Consumer and Laboratory Panel Evaluation of Good and Choice Beef Loins.

H. D. NAUMANN, CURTIS BRASCHLER,

MARGARET MANGEL, AND V. JAMES RHODES

CONTENTS

Consumer and Laboratory Panel Evaluation of Good and Choice Beef Loins.

H. D. NAUMANN, CURTIS BRASCHLER, MARGARET MANGEL, AND V. JAMES RHODES

SUMMARY

The major purpose of this report is to provide a definitive analysis of the relationship of consumer acceptability to:

1. grade
2. marbling
3. season of slaughter

To accomplish that purpose, a 400 member consumer panel in St. Louis County, Missouri, ate and evaluated 5,600 loin steaks from 560 carcasses. Consumers in their normal home environment evaluated the steaks for acceptance on the nine-point scale used several times previously in our research. The Good and Choice grade carcasses, from which these 560 loins were obtained, weighed 600 to 650 pounds and were identified by third of grade by regular federal graders. Degree of marbling was measured visually by University technologists on the basis of USDA photographic standards.

The basic experiment was replicated four times between March, 1959, and January, 1960, to provide an analysis of the effect of season of slaughter on acceptance and to provide an exceptionally large sample of loins.

CONCLUSIONS

(1) Acceptance-grade:

Since the basically weak relationship of grade to consumer acceptance has been established already by a series of studies,* the contribution of this large study is the careful specification of the relationship by thirds of a grade. The distribution of loin acceptance means is described in Table 1; for example, the first number in column (1) indicates that 40 of the 80 High Choice loins had mean ratings poorer than 2.3 (i.e. had *larger* numbers on the 9 point scale) while the other 40 mean ratings were as good as or better than 2.3.

Since most retailers are interested in maintaining a satisfactory minimum quality of beef, columns (4) to (6) of Table 1 indicate the percentages of loins falling below certain quality ratings. For example, column (5) indicates that only 2.5 percent of the 80 High Choice loins had an average

*See Review of Literature.

TABLE 1-CONSUMER ACCEPTABILITY RATINGS OF LOINS BY THIRD OF GRADE

	Median	Mean	Upper and Lower Boundaries of two Middle Quartiles	Percentage of Loins Rating Poorer than:		
	(1)	(2)	(3)	(4)	(5)	(6)
				4.00	3.75	3.50
H. Choice	2.3	2.43	2.0 - 2.8	2.5	2.5	3.8
M. Choice	2.25	2.33	2.0 - 2.7	0.0	0.0	0.0
L. Choice	2.5	2.62	2.3 - 2.9	0.8	2.5	4.2
H. Good	2.7	2.72	2.3 - 3.1	1.7	5.0	10.0
M. Good	2.8	2.92	2.4 - 3.3	6.0	10.0	17.5
L. Good	3.0	3.02	2.7 - 3.4	3.8	12.5	20.0

rating poorer than 3.75 as contrasted with 12.5 percent of the Low Good grade. It is our judgment that the quality minimum which now can be the realistic objective of retailers will fall somewhere in the range of acceptance ratings 3.5 to 4.0 of columns (4) to (6).

(2) **Acceptance—marbling.**

This relationship was quite weak; the linear correlation coefficient was 0.21 (N = 380).

(3) **Seasonal**

While some differences in acceptance by seasons were found by analysis of variance, the distributions of acceptance ratings were actually very similar. The eating quality of Good and Choice Beef was apparently quite stable throughout the year.

(4) **Homogenity of Consumers' Preferences**

This study did not substantiate the rather popular idea that Middle Choice is logically separated in the market from Middle Good *because* one group of consumers "likes" or "prefers" Middle Choice while another group "prefers" Middle Good. The overwhelming majority of these 400 consumers revealed no such "preference" for any single grade division.

IMPLICATIONS

At a time when the livestock and meat industry's thinking about grading, grade-labeling, and quality is in an uncertain state of transition, the following discussion of implications is undoubtedly incomplete.

Excessive finish is no longer regarded as an indicator of eating quality and is being discounted because it reduces trimmed retail yields. It is now apparent that marbling is not a sure indicator of eating quality. While there is certainly no implication that marbling should be avoided, it is possible that marbling could be overemphasized in certain breeding and feeding programs. What is even clearer is the need for an intensification of the breeding and meats research

on heredity and other factors to determine which are basic to beef quality. It is now apparent that eating quality of beef largely depends on different and probably more complicated factors than those involved in the grading standards' definition of "quality."

All kinds of considerations—both public and private—bear upon the question of the appropriate location of grade boundaries. This study suggests that the satisfactory consumer acceptance of High Good and Low Choice explains why the bulk of retailer demand appears to center on those two grade segments. It then follows that, if a grade boundary is drawn through this area of popular demand, the possible number of "liner" controversies will be much larger than if the grade boundary were drawn outside the area. Those who argue that changing the grade boundary will likewise change the area of popular demand are arguing, in effect, that grades determine demand rather than reflect it. On the basis of this acceptance study, it seems extremely improbable that the bulk of retail demand would be "astraddle" the Choice boundary if that boundary were located between, say, the present segments of Low and Middle Good. Of course, reduction of the liner controversies by moving the grade boundary might be accompanied by the emergence of other problems.

Since a majority of loins in the Good grade are just as palatable as the typical Choice loin, there are powerful market incentives for development of improved sorting methods which would enable the merchandising of such loins at a better price. The potential welfare gains from development of improved grading and/or production methods which would encourage the more efficient production of lean, palatable, beef would accrue to both the beef producer and consumer.

Consumer Acceptance and Laboratory Evaluation Introduction

Variability in eating characteristics of Choice and Good short loins is thought to be considerable and to be directly related to the price variation found among such loins in the market. The belief in considerable variability has led to the development of various devices for classifying carcasses and wholesale cuts as to their degree of homogeneity of eating characteristics. The following sections describe this variability in eating characteristics as it is associated with grade and marbling.

Beef steaks are eaten by millions of consumers who put different emphasis on various eating characteristics. For best relating of research results to the market, it is the general belief of most researchers that a sample of consumers is a better "measuring stick" than a handful of "expert appraisers." There is a very real possibility that the "experts" will concentrate upon some eating characteristic in a manner unlike consumers. However, the expert panel conducted under controlled laboratory conditions can show whether there are differences in various eating characteristics.

Consumer Acceptance—Grade

Consumer acceptance of short loins increased slightly by thirds of grades moving from Low Good to Middle Choice (Figure 1). The relationship was found in each of the four seasonal tests. While a positive relationship exists between grade and consumer acceptance, it is also obvious that the relationship is slight. There are various ways of describing the relationship, Table 1 and Figure 1.

Variation within each grade was about the same as total variation regardless of grade. Mean ratings of Choice loins ranged from 1.5 to 4.8. However, division of loins into thirds of grades generally reduced variation within as compared to total variation. Within a third of a grade, there was, on the average, about 80 percent as large a range of variation as total variation between the best and poorest loins in the test. Whether the diversity in eating quality within grades detracts significantly from their usefulness to buyers is a question which each buyer can answer.

Based on the ratings and comments in this and previous studies, it appears that the lower boundary of fair satisfaction lies at about 3.75. This score is admittedly arbitrary; some retailers may suggest lower boundary limit of 3.5, while others in more price conscious markets may settle for 4.0. To insist on any boundary higher (smaller numbers on vertical scale in Figure 1) than 3.5 is to say that a significant percentage of Choice is unsatisfactory. Such insistence would contradict the widespread satisfaction among consumers and retailers with the eating qualities of Choice beef. Table 1 indicates the percentages of "unsatisfactory loins" in each third of grade by three different levels of minimum satisfaction.

FIGURE 1-SUMMARY OF ST. LOUIS PANEL DATA - ALL 560 LOINS - ACCEPTANCE
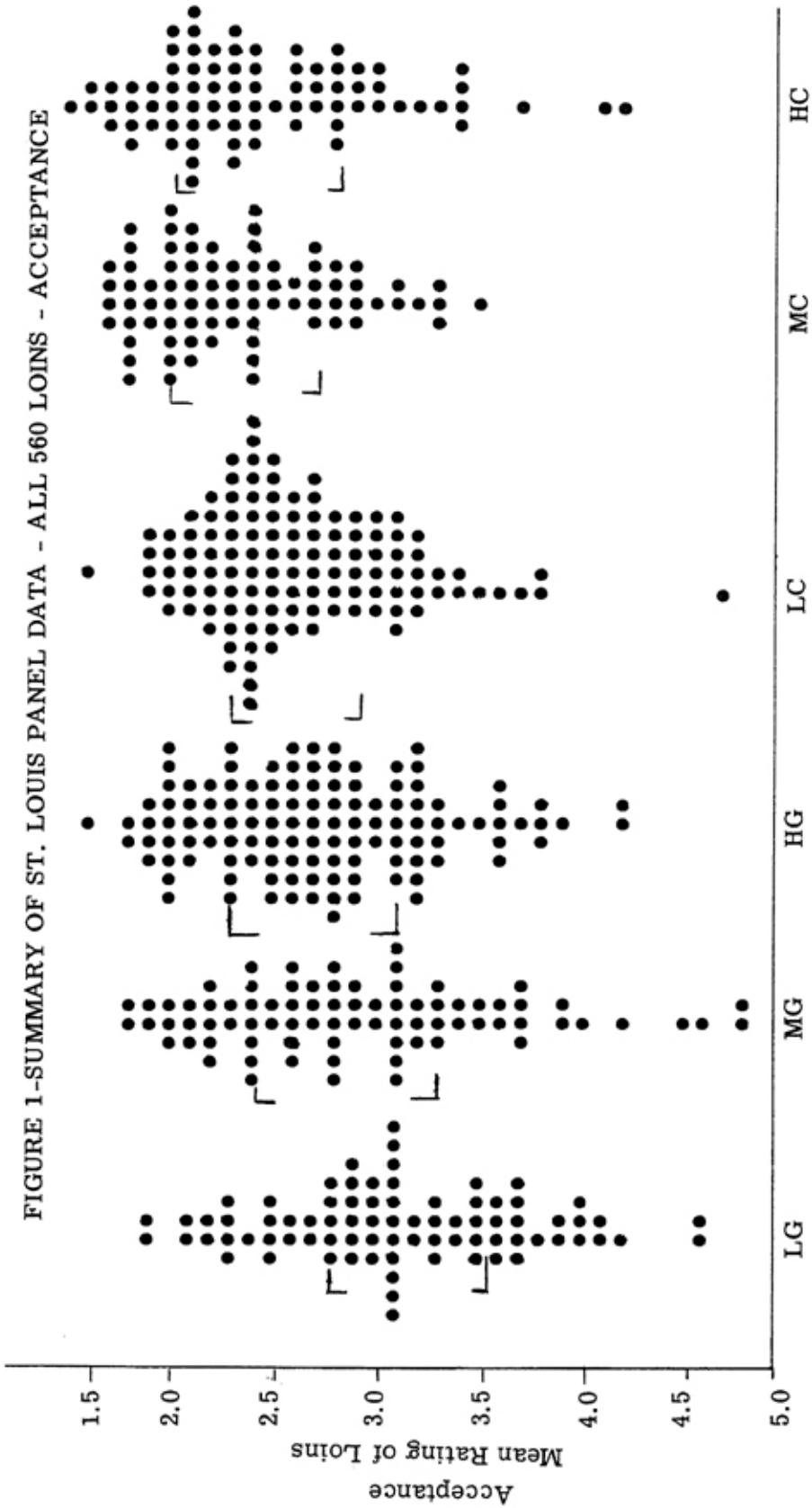
*Brackets indicate two middle quartiles of distribution in each grade category

Comparison of High Good, and Low Choice

Among retailers, the two most popular segments of the Choice and Good grades are apparently Low Choice and High Good. A price difference of $2 per cwt. for carcasses and of much more for loins is not unusual between these two adjacent grades. This value difference leads to many controversies about the grade of "liner" carcasses. Therefore, it seems relevant to enlarge the samples of High Good and Low Choice and to make a special comparison of their eating acceptance. As shown in Figure 2, these adjacent grade segments have similar and largely overlapping distributions; the absence of a natural "break" in acceptance is not an argument for, or against, a grade boundary being located at this point. Various economic factors influence the location of grade boundaries; natural "breaks" in acceptance are not found in the whole of the Choice-Good range. Perhaps, the location of the grade boundary at this point leads to a greater discounting of the prices of High Good than can be justified by the extremely small quality differential and the relatively high acceptability of High Good loins.

The distribution of individual consumer ratings for one test by third of grade is illustrated in Figure 3. Most ratings were better than 4.0, of course. An impression of loin and panel variability can be obtained from comparing Low Choice and Low Choice Extra. Loins of that grade were randomly divided into those two groups and eaten by two groups of consumers; the ratings indicate the degree of difference in acceptance.

Favorable and Less Favorable Mean Ratings

Since consumers differ as to their ratings on the same loin,* the grades were compared by computing two means for each loin. The "more favorable" mean included the five best ratings of the loin while the "less favorable" mean included the other five ratings.

The great similarity by thirds of a grade of the distribution of the "more favorable" means is indicated in Figure 4. Note, also, the high degree of satisfaction of those consumers with almost all of the loins. The diversity in satisfaction of the loins is much larger in the distributions of "less favorable" means with a range from 2.0 to 6.5. Stated in another way, the diversity in mean ratings of loins stems mainly from the ratings of the less satisfied consumers. The distributions of "less favorable means" are fairly similar by thirds of grades, but there is a small and fairly consistent shifting toward better ratings moving "up the grades" from Low Good to Middle Choice.

Consistency of Consumer Acceptance Ratings

Questions are often raised about the consistency of consumer ratings. An experiment reported previously† indicated that the same individuals rate the

*This difference is probably in part a different use of the rating scale. On more than one occasion, we have found that high-income people rated *all* steaks a little lower than other people.

†Numbers refer to references in the back.

Figure 2   Distribution of Acceptance Loin Means by Grade, St. Louis County Panel
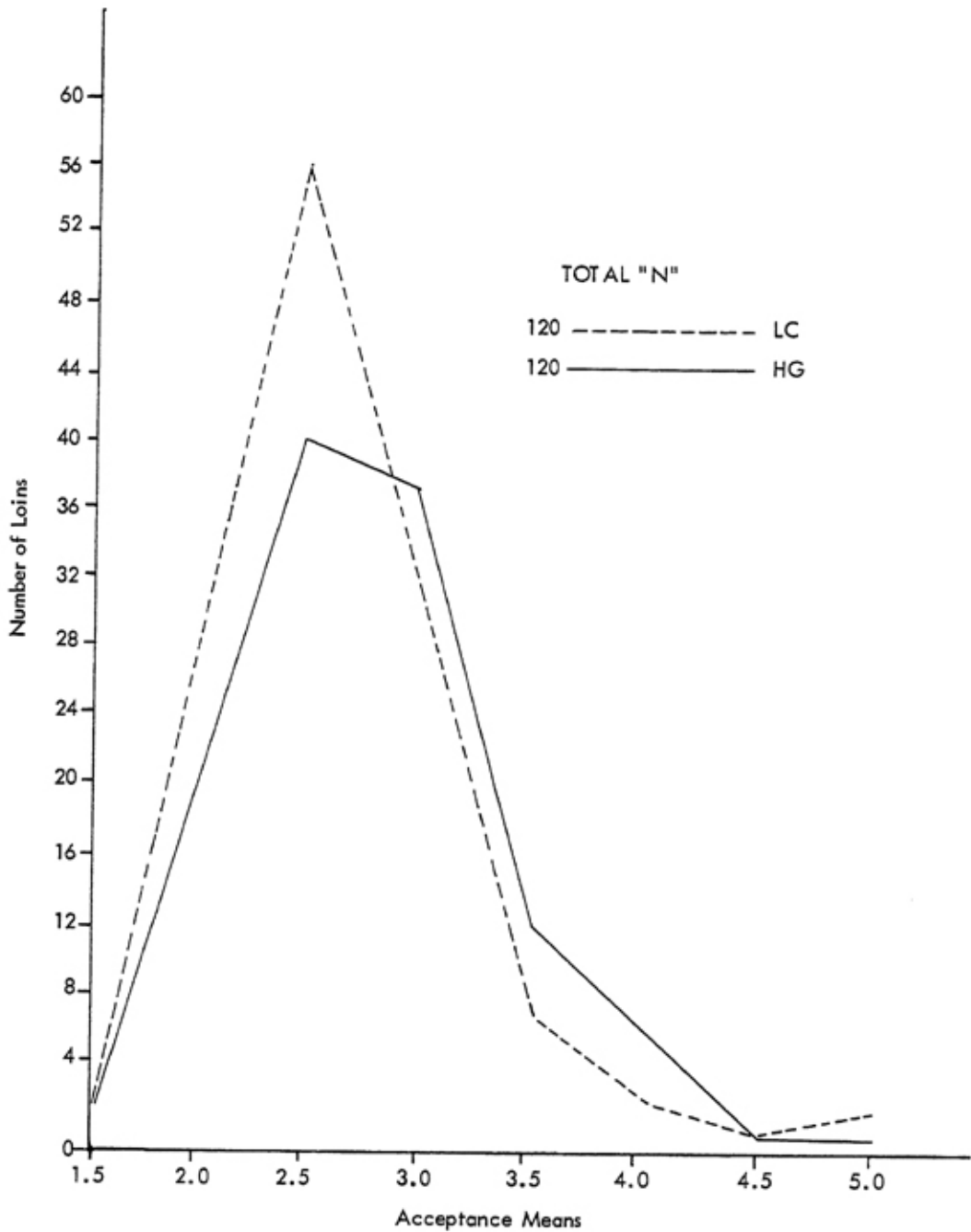
same set of products consistently. There are a number of other problems of consistency, which are discussed in this section. In general, these problems are concerned with whether there are discernible groupings of consumers who con-

Figure 3   Individual Ratings, Third Replicate
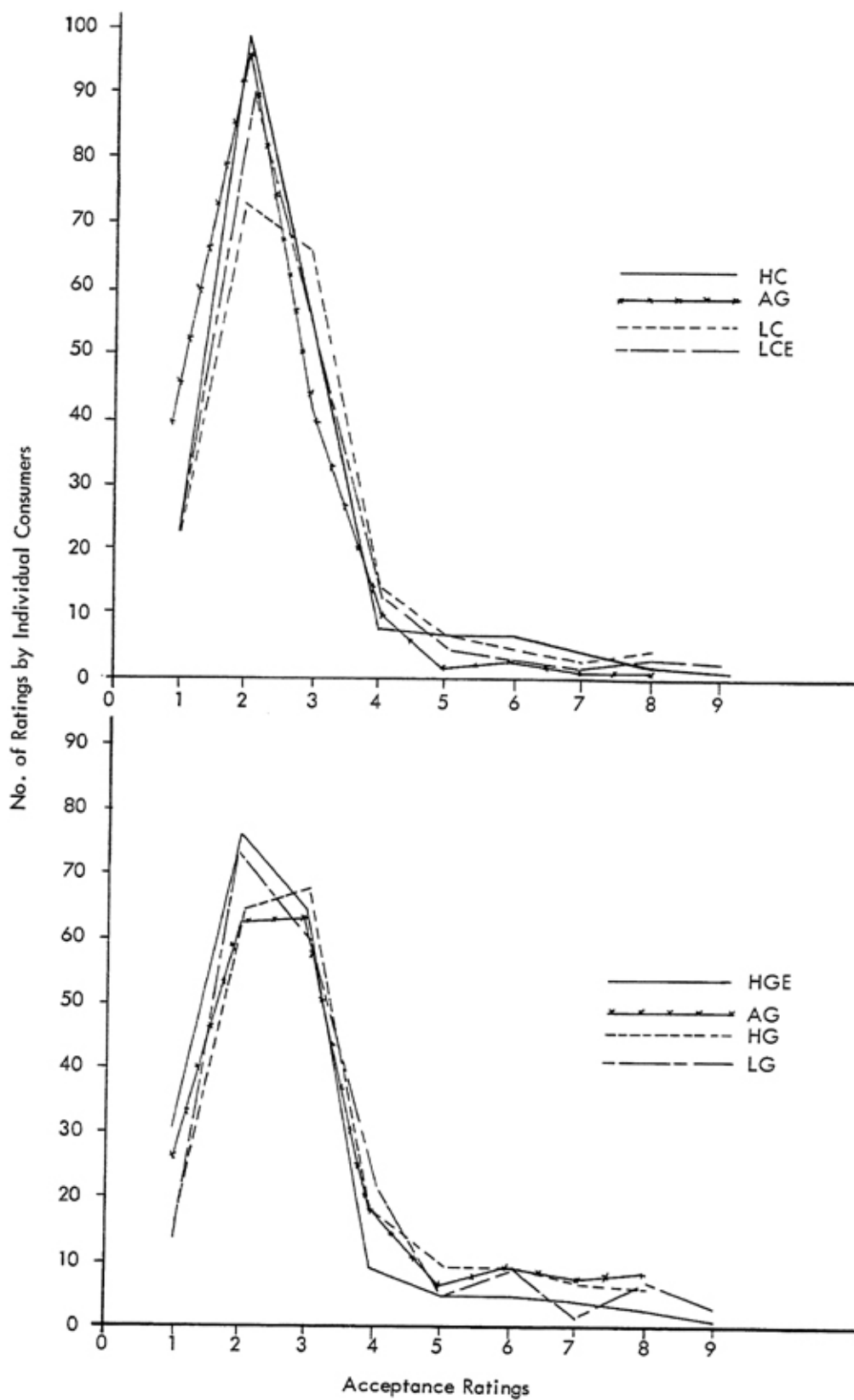
Figure 4   Distributions of Loin Means of the More and Less Favorable Consumers by Grades
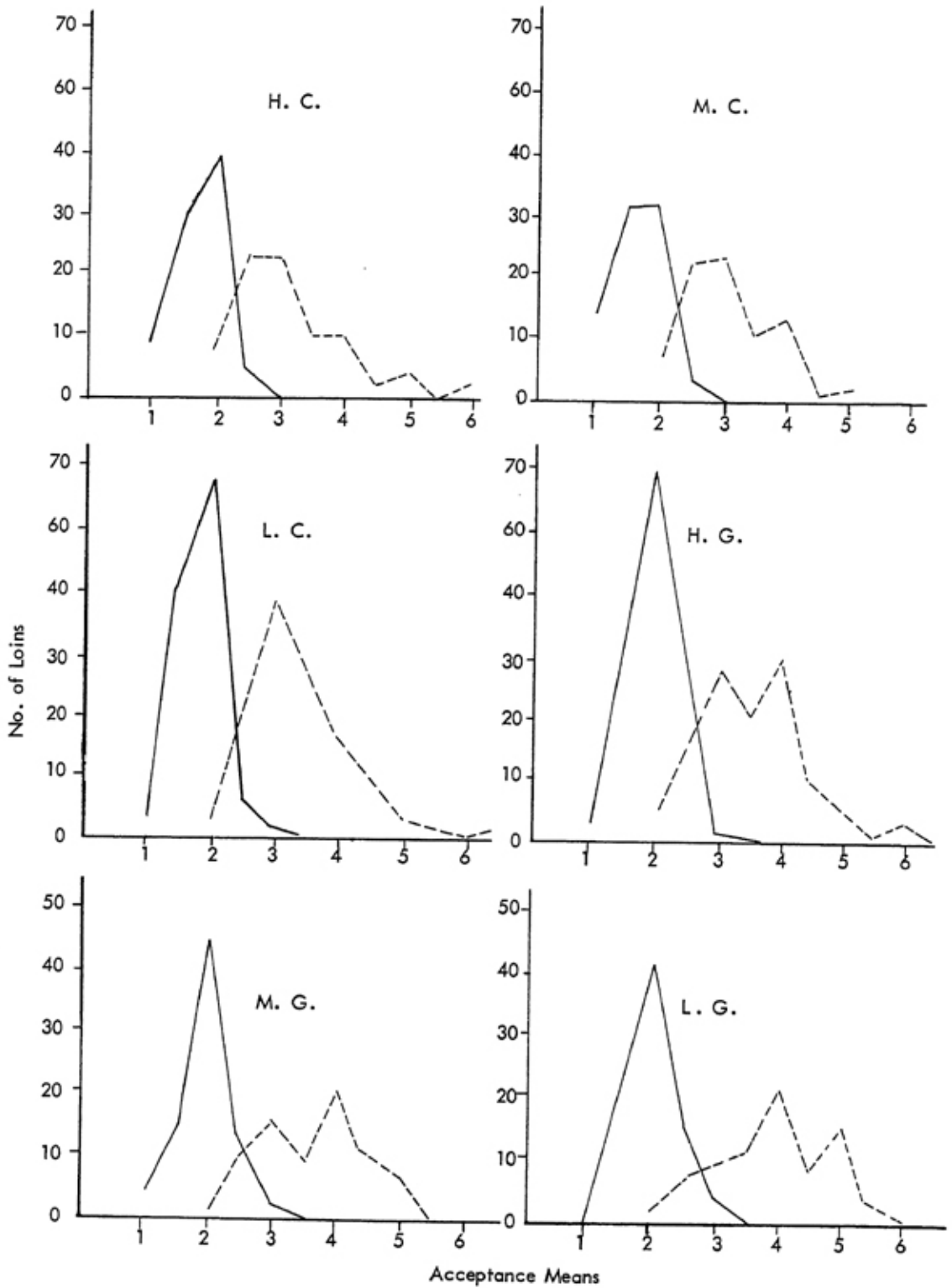
sistently rate differently, and/or indicate differing degrees of acceptence of steaks by grades or by other characteristics.

Hypothesis A: All of the "less favorable" ratings were given by one group of consumers, while all of the "more favorable" ratings were given by the other consumers.

Hypothesis A, as stated, could hardly be expected to be completely true. Hypothesis A was not supported by the data. About one-fourth (49 + 49 out of 400) of the consumers were consistently "more favorable" or "less favorable" raters of all steaks, while the other three-fourths were often in each group. (Table 2).

Consistently high or low "raters" were not necessarily insensitive to variations in eating quality between loins. For example, high "raters" may have varied somewhat in their ratings, but they have always used high ratings, and, thus, were high relative to other raters. In a preceding section, it was shown that most of the differentiation in loin means was in the "less favorable" ratings. This data indicates clearly that this "less favorable" group was not a select group of especially discriminating consumers, but rather it included all consumers occasionally and it included seven-eighths of consumers as often as four or more out of 14 times. (Table 2).

Hypothesis B: There was no difference among consumers over time as to the grades they rated best.

Consumers differ considerably as to their evaluation of the use and merits of pot and oven roasts, round and loin steaks, ham and chicken, and well-done and rare steaks. However, there is no conclusive evidence that consumers differ significantly as to the desired basic eating qualities of a loin steak—aside from differences in cooking, doneness, and added flavoring. It appears reasonable to assume that all consumers like tender steaks, although there may be minor differences of tolerance as to what degree of tenderness is satisfactorily tender. Much the same reasoning may be applied to flavor, juiciness, and texture, as found in the meat. This argument does not contradict the fact that some consumers may prefer more "aged flavor" than others, some like and some dislike charcoal flavor, etc.

TABLE 2- CONSUMERS AND NO. OF LESS FAVORABLE RATINGS

No. of "Less Favorable" Ratings (14 possible in total test)	No. of Consumers (out of 400)
1 or more	400
1 to 3	49
4 to 10	302
11 to 14	49

There is no completely satisfactory way of learning whether consumers preferences are alike or different as to the eating characteristics of loin steaks. Two problems limit the researcher: first, the supply of steaks within one loin, or a

loin pair, is too limited to test many consumers; second, there is no independent, reliable, appraisal of the loins eating characteristics aside from the average opinion of those eating. In other words, to test differences in preferences for certain eating qualities, we must have certain known variations in eating qualities identified independently of the tested consumers. Such identification can be obtained for a manufactured product like ground beef or orange juice, but not for steaks.

An approximation was used to test the similarity of consumers' preferences for these products classified as alike in grade, or marbling, or shear force, and other characteristics. It is true that the products within any such classification are probably *not* homogeneous. Therefore, we cannot demonstrate conclusively whether consumers are alike as to preferences for basic eating qualities, but a test of Hypothesis B can indicate whether consumers appear to be alike or not as to preferences for grades. The evidence of the entire test supports Hypothesis B that there was generally no difference among consumers as to the grades rated best. While this conclusion contradicts some popular ideas about the great variation in consumers' preferences, it is supported by a detailed analysis of the data.

The following rather complicated method was used to test Hypothesis B. Each consumer rated a set of three grades on each of four seasonal runs. One group of 200 consumers rated H. Choice, L. Choice and M. Good, while the other 200 rated M. Choice, H. Good and L. Good. The ratings of each consumer of the three grades on each run were ranked, and it was determined how many times (out of four) each consumer "ranked" each grade as best. There were a great many "tie rankings," of course, since two or even all three grades were sometimes rated the same. The distinction between the *number of times* a grade was rated best and the *number of people* who rated each grade best must be kept carefully in mind.

Table 3 indicates that each grade in the High Choice set rated best over 295 times out of a possible 740 times (185‡ people multiplied by 4 possible times per person). The relative number of times rated best per grade in table 3 indicates something of the relative popularity of grades, but it does not show

TABLE 3 - NUMBER OF FIRST RANKS (BOTH INDIVIDUAL AND TIES) BY GRADE

No. OF TIMES RANKED FIRST

	Individual	Ties, with one other Grade	Ties with Both other Grades	Totals[a]
I. M.G.-L.C.-H.C. Set				
M.G.	102	129	98	329
L.C.	119	165	98	373
H.C.	178	192	98	468
II. L.G.-H.G.-M.C. Set				
L.G.	85	126	87	298
H.G.	125	140	87	352
M.C.	215	190	87	492

a/ Maximum possible totals per Grade = 740

‡Because of replacements during the panel, complete ratings were available for 185 rather than 200 people.

which people did what. The large number of ties indicates that there were many loins of similar nature from the different grades. The randomness of experimental design affected the number of times that any consumer happened to get similar loins during the four seasonal tests. There appeared to have been more eating differences in the No. II set (Table 3) than in the No. I set as evidenced by fewer ties and greater differences between grades on individual rankings in the No. II set.

Table 4 indicates that in the M. G.—L. C.—H. C. set no one consistently ranked a single grade as best. That is, there was no single grade ranked first all four times. For example, 78 people ranked H. Choice best on one occasion, 32 twice and 12 people three times, leaving a total of 63 people not ranking it best any time. However, a number of these 63 people tied H. Choice and some other grade or grades as first. As indicated in Table 5, only 3 persons never ranked H. Choice as first (neither individually nor as a tie with another grade or grades). Only 10 persons never ranked L. Choice as first, and only 14 never ranked M. Good as first (Table 5). Thus, the empirical evidence was insufficient for rejection of Hypothesis B, for almost everyone ranked each grade as best one or more times of the four times. While the grades differed as to popularity, there was no one group of people specially favoring one grade and not others.

Results are essentially the same for the L. G.—H. G.—M. C. set as for the grade set discussed above. (Table 4 and 5).

There were only 18 people who gave no tie ratings on the M. G.—L. C.—H. C. set (Table 6). As indicated in Table 6, none gave more than three first ratings to one grade and about two-thirds gave one or two first ratings to each grade.

TABLE 4-DETAILED DISTRIBUTION OF "RANKING" BY GRADE

I.   M.G.-L.C.-H.C. Set

No. of Times "Ranked" First	M.G.	L.C.	H.C.	M.G.-L.C.	M.G.-H.C.	L.C.-H.C.	M.G.-L.C.-H.C.
0	99	93	63	140	115	95	109
1	70	68	78	39	63	68	58
2	16	21	32	6	6	20	14
3	0	3	12	0	1	2	4
4	0	0	0	0	0	0	0
Totals	185	185	185	185	185	185	185

II.   L.G.-H.G.-M.C. Set

No. of Times "Ranked" First	L.G.	H.G.	M.C.	L.G.-H.G.	L.G.-M.C.	H.G.-M.C.	L.G.-H.G.-M.C.
0	110	93	49	148	110	104	117
1	65	65	75	36	63	60	52
2	10	21	44	1	11	21	14
3	0	6	16	0	1	0	1
4	0	0	1	0	0	0	1
Totals	185	185	185	185	185	185	185

TABLE 5-NUMBER OF TIMES EACH GRADE RATED BEST BY NUMBER OF
CONSUMERS

No. of times Rated Best (or tied best)	H.Ch.	L.Ch.	M.G.	Grades a/	M.Ch.	H.G.	L.G.
0	3b/	10	14		2	24	21
1	22	45	65		22	41	68
2	66	68	59		45	61	64
3	62	47	42		84	47	26
4	32	15	5		32	12	6
Total	185	185	185		185	185	185

a/ Each set of three grades was eaten by a group of 200 consumers. However,
complete responses were available on only 185 for each of grades.

b/ Number of people rating that grade best zero times. There were many tie-ratings.

TABLE 6-NUMBER OF TIMES EACH GRADE RATED BEST BY THOSE
CONSUMERS GIVING NO TIE RATINGS

No. of times Rated Best	H.C.	Grades L.C.	M.G.
0	2	5	6
1	5	4	9
2	6	6	3
3	5	3	0
4	0	0	0
Totals	18	18	18

SOCIAL ECONOMIC FACTORS IN RELATION
TO MEAN GRADE ACCEPTANCE

The effect of income upon the mean acceptance ratings of grades was evalu-
ated for the October and January replications. The mean acceptance ratings for
the different grades did not appear to be related to income for either replica-
tion (Figures 5 and 6). In other words, the rate of increase in the mean accept-
ance rating was about the same for all income groups with increases in grades
(Figure 5 and 6). The higher income groups did appear to rank all grades slight-
ly lower than the lower income groups (Figures 5 and 6). This was consistent
with results of previous studies.[2]

The education of the housewife did not appear to affect the improvement
in the mean acceptance rating with respect to movements from lower to higher
grades for either replication (Figures 7 and 8). But there was a tendency for the
less educated groups to rate all grades higher for both replications (Figures 7
and 8), since income and education are generally positively correlated.

The age of the husband was not a significant variable in affecting the rate
of improvement in acceptance rating from lower to higher grades (Figures 9 and
10). Neither was there a clear tendency for one age group to rate all grades con-
sistently higher than other age groups (Figures 9 and 10).

Figure 5   Mean Rating of Grade by Income for October Replication

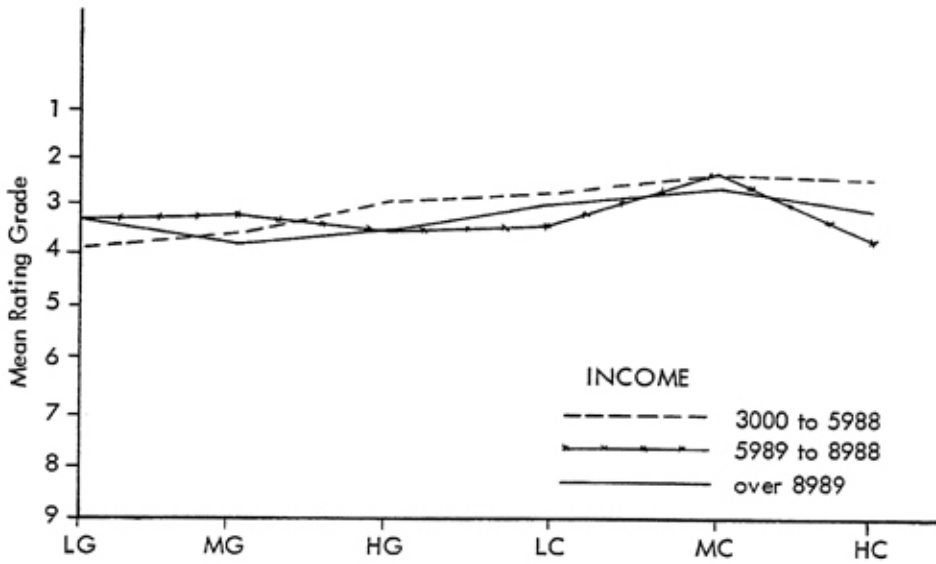Figure 6   Mean Rating of Grade by Household Income for January Replicatior
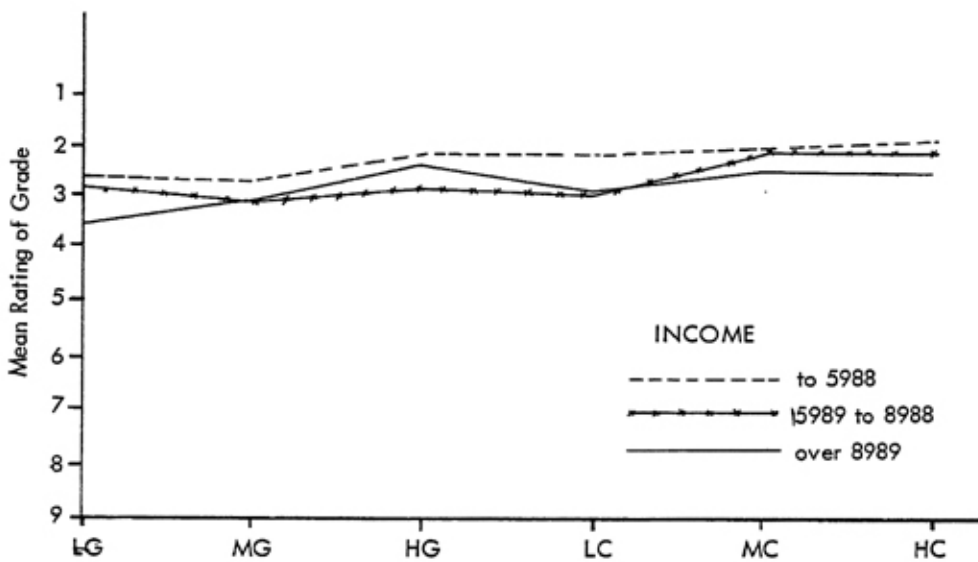
Figure 7   Mean Rating of Grade by Education of Housewife for October Replication
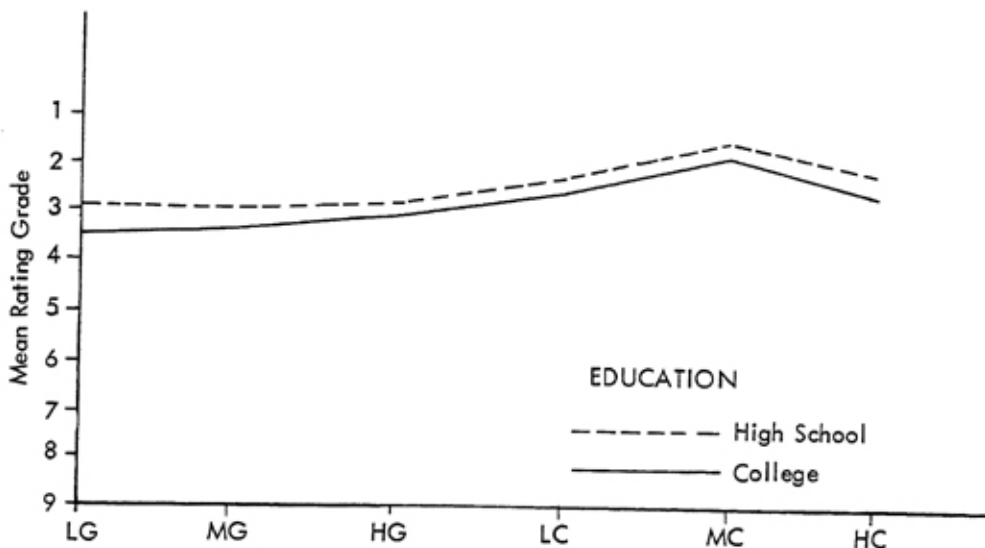

Figure 8   Mean Rating of Grade by Education of Housewife for January Replication
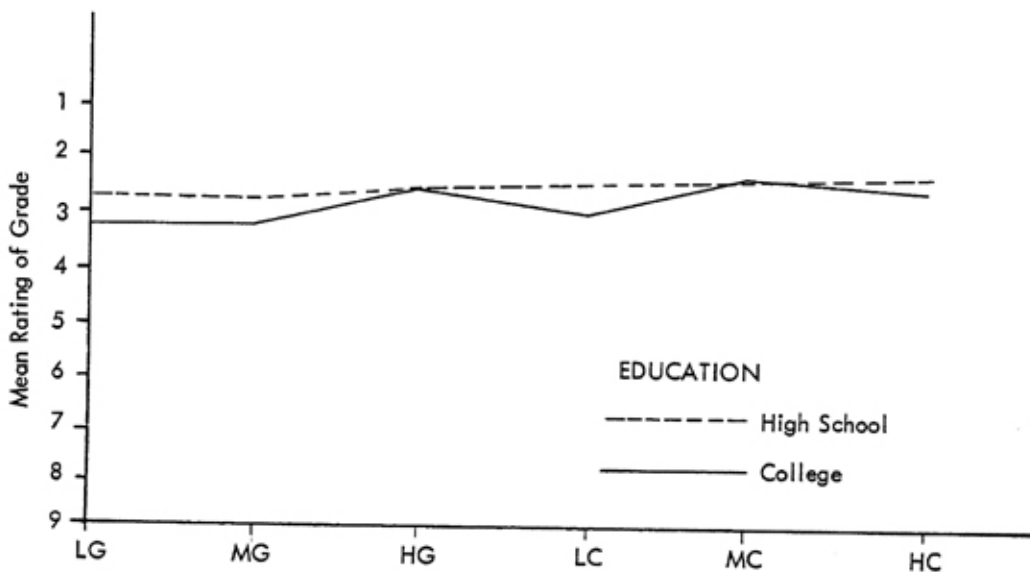
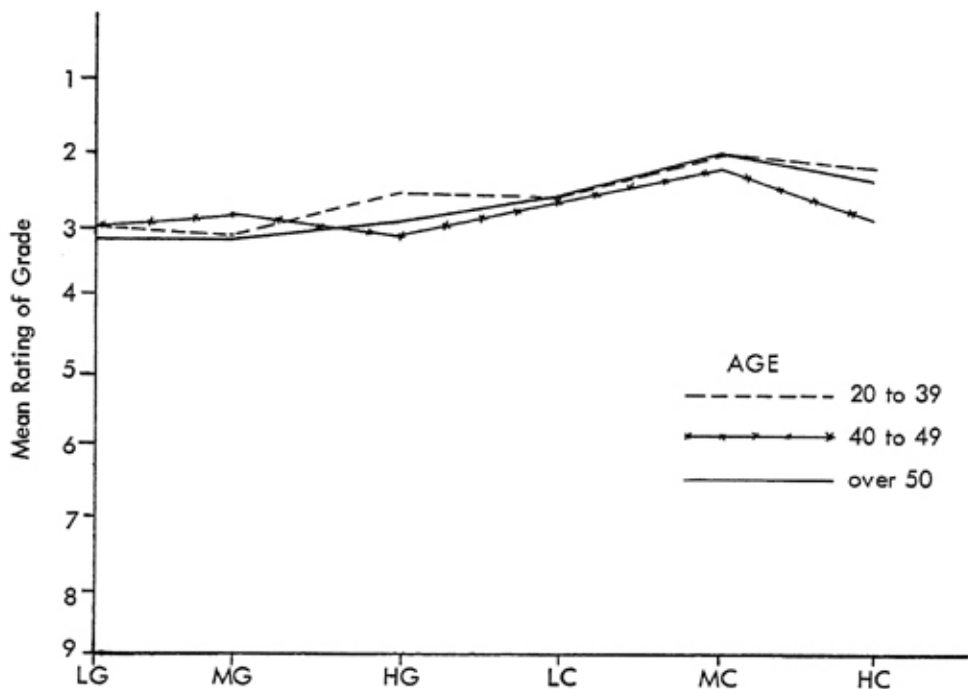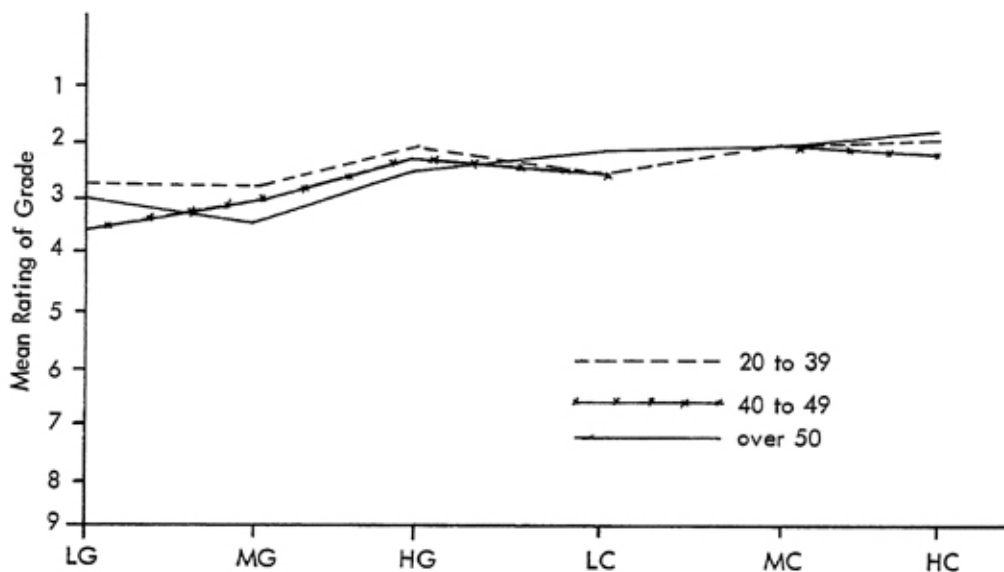Figure 9   Mean Rating of Grades by Age of Husband for October Replication

Figure 10   Mean Rating of Grades by Age of Husband for January Replication

In general, it was concluded that improvements in mean acceptance ratings associated with changes in grades were unaffected by the socio-economic variables delineated in this study. This finding is consistent with the finding above that particular grades were not specially favored by any particular group of consumers.

ACCEPTANCE-MARBLING

Marbling and grade are usually considered to be highly related. Since there was a slight relationship between acceptance and grade, a similar relationship might be expected between acceptance and marbling (Figure 11). Such was the case. The linear correlation coefficient between acceptance and marbling was .21 (N = 380).†† A visual examination of the data indicated that the relationship was essentially linear.

The average marbling score of the loins by grade and replicate appear in Table 7. There appeared to be little difference among average marbling scores of the three replications.

The overall marbling means by thirds of grades fell in an interesting pattern. Variation in average marbling was considerably less within the Good grade than within the Choice grade, while differences in average consumer palatability were slightly more within the Good grade than within the Choice (Table 1). Overall mean marbling differences between adjacent thirds of grades in this particular sample of loins ranged from 1.5 between High and Middle Choice to 0.3 between High and Middle Good. Note the 1.2 difference between Low Choice and High Good.

CONSUMER TENDERNESS—JUICINESS—FLAVOR BY THIRDS OF GRADES

Consumers evaluated each steak eaten on a four point scale (Appendix A) with respect to its relative tenderness, juiciness and flavor.

In general, the results of the evaluation were similar and consistent with general consumers acceptance by thirds of grades (Figures 12, 13, and 14). The variation in all three evaluated attributes was large within grades (Figures 12, 13, and 14). However, there was a tendency for higher relative ratings in all three attributes moving from the lower to the higher thirds of grade categories (Figures 12, 13, and 14). This same tendency appeared in the overall acceptance ratings previously noted (Figure 1).

The simple correlation coefficients between consumer acceptance and consumer evaluation of tenderness, juiciness, and flavor, were .53, .34, and .35, respectively. The simple correlation between consumer acceptance and marbling was .21. The simple correlation coefficients between marbling and consumer evaluation of tenderness, juiciness, and flavor, were .24, .30, and .16, respectively.

††Marbling measurements were not secured on the first replicate.

FIGURE 11- RELATION BETWEEN MARBLING AND ACCEPTANCE - 380 LOINS
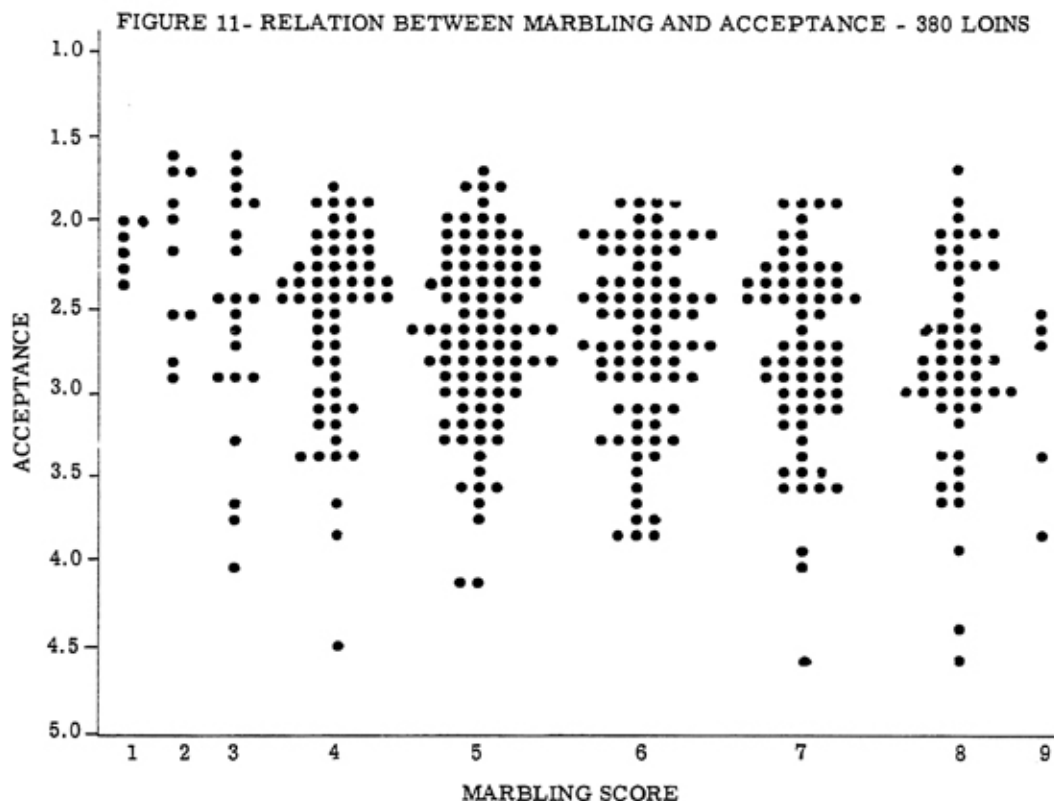
The tenderness factor is apparently somewhat better related to consumer acceptance than is marbling (marbling on acceptance, .21) or juiciness, and flavor.

TABLE 7-AVERAGE MARBLING SCORE BY GRADE AND REPLICATE

Slaughter Month	Repli- cate	Low Good	Middle Good	High Good	Low Choice	Middle Choice	High Choice
April-May	2	8.2	7.9	7.6	6.2	6.0	4.8
Aug.-Sept.	3	8.2	7.1	6.4	5.6	5.1	3.6
Dec.	4	8.6	8.4	8.6	7.0	6.4	4.6
Overall Mean		8.3	7.8	7.5	6.3	5.8	4.3

FIGURE 12- SUMMARY OF ST. LOUIS PANEL DATA - ALL 560 LOINS-TENDERNESS

FIGURE 13-SUMMARY OF ST. LOUIS PANEL DATA - ALL 560 LOINS-JUICINESS
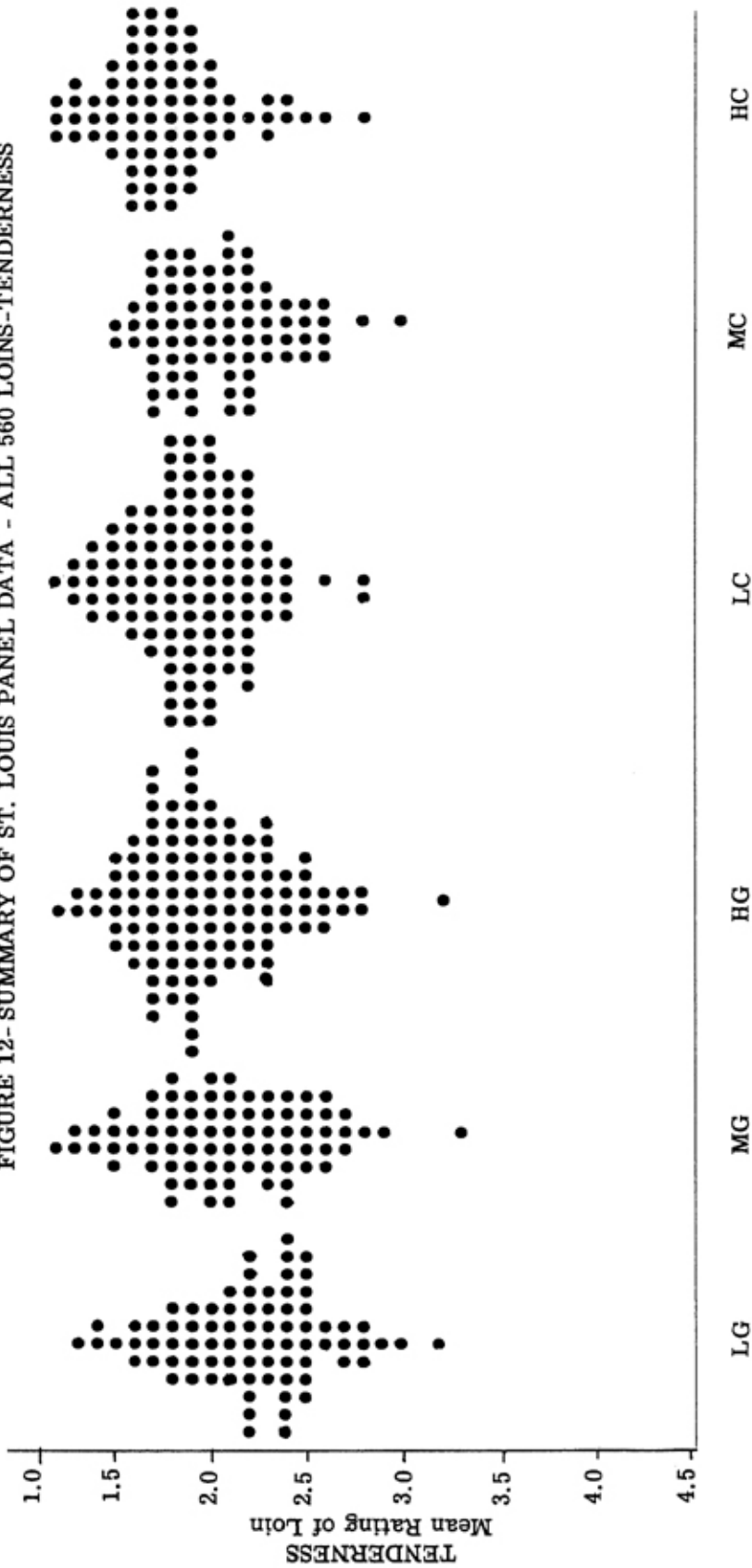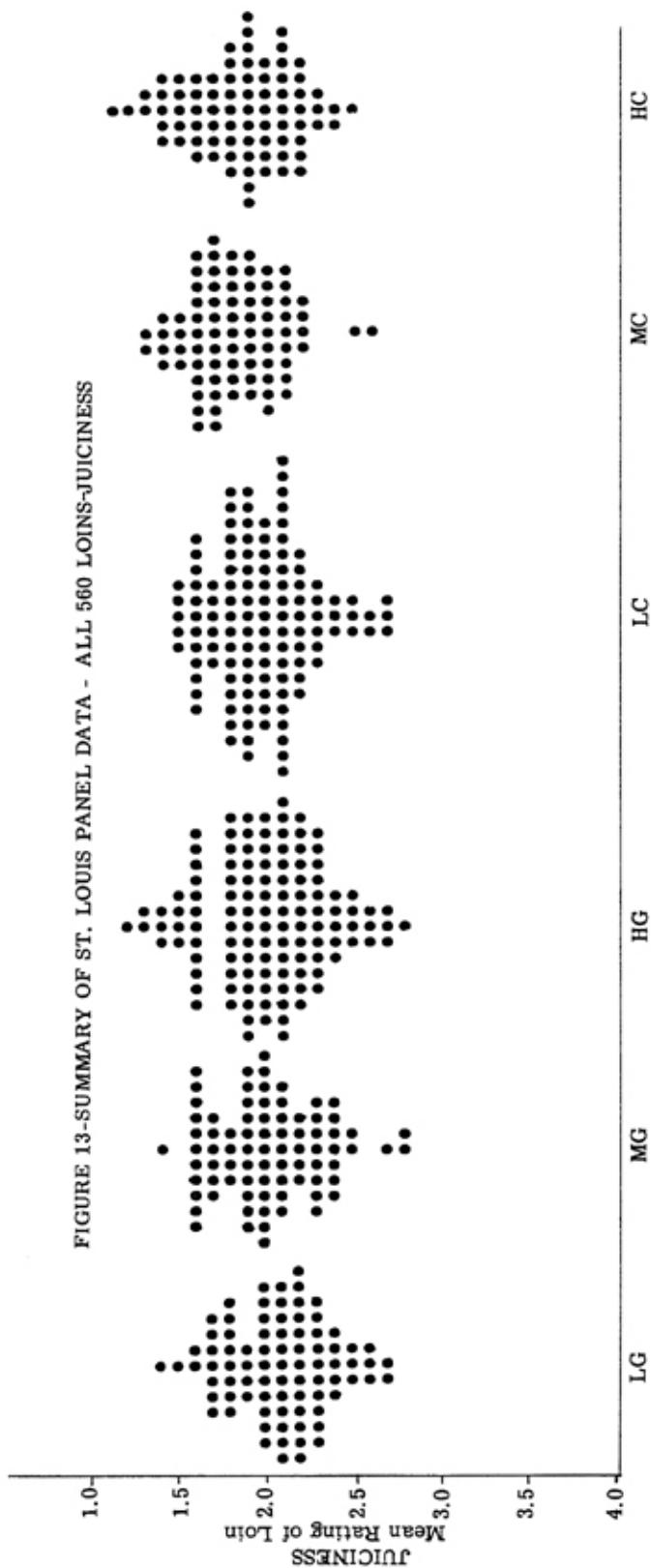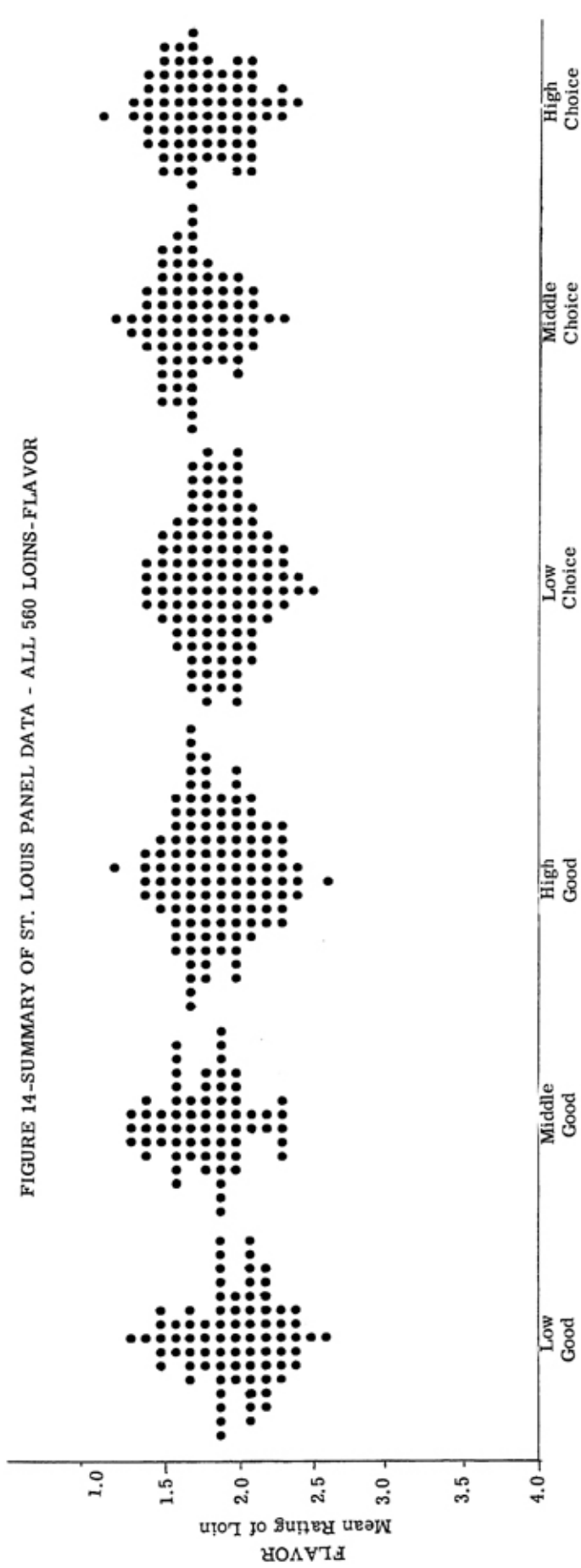
FIGURE 14—SUMMARY OF ST. LOUIS PANEL DATA – ALL 560 LOINS-FLAVOR

COMMENTS—GRADE

Consumers were asked to make comments concerning each steak. These comments were classified into 24 categories for the October and January replications.

The percentage of comments by category and grade classification appear in Tables 8 and 9. The lack of relation between comments and grades was apparent. Lack of tenderness was the most frequent single complaint.

TABLE 8 - PERCENT OF COMMENTS BY CATEGORIES FOR EACH GRADE (OCTOBER REPLICATION)

		Low Good	Middle Good	High Good	Low Choice	Middle Choice	High Choice
1.	General Approval	23.9	25.2	20.1	22.0	32.3	12.0
2.	Average	6.5	1.8	2.7	2.2	3.0	24.1
3.	General Disapproval	4.3	8.1	7.3	7.5	0	1.4
4.	Good Flavor	13.0	9.0	11.9	10.1	13.5	1.8
5.	Flavorless	5.4	6.3	7.3	9.3	2.3	13.3
6.	Objectionable Flavor	4.3	2.7	0.5	2.6	3.0	7.0
7.	Tender	10.9	17.1	11.0	11.0	17.3	1.8
8.	Tough	13.0	13.5	16.4	13.7	7.5	12.6
9.	Good Texture	1.1	0.9	1.0	0.4	0.8	7.2
10.	Poor Texture	2.2	1.8	5.9	2.2	0.8	5.4
11.	Juicy	1.1	1.8	3.2	3.1	5.3	2.4
12.	Dry	4.3	4.5	4.6	7.0	.8	1.2
13.	Too Fat	0	0	0.9	1.8	3.0	2.4
14.	Nice Leaness	0	0	0.9	1.3	2.3	0
15.	Too Lean	0	0.9	0	0.4	0	0
16.	Too Large	0	0.9	0	0	0	0
17.	Nice Size	0	0	0.5	0	0.8	0
18.	Too Small	1.1	0	0.5	0	0.8	0
19.	Too thick	0	0	0	0	0	0.6
20.	Nice Thickness	0	0	0.5	0.4	0.8	0
21.	Too Thin	1.1	0	1.4	1.8	1.5	2.4
22.	Well-Marbled	0	0	0	0	1.5	2.4
23.	Lacked Marbling	0	0	0	0	0	0.6
24.	Miscellaneous	7.6	5.4	3.6	3.1	3.0	2.6
	Totals	100.0	100.0	.00.0	100.0	100.0	100.0
	Number of Comments	(92)	(111)	(219)	(227)	(133)	(166)

DESCRIPTION OF ATTRIBUTES OF LOINS
WITH CONSUMER MEAN ACCEPTANCE RATINGS BELOW 3.75

An effort was made to determine the significant attributes of loins with unsatisfactory mean acceptance ratings. The factors considered were marbling ratings, grade category, Warner-Bratzler cooked shear value, and the consumer ratings for tenderness, juiciness, and flavor. Laboratory panel ratings, consumer cooking methods, and feeding data, where available were also examined.

TABLE 9-PERCENT OF COMMENTS BY CATEGORY FOR EACH GRADE
(JANUARY REPLICATION)

	Low Good	Middle Good	High Good	Low Choice	Middle Choice	High Choice
1. General Approval	36.1	29.5	50.5	38.3	52.2	58.0
2. Average	1.6	6.6	4.9	2.8	2.7	2.5
3. General Disapproval	6.6	4.9	4.9	2.8	0.9	1.7
4. Good Flavor	7.4	9.1	7.0	8.4	4.4	10.1
5. Flavorless	8.2	6.6	1.0	5.6	4.5	5.0
6. Objectionable Flavor	0.8	1.6	0	0	0.9	0.8
7. Tender	4.1	7.4	6.9	6.6	10.6	6.8
8. Tough	17.2	15.6	7.9	8.4	3.5	5.0
9. Good Texture	1.6	1.6	1.0	0	0	0.8
10. Poor Texture	2.5	1.6	0	2.8	0	0
11. Juicy	0.8	4.9	4.0	1.9	6.2	3.4
12. Dry	5.7	5.8	4.9	2.8	2.6	1.7
13. Too Fat	1.6	0	0	0.9	2.6	3.4
14. Nice Leanness	0	0	0	1.9	0	0
15. Too Lean	0	0	0	0	0	0
16. Too Large	0	0	0	0	0	0
17. Nice Size	0	0	0	0	0	0.8
18. Too Small	0	0	0	0	0	0
19. Too Thick	0	0	0	0	0	0
20. Nice Thickness	0	0.8	0	0.9	0	0
21. Too Thin	0	1.6	0	0	0	0
22. Well Marbled	0	0.8	1.0	0	1.8	0
23. Lacking Marbling	0	0	0	0	0	0
24. Miscellaneous	5.8	1.6	6.0	15.9	7.1	0
Totals	100.0	100.0	100.0	100.0	100.0	100.0
Number of Comments	(122)	(122)	(101)	(107)	(113)	(119)

Twenty-nine loins or approximately five percent of the 560 loins tested received consumer acceptance ratings lower than 3.75. The wide variation in marbling scores associated with the lower rated loins indicates that there was little relation between the unpalatability of these loins and degree of marbling. The mean marbling score of 7 for these poor loins fell between the mean scores of High Good and Low Choice in these tests. It may be significant that no highly marbled loins fell in this unsatisfactory group; however, only 4.2 percent of all loins had marbling scores above 4. Most of the unsatisfactory loins were of Good grade with only 5 of the 29 being Choice. Warner-Bratzler cooked shear values were generally higher than 18 pounds, which is believed to be a critical point for tenderness necessary for highly acceptable steaks. However, 9 of the loins had shears under 18 pounds. Shear was also an imperfect indicator of consumer palatability. The laboratory ratings suggested a lack of consumer acceptance in most cases, not in all. For example, 20 of these 26 loins fell in the lowest quartile of laboratory ratings for tenderness. The consumer attribute ratings parallel the consumer acceptance ratings closely with a bit more criticism of tenderness than the other attributes.

LABORATORY EVALUATION OF GRADES

A laboratory panel rated the steaks from each loin in a manner similar to the consumer panel. The attributes studied were tenderness, juiciness, and flavor.

The loin mean rating for all loins†‡ for each attribute by grade appear in Figures 15, 16, and 17. The result indicates that a slight, but positive, relation exists between grades and the three attributes tested (Figures 15, 16, and 17). This was in general consistent with the results of the consumer panel. The differences among grades were the most pronounced in the tenderness attribute. The development of a better tenderizer would probably reduce the general acceptance differential between grades.

The simple correlation between consumer acceptance and laboratory tenderness was .47. This value was relatively low, but was consistent with previous studies.

The simple correlation between laboratory tenderness and marbling was .27. The correlations of laboratory juiciness and flavor with marbling were .43 and .22, respectively.

FIGURE 15-SUMMARY OF LABORATORY PANEL DATA - ALL 520 LOINS-TENDERNESS
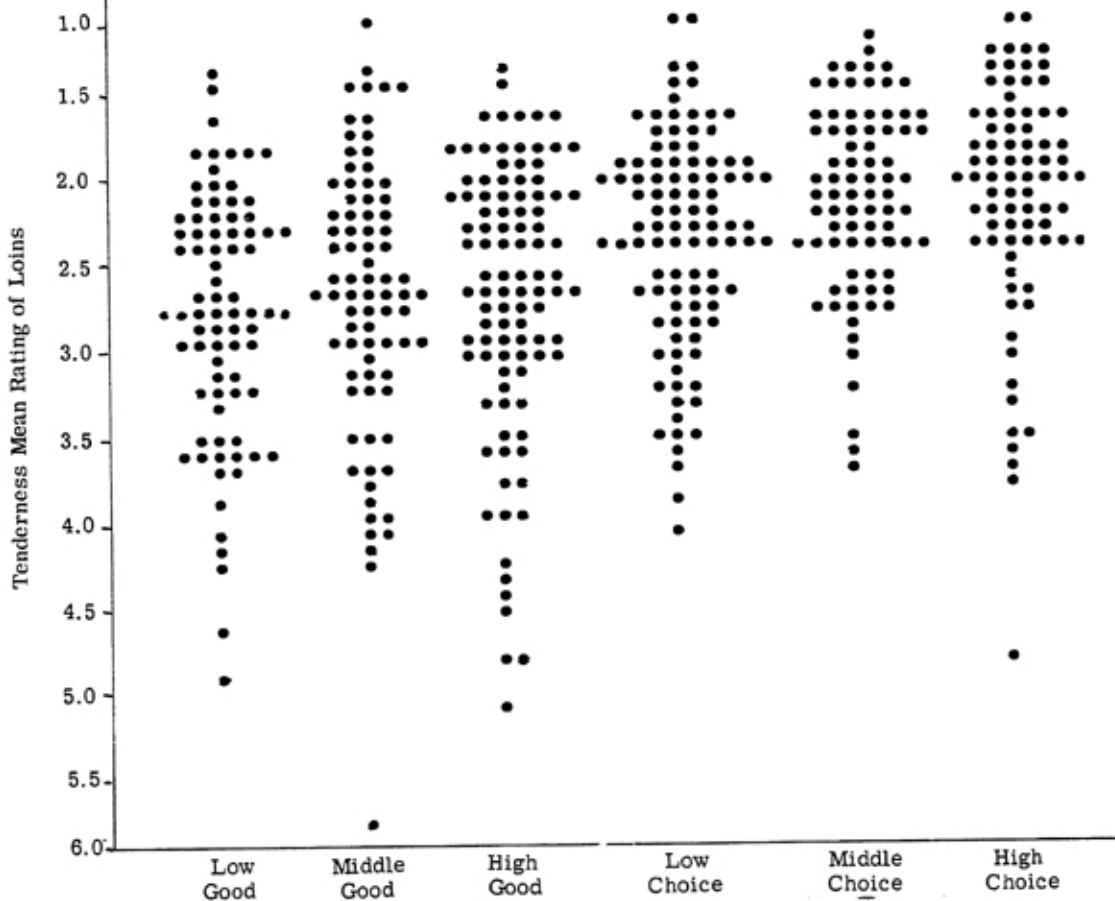
†‡Laboratory evaluations were not made on the first 40 "pretest" loins.

FIGURE 16–SUMMARY OF LABORATORY PANEL DATA - ALL 520 LOINS–JUICINESS
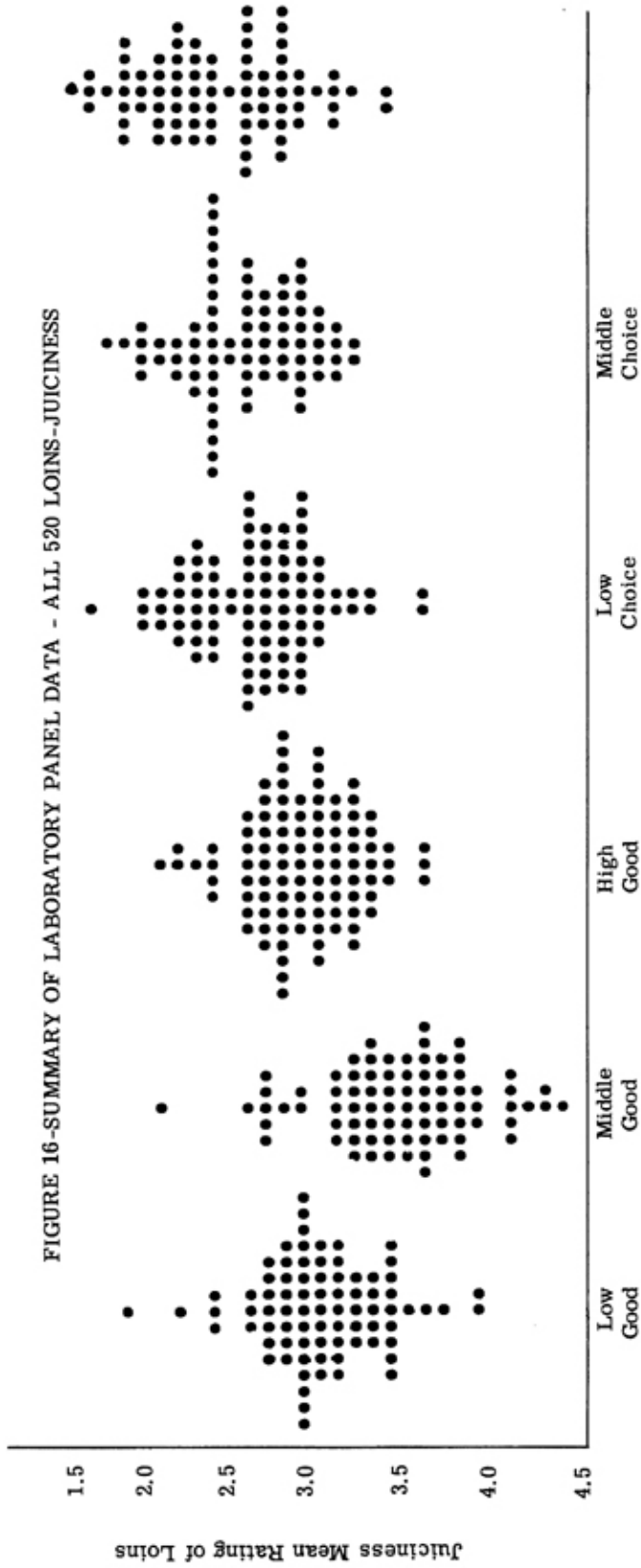
FIGURE 17-SUMMARY OF LABORATORY PANEL DATA - ALL 520 LOINS-FLAVOR

MISSOURI AGRICULTURAL EXPERIMENT STATION
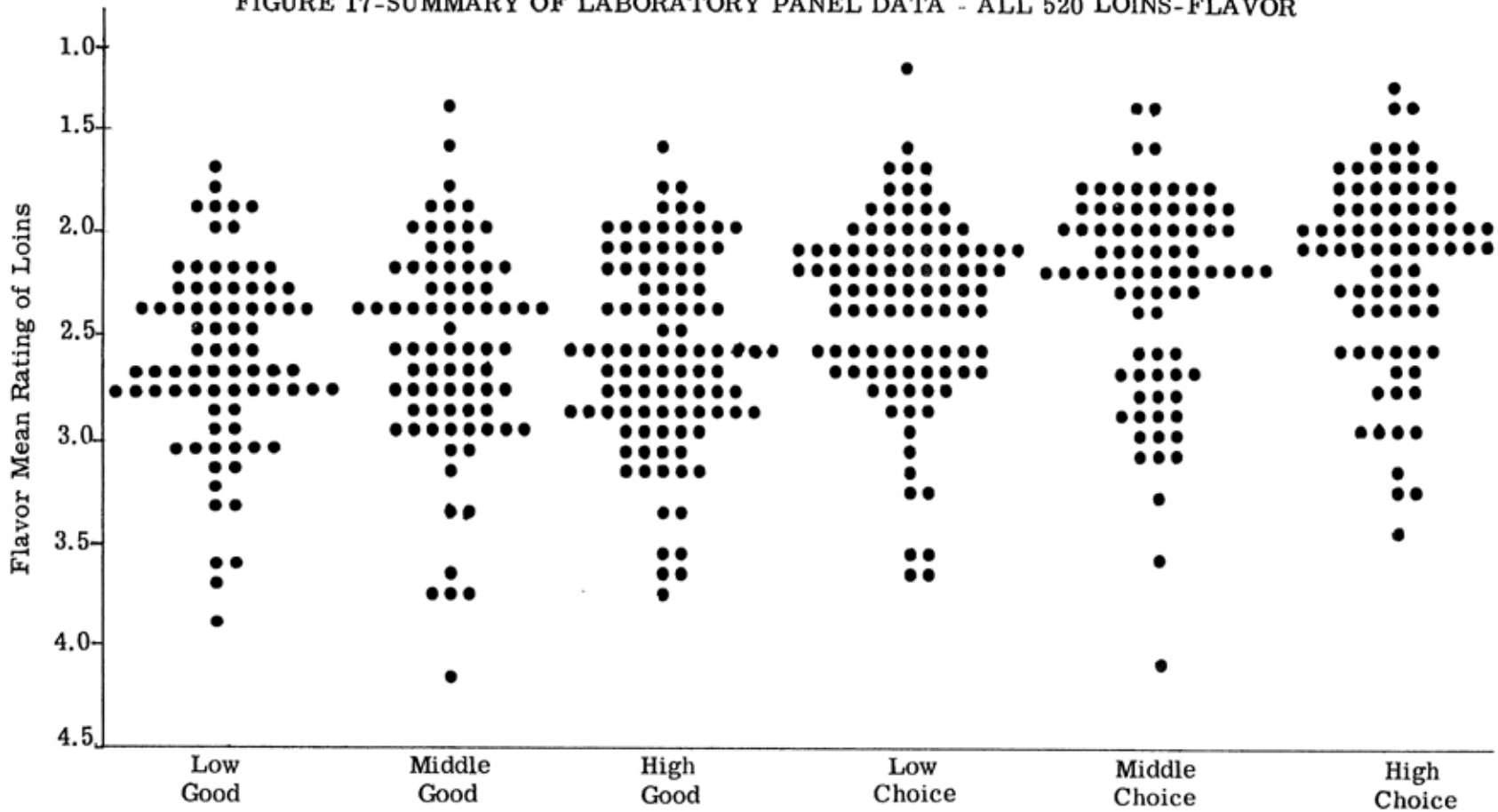
COMPARISON OF ACCEPTANCE AND LABORATORY EVALUATION

Consumer and laboratory panels rating beef loins both indicated a weak, but positive, relationship of palatability to grade and marbling categories. The general results of laboratory and consumer ratings of steak attributes were consistent. But the correlation between laboratory and consumer evaluations was relatively low. This aspect of the study was consistent with previous studies. While the error in measuring both variables explains part of the difficulty, the differences in size of sample, nature of cooking, personal preferences, and in general frame of reference are probably other factors affecting these low correlation coefficients.

REVIEW OF LITERATURE

Eating Quality of Choice and Good Beef

A number of researchers have studied the comparative merits of Choice and Good grades as to eating qualities, yields, and cooking losses. For convenience, the postwar studies on eating quality are summarized in Table 10. The reader will be aware that factors such as sex and age of animal, segments of grade, aging of carcass, cooking, etc. vary among the studies.‡‡ While the experimental details vary, the pattern of results is consistent and clear. Most Choice and Good carcasses have similar eating qualities, but a few Good carcasses are inferior to the Choice. Therefore, some studies find no difference, while others find a small average difference in favor of the Choice. In all reports in which individual carcass results are given, there is the same overlapping pattern found in this study.

Marbling and Eating Quality

A few prewar studies with somewhat conflicting evidence as to the relationship of marbling and eating quality are reviewed by Wellington and Stouffer. Postwar studies of this relationship have had the following results:

1. Doty[16] reported a linear correlation coefficient of 0.56 between marbling and tenderness as judged by a laboratory panel. He used rib-eyes from 42 carcasses of Prime, Good, and Commercial Cow.
2. Wellington and Stouffer[17] reported an r of 0.26 between marbling and tenderness on loin steaks from 121 carcasses of Prime, Choice, Good and Standard grades. Tenderness was appraised by a laboratory panel.

Trimmed Retail Yield of Choice and Good

Research is still in progress in this area. It is now clear that Good grade
‡‡A large laboratory study at Ohio[3] is omitted because the Choice animals were bulls. The authors concluded that age and sex were important indicators of tenderness and that marbling was not important except as it was an indicator of a sex difference.

TABLE 10-SUMMARY OF RECENT COMPARISONS OF PALATABILITY OF CHOICE AND GOOD

Research Institution	Number of Carcasses		Cuts Used	Consumers or Laboratory Panel	Palatability Results
	Good	Choice			
Missouri 4/	6	2	Loin, sirloin, round and rib steaks; blade and arm roasts	32 households	No difference
Missouri 1/	23	13	Loin steaks	60 households	Choice slightly better
Missouri 2/	21	21	Loin steaks	266 households	Choice slightly better
Quartermaster 5/6/	100	100	Loin, rib & sir-butt steaks round, knuckle & rump roasts	Lab panel & hundreds of soldiers	No difference (lab panel found Choice tenderness slightly better.)
Arizona 7/	a/	a/	Loin steaks and chuck roasts	520 customers	No difference
Texas 8/	54	15	Loin steaks	Lab panel	Choice slightly better
California 10/	3	3	Loin and sirloin roasts	151 consumers	Choice slightly better
California 9/	9	9	Rib and round roasts	Lab panel	Choice slightly better
Tennessee 11/	9	9	Loin and round steaks	Lab and consumers	Choice slightly better
Tennessee 12/	20	20	Rib roasts	Lab panel	Choice slightly better
Missouri 13/	4	4	Rib, round & loin steaks	Lab panel	No difference
Oklahoma 14/	5	5	Loin steaks	Lab panel	No difference
New York 15/	14b/	42	Rib roasts	Lab panel	Choice better

a/ Store sales experiment.
b/ These varied from Standard to Good (present grade equivalent) while the 42 "Choice" varied from Choice to Prime.

carcasses, on the average, yield a somewhat higher percentage of trimmed retail cuts than do Choice carcasses. Since a number of factors affect this yield relationship, it is not clear exactly what percentage better yield of trimmed cuts can be expected from Good. Quartermaster researchers report percentage differentials of 0.79, 1.54, and 4.62, in favor of Good over Choice in 3 lots of steers, and 2.10, and 4.17, in favor of Good over Choice in 2 lots of heifers. Each lot was composed of 20 Good and 20 Choice carcasses of similar weights. These Quartermaster results were yields of boneless beef cuts trimmed to a uniform half inch of external fat.[18]

Published fragmentary results of USDA researchers indicate that trimmed retail yield is primarily an inverse function of finish and, to a lesser extent, is also a function of conformation.[19,20]

Butler has also reported that fatness of the carcass is the main determinant of percentage yield of trimmed cuts.[21] Thus, it follows that Good grade carcasses with generally less external fat than Choice carcasses also have better yields to the retailer of trimmed retail cuts.

Cooking Losses of Choice and Good

A review of literature by Alexander and Clark indicates that cooking losses are mainly related to the cooking temperature and time and to the fat/lean composition of the meat.[22] Generally, the higher the grade, the less the evaporation loss and the higher the drippings loss. Cooking losses in Missouri experiments on a large number of loin steaks broiled to well done were very slightly smaller for Good than Choice.[23] Cooking losses in a recent Tennessee experiment[12] averaged 0.21 percent larger for 20 Good rib roasts than for 20 Choice rib roasts. It appears that cooking losses are too similar between Choice and Good grades to influence their relative value to consumers.

APPENDIX A—RESEARCH PROCEDURE, SELECTION OF PANEL

A representative sample of middle and upper income white families was desired. Therefore, the population was limited to a large section of urban St. Louis County. This area is largely made up of the better income type of suburban developments found around many of our larger cities. The area was bounded on the east by the City of St. Louis, on the south by City Highway 66, on the west by the circumferential highway known as Lindbergh Boulevard, and on the north by St. Charles Rock Road.

A two-stage probability sample was drawn. Twenty cluster areas or neighborhoods were drawn and a serial sample of 10 households was drawn from each neighborhood. A sample of 20 tracts were drawn from the 42 census tracts making up the area. The probability of drawing a tract was weighted by its proportionate estimated population. Population estimates were developed from adjustments of 1950 census counts by building permit data.

A starting point within each of the 20 tracts was randomly selected by a 0.25 inch grid overlay. The experienced interviewer began at the designated starting point and, circling each block clockwise, interviewed every other household. A set of contiguous blocks was randomly specified. The selection of households to be interviewed required no exercise of judgment by interviewers.

Of 763 households contacted, 206 became cooperators, while 229 were not at home, 198 were ineligible, 130 did not wish to participate. Since only one call-back was made, the percentage of not-at-home was relatively high, and the proportion of cooperating families with working wives was probably lower than in the population.

At the initial interview, certain restrictions were placed on participation to ensure that panel members were members of the "steak market" and to facilitate the practicable operation of the panel. Restrictions were as follows:

1. Monthly household income of $250 or more.
2. Minimum of two adults under age 65 in household and regularly eating beef at home.
3. Cooperators who expected to reside in the area for the duration of the project.
4. Both adults could not be males.
5. The two cooperators in each household agreed to eat the steaks provided, to use similar methods of preparation throughout the study, and to complete the evaluation schedules provided.
6. Persons cooperating were required to have no previous training in meats or meat cutting.

Product Specifications
and
Handling Procedures

Short loins were obtained from 560 carcasses weighing 600 to 650 pounds (chilled) and identified by third of grade by regular federal graders in the packing plants. The loins were aged 10 to 11 days at 36 to 38°F. Thirteen steaks—three for the laboratory panel and ten for the consumer panel—were cut from each loin. Each steak was ¾ inch in thickness and was trimmed to a maximum of ⅜ inch of outside fat. Two 1½ inch thick steaks were used for shear tests.

Steaks for the consumer panel were individually wrapped in laminated freezing paper and assigned code numbers based on delivery, grade, neighborhood, steak position, household and loin. Individual steaks were paired and packaged for household delivery. The steaks were delivered to frozen-food lockers in St. Louis in preparation for distribution to consumers.

Operation of Panel

The three interviewers who had recruited the panel made regular weekly deliveries of two steaks per household. The two steaks were from adjacent thirds of grades, but were from the same positions on the loins and so were of the

same shape. The husband's steak was always identified with a metal ring clamped on the bone.

To maximize the socio-economic diversity of cooperators testing each loin, its ten steaks went to 10 different neighborhoods, where they were eaten by 5 men and 5 women.

Each person received nearly the full range of grades during each test "month." For example, each of 200 persons ate High Choice, Low Choice, and Middle Good loins during each and every test "month." At the same times, their 200 spouses ate Middle Choice, High Good and Low Good loins. During the October test, the regular testers of Low Choice also tested High Good loins, and the regular testers of High Good also tested Low Choice Loins.

No information was given cooperators concerning the grade or identity of the steaks tested. All steaks packages were coded for the purpose of delivery and analysis.

Each of the 400 cooperators filled out an evaluation sheet. Cooperating families were told repeatedly that the husband's steak was from a different carcass than the wife's, so that they would feel free to rate them independently.

Four deliveries were made in March, 3 in May, 4 in September-October, 1959, and 3 in January, 1960.

Panel cooperation was excellent during the 12 month period. A few families were replaced because of moving out of the area, ill health and other reasons. However, 91.5 percent of the original panel completed the project.

APPENDIX B—SEASONAL PATTERN OF ACCEPTANCE

The general experiment was replicated by seasons to determine whether there were seasonal differences in consumer acceptance of the beef steaks.

The general distribution of acceptance ratings by months of *slaughter* appear in Figure 18. There was a slight difference in favor of the distribution of ratings for December as compared to the other months (Figure 17). The distributions of ratings for the other three periods were quite similar.

More detailed distributions by seasons for adjacent two thirds of grades appear in Figures 18, 19, and 20. The distribution of Low Good and Middle Good rating appears to favor cattle slaughtered in April and May (Figure 18). The distribution of ratings for High Good and Low Choice and for Middle Choice and High Choice appears slightly better for cattle slaughtered in December (Figure 19 and 20).

Marbling of loins was rated from 1 through 10 with 1 being very abundant, and 10 indicating only traces of marbling. The loins for the second, third, and fourth replicate were rated according to this scale.

The average marbling score of the loins by grade and replicate appear in Table 7. There were small differences among the average marbling scores of the three replicates. Since the December replicate had somewhat less marbling on the average than the other two replicates, its slightly superior ratings were not explained by greater marbling.

Figure 18   St. Louis County Beef Panel Mean Ratings of 120 Short Loins Each Season
Months Indicate the Slaughter Dates of the Cattle Tested
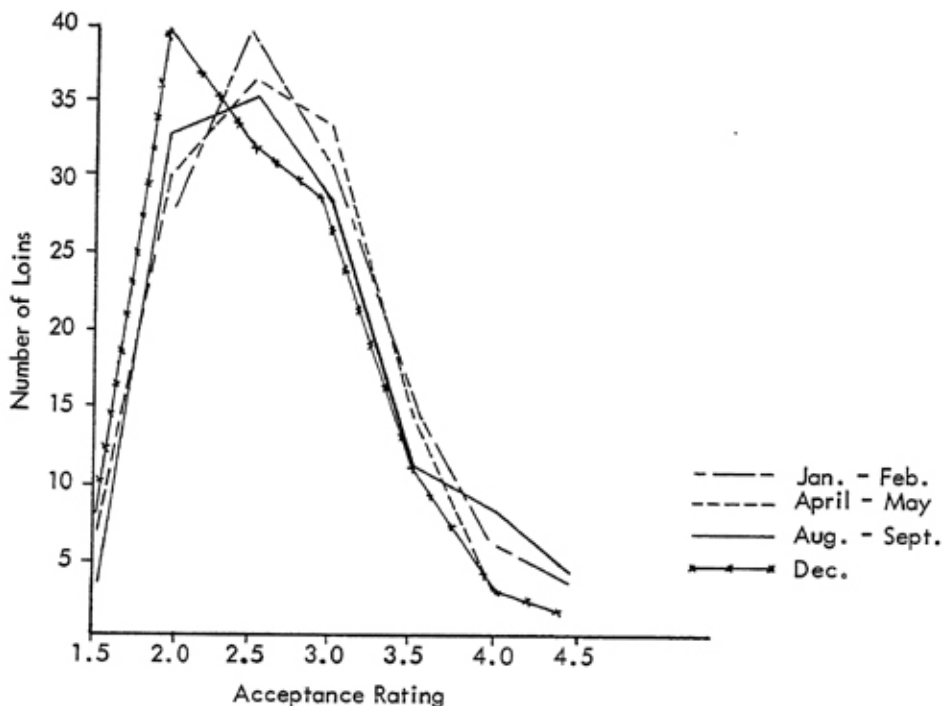
APPENDIX C—COOKING METHODS

Consumers classified the method of cooking steaks as to degree of doneness
and use of moist or dry heat. The number and percent of steaks in each category
was computed in terms of the rating for tenderness, juiciness, and flavor. This
method of analysis was performed for both the October and January (3rd and
4th) replications.

The results of the method of cooking analysis appear in Tables 11, 12,
13, 14, 15, and 16. The tenderness distributions appeared better for the moist
heat cooking method for both replications (Tables 11 and 12). There was a
slight difference in the juiciness distributions in favor of the moist heat cooking
method for both replications. (Tables 13 and 14). The distribution of flavor
ratings also favored the moist heat method of cooking (Tables 15 and 16).

In general, the distribution of attribute ratings favored rare steaks for both
methods of cooking. However, the moist-heat-rare category must be interpreted
cautiously as the number of steaks was very small and the degree of rareness was
probably a very minimum amount of pink. It is also interesting to note that for
all attributes the October ratings were less favorable than the January ratings.
This result was consistent with the seasonal acceptance results.

TABLE 11—PERCENT AND NUMBER OF STEAKS RATED AS TO TENDERNESS BY MOIST AND DRY COOKING METHODS AND DEGREE OF DONENESS

October Replication

Method of Cooking	Rating	More Tender		Average Tender		Less Tender		Tough		Total Steaks
		No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	
MOIST	Well Done	54	28.72	95	50.53	22	11.70	17	9.05	188
	Rare	10	22.20	23	51.10	9	20.00	3	7.70	45
Sub-total		64	27.46	118	50.64	31	13.30	20	8.60	233
DRY	Well Done	169	25.40	277	41.60	164	24.60	56	8.40	666
	Rare	235	34.97	291	43.30	104	15.48	42	6.25	672
Sub-total		404	30.19	568	42.45	268	20.03	98	7.33	1338
Grand Total or Mean		468	29.79	686	43.67	299	19.03	118	7.51	1571

TABLE 12—ERCENT AND NUMBER OF STEAKS RATED AS TO TENDERNESS BY MOIST AND DRY COOKING METHODS AND DEGREE OF DONENESS

January Replication

Method of Cooking	Rating	More Tender		Average Tender		Less Tender		Tough		Total Steaks
		No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	
MOIST	Well Done	61	41.8	61	41.8	18	12.3	6	4.1	146
	Rare	19	52.8	14	38.9	2	5.5	1	2.8	36
Sub-total		80	44.0	75	41.2	20	11.0	7	3.8	182
DRY	Well Done	162	31.6	247	48.3	84	16.4	19	3.7	512
	Rare	183	37.4	207	42.3	86	17.6	13	2.7	489
Sub-total		345	34.5	454	45.3	170	17.0	32	3.2	1001
Grand Total or Mean		425	35.9	529	44.7	190	16.1	39	3.29	1183

TABLE 13-PERCENT AND NUMBER OF STEAKS RATED AS TO JUICINESS BY MOIST AND DRY COOKING METHODS AND DEGREE OF DONENESS

October Replication

Method of Cooking	§ Rating	More Juicy		Average Juicy		Less Juicy		Dry		Total Steaks
		No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	
MOIST	Well Done	45	24.06	103	55.08	29	15.51	10	5.35	187
	Rare	10	22.22	30	66.67	4	8.89	1	2.22	45
Sub-total		55	23.70	133	57.34	33	14.22	11	4.74	232
DRY	Well Done	123	18.55	337	50.83	156	23.53	47	7.09	663
	Rare	202	30.10	340	50.67	107	15.95	22	3.28	671
Sub-total		325	24.36	677	50.75	263	19.71	69	5.18	1334
Grand Total or Mean		380	24.27	810	51.72	296	18.90	80	5.11	1566

TABLE 14-PERCENT AND NUMBER OF STEAKS RATED AS TO JUICINESS BY MOIST AND DRY COOKING METHODS AND DEGREE OF DONENESS

January Replication

Method of Cooking	Rating	More Juicy		Average Juicy		Less Juicy		Dry		Total Steaks
		No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	
MOIST	Well Done	50	34.48	69	47.59	21	14.48	5	3.45	145
	Rare	16	44.5	15	41.7	4	11.1	1	2.7	36
Sub-total		66	36.46	84	46.41	25	13.81	6	3.31	181
DRY	Well Done	120	23.52	267	52.35	101	19.80	22	4.3	510
	Rare	197	40.62	208	42.89--	65	13.40	15	3.09	485
Sub-total		317	31.86	475	47.74	166	16.68	37	3.72	995
Grand Total or Mean		383	32.67	559	47.53	191	16.24	43	3.66	1176

TABLE 15-PERCENT AND NUMBER OF STEAKS RATED AS TO FLAVOR BY MOIST AND DRY COOKING METHODS AND DEGREE OF DONENESS

October Replication

Method of Cooking	Rating	Better Flavor		Average Flavor		Poor Flavor		Terrible		Total Steaks
		No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	
MOIST	Well Done	51	25.42	110	59.14	23	12.36	2	1.08	186
	Rare	17	38.64	22	50.00	5	11.36	0	0	44
Sub-total		68	29.57	132	57.39	28	12.17	2	.87	230
DRY	Well Done	181	27.42	336	50.91	129	19.55	14	2.12	660
	Rare	213	31.79	342	51.04	106	15.82	9	1.35	670
Sub-total		394	29.62	678	50.98	235	17.67	23	1.73	1330
Grand Total or Mean		462	29.62	810	51.92	263	16.86	25	1.60	1560

TABLE 16- PERCENT AND NUMBER OF STEAKS RATED AS TO FLAVOR BY MOIST AND DRY COOKING METHODS AND DEGREE OF DONENESS

January Replication

Method of Cooking	Rating	Better Flavor		Average Flavor		Poorer Flavor		Terrible		Total Steaks
		No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	No. of Steaks	Percent of Steaks	
MOIST	Well Done	64	44.14	68	46.90	12	8.28	1	.69	145
	Rare	17	48.57	16	45.71	2	5.71	0	0	35
Sub-total		81	45.00	84	46.67	14	7.78	1	.56	180
DRY	Well Done	156	30.59	279	54.70	71	13.92	4	.78	510
	Rare	173	35.38	242	49.49	67	13.70	7	1.43	489
Sub-total		329	32.93	521	52.14	138	13.81	11	1.10	999
Grand Total or Mean		410	34.77	605	51.31	152	12.89	12	1.02	1179

APPENDIX D—ANALYSIS OF VARIANCE OF LOIN WITHIN GRADES, SEASONS, AND BETWEEN GRADES

Since the basic consumer experiment was repeated 4 times, the first hypothesis tested was equality of all mean attribute ratings of steak by grades between replications. An analysis of variance model was utilized to test the hypothesis of seasonal equality of the thirds of grades tested in the study. The use of such a preliminary analysis also provided criteria for decisions concerning the aggregation of replicates for further analysis.

The observed F ratio and degrees of freedom concerning the test of seasonal differences in loin mean attribute ratings for all 6 grades appear in Table 17. All grades except Middle Good produced F ratios significant to the 5 percent level for at least one of the attributes tested on a seasonal basis (Table 17). F ratios tended to increase with higher grades (Table 17). High Choice produced significantly different seasonal ratings for all attributes at the 5 percent level or less (Table 17).

The F tests for seasonal differences among the same grade suggested some possibility for seasonal variation among these grades. Thus, the succeeding analysis of variance tests were performed on an individual seasonal replicate basis.

TABLE 17-OBSERVED F RATIOS FOR TEST OF SEASONAL DIFFERENCES ATTRIBUTES FOR LOIN MEAN RATINGS

	Low Good	Middle Good	High Good	Low Choice	Middle Choice	High Choice
Acceptance	1.80	.83	3.17*	1.20	2.86*	4.04**
Tenderness	.69	.30	3.83*	4.25**	3.44	3.33*
Juiciness	1.38	.89	3.67*	2.25	2.33	4.38**
Flavor	3.14*	.62	1.43	.50	2.20	4.00*
DFb	3	3	3	3	3	3
DFw	76	76	96	96	76	76
DFt	79	79	99	99	79	79

*Significant at the five per cent level
**Significant at the one per cent level

ANALYSIS OF VARIANCE OF DIFFERENCES IN ATTRIBUTE RATINGS BETWEEN LOINS WITHIN GRADES BY SEASONAL REPLICATIONS

The analysis of variance was used to test hypothesis of equality of loin mean attribute ratings of loins within the same grade. The F ratios pertaining to the various attribute ratings within loins of the same grade for the four replications appear in Tables 18, 19, 20, and 21. These tests indicate some hetero genity of attribute ratings within grades although the tests were not all significant for all attributes in all grades for the four replications. The third replicate appeared to be the most heterogeneous within grades, while the fourth replicate appeared to be the least heterogeneous within grades for the various steak attributes.

Figure 19   Acceptance Ratings by Season of Slaughter

Jan. - Feb.
Apr. - May
Aug. - Sept.
Dec.

Percentage of Loins

Mean Acceptance Rating (Low Good and Middle Good)

52.5

Figure 20   Acceptance Ratings by Season of Slaughter

Jan. - Feb.
Apr. - May
Aug. - Sept.
Dec.

Mean Acceptance Ratings (High Good - Low Choice)
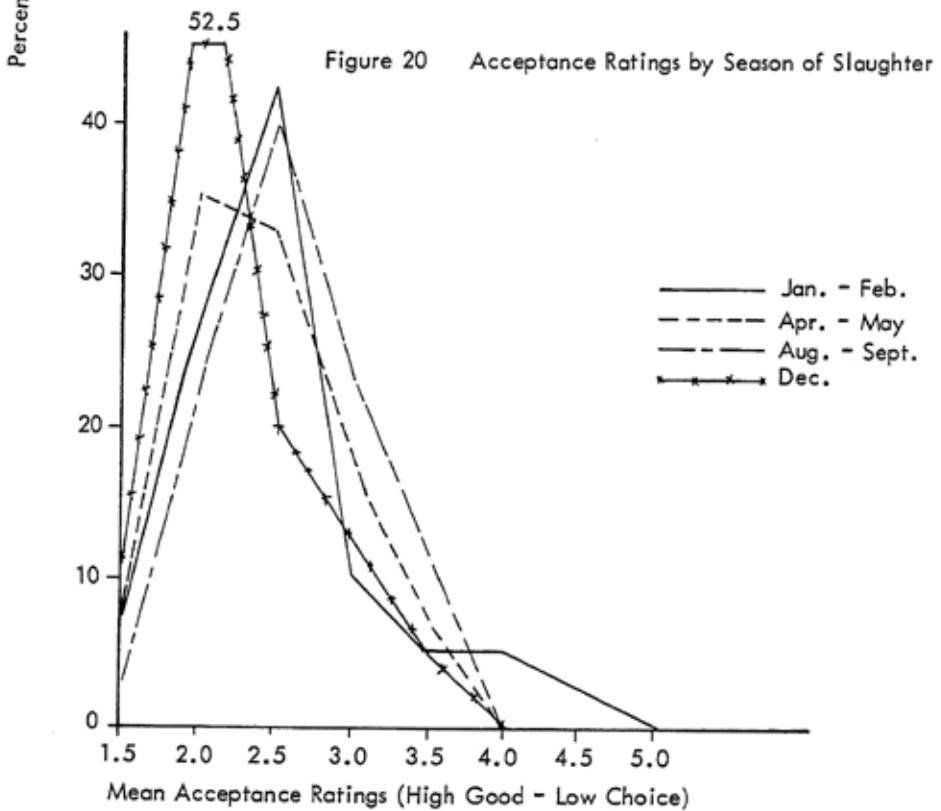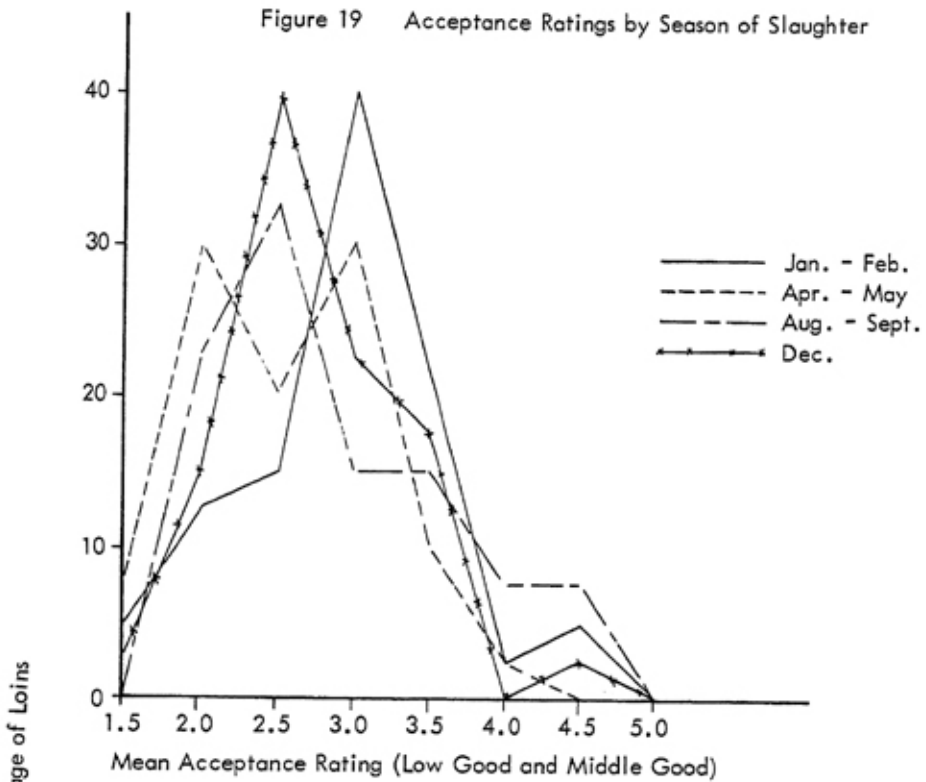
TABLE 18-OBSERVED F RATIOS FOR TESTS OF DIFFERENCE WITHIN LOIN GRADE ATTRIBUTE RATINGS, FIRST REPLICATION

	Low Good	Middle Good	High Good	Low Choice	Middle Choice	High Choice
Acceptance	1.29	1.64	2.04*	2.71	1.22	1.26
Tenderness	3.27*	.50	1.36	1.56*	1.03	1.10
Juiciness	1.75*	1.76*	1.62*	1.84*	1.09	.80
Flavor	2.26*	1.98*	1.73*	1.86*	.61	1.14

*Significant at the five per cent level or lower

DFb	19	19	19	19	19	19
DFw	180	180	180	180	180	180
DFt	199	199	199	199	199	199

TABLE 19-OBSERVED F RATIOS FOR TEST OF DIFFERENCES WITHIN LOIN GRADE ATTRIBUTE RATINGS, SECOND REPLICATION

	Low Good	Middle Good	High Good	Low Choice	Middle Choice	High Choice
Acceptance	1.02	1.63	.39	1.28	2.47*	2.62*
Tenderness	1.50	1.59	2.02*	.54	1.76*	1.75*
Juiciness	1.69	1.38	1.83*	.95	1.48	1.79*
Flavor	1.71*	1.28	1.63*	1.38	1.88*	1.98*

*Significant at the five per cent level or lower

DFb	19	19	19	19	19	19
DFw	180	180	180	180	180	180
DFt	199	199	199	199	199	199

'TABLE 20-OBSERVED F RATIOS FOR TEST OF DIFFERENCE WITHIN LOIN GRADE ATTRIBUTE RATINGS, THIRD REPLICATION

	Low Good	Middle Good	High Good	Low Choice	Middle Choice	High Choice
Acceptance	1.52	3.00*	1.35	.99	1.25	2.15*
Tenderness	2.36*	3.42*	2.49*	1.57*	2.13*	2.74*
Juiciness	1.67	1.88*	1.71*	1.28	2.02*	3.46*
Flavor	1.82*	2.18*	1.96*	1.06	1.70	3.12*

*Significant at the five per cent level or lower

DFb	19	19	39	39	19	19
DFw	180	180	360	360	180	180
DFt	199	199	399	399	199	199

TABLE 21-OBSERVED F RATIOS FOR TEST OF DIFFERENCE WITHIN LOIN GRADE ATTRIBUTE RATINGS, FOURTH REPLICATION

	Low Good	Middle Good	High Good	Low Choice	Middle Choice	High Choice
Acceptance	1.09	1.81*	1.85*	1.48*	.89	1.26
Tenderness	1.71*	2.10*	1.63*	1.09	1.62	.60
Juiciness	.88	1.12	2.25*	1.14	1.34	1.30
Flavor	1.02	1.30	1.88*	1.04	1.02	1.00

*Significant at the five per cent level or lower

DFb	19	19	19	19	19	19
DFw	180	180	180	180	180	180
DFt	199	199	199	199	199	199

ANALYSIS OF VARIANCE OF DIFFERENCES BETWEEN GRADES ATTRIBUTE RATINGS BY SEASONAL REPLICATION

The analysis of variance was used to test the hypothesis of equality of loin mean attribute ratings between grades for the four attributes.

The F ratios for the different attributes by season appear in the Table 22. The fourth replicate indicates the strongest evidence for differences in attribute ratings between grades (Table 22). The second replicate indicated little difference in the attributes between grades since the tenderness rating was the only significantly different variable between grades (Table 22). All attributes were significantly different between grades for the first replicate but F ratios were considerably lower than for the fourth replicate (Table 22). Flavor was not significantly different between grades for the second and third replicates.

The large variation in F ratios for attributes between grades by seasonal replication suggests further evidence of discernible differences in loin attributes between seasons. Possible reasons for such variation are still to be delineated.

TABLE 22-OBSERVED F RATIO FOR TEST OF DIFFERENCE BETWEEN GRADE MEAN ATTRIBUTE RATINGS BY SEASONAL REPLICATION

	1st Rep	2nd Rep	3rd Rep	4th Rep
Acceptance	4.73	1.84	5.65*	12.17*
Tenderness	5.38*	1.30*	5.60*	13.11*
Juiciness	3.00**	1.75	4.22*	14.14*
Flavor	5.00*	.57	.47	9.20*

*Significant at the one per cent level.
**Significant at the five per cent level.

DFb	5	5	5	5
DFw	114	114	154	114
DFt	119	119	159	119

Summary of the Analysis of Variance

1. In general, there appears to be rather strong evidence of discernible consumer differences in steaks between thirds of grades.
2. Some evidence of seasonal difference in steaks within grades was delineated by the analysis of variance tests. Possible reasons for these differences are too numerous to allow for more than speculation at this stage of research development.
3. Differences in loins within grades as measured by analysis of variance was not as large as might have been expected. This was apparently the result of large variation in consumer use of the hedonic scale in evaluating steaks cut from the same loin.

APPENDIX E—METHODOLOGY

I. Comparison of High Good and Low Choice by all 400 consumers.

The experimental design provided for each of two groups of 200 consumers to test alternate thirds of the grade. To enable a more direct comparison of

High Good and Low Choice, an extra third of a grade was tested in the October test. Those normally receiving High Good also tested Low Choice and those normally receiving Low Choice also received High Good loins. Ratings are shown in Table 23.

The difference in ratings of the two grades is almost the same as in all the other tests. However, both grades were rated a little less favorably than was usually the case.*

TABLE 23-RATING OF HIGH GOOD AND LOW CHOICE BY ALL 400 CONSUMERS

	Mean	Median	Upper & Lower Boundaries of Middle Quartile
High Good	2.89	2.85	2.55=3.25
Low Choice	2.71	2.65	2.30-3.05

Figure 21

Jan. - Feb.
Apr. - May
Aug. - Sept.
Dec.

Percentage of Loins

Mean Acceptance Rating (Middle Choice and High Choice)

TABLE 24—ARRAY, RANGE, AND AVERAGE OF CARCASS MEANS BY "GRADE"

LG 1	LG 2	MG 1	MG 2	HG P	HG 1	HG 2	LC P	LC 1	LC 2	MC 1	MC 2	HC 1	HC 2
2.1	1.8	1.9	1.8	2.0	1.9	1.9	2.0	1.9	1.6	1.8	1.7	1.9	1.7
2.2	1.8	1.9	2.0	2.0	1.9	2.1	2.1	2.0	1.9	2.0	1.7	2.0	1.8
2.2	2.2	2.1	2.1	2.0	2.1	2.2	2.1	2.0	2.1	2.1	1.8	2.1	1.9
2.7	2.3	2.4	2.2	2.1	2.2	2.3	2.1	2.1	2.2	2.1	1.8	2.1	1.9
2.7	2.4	2.6	2.4	2.3	2.3	2.3	2.3	2.2	2.2	2.1	2.0	2.1	2.0
3.0	2.4	2.8	2.4	2.3	2.4	2.4	2.3	2.4	2.3	2.1	2.0	2.2	2.1
3.1	2.7	2.9	2.4	2.4	2.5	2.4	2.4	2.4	2.3	2.3	2.1	2.3	2.2
3.2	2.8	2.9	2.4	2.4	2.6	2.5	2.4	2.4	2.3	2.3	2.1	2.3	2.2
3.2	2.9	3.0	2.5	2.5	2.6	2.7	2.4	2.5	2.4	2.3	2.1	2.4	2.3
3.2	2.9	3.0	2.6	2.5	2.6	2.7	2.5	2.5	2.4	2.4	2.3	2.4	2.3
3.3	2.9	3.1	3.1	2.7	2.6	2.7	2.5	2.5	2.5	2.4	2.4	2.4	2.3
3.4	3.0	3.1	3.1	2.8	2.6	2.8	2.5	2.6	2.6	2.6	2.4	2.6	2.4
3.4	3.0	3.1	3.1	2.8	2.7	2.8	2.7	2.6	2.8	2.6	2.4	2.6	2.4
3.5	3.0	4.1	3.1	2.9	2.7	3.1	2.7	2.7	2.8	2.7	2.7	2.7	2.5
3.6	3.0	3.2	3.2	3.0	2.9	3.1	2.8	2.8	2.9	2.8	2.8	2.8	2.6
3.6	3.0	3.3	3.3	3.2	2.9	3.2	2.9	3.2	2.9	2.9	2.8	2.8	3.1
3.8	3.0	3.5	3.3	3.2	3.0	3.3	3.0	3.2	2.9	3.0	2.9	2.8	3.2
3.9	3.5	3.6	3.7	3.2	3.1	3.3	3.1	3.6	3.0	3.1	3.2	3.0	3.4
4.0	3.5	3.6	4.0	3.2	4.2	3.5	3.1	3.8	3.0	3.3	3.3	3.3	3.4
4.5	3.9	4.5	2.2	3.8	4.2	3.6	3.2	4.7	3.5	3.5	3.3	3.4	4.2
													Range
2.4	2.1	2.6	2.2	1.8	2.3	1.7	1.2	2.8	1.9	1.7	1.6	1.5	2.5
													Average
3.23	2.80	2.98	2.73	2.67	2.70	2.75	2.56	2.71	2.53	2.52	2.39	2.50	2.48

LG = Low Good
MG = Middle Good
HG = High Good

LC = Low Choice
MC = Middle Choice
HC = High Choice

P = Pre-test
1 = March series
2 = May series

TABLE 25-ARRAY, RANGE AND AVERAGE OF CARCASS MEANS BY "GRADE" October and January Rep.

LG 3	LG 4	MG 3	MG 4	HG 3		HG 4	LC 3		LC 4	MC 3	MC 4	HC 3	HC 4
2.0	2.0	2.0	1.8	1.8	2.9	1.6	1.9	2.7	1.9	1.7	1.7	1.6	1.5
2.2	2.1	2.1	2.2	2.0	2.9	1.8	2.1	2.7	2.0	1.8	1.8	1.8	1.6
2.4	2.4	2.2	2.3	2.0	2.9	1.9	2.2	2.7	2.0	1.8	1.8	2.0	1.7
2.5	2.5	2.2	2.4	2.1	3.0	2.0	2.2	2.7	2.1	1.9	1.9	2.0	1.7
2.6	2.6	2.3	2.6	2.3	3.1	2.0	2.2	2.8	2.3	1.9	2.0	2.1	1.8
2.7	2.7	2.4	2.6	2.3	3.1	2.0	2.2	2.8	2.3	2.0	2.0	2.1	2.0
2.7	2.8	2.5	2.7	2.3	3.2	2.1	2.3	2.9	2.3	2.0	2.0	2.3	2.0
2.8	2.8	2.6	2.7	2.5	3.2	2.1	2.3	3.0	2.4	2.1	2.0	2.4	2.0
2.8	2.8	2.6	2.8	2.5	3.2	2.2	2.3	3.0	2.4	2.2	2.0	2.6	2.0
2.9	3.0	2.7	2.8	2.5	3.3	2.2	2.3	3.0	2.4	2.2	2.0	2.6	2.1
2.9	3.0	2.8	2.8	2.6	3.3	2.3	2.4	3.1	2.4	2.2	2.2	2.7	2.1
3.0	3.1	3.1	2.8	2.6	3.4	2.3	2.4	3.1	2.4	2.3	2.2	2.8	2.1
3.0	3.2	3.3	2.9	2.6	3.6	2.5	2.5	3.1	2.4	2.4	2.2	2.9	2.1
3.3	3.4	3.5	2.9	2.6	3.6	2.8	2.5	3.2	2.5	2.4	2.3	2.9	2.2
3.4	3.4	3.6	3.0	2.6	3.6	2.9	2.5	3.2	2.7	2.4	2.4	2.9	2.3
3.6	3.5	3.7	3.4	2.7	3.7	2.9	2.6	3.3	2.7	2.5	2.5	3.0	2.3
3.6	3.7	4.2	3.7	2.7	3.8	3.1	2.6	3.3	2.9	2.8	2.7	3.0	2.4
4.0	3.7	4.2	3.7	2.7	3.8	3.1	2.6	3.4	3.1	2.9	2.7	3.4	2.6
4.1	3.8	4.6	3.7	2.8	3.9	3.2	2.6	3.7	3.4	3.1	2.8	3.7	2.7
4.5	3.9	4.8	4.8	2.8		3.6	2.6	3.8	3.6	3.5	2.8	4.1	2.8
2.5	1.9	2.8	2.0	2.1		2.0	1.9		1.7	1.4	1.1	2.5	1.3 Range
3.05	3.02	3.06	2.91	2.89		2.43	2.72		2.51	2.25	2.15	2.65	2.09 Average

LG = Low Good
MG =Middle Good
HG = High Good

LC = Low Choice
MC = Middle Choice
HC = High Choice

3 = October series
4 = January series

BIBLIOGRAPHY

1. V. James Rhodes, Max F. Jordan, H. D. Naumann, Elmer R. Kiehl and Margaret Mangel, *The Effect of Continued Testing Upon Consumer Evaluation of Beef Loin Steaks,* Mo. Res. Bul. 676, 1958.

2. V. James Rhodes, Elmer R. Kiehl, D. E. Brady and H. D. Naumann, *Predicting Consumer Acceptance of Beef Loin Steaks,* Mo. Res. Bul. 651.

3. Eugene Wierbicki, V. R. Cahill, L. E. Kunkle, E. W. Klosterman and F. E. Deatherage, "Effect of Castration on Biochemistry and Quality of Beef," *J. of Agricultural and Food Chemistry,* 3:244-49, 1955.

4. V. James Rhodes, H. D. Naumann, Elmer R. Kiehl, D. E. Brady, and Ruth Cook, *A New Approach to Measuring Consumer Acceptability of Beef,* Mo. Res. Bul. 677, 1958.

5. Robert Graf, "Beef Studies," *Reciprocal Meat Conference Proceedings,* 9:64-67, 1956.

6. D. H. Kropf and R. L. Graf, "The Effect of Grade, Weight and Class of Beef Carcasses upon Certain Chemical and Sensory Evaluations of Beef Quality," *Food Technology,* 13:719-21, Dec. 1959.

7. George W. Campbell, *Consumer Acceptance of Beef: A controlled Retail Store Experiment, Phoenix Arizona,* Ariz. Ag. Exp. Sta. Report 145, 1956.

8. Sylvia Cover, G. T. King and O. D. Butler, *Effect of Carcass Grades and Fatness on Tenderness of Meat from Steers of Known History,* Texas Ag. Exp. Sta. Bul. 889, 1958.

9. Marion Simone, Floyd Carroll and M. T. Clegg, "Effect of Degree of Finish of Differences in Quality Factors of Beef," *Food Research,* 23:32-40, 1958.

10. Marilyn Dunsing, "Visual and Eating Preferences of Consumer Household Panel for Beef of Different Grades," *Food Research,* 24:434-444, 1959.

11. J. W. Cole, D. D. Brannan, and C. S. Hobbs, "Consumer Preferences and Organoleptic Studies as Related to Federal Beef Grades and Selected Beef Carcass Characteristics," *Journal of Animal Science,* (abstract) vol. 16, 1074.

12. J. W. Cole, W. R. Backus and L. E. Orme, "Specific Gravity as an Objective Measure of Beef Eating Quality," *J. of Animal Science,* 19:167-74, Feb., 1960.

13. R. B. Sleeth, G. G. Kelley and D. E. Brady, "Shrinkage and Organoleptic Characteristics of Beef Aged in Controlled Environments," *Food Technology,* 12:88, 1958.

14. P. J. Lane and L. E. Walters, "Acceptability Studies in Beef," Processed Series P. 305, Okla. Ag. Exp. Sta., Sept. 1956.

15. J. J. Wanderstock and J. I. Miller, "Quality and Palatability of Beef as Affected by Method of Feeding and Carcass Grade," *Food Research* 13:291-303.

16. D. M. Doty, "Laboratory Characteristics of Graded Beef Carcasses," *Recip. Meat Conf. Proc.,* 9:10-18, 1956.

17. G. H. Wellington and J. R. Stouffer, *Beef Marbling—Its Estimation and Influence on Tenderness and Juiciness,* Cornell Ag. Exp. Sta. Bul. 941, 1959.

18. D. H. Kropf and R. L. Graf, "The Effect of Carcass Grade, Weight and Classification Upon Boneless Beef Yield," *J. of Animal Science,* 18:95-103, Feb., 1959.

19. J. C. Pierce *et al,* "Some Factors Influencing Yields of Wholesale and Retail Cuts from Beef Carcasses," paper at Am. Soc. of Anim. Prod., 1956.

20. David Pettus, "New Developments in Livestock Marketing Services," USDA Mimeo, Dec. 5, 1958.

21. O. D. Butler, "Type and Quality in the Live Animal and in the Carcass," paper at the National Beef Industry Conference, Purdue, 1959.