



109
448
THS

PHOTOGRAPHS TO EVALUATE CHANGES IN TYPE
OF DAIRY CATTLE

Thesis for the Degree of M. S.

MICHIGAN STATE UNIVERSITY

CHARLES LEE GOEKE

1973



LIBRARY
Michigan State
University

ABSTRACT

PHOTOGRAPHS TO EVALUATE CHANGE IN TYPE OF DAIRY CATTLE

By

Charles Lee Goeke

Photographs might offer partial relief to changes in scale, ideals, and appraisers hampering measures of trends in type. I examined variation of scores of 161 Jersey cows from 344 7 x 7 cm color transparencies taken 3 months after calving, each rated twice by four judges in five half-day sessions over 2.5 weeks. The Jerseys were selected for single traits in closed herds from 1956 to 1967. Correlations between ratings from duplicate photographs ranged from .72 to .91 for individual judges. Nearness in time of second rating to first of the same photograph influenced correlations (.91 in same session; .81 in different sessions).

The average score was 9.0 on a scale of 0 to 17. Cow within sire, judge, year, sire within line, line, interaction of year and line, parity number, inbreeding, and errors of measurement were sources of differences.

Charles Lee Goeke

Variances were 2.18 between means of cows, .07 between scales of judges, .03 between ratings of the same cow in different years, and 1.45 residual. Product correlations between judges averaged .68. This agreement of judges is near the .74 previously experienced with ratings of the actual cow. Average yearly scores rose to a high in 1961, then decreased slightly. Trends in type can be evaluated by photographs to hold ideals at time of rating constant.

PHOTOGRAPHS TO EVALUATE CHANGES IN TYPE
OF DAIRY CATTLE

By

Charles Lee Goeke

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Dairy Science

1973

G-7890

DEDICATION

To my wife Dolores

ACKNOWLEDGMENTS

I express my appreciation to Dr. Lon D. McGilliard for his interest, guidance, and critical review of this manuscript.

Financial assistance offered by the Department of Dairy Science is gratefully acknowledged.

Use of the Michigan State University computing facilities was possible in part through support from the National Science Foundation.

I thank my parents, Mr. and Mrs. Arthur Goeke, Jr., for their encouragement and interest in education. Most of all, thanks to my wife, Dolores, for her encouragement, interest, and understanding.

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
INTRODUCTION	1
REVIEW OF LITERATURE	3
SOURCE OF DATA	8
Herd Composition	8
Management	9
Selection Procedures	10
Type Scores	13
METHODS AND RESULTS	15
Introduction	15
Duplicate Ratings	16
Judges, Cows, Years	21
Age, Parity	22
Sire, Line	25
Agreement among Judges	33
Trends	35
CONCLUSION	38
REFERENCES	40

LIST OF TABLES

Table	Page
1. Correlations of duplicate ratings in session combinations for each judge	18
2. Correlations of duplicate ratings by judge and proximity of sessions	19
3. Least squares deviations of type ratings of photographs for judge, session, and judge-session subclasses	21
4. Analysis of variance in type ratings 0 to 17 from photographs	22
5. Yearly and overall type ratings by ages of photographs of Jersey cows, combined herds . .	23
6. Number of calvings by age and year for Jersey cows	24
7. Mean type ratings of photographs of Jersey cows for each lactation	26
8. Average type scores of photographs of daughters for bulls coded by line and sequence of use	27
9. Mean yearly type of photographs of Jersey cows for each line of selection	28
10. Mean type of photographs of Jersey cows for lactation-line combinations	29
11. Analysis of variance in type scores from photographs of Jersey cows	30
12. Analysis of variance and variance components of type ratings from photographs of Jersey cows	31

Table

Page

13.	Correlations of judges' ratings of type from photographs	34
14.	Overall yearly averages and average yearly type scores of photographs of Jersey cows for each judge	36

INTRODUCTION

Measurement of type, body conformation, is a component of dairy cattle evaluation. Classification and the showing are two formal current methods to measure type. Most type measurements are on actual animals; however, pictures are important in advertising, in some forms of showing competition as in All American programs of breed societies, and could be useful for measurement of type.

Measuring trends in type of dairy cattle to evaluate programs in breeding, nutrition, or management involves difficulties not usually experienced with characteristics more objectively measured. Previous attempts to measure type and its changes have been with ratings of live animals by one or more judges. The study examines the possibility of rating pictures (color transparencies) to measure type and trends in type to relieve partially complications caused by change in judges' scales, change in ideals, and change of judges. All three hamper evaluating trends or changes in type. If relationships between variations are suitable, another purpose is to apply the method to data of a breeding project.

Questions asked in fulfilling the purpose were:

(1) Can judges find common factors in pictures upon which to rate animals? (2) What are possible causes of discrepancies between ratings of color transparencies? (3) Are these small enough to permit detection of trends?

REVIEW OF LITERATURE

Dairy cattle conformation or type has been a part of evaluating cattle for utility, for aesthetics, and in merchandising. As early as 1814 the Highland Agricultural Society was awarding prizes to Ayrshires at local shows. In 1834 a scorecard or scale of points to guide showing judging was developed on the Island of Jersey (9). The scorecard has changed several times to its present form, the Uniform Dairy Scorecard, providing a uniform guide to type evaluation of five dairy breeds (1, 9).

Breed organizations have developed programs to encourage improvement of type. True type models of Holstein bull and cow were developed in 1924 (9). Today each breed has models which set standards to aid type evaluation. Each dairy breed organization of the United States has type classification, the earliest started in 1928 by the Holstein organization. These programs were to provide more widespread and uniform type evaluation than was provided by showing competition (9). Selection of All-American animals is another program of breed organizations to promote improvement of type by recognition of superior show animals. Photographs of an animal and information about its season's show record are submitted to a panel

of judges to choose All-American, Reserve All-American, and Honorable Mention winners (9). Breed organizations have other programs to improve type that differ between organizations and change over time. Artificial insemination organizations also have programs for type improvement. These programs vary from one organization to another.

Photographs are used in merchandising animals, in competition, for identification, and for other purposes. Research on the effectiveness of measuring type through photographs has been limited, but sources of variation in actual ratings of live animals have been investigated more thoroughly. My review is of these studies of ratings of live animals.

Potential sources of variation in type scores are animals, judges, time of rating, animal's parents, animal's age, and interactions among these. Variation among animals is expected to be large since cows differ and type ratings intend to reflect differences in body conformation. Different judges should contribute little variation in ratings where they, with similar ability, work from the same model and with the same scale. Large variation from judges would indicate discrepancy from these conditions.

A change or difference in scale from one rating to another would be reflected in variation from cows being rated at different times. Another source of variation is

groups of paternal half sisters; this can be large when one bull's daughters differ substantially in type from daughters of another bull. Age could influence variation also if type changes appreciably with increased age of animal.

Variation from interaction between cow and date of measuring is expected to be large where cows change differently in type with time (8). A small variance from interaction of cow with judge indicates judges prefer the same cows, or have similar ideals (8). Variation of year x judge is also small if change in judges' scales is similar from one year to another (8).

Examination of variances of ratings on live animals may indicate sources of variance in type scores from photographs. Total variance in type ratings of live animals was 5.45 with a mean score of 6.64 points on a 0 to 17 scale for Holstein cows (8). A comparable total variance was 5.26 for Holstein and Guernsey cows (12). Variation in ratings due to cows was 3.02 or 55% of the total (8) and 1.64 or 31% (12). Variations between cows were .484 and .727 when Ayrshire cows were rated two or more times by different inspectors, or rated by one judge three times in a year on scale 1 to 5 (5).

Judges have contributed little to total variation of type ratings, variation among them being insignificant in three studies (8, 11, 3) and accounting for 2.5% in

another (10). Johnson and Lush (7) also stated that judges contributed a small amount of variance; however, their experiment was not designed to separate this information.

Year or time of classification has been a source of small variance, 3% and 6% of the total (8, 12). Both rounds of classifiers and time of classification were insignificant when official classifiers rated New York herds at least three times between 1950 and 1963 (3).

Contribution from interactions of cow with date, cow with judge, and year with judge to the total variation was also examined. Variance of cow with date was largest, .65 or 12% and 1.45 or 27.4% of the total (8, 12). Variation from interaction of cow with judge and judge with year accounted for 1% and 5% of the total (8). Product-moment correlations confirm the large interaction between cow and year or cow and date and the small interaction of cow with judge. The animal's change in appearance was an important source of variance while the difference or shift in judge's ideals with time was unimportant (8).

Residual variation, left unexplained, depends on what sources were examined or the completeness of the model. Residual variance was 1.27 or 23% of the total which included interaction of cow x judge x year plus anything not accounted for by these sources individually or in any two-way interaction (8). A residual variance

of .735 or 14% of the total included all not accounted for by cow, judge, or cow x date interaction (12). Variation within cow including everything but differences among cows was .403 (5).

Variance attributed to an animal's parents, age, and stage of lactation has been examined previously. Animal's sire accounted for 8.5% of the total variation (10). Changes in type of cows with advancing age were not large but were significant between ages 4 and 5; variance due to age was small (5). Differences between scores at different stages of lactation were small but significant when early or late part of the lactation was compared with the middle segment. Stage of lactation was a small source of variation (2, 5).

In this thesis I intend to examine variation and sources of variation of type ratings where ratings are from pictures of cows. I shall compare these with studies of direct ratings of live animals.

SOURCE OF DATA

Type ratings were from photographs of animals in three Michigan State University Jersey herds. Originally two of the herds were closed in an inbreeding experiment while the other was a teaching herd. Herd composition, management information, selection procedures, and method of obtaining type scores follow.

Herd Composition

Foundation animals for the inbreeding study were brought to Michigan State University in 1951 from three California dairies which had been using a series of bulls from inbreeding experiments of the University of California. In 1954 two mating groups were formed according to closeness of animal's relationship to each of two sets of full brothers available for service.

Breeding plans changed in 1955 shifting the groups so animals with the closest genetic covariance with the dam of one set were in one group and those most closely related to the dam of the other set of sires were in a second group. One of the two groups was chosen by chance to be selected for milk production and was designated the

select herd. The other group with mating choices by chance was the control herd. These two herds were kept together and managed as one herd with about 15 cows in milk in each (4).

In 1955 the teaching Jersey herd which was housed at a different location was incorporated into the plan. This group was unrelated to select and control herds and was selected for type. In 1961 all animals were moved to the same location and managed as one herd.

Management

Standard good herd management procedures were followed as closely as possible with available facilities. Environmental conditions for the cows were to be as constant as practical with moderations for animals to be in nutrition and management experiments. During summer, lactating cows were on pasture whenever it was available and practical, with grass silage and grain as a supplement. Corn silage, good quality legume hay, and grain were fed cows during winter months. Grain was fed according to milk production with adequate added amount to younger cows for growth.

Heifers were bred at first regular heat after reaching either 15 mo or 250 kg body weight, whichever came first. A heifer not pregnant at 2 yr of age was discarded as a nonbreeder.

Cows were bred at the heat which was nearest 60 days postpartum. A cow still open after 10 mo was discarded. Palpations of genital organs were postpartum, before breeding, 60 to 90 days after breeding, and whenever there were irregularities (4).

Selection Procedures (4)

I. Removal of cows from herd

A. Control

1. Retain all females until one record is complete.
2. Maintain herd size at about 15 cows in milk.
3. When it is necessary to eliminate cows to maintain herd size, excess cows shall be removed at random from among cows which have completed at least one lactation.

B. Select herd

1. Maintain herd size at 15 cows in milk.
2. When it is necessary to eliminate animals to maintain herd size, animals ranking lowest in index value based on milk production of the individual and her relatives shall be removed from the herd.

C. Type herd

1. Maintain herd size at 15 cows in milk.
2. When it is necessary to eliminate animals to maintain herd size, animals with the lowest

average type score from ratings three times a year shall be removed from the herd.

D. All three herds

1. Any heifer which has not conceived after eight services or by 24 mo of age shall be removed from the herd.
2. Any cow which has not conceived within 10 mo from the previous calving date shall be removed from the herd.
3. Females which cannot be milked by machine because of udder structure, injury, etc., shall be removed from the herd.
4. Animals shall be removed from the herd for any disease or injury for which removal is the recommended veterinary control.

II. Choice and removal of bulls.

A. Control herd

1. After the birth of each bull, determine by chance whether he shall be retained or discarded.
2. When the number of bulls 6 mo to 12 mo of age exceeds three, or the number over 12 mo of age, exclusive of herd sire, exceeds two, discard excess bulls from each age group at random.
3. Choose herd sire by chance from bulls of

breeding age which have not previously served as herd sire.

B. Select herd

1. Designate periodically the best cows from which to save bulls on basis of milk production of them and their close relatives.
2. When it is necessary to discard young bulls to maintain proper numbers for replacement, discard bulls scoring lowest in index value of milk production of close relatives.
3. Use the bull with the highest index value as herd sire.

C. Type herd

1. Designate periodically the best cows from which to save bulls on their average type classification from evaluations every 4 mo.
2. When it is necessary to discard young bulls to maintain proper numbers for replacement, discard bulls whose dams have lowest average type scores.
3. Use the bull whose dam has the highest average type score as the herd sire.
4. When a bull is chosen as a herd sire, continue to use him as herd sire until either another bull has a dam with a higher average score, or

the herd sires in the other two herds have been replaced.

D. All three herds

1. When a bull is chosen to be herd sire, use him as soon as he is able to serve. Retain the previous herd sire or a substitute until the fertility of the new herd sire is ascertained.
2. Discard infertile bulls as well as those unable or unwilling to serve.

Type Scores

The leader of the project took color transparencies 7 x 7 cm of the right side of each animal in the three Jersey milking herds approximately 3 mo after she calved. Cows were weighed, measured, scored, posed and photographed usually within 3 days of 3 mo after calving. Photographs accumulated from December 1956 through 1964 with a total of 344 pictures from 131 cows. Each slide was randomly assigned two of 688 possible numbers from 1 to 689 excluding 234 for sequential presentation for scoring.

During January, 1965 a panel of four judges simultaneously and independently evaluated the type conformation of the Jersey cows from 2 x 2 m projection of each transparency. The scale was 0 (low poor) to 17 (high

excellent) for overall score and 0 (poor) to 5 (excellent) for each subclass of type. Subclasses included mammary system, body capacity, dairy character, and general appearance along with teat placement, rear udder, fore udder, barrel, chest, hind feet and legs, fore feet and legs, rump, and breed character.

The presentation included each transparency twice in the order arranged by procedure to provide randomness. Three morning sessions and two afternoon sessions with about 2.5 wk between the first and last sessions were required to evaluate the slides. The panel has previous judging experience but had not evaluated type from slides. Each judge recorded his own scores and was uninformed of duplication of photographs.

METHODS AND RESULTS

Introduction

Analysis of variation in scores was separated into several sections because of the dependency of later portions on earlier results. Relationships of first and second ratings of the same photograph were first examined by two methods. The absolute difference between the two ratings was regressed on the difference between sequence numbers to learn if discrepancies were larger between duplicate ratings separated by more time than between pictures rated closer in time. Correlations of first and second ratings of the same slide were also examined for consistency of ratings, especially those widely separated by time. These are sources of potential error in scoring.

Another section contains examination of whether judges' scales of scoring changed from one session to another. A substantial change would require adjusting scores to a common session for the remainder of the study. Scores of same pictures rated in each of two sessions were compared for change in scales.

Other sources of variation in type scores were examined after the relationship of duplicate ratings and

change in judges' scales. Variances due to judges, cows, year of picture, and all two-way interactions were investigated for each judge and for judges combined. Average ratings by each judge for each year of photograph were also compared. Age at photographing, parity number of animal, sire, line of selection, and inbreeding were examined for variation by comparing average scores and analyzing variance. Variances were partitioned into components for cow, year, and judge. Agreement among judges was examined by comparing product-moment correlations of simultaneous ratings of a slide by different judges to base correlations of two ratings of a slide by one judge.

Duplicate Ratings

Time between ratings of a slide as a source of discrepancies between duplicate ratings was examined by regression of differences in scores on differences in sequence number and by correlations of duplicate ratings. Large regressions or small correlations indicate time is a large source of variation, and causes would need investigating.

Absolute differences between scores of duplicate ratings by the same judge were regressed on positive differences in sequence numbers, a nonlinear measure of interval between repeated ratings by same judge. Slopes of line for each judge were .0005, .0007, .0002, and .0003 with F values for the null hypothesis $b=0$ of 3.3,

12.6, .5, and 1.9. Only for judge 2 was the slope significantly different from zero at $P < .01$. With a maximum difference between sequence numbers of two ratings of 688 and a slope of .0007, scores of a photograph rated first and subsequently last of all by judge 2 would be expected to differ by less than .5 point, an unimportant difference. As time between duplicate ratings of a slide increased, differences between ratings remained almost nil.

Correlation of ratings separated by approximately equal time appear in columns in Table 1. Groups of session combination correspond to sessions in which first and second ratings occurred. A picture rated both times in session 1 was in group (1/1), but a slide appearing first in session 2 and second in session 5 was in group (2/5).

Approximately half the correlations of duplicate ratings in the same session were larger than other correlations for each judge. The high correlations were in session combination (1/1) for judges 2 and 4; (2/2) for judges 1, 2, and 3; (4/4) for judges 1 and 2; and (5/5) for judges 1, 2, and 3. There were also three correlations of duplicate ratings in the same session which were lower than other correlations. These were in session combinations (1/1) and (4/4) for judge 3, and (1/1) for judge 1. Four of the five correlations of duplicate ratings in the same session were larger than remaining correlations for judge 2; however, only one was higher for judge 4.

Table 1. Correlations of duplicate ratings in session combinations for each judge.

Judge	Session Combinations														
	1/1	1/2	1/3	1/4	1/5	2/2	2/3	2/4	2/5	3/3	3/4	3/5	4/4	4/5	5/5
1	.65	.87	.83	.84	.79	.95	.85	.84	.78	.86	.86	.86	.97	.86	.92
2	.95	.89	.91	.92	.89	.99	.90	.85	.85	.92	.93	.93	.96	.93	.97
3	.66	.61	.40	.85	.73	.94	.83	.79	.69	.82	.78	.80	.71	.81	.90
4	.91	.77	.74	.77	.73	.85	.81	.81	.79	.87	.87	.89	.87	.80	.89
No. of Pairs	8	26	24	16	26	15	37	30	38	25	18	33	9	26	13

Duplicate ratings can be regrouped into "difference groups" further to investigate consistency of ratings. Difference groups are composed of one or more session combinations, and correlations in Table 2 are averages of appropriate session combination correlations. Averages of correlations were by z-transformations for normality of distribution. The 0 group contains ratings of all slides rated twice in one session while the 1 Difference group contains ratings of slides evaluated in adjacent sessions and 2 Difference has one session between sessions of paired ratings.

Table 2. Correlations of duplicate ratings by judge and proximity of sessions.

Difference	Judge				Average	Number of Observations
	1	2	3	4		
0	.90	.96	.85	.88	.91	70
1	.86	.90	.72	.77	.83	107
2	.84	.90	.62	.81	.81	87
3	.81	.88	.70	.77	.80	57
4	.79	.89	.73	.73	.80	26

The average correlation of duplicate ratings in the same session was .91, slightly larger than correlations of duplicate ratings in different sessions whose averages

ranged from .80 to .83 and were not significantly different. Ratings of a slide appearing twice in the same session were more consistent than those in different sessions, but once in different sessions, consistency of ratings was similar. This with variations of degree was typical for each judge.

Negligible regression of differences in scores on difference in sequence number indicated time between ratings was not important. Correlations of duplicate ratings indicated time introduced error but even when duplicate ratings were in different sessions, error was small. I concluded that time is a minor source of discrepancies between ratings.

Large average differences between duplicate ratings in different sessions could indicate change in judge's scale over time, requiring adjustment of ratings to a common session for analysis of differences between cows or trends in type. Lack of apparent reason for time between ratings being a source of discrepancies in ratings would necessitate considering it in subsequent analysis.

I compared average scores of identical pictures in different sessions. These duplicate pictures were in judge sessions between which deviations in Table 3 differed substantially. Large changes of scale seemed more likely between these cells. Differences between duplicates in sessions 1 and 2 for judge 1, sessions 2 and 5

Table 3. Least squares deviations of type ratings of photographs for judge, session, and judge-session subclasses.

Judge	Session					Mean
	1	2	3	4	5	
1	.37	-.29	-.04	.05	-.08	-.17
2	-.03	-.14	-.03	-.12	.32	-.39
3	.11	.32	-.24	-.09	-.10	.42
4	-.45	.11	.31	.17	-.14	.13
Mean	-.16	-.06	.11	.28	-.17	

Scale of type ratings 0 to 17, low poor to high excellent; mean = 9.01, low good plus.

for judge 2, sessions 2 and 3 for judge 3, and sessions 1 and 3, 1 and 4 for judge 4 were small enough to deny changes in scale important enough to need adjustment. Thus, I did not adjust scores for subsequent analyses.

Judges, Cows, Years

I partitioned variation in type scores of color transparencies into differences attributed to cow, year, and interaction between cow and year for each judge. Means from all three sources differed significantly at $P < .01$ for each of the four judges. The analysis of variance with judges combined is in Table 4. Judge, year, cow, and

Table 4. Analysis of variance in type ratings 0 to 17 from photographs.

Source	df	MS	F**
Judge	3	87.92	124
Year	8	49.63	70
Cow	160	38.77	55
Year X Judge	24	7.75	11
Cow X Year	175	4.85	7
Cow X Judge	480	2.82	4
Error	1901	.71	

**p < .01 for all classifications.

all two-way interactions are significant at $P < .01$. The three-way interaction is included in the error term.

Age, Parity

Differences in ratings because of age might be reflected by different average scores for groups by age of cow. Average scores in years 1957 through 1964 for each age and an overall average are in Table 5. Age seemed to be an important source of variation with average scores generally rising with each age group. A regression equation limited to linear and quadratic terms

Table 5. Yearly and overall type ratings by ages of photographs of Jersey cows, combined herds.

Age	Year							Mean	Number
	1957	1958	1959	1960	1961	1962	1963		
2	8.36	8.46	8.82	8.82	9.03	7.90	8.77	8.68	944
3	7.38	9.05	9.06	9.67	9.87	9.24	8.17	9.25	640
4	8.00	9.08	9.43	8.97	9.46	9.43	9.13	8.93	360
5	11.25	8.63	10.19	9.92	9.35	8.84	9.91	9.31	280
6	8.06	11.50	10.38	10.44	7.87	9.10	8.44	9.32	152
7	8.68		8.75	8.75	10.37	10.25	11.44	9.62	112
8	9.37	9.20			8.06	9.81	9.87	9.39	104
9	5.37	8.13	9.13			7.13	11.37	8.31	72
10	7.75	12.13						9.94	16
11		9.87		8.25			9.37	9.16	24
>12			9.25	9.50	9.25	8.62		9.05	40

Scale of type ratings 0 to 17, low poor to high excellent; mean = 9, low good plus. Number of cows per cell are in Table 6.

of mean scores on age shows how average type scores changed with age.

$$\text{Average score} = 8.20 + .34 (\text{age}) + .02 (\text{age})^2$$

Age is age at calving to nearest year. Average type score rose to a peak around age 8 or 9 yr, then slowly decreased.

The increase in average score with age and the seeming importance of age to type can in part be explained by selection and consequential reduction in numbers. The number of animals in each age category decreased rapidly (Table 6), and since one line was selected on type, some animals left the herd for inferior type, increasing the average score with age.

Table 6. Number of calvings by age and year for Jersey cows.

Age at Calving yr	Year							
	1957	1958	1959	1960	1961	1962	1963	1964
2	8	16	28	23	14	12	7	10
3	1	16	11	16	9	13	6	8
4	4	6	7	5	7	5	3	8
5	1	4	2	3	5	8	4	8
6	2	1	2	2	3	5	2	2
7	4		1	1	2	1	2	3
8	1	5			2	2	1	2
9	1	2	2			2	1	1
10	1	1						
11		1		1				1
<u>>12</u>			1	1	1	2		

Aging can also be measured by parity number since one picture was taken per lactation. Table 7 contains means for parities and numbers of observations per lactation. Scores of lactations 1 through 4 compare favorably with ages 2 through 5, Table 5. Discrepancies are probably due to calving intervals more than 12 mo.

Differences in ratings because of interaction between age and year are reflected by differences between average ratings by age not being the same from one year to another. Differences between the average scores by age in Table 5 change from year to year; hence, interaction of age and year seemed another important source of variation in type scores. However, age can be ignored as a source of variation in yearly herd averages if distribution of cows by age is similar for each year.

Sire, Line

Significant differences in scores between cows suggested animal's sire and sequence of sire usage were likely sources of variation. Too, sire accounted for 8.5% of the total variation of type scores in one study (10). I coded bulls by line or herd and sequence of use. Line 1 was selected for milk production, line 2 control, and line 3 was selected for type. All bulls used before the breeding project began were pooled to make up sequence 1. Bulls were used from September, 1956 through 1966 with

Table 7. Mean type ratings of photographs of Jersey cows for each lactation.

Lactation Number	Number of Observations	Mean
1	984	8.64
2	656	9.24
3	408	9.06
4	240	9.58
5	136	8.89
6	120	9.44
7	104	9.21
8	48	9.19
9	8	9.37
10	8	9.87
11	8	9.25
12	8	9.50
13	8	9.25
14	8	9.00

Scale of type ratings 0 to 17, low poor to high excellent; mean = 9, low good plus.

changes approximately yearly, but as specified in the breeding plan, time was not primary in deciding when to change. Occasionally bulls in line 3 were switched to keep the number of bulls used per line about equal.

Average type scores for daughters in each sire sequence in Table 8 differ, sometimes substantially, from one bull to another in the same line. Differences in average scores indicate animal's sire is a source of variation in type ratings and differences between sires can be detected in ratings of type from photographs. Three of four bull's daughters in the type line have higher average ratings than daughters of bulls in the milk line.

Table 8. Average type scores of photographs of daughters for bulls coded by line and sequence of use.

Sequence Number	1(Milk)	Line 2(Control)	3(Type)
1	9.01	8.05	10.04
2	8.91	9.26	9.97
3	9.01	8.32	8.29
4	8.38	8.75	11.33
5		7.59	
6	8.13		

Scale of type ratings 0 to 17, low poor to high excellent; mean = 9, low good plus.

Differences in daughter averages between lines, and selection of one line for type suggests line as a source of variation. Means of years and lactations for each line in Tables 9 and 10 show scores for animals

Table 9. Mean yearly type scores of photographs of Jersey cows for each line of selection,

Year	1(Milk)	Line 2(Control)	3(Type)
1957	8.31	7.61	9.65
1958	8.82	8.10	9.80
1959	9.07	8.32	9.95
1960	8.99	8.77	9.94
1961	8.96	8.46	10.32
1962	8.73	8.46	9.20
1963	9.58	8.26	9.56
1964	9.28	8.47	9.73

Scale of type ratings 0 to 17, low poor to high excellent; mean = 9, low good plus.

selected for type are higher than both other lines except in 1963. Average scores of animals selected for milk production are above control animals for all years and for all lactations except lactation 6. These averages provide further evidence that line or herd is a source of differences in type ratings. Interaction of year and line is suggested by differences between average scores of lines differing from year to year.

Age, sire, and line as possible sources of variation as well as inbreeding were more completely analyzed, Table 11. This analysis is on data from 1957 through 1964

Table 10. Mean type ratings of photographs of Jersey cows for lactation-line combinations.

Lactation Number	1 (Milk)	Line 2 (Control)	3 (Type)
1	8.57	8.09	9.20
2	9.40	8.53	10.01
3	9.11	8.27	10.12
4	9.43	9.07	10.41
5	8.95	6.88	10.16
6	8.83	9.00	10.27
7	8.96	8.33	10.25
8	8.44		10.69
9	9.37		
<u>>10</u>	9.50		

Scale of type ratings 0 to 17, low poor to high excellent; mean = 9, low good plus.

and uses lactation numbers to measure age. Judge, line, sire within line, cow within sire, year, year by line interaction, lactation number, and inbreeding were all significant at $P < .01$. This analysis of variance confirms that variations from different ages, sires, and lines in type scores are reflected through photographs.

The previous sections of the analysis answered one of the study's major questions: what were possible

Table 11. Analysis of variance in type scores from photographs of Jersey cows.

Source	df	MS	p**
Judge	3	87.5	171
Line	2	487.9	953
Sire/line	11	55.6	109
Cow/sire	147	29.8	58
Year	7	23.6	46
Year x line	14	5.6	11
Lactation Number	13	22.5	44
Inbreeding	42	54.9	107
Residual	2504	.5	

**Significant at .01.

causes of differences between ratings of color transparencies. Judges, years, cows, their two-way interactions, sires, lines, ages, inbreeding, year x line interaction and errors of measurement were sources of differences.

Variance components in Table 12 show the contribution of judges, years, and cows to the total variation in ratings of color transparencies. The mean of the ratings was 9.0 with a variance 3.74. Percentage of the total each component accounts for is also given.

Table 12. Analysis of variance and variance components of type ratings from photographs of Jersey cows.

Source	df	MS	EMS	Component	% of total
Judge	3	87.92	$\sigma_R^2 + K_2\sigma_C^2 + K_1\sigma_J^2$	.07	1.9
Year	8	49.63	$\sigma_R^2 + K_2\sigma_C^2 + K_3\sigma_Y^2$	.04	1.1
Cow	160	38.77	$\sigma_R^2 + K_2\sigma_C^2$	2.18	58.3
Rep.	2580	1.45	σ_R^2	1.45	38.7
Total	2751			3.74	100.0

K_1 = Number of observations per judge = 688.

K_2 = Average number of observations per cow = 17.1.

K_3 = Weighted number of observations per year = 298.3.

Constants to estimate components are K_1 , the number of observations per judge, K_2 , the arithmetic mean number of observations per cow, and K_3 , a weighted mean number of observations per year. The equation to calculate K_3 was:

$$K_3 = [1/(n-1)][K.(\sum K_i^2/K.)] \text{ where } n = \text{number of years} = 9$$

K_i = number of observations in year i
 $K.$ = total number of observations.

These K terms are approximations because of unequal subclass numbers and expected values of means square were used for simplicity rather than least squares methods.

Variance due to cows, 2.18 or 58.3% of the total is largest. Standard error of this component was .06. Studies of ratings on live animals found variance due to cows of 3.0 or 55% of the total variation (8) and 1.64 or 31% (12). The .07 for variation in judges' scales also agrees with 0 in other studies (8, 12). Variance due to year of picture was .04 or 1.1% of the total. This compares with .2 (8) and .3 (12). The residual variance was 1.45 or 38.7% of the total.

Components of variance for type ratings from photographs are similar in magnitude and percentage of the total to ratings on live animals. Cow is the largest source of differences in scores as was hoped because ratings are to detect differences between cows. The magnitude of variation of cow compared to other components and the residual indicates real differences between cows are detectable in photographs. Variation of judges was small indicating all judges had similar scales and ideals. This was also expected since judges should be working from the same model. Variation in scores from years was also small. Reasonably large variation from years would be expected if selection on all animals increased type scores, but only

one-third of the cows were being selected for type, and management was reasonably constant.

Agreement among Judges

Correlations to investigate further sources of variation to confirm earlier results, and to help answer the question, "Could judges find differences independently observed and agreed upon in pictures upon which to rate animals?" was the final portion of the analysis. Correlations between two ratings of a slide by each judge, along the diagonal in Table 13, shows agreement of judges with themselves. Judge 2 was more consistent in his ratings with a correlation of .91 than judge 3 with .72. All judges were reasonably consistent in scoring as they could recognize similarities in the photographs from one rating to another.

Off-diagonal elements in Table 13 are correlations between simultaneous ratings by different judges. These range from .59 to .77 averaging .68. Similar types of correlations between ratings of live animals rated almost simultaneously by different judges were .74 (8), .61 to .76 (2) and .73 and .76 (12). Agreement of judges rating type from photographs is similar to agreement of judges rating live animals.

Correlations between ratings from photographs by the same judge at different times are only slightly, if

Table 13. Correlations of judges' ratings of type from photographs.

	Judge 1	Judge 2	Judge 3	Judge 4
Judge 1	.85	.77	.59	.72
Judge 2		.91	.61	.71
Judge 3			.72	.69
Judge 4				.81

Average correlation of different judges = .68.

any, higher than correlations of simultaneous ratings by different judges. This similarity of moderately large correlations indicates lack of influence from the significant but small interaction between cow and judge found earlier in this study. Judges seem to find common factors in photographs of animals to rate body conformation.

The correlations analysis confirmed earlier results and indicated cows differ in ways that judges are able to observe these differences from photographs and agree relatively well on how to score these differences. The consistent correlations wouldn't have existed otherwise. Agreement of correlations of ratings on photographs with correlations of live animal ratings also suggest judges find common factors.

Trends

The study's second purpose was to apply the method of using ratings of photographs in measuring type and trends in type to a set of data if relationships between variations were suitable. Since variance components of ratings from photographs resembled those when actual animals were rated, the second purpose was attempted. Data collected for the study were reexamined because they were the only data available.

Average yearly type scores for the complete herd, Table 14, rise until 1961. Average scores for each judge follow a similar pattern. Scores for each line, Table 9, show average ratings of animals selected for milk production, line 1, increase until 1959 then are slightly lower until a peak in 1963. Average scores of control animals, line 2, are fairly consistent from 1958 through 1964 with the highest average in 1960. Type selected animals, line 3, average scores are reasonably consistent also, peaking in 1961. Average yearly type scores in Table 5 for ages 2, 3, and 4 peak in 1961, with yearly changes similar to those for overall average. Scores for older animals do not follow this pattern.

Ratings of photographs show average ratings of the lines followed similar patterns over the years and rank as expected. Averages of animals selected for type are generally above other lines while the herd with chance

Table 14. Overall yearly averages and average yearly type scores of photographs of Jersey cows for each judge.

Year	Judge				Ave.	Ratings per judge
	1	2	3	4		
1956	3.50	2.50	3.00	5.50	3.63	2
1957	7.87	7.65	9.41	8.26	8.30	46
1958	8.64	8.03	9.90	9.18	8.94	104
1959	8.90	8.56	9.66	9.17	9.07	108
1960	9.36	8.84	9.43	9.27	9.23	104
1961	9.06	9.22	9.53	9.22	9.26	86
1962	8.61	8.62	8.96	8.99	8.79	100
1963	8.90	9.13	9.32	9.31	9.17	52
1964	8.77	8.92	9.17	9.62	9.12	86

Scale of type rating 0 to 17, low poor to high excellent; mean = 9, low good plus.

matings generally has the lowest average scores. Averages of ages 2 through 4 also followed similar patterns. These averages indicate trends in type can be evaluated by photographs to hold ideals constant during rating. This does not prevent changes in ideals with time altering measurement of trends as the measurement is in accord with the ideal at time of rating. At another time measurement might be by a different ideal and produce a different trend even though ideals were fixed during either measurement.

Comparing ratings of animals from photographs with ratings of live animals at the time of photographing was not attempted in this study. Analyses of variance and correlations indicate the two ratings would be comparable, but further work is needed. Comparison of ratings from photographs taken at the time of classification with classification scores may indicate how alike ratings from live animals and photographs are. Precautions to limit variation in scores due to raters, time between ratings, and changes in scale are needed. If ratings from photographs are comparable with ratings of live animals, photographs might be useful in type classifying.

CONCLUSION

Separation of variance into its components resembled reasonably well those when actual animals were rated. Color accounted for 57% of the variance in this study while texture accounting for 55% and 31% in studies with live animals (8, 12). In all three studies judge and year accounted for little of the total variation.

Agreement of judges ratings in this study also resembled reasonably well analysis with actual animals. Correlations of simultaneous ratings by different judges ranged from .59 to .77 averaging .68. Correlations of ratings on live animals rated almost simultaneously by different judges were .74 at Iowa, .61 to .76 at West Virginia, and .73 and .76 at Illinois.

Color transparencies to evaluate type, or trends in type, may be useful. Subsequent rerating of these same photographs may in time hint of changes in ideals. If these do not exist, photographs are probably not needed to assure a single ideal.

I recommend taking color transparencies of animals in breeding projects which seek to monitor changes in type. Photographs should be taken at least yearly at

the same stage of lactation. Management of animals should be as constant as practical throughout the study. Ratings over a short interval reduce changes in scales and ideals from first to last rating. Ratings, however, are on the scales and ideals accepted at the time of rating. If ideals change over time, rating the photographs some time after project completion may help eliminate influence from ideals at time of rating being more like ideals existing at the project's completion than like ideals existing at the project's beginning.

REFERENCES

1. Atkeson, G. 1967. Weighting components of type in classifying Holsteins. M. S. Thesis, Michigan State University, East Lansing.
2. Benson, R. H., W. J. Tyler, and G. Hyatt, Jr. 1951. Some causes of variation in type ratings of Ayrshire cows. J. Dairy Sci. 34:502.
3. Carter, H. W., J. C. Rennie, and E. B. Burnside. 1965. Causes of variation in type classification data. J. Dairy Sci. 48:790 (Abstr.).
4. Dayton, A. D. 1969. The effects of inbreeding on heritable traits in a herd of Jersey cattle. Ph.D. Thesis, Michigan State University, East Lansing.
5. Hyatt, G., Jr., and W. J. Tyler. 1948. Variation in type ratings of individual Ayrshire cows. J. Dairy Sci. 31:71.
6. Hyatt, G., Jr., W. J. Tyler, and C. T. Conklin. 1949. The relationship between type ratings of Ayrshire females as young heifers and as cows. J. Dairy Sci. 32:375.
7. Johnson, L. E., and J. L. Lush. 1942. Repeatability of type ratings in dairy cattle. J. Dairy Sci. 25:45.
8. Mc Gilliard, L. D., and J. L. Lush. 1956. Changes in type classifications of dairy cattle. J. Dairy Sci. 39:1015.
9. Porter, A. R., J. A. Sims, and C. F. Foreman. 1965. Dairy cattle in American agriculture, Iowa State University Press, Ames.
10. Specht, L. W., H. W. Carter, and L. D. Van Vleck. 1967. Type classification scores of New York Holstein Cattle. J. Dairy Sci. 50:1690.

11. Touchberry, R. W., and K. R. Tabler. 1951. The changes in type ratings of Holstein and Guernsey cows when rated by the same three judges at two consecutive times. J. Anim. Sci. 10:1029.
12. Touchberry, R. W., and K. R. Tabler. 1951. The changes in the type ratings of Holstein and Guernsey cows when rated by the same three judges at two consecutive times. Mimeo.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03061 4915