Investigation of Linear Type Evaluation for Collegiate Judging Contests

J. M. WANNER,* D. R. OLVER,* G. W. ROGERS,*
J. E. UMPHREY,† and D. P. DICKSON‡
*Department of Dairy and Animal Science, The Pennsylvania State University, University Park 16802
†Department of Dairy Science, University of Florida, Gainesville 32611
‡Department of Dairy Science, University of Wisconsin, Madison 53706

ABSTRACT

Various time constraints and cow numbers were used to identify the minimum time per cow and the minimum number of cows needed to permit ranking of students for their skills in linear type evaluation. Three separate trials were conducted involving 23, 15, and 12 students from three universities who evaluated 15 traits on 30, 20, and 27 cows, respectively, using a 50-point scale. Time constraints included 1.5, 2, 3, and 4 min per cow as well as unlimited time per cow. The mean of the absolute difference between official and student scores was used to rank students. Lower means indicated greater student accuracy. Mean differences and (standard errors) across trials were 7.28 (0.24), 7.41 (0.11), 7.18 (0.11), and 7.41 (0.18) for 1.5, 2, 3, and 4 min of evaluation, respectively. The overall effect of time was not significant, and no pairs were significantly different. Rank correlations between student means suggested that 2 min of evaluation were adequate, and no advantage was found for evaluations of >3 min per cow. Rank correlations for student evaluation means based on 5 versus 10 cows were 0.85 to 0.90. Evaluation of <4 cows dropped rank correlations, but evaluation of >5 cows did not improve correlations greatly. Results suggested that use of 4 or 5 cows, evaluated at 2 to 3 min per cow, permitted accurate ranking of students for their ability to evaluate linear type traits.
(**Key words**: dairy cattle, linear type evaluation, collegiate judging contest)

INTRODUCTION

Dairy cattle judging contests have played an important role in developing decision-making abilities, building self-confidence, and enhancing communication skills of young people over the last 87 yr. Since the inception of the National Intercollegiate Dairy Cattle Judging Contest in 1908 in conjunction with the National Dairy Cattle Show, over 6000 students have participated in the national event, and many more have judged in regional and local contests. There have been very few changes in the format of the national contest. The most recent major change occurred 29 yr ago when the oral reasons component was added.

Because the dairy industry places more emphasis on forms of evaluation other than show results, some think that judging contests should become more reflective of the entire dairy cattle industry. At the 1991 national meeting of collegiate coaches, a committee was established to review present formats of contests and to determine potential changes. Linear evaluation of type has been a heavily discussed aspect that many people think could become a standard part of judging contests when formats are revised.

Some contests currently include a linear component of type in their events. However, formats often differ regarding contest scoring procedures, number of cows evaluated, and evaluation time allotted per cow. The objectives of this study were to determine the appropriate constraints for evaluation time and the number of cows that need to be evaluated to rank the ability of students to appraise cows using linear type evaluation.

MATERIALS AND METHODS

A pilot trial conducted at The Pennsylvania State University involved 27 students evaluating 8 cows at 3 min per cow. Results of the pilot trial were used to design three subsequent trials.

Three separate trials were conducted at three universities: The University of Florida, The University of Wisconsin, and The Pennsylvania State University. Students with different levels of experience and training, ranging from beginners to those who previously completed class work or competition in linear type evaluation, evaluated cows for 15

Received September 18, 1995.
Accepted February 26, 1996.

TABLE 1. Numbers of students, breed, number of cows, and evaluation time for three separate trials of linear type evaluation.

Trial	Students	Breed	Cows	Time
	(no.)		(no.)	(min)
The Pennsylvania State University	23	Holstein	10	4
		Holstein	10	2
		Holstein	10	3
University of Florida	15	Holstein	10	3
		Holstein	10	2
University of Wisconsin	12	Holstein	7	3
		Holstein	7	2
		Brown Swiss	7	2
		Holstein	6	1.5[1]
		Holstein	6	Unlimited

[1]The same 6 cows were reevaluated with unlimited time to change scores given during 1.5-min evaluation.

traits using a 50-point scale with Holstein Association benchmarks (1). Cows were chosen to provide a wide range of scores for the linear type traits. A professional linear evaluator of type provided a brief review of benchmarks for the traits at the start of each trial. The official evaluated all cows simultaneously with students.

Table 1 summarizes the number of students, cows, breeds, and evaluation time constraints for the three separate trials. The Pennsylvania and Wisconsin trials were conducted in stanchion barns. Groups of 3 to 4 students evaluated the same cow simultaneously. The Florida trial was conducted in a corral holding area with all students evaluating the same cow simultaneously. The three separate trials placed constraints on evaluation time and the number of cows evaluated. Participants in the Penn State trial evaluated groups of cows at 4 min per cow, 2 min per cow, and then 3 min per cow. During the Florida and Wisconsin trials, students initially evaluated cows under the longest time constraint and worked toward the shortest time per cow as they became more familiar with the scorecard. Additionally, the Wisconsin trial had a group of 6 Holsteins evaluated at 1.5 min per cow to determine how quickly students could accurately evaluate a cow. After all 6 cows were evaluated at 1.5 min per cow, students were offered additional unlimited time per cow to reevaluate the cows and change their 1.5-min scores by writing and circling an unlimited time score. This procedure provided a check for the change in evaluation accuracy with unlimited time per cow. A group of 7 Brown Swiss cows was also included in the Wisconsin trial to evaluate whether students would rank the same when time constraints and cow numbers of different breeds varied.

TABLE 2. Mean deviations from official scores for student evaluation at various time constraints and overall mean for all Holstein and Brown Swiss (BS) cows evaluated, regardless of time constraint.

Trial	Time constraint	Mean deviations	Overall mean
	(min)		
The Pennsylvania State University	2	6.40	
	3	6.27	
	4	6.45	
			6.37
University of Florida	2	8.79	
	3	8.55	
			8.67
University of Wisconsin	1.5	6.87	
	2	7.08	
	2 (BS)	7.62	
	3	6.70	
	Unlimited[1]	6.69	
			7.02[2]

[1]Additional unlimited time was allotted to change scores for cows evaluated at 1.5 min.

[2]The information obtained with the 1.5-min time constraint was excluded because the same cows were evaluated under the unlimited time constraint.

All data were analyzed using SAS procedures (2). The Holstein data were analyzed with the model:

$$Y_{ij} = \text{location}_i + \text{time}_j + e_{ij}$$

where Y_{ij} = student deviation, location_i = location of trial i, time_j = time constraint j, and e_{ij} = random error. Data obtained from student evaluation of the Brown Swiss cows were not analyzed with this model. Data from the unlimited time constraint at Wisconsin were not analyzed with the model because the data

TABLE 3. Least squares means for student deviations from official scores (absolute values) at various time constraints for Holsteins.

Time constraint	Mean deviation[1]	SE
1.5	7.29	0.25
2	7.41	0.11
3	7.19	0.11
4	7.41	0.18

[1]Mean deviation excludes data from cows evaluated under unlimited time constraint.

resulted from an extension of the 1.5-min time constraint.

Student accuracy was the difference between official and student scores. Absolute values were used for negative deviations. Analyses were done using student means of deviations from the official scores. Time allotted per cow was evaluated based on the magnitude of the mean difference from the official score. Least squares means for the student deviations at various time limits were compared. Rank correlations for student means were used to determine how well student rankings agreed across groups of cows that differed by evaluation time and breed. Rank correlations for student means, using progressively fewer cows randomly removed from remaining cows, were used to determine the number of cows that should be evaluated. Rank correlations for student means using progressively fewer cows, eliminated in reverse order of random removal, were also calculated.

RESULTS AND DISCUSSION

Students whose evaluations of cows scored closest to the official score had lower mean deviations. Mean deviations within a trial were 0.13 to 0.92 points lower for the 3-min constraints than for the constraints of 1.5, 2, or 4 min per cow (Table 2). Mean deviations, excluding the Brown Swiss cows, were only 0.13 to 0.38 points lower for the 3-min constraints than for the constraints of 1.5, 2, or 4 min per cow. Unlimited evaluation time per cow improved mean deviations by only 0.01 points over the 3-min constraint for the Wisconsin trial.

Average deviations from official and corresponding (standard errors) across trials were 7.29 (0.25), 7.41 (0.11), 7.19 (0.11), and 7.41 (0.18) for evaluations of 1.5, 2, 3, and 4 min, respectively. The overall effect of time was not significant (Table 3), and pairwise comparisons of the various time limits showed no significant differences between time constraints (Table 4). Students rarely needed or used >3 min per cow, even when unlimited time was allowed.

Rank correlations for student means showed how well student rankings agreed across groups of cows that differed in evaluation time. The rank correlations between student means for the 2-min time constraint and the average of student means for all cows evaluated were as high as or higher than other comparisons (Table 5). The rank correlation between 1.5 min and unlimited time in the Wisconsin replicate was extremely high because these were the same cows evaluated under different time limits. The rank correlations could also have been high because some students did not change scores for a cow during the reevaluation period. These results suggested that, for students with varying degrees of experience and training, evaluation time allotted per cow should be limited to 2 or 3 min. If the students participating in collegiate judging contests have a high degree of training, experience, and an unobstructed view of the cow, 2 min per cow might be an adequate evaluation time.

The minimum number of cows that are needed for evaluation is important for contest logistics. Enough

TABLE 4. Probability of pairwise significance between different levels of the time constraint and trial locations.

	2 min	3 min	4 min
Time constraint			
1.5 min	0.6562	0.7164	0.6927
2 min		0.1391	0.9744
3 min			0.2658
		The Pennsylvania State University	University of Wisconsin
Location			
University of Florida		0.0001	0.0001
The Pennsylvania State University			0.0040

TABLE 5. Rank correlations of student mean deviations from official scores for various time constraints and Holstein or Brown Swiss (BS) cows.

Trial			3 min	2 min	All
The Pennsylvania State University					
4 min			0.43	0.54	0.73
3 min				0.46	0.80
2 min					0.82
University of Florida			2 min	All	
3 min				0.75	0.86
2 min					0.96
University of Wisconsin	2 min	2 min (BS)	1.5 min	Unlimited	All[1]
3 min	0.82	0.71	0.53	0.69	0.90
2 min		0.83	0.54	0.70	0.90
2 min (BS)			0.52	0.66	0.90
1.5 min				0.90	0.66
Unlimited					0.85

[1]The information obtained with the 1.5-min time constraint was excluded because the same cows were evaluated under unlimited time constraint.

cows must be evaluated to permit the correct ranking of student evaluators. Table 6 has rank correlations for the 2-min time constraints from the Pennsylvania and Florida trials. Rank correlations of student mean scores for 5 versus 10 cows were 0.85 and 0.90, depending on the trial. Evaluating <4 cows dropped rank correlations. The correlations increased slightly when >5 cows were evaluated. Although rank correlations of student means based on 9 cows were highest, those correlations were only 2 to 14% higher than student means based on 5 cows. Reversing the order of random removal showed that rank correlations changed little as long as ≥4 cows were evaluated. Wisconsin trial data were analyzed separately because only 7 cows were evaluated during each time category. Results (not shown) supported evaluation of 4 or 5 cows to rank students accurately for their ability to score cows using linear type evaluation.

Linear scoring of type was conducted on cows that were restrained for the Pennsylvania and Wisconsin trials. Conversely, individual cows were evaluated on the move in a corral for the Florida trial. Comparison of the mean deviations among trials (Table 2) and the significant effect of trial location (Table 4) suggested that scoring of restrained animals or having <15 students evaluating the same cow simultaneously might provide a more unobstructed view of the cow and improve evaluator accuracy. Within trials, there was no indication that scoring environment affected the number of cows needed to accurately rank students for their skills in linear type evaluation.

TABLE 6. Rank correlations for student mean deviations using the progressively fewer cows remaining after random removal of and reverse elimination of random removal (within the cows evaluated at 2 min per cow).

	The Pennsylvania State University		University of Florida	
Cows	Random removal	Reverse removal	Random removal	Reverse removal
10				
9	0.99	0.97	0.92	0.93
8	0.98	0.97	0.90	0.94
7	0.95	0.97	0.87	0.95
6	0.93	0.95	0.90	0.94
5	0.85	0.91	0.90	0.88
4	0.87	0.87	0.89	0.94
3	0.71	0.79	0.65	0.95
2	0.57	0.72	0.64	0.85
1	0.48	0.73	0.52	0.71

CONCLUSIONS

The use of linear type evaluation in the dairy industry has prompted discussion about addition of a class for linear type evaluation to the traditional format of judging contests. Results from three separate research trials at three universities, using students with various levels of training and experience, suggested that students can be effectively ranked on their ability to perform linear type evaluation. Apparently, an unobstructed view of 4 or 5 cows evaluated with a time constraint of 2 to 3 min per cow provides the opportunity to rank students accurately. A class of cows of this size and time constraint of this length would work well logistically in traditional formats for judging contests.

ACKNOWLEDGMENTS

Appreciation is expressed to Arthur and Donna Nesbitt and Robert Rumler for support; Lloyd Ebersole, Gary Bicksler, Charles Howe, and Cheryl Marti for their time and linear evaluation expertise; and to students from The University of Florida, The University of Wisconsin-Madison, and The Pennsylvania State University for their time and cooperation.

REFERENCES

1 Holstein-Friesian Association of America. 1994. Linear Classification System. Holstein-Friesian Assoc. Am., Brattleboro, VT.
2 SAS® User's Guide: Statistics, Version 6.04 Edition. 1990. SAS Inst., Inc., Cary, NC.