.05$).

¹ Sensory scores: 0 = extremely dry/tough/bland/none, 50 = neither juicy/dry/tough/tender, 100 = extremely juicy/tender/abundant/intense.

² Marbling scores: Super Prime-moderately abundant, abundant, and very abundant, Low Prime-slightly abundant, Top Choice-moderate and modest, and Low Choice-small.

³ Degrees of doneness: Rare = 140°F, Medium = 160°F, and Well-done = 170°F.

⁴ SEM (largest) of the least square means.

Table 2. Least square means for the interaction ($P < 0.05$) between quality¹ treatment and degree of doneness² for Warner-Brazler shear force (WBSF), slice shear force (SSF), pressed juice percentage (PJP), and cook loss

Degree of doneness/ quality treatment	WBSF, lb	SSF, lb	PJP, %	Cook loss, %
Rare				
Super Prime	4.6 ^c	28.0	23.5	15.6
Low Prime	5.5 ^b	30	21.1	14.6
Top Choice	5.7 ^b	29.8	21.9	15.6
Low Choice	6.4 ^a	33.1	21.7	14.4
SEM ³	0.2	1.5	0.7	0.8
<i>P</i> -value	< 0.01	0.11	0.11	0.56
Medium				
Super Prime	4.4 ^c	25.4 ^b	20.3 ^a	20.0 ^b
Low Prime	6.0 ^b	29.5 ^{ab}	15.0 ^b	24.6 ^a
Top Choice	6.6 ^a	32.4 ^a	16.9 ^b	23.4 ^a
Low Choice	6.6 ^a	32.4 ^a	17.0 ^b	23.1 ^a
SEM ³	0.2	1.5	0.7	0.8
<i>P</i> -value	< 0.01	< 0.01	< 0.01	< 0.01
Well-done				
Super Prime	4.4 ^c	24.3 ^c	16.8 ^a	24.3 ^b
Low Prime	6.4 ^b	28.4 ^b	14.9 ^{ab}	26.2 ^{ab}
Top Choice	7.1 ^a	29.1 ^b	13.3 ^b	27.2 ^a
Low Choice	7.3 ^a	36.6 ^a	14.8 ^{ab}	25.3 ^{ab}
SEM ³	0.2	1.5	0.7	0.8
<i>P</i> -value	< 0.01	< 0.01	< 0.01	0.05

^{abc} Least square means within the same section of the same column without a common superscript differ ($P < 0.05$).

¹ Marbling scores: Super Prime-moderately abundant, abundant, and very abundant, Low Prime-slightly abundant, Top Choice-moderate and modest, and Low Choice-small.

² Degrees of doneness: Rare = 140°F, Medium = 160°F, and Well-done = 170°F.

³ SEM (largest) of the least square means.