

Factor Analysis of Genetic Evaluations For Type Traits of Canadian Holstein Sires and Cows

A. K. Ali¹, K. R. Koots² and E. B. Burnside³

Animal Production Department, Colloege of Agriculture, King Saud University

P. O. Box 2460, Riyadh, Kingdom of Saudi Arabia

ABSTRACT : Factor analysis was applied as a multi-variate statistical technique to official genetic evaluations of type classification traits for 1,265,785 Holstein cows and 10,321 sires computed from data collected between August 1982 and June 1994 in Canada. Type traits included eighteen linear descriptive traits and eight major score card traits. Principal components of the factor analysis showed that only five factors explain the information of the genetic value of linear descriptive traits for both cows and sires. Factor 1 included traits related to mammary system, like texture, median suspensory, fore attachment, fore teat placement and rear attachment height and width. Factor 2 described stature, size, chest width and pin width. These two factors had a similar pattern for both cows and sires. In contrast, Factor 3 for cows involved only bone-quality, while in addition for sires, Factor 3 included foot angle, rear legs desirability and legs set. Factor 4 for cows related to foot angle, set of rear leg and leg desirability, while Factor 4 related to

loin strength and pin setting for sires. Finally, Factor 5 included loin strength and pin setting for cows and described only pin setting for sires. Two factors only were required to describe score card traits of cows and sires. Factor 1 related to final score, feet and legs, udder traits, mammary system and dairy character, while frame/capacity and rump were described by Factor 2. Communality estimates which determine the proportion of variance of a type trait that is shared with other type traits via the common factor variant were high, the highest $\geq 80\%$ for final score, stature, size and chest width. Pin width and pin desirability had the lowest communality, 56% and 37%. Results indicated shifts in emphasis over the twelve-year period away from udder traits and dairy character, and towards size, scale and width traits. A new system that computes final score from type components has been initiated.

(Key Words: Factor Analysis, Communalities, Principal Component, Genetic Evaluation, Type Traits)

INTRODUCTION

Many dairy conformation systems have evolved over time, and involve highly correlated traits which need to be simplified. Some type traits like basic form and strength are similarly defined, represent the same part of the body and have high genetic correlation ($r_g = .91$, Foster et al., 1988). Schaeffer et al. (1985), found large genetic correlations among udder traits, ranging from .75 between udder support and teat placement and .72 between rear udder height and width to .27 between rear udder height and teat placement. Similarly Short and Lawlor (1992) found high genetic correlations between

udder depth and fore udder ($r_g = .79$) and fore udder attachment and udder width ($r_g = .90$). Large genetic correlations were also found among non udder traits, for example, Lin et al. (1987) and Misztal et al. (1995) found high genetic correlations between stature, body depth and strength, ranging from .75 to .95.

Multi-trait analysis of type traits also requires that traits with similar biological meaning and high correlations be grouped together or simplified, since several researchers used type traits separately for different prediction purposes; Moore et al. (1981), Keller and Allaire (1987) and Van Raden et al. (1990) predicted milk production; Honnette et al. (1980), Sullivan et al. (1996), Van Doormal et al. (1986), and Weigel et al. (1992) predicted stayability, herd life, and life-time index and relative net income.

The shortcoming of using large numbers of traits in genetic evaluation has been reported by some researchers. Hill and Thompson (1978) have shown that the

¹ Address reprint requests to A. K. Ali.

² Centre for Genetic Improvement of Livestock, Department of Animal & Poultry Science, University of Guelph, Guelph, Ontario N1G 2W1 Canada.

³ Nova Scotia Agricultural College, PO Box 550, Truro, NS B2N 5E3 Canada

probability of an estimated genetic correlation matrix is not positive definite but increases rapidly as the number of traits included in the analysis increases. Moreover, Visscher (1994) found a severe overestimate for the accuracy of Estimating Breeding Value (EBV) of herd life based on a large number of conformation traits. To avoid this overestimation, indirect EBV for herd life should be computed based on only a limited number of traits with a known biological relationship with survival ability.

Applying factor analysis to type trait data is important for several reasons. It enables: (1) summarizing information from the observed type traits into a few unobserved and relatively uncorrelated derived factors; (2) partitioning each trait response into a covariant and a specific part and the variance of each component can be estimated; (3) grouping type traits such that correlated traits could be isolated in the same factor and each factor will include traits with common biological and/or physiological characters. (4) The magnitude of each loading in the factor pattern reflects the importance of each type trait within the derived factor. Factor analysis also involves an orthogonal transformation of the loading of type traits which may have simpler interpretations than the original solutions.

Factor analysis can be implemented on: 1) phenotypic values of type traits as described by Seiber et al. (1987, 1988) and Ali et al (1995), or 2) on genetic values to obtain factors free from the non-genetic deviations, Seiber et al. (1988). The objectives of this study were two-fold: 1) to carry out factor analysis on the estimated genetic values of linear and scorecard type traits of Canadian Holstein sires and cows, and 2) to investigate the time trend in emphasis on certain genetic factor patterns for type trait data over four periods from 1882 to 1994.

MATERIALS AND METHODS

Data available for this study consisted of type classification records on 1,265,785 Holstein cows. Classification records on 25 conformation traits were collected by the Holstein Association of Canada between August 1982 and June 1994. The number of records for four periods 1982-1985; 1985-1988; 1988-1991 and 1991-1994 were 301,936, 299,865, 314,941 and 349,043 respectively. Estimated transmitting abilities (ETA), the unit of genetic evaluation, were computed from these data following the procedures generally used for genetic evaluations for conformation in Canada (Koots et al., 1994).

Briefly, the data were preadjusted for age at classification, stage of lactation and heterogeneous herd variance. For each trait, ETA were computed from a

single trait model, which included the effects of herd-round-classifier (A round is the length of time required for classifiers from Holstein Association of Canada to visit all herds, usually 9 to 10 mo.) and additive genetic value of animal. After adding the pedigree file, a total of 1,880,129 animals were used in the animal model analysis, including 42,582 sires. There were 119,017 used in the factor analysis.

The sire data set included 10,321 records which represented only sires with at least 60% repeatability. Unknown parents were assigned to 83 groups based on year and region. The ETA on the 1,265,785 cows with records formed the data set which was analyzed for the twenty-five traits. Type traits included eight major score card traits: (1) final score; (2) frame/capacity; (3) rump; (4) feed and legs; (5) fore udder; (6) rear udder; (7) mammary system; and (8) dairy character; while the seventeen descriptive traits included: (1) stature; (2) size; (3) chest width; (4) loin strength; (5) pin setting; (6) pin width; (7) foot angle; (8) bone quality; (9) set of rear leg; (10) udder texture; (11) median suspensory; (12) fore attachment height; (13) fore teat placement; (14) rear attachment height; (15) rear attachment width; (16) slope hooks to pins desirability; and (17) legs - desirability. Major score-card traits and descriptive traits were linear and scored into 18 and 9 categories, respectively. Prior to analyses the data was transformed using Snell's (1964) objective scoring procedure as is common for genetic evaluations for conformation in Canada.

Principal components (PC) of factor analysis (SAS-PROC Factor) were computed using genetic values of cows and sires. The analysis was carried out separately on the eight major score card traits and on the seventeen descriptive type traits. Final score which is a subjective function of type traits was included with descriptive type traits, to determine which factor would include it. Only those PC's were retained which were above the cut-off point when any additional PC would account for less variance than any standard type trait, which is equivalent to retaining principal components with Eigen values greater than 1.00, to represent the minimum number of factors which explain relationships among the type trait data. Factor analysis with the varimax rotation as described by Kaiser (1958) was carried out to obtain factors such that each factor has only the minimum number of traits with large absolute value of loadings, scaled by multiplying them by one hundred.

Factor model

According to Mulaik (1972), the factor model can be written as

$X = AY + E$, where

X is a vector of p responses described by observed random variables having a non-singular multinormal distribution, $A = ((\lambda_{ij}))$ is a matrix of loadings of the i^{th} response on the j^{th} common factor Y and E are normally distributed with mean zero and variance $\text{Var}(E) = \Psi = \text{diag}(\psi_1, \dots, \psi_p)$ where ψ_i is the specific variance of the i^{th} response. $\text{Var}(X) = \Sigma$ and $\Sigma = A'A + \Psi$. The diagonal elements of $A'A$ are called the communality.

RESULTS AND DISCUSSION

Eigen values and the percentage of total variance explained by principal components of genetic values of type traits as shown in table 1, indicate that the first five principal components of cows and sires generally have Eigen values greater than one. Five PC were retained for the analysis and these accounted for 71% of the total variance in all type traits. Seiber et al (1988) found that the first seven components accounted for 73.6% of the total variance in type score.

Table 1. Eigen Values and percentage of total variance explained by principal components (PC) of genetic values of type traits of cows and sires

PC	Cows		Sires	
	Eigen Values	%	Eigen Values	%
1	6.37	35.41	6.10	33.87
2	2.17	12.07	2.45	13.61
3	1.69	9.39	1.78	9.88
4	1.34	7.42	1.31	7.29
5	1.06	5.90	1.13	6.28
6	.97	5.39	.97	5.38
7	.81	4.53	.82	4.58
8	.63	3.50	.54	3.00
9	.53	2.93	.51	2.86
10	.49	2.73	.49	2.74
11	.42	2.35	.42	2.34
12	.37	2.07	.39	2.17
13	.32	1.77	.32	1.79
14	.28	1.56	.27	1.51
15	.23	1.29	.23	1.26
16	.15	.82	.13	.71
17	.09	.49	.08	.46
18	.07	.38	.05	.26

Table 2 shows that Factor 1 is related to the cow's mammary system; including udder texture and median suspensory ligament strength, as well as width and height of reat attachment and strength of fore attachment. Final classification in also included in Factor 1, as it is most affected by scores for udder characteristics. Rear attachment height and width, as well as median suspensory ligament strength had the higher coefficients, excluding final classification. The animal with a high score of this factor would have a fine texture and strongly supported udder with a high wide rear attachment and fore teat placement towards the inside of the quarters. Factor 2 clearly describes the size, stature and capacity of the cow, including width of chest and pins, with size of primary significance. So the cow with a large score for this factor is tall and large, with deep chest and wide pins. Factors 3 and 4 for cows, and Factors 3, for sires, related to bone quality, rear leg set, and foot angle; the latter being least significant for cows, and most for sires, which threw all three traits into one factor. High scores for Factor 3 and 4 for cows is an indication of smooth bone quality, steep foot angle and intermediate setting of rear legs. However, high scores for factor 3 for bulls would distinguish them with the most desirability for these three traits. Factor 5 for cows, and Factors 4 and 5, for sires, related to loin strength, pin width and slope from hooks to pins. The top scoring cows and bulls for these factors would exhibit strong loins, wide pins with a definite slope from hooks to pins.

Factor analysis of bulls linear type traits shows that Facotor 1 and Factor 2 have the same pattern of type traits as cow's data (table 2). However, Factors 3, 4 and 5 show some subtle differences since Factor 3 included foot angle, bone equality, rear legs set and legs desirability, Factor 4 relates to loin strength and pin setting while Factor 5 describes the hook to pin slope desirability. The similarity in the distribution of type traits by factors for sires and cows is expected ($F^1_{\text{cow}} = F^1_{\text{sire}}, F^2_{\text{cow}} = F^2_{\text{sire}}, (F3 + F4)_{\text{cow}} = F3_{\text{sire}}$ and finally $F5_{\text{cow}} = (F4 + F5)_{\text{sire}}$) since sire and cow genetic evaluations are not independent under the animal model. However, the model results in genetic evaluation for sires with higher accuracy than for cows, consequently factor analysis on sire genetic evaluations should confirm the analysis on cow's genetic evaluations.

Seiber et al. (1988) found that 18 type traits and milk our are described by seven factors, the first of which included the basic form, strength, rump width and rump side view. However, udder traits are included in Factor 2 and Facotor 3 and in particular, udder depth and fore udder attachment included in three factors (factors 2, 3

Table 2. Factor pattern co-efficients (*100) for five genetic factors and communality estimates of cows and sires linear type traits

Trait*	Cows					Communality	Sires					Communality
	F1	F2	F3	F4	F5		F1	F2	F3	F4	F5	
1 FS	73					90.33	79					91.16
2 Stat		84				80.75		81				83.30
3 Size		92				91.09		92				92.48
4 Chst		85				79.87		87				80.28
5 Loin					72	75.18				64		77.11
6 Pins					73	67.93				80		69.96
7 Pin W		65				55.52		68				57.27
8 Foot				65		72.14			76			71.15
9 BoneQ			80			71.68			33			63.71
10 RLS				77		67.41			65			59.86
11 Text	67					74.13	83					73.89
12 Med SL	72					78.63	88					79.41
13 FA	69					57.00	58					55.49
14 FTP	57					56.88	78					62.44
15 RAH	85					76.46	75					73.59
16 RAW	77					63.07	46					51.23
17 PinD					53	37.35					74	60.91
18 LegD				73		67.93			84			73.65

* FS = final score, St = stature, Chst = chest width, Loin = loin strength, Pins = pin setting, Pin W = pin width, Foot = foot angle, BoneQ = bone quality, RLS = set of rear legs, Text = dudder texture, MedSL = median suspensory ligament, FA = fore attachment, FTP = fore teat placement, RAH = rear attachment height, RAW = rear attachment width, PinD = Pin desirability, LegD = legs desirability.

and 7). Legs and foot characteristics are located at Factor 4 and body depth and stature are described by Factors 5 and 7. The discrepancy between the Seiber et al. (1987, 1988) studies and this study might be due to the following reasons: 1) Different source of data, since Seiber et al. (1987, 1988) used data from the twenty-First Century Genetics Mating Appraisal for profit program, 2) Different type traits because in addition to the sixteen type traits two other traits have been incorporated, milkout and disposition, which were isolated in factor 6. 3) Differences in scoring system, whereas type traits were scored from 1 to 50. Vinson et al. (1982) found differences in linear scores which seemed to identify conformational differences found by measurement and those evaluators were not using the entire scale for scoring, possibly because of inexperience.

The communality estimates (table 2) are the proportion of variance of a type trait that is shared with other type traits via the common factor variant. Also, as indicated by Mulaik (1972), communality estimates can be interpreted as the square of the multiple correlation coefficient between a trait and a linear function of other

type traits. Communality estimates of final score, stature, size and chest width for cows were ≥ 80 and were the highest. Perhaps classifiers had a final score in mind as they score these closely related traits. It may be wise to remove them altogether from the classification system. Low communality values ($< 60\%$) are found for bone quality, set of rear legs, fore attachment, fore teat placement and rear attachment height as well as for pin width and desirability of slope from hooks to pins for either cows or bulls (table 3). Since the variance of a trait is the sum of the communality and specificity, one can argue that traits with low communality that are similar do not share much variability with the major score card traits.

Factor analysis across periods (table 3) shows that in the first and third periods, Factor 1 describes udder traits but the same factor describes body form (stature, size and chest width) for the second and third periods. In contrast, Factor 2 involves udder traits in the second period and body form in the first and third period. In the later period, udder and teat traits were distributed among Factors 2 and 3. Foot angle, rear leg set and legs desirability were

Table 3. Factor Pattern Co-efficients (*100) for six genetic factors across periods for cows

Trait*	< 1985						1985-1988						1988-1991						1991-1994						Overall					
	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6
1 FS	82						69						71						60						73					
2 Stat	78						81						85						85						84					
3 Size	91						91						92						92						92					
4 Chst	86						87						85						84						85					
5 Lion				70						68						76						72						72		
6 Pins				88						88						73						71						73		
7 PinW																														
8 Foot				73						75						73						71						65		
9 BoneQ				53												78						77						80		
10 RLS				69						69						88						79						77		
11 Text				76						70						67						65						67		
12 Med SL				88						84						72						71						72		
13 FA				69						51						65						72						69		
14 FTP				75						81						54						59						57		
15 RAH				79						68						86						81						85		
16 RAW				56						86						77						82						77		
17 PinD																														
18 LegD				81						81						74						68						73		

* FS = final score, St = stature, Chst = chest width, Loin = loin strength, Pins = pin setting, PinW = pin width, Foot = foot angle, BoneQ = bone quality, RLS = set of rear legs, Text = udder texture, MedSL = median suspensory ligament, FA = fore attachment, FTP = fore teat placement, RAH = rear attachment height, RAW = rear attachment width, PinD = Pin desirability, LegD = legs desirability.

located at Factor 3 and Factor 4. Factor 4 and Factor 5 included loin strength and pin-setting. All periods had five factors except for the first period which showed that pin desirability was found at Factor 6. These findings clearly illustrated shifts in emphasis on udder vs. size/scale over time.

Eigen values and percentage of total variance explained by principal components (PC) of genetic values of score and traits of cows and sires are shown in table 4. Only the first two principal components had eigen values greater than one and these two principal components accounted for 69% of the variance for both cows and sires.

Table 4. Eigen values and percentage of total variance explained by principal components (PC) of genetic values of major score card traits of cows and sires

PC	Cows		Sires	
	Eigen Values	%	Eigen Values	%
1	4.46	55.78	4.40	54.99
2	1.03	12.89	1.17	14.63
3	.94	11.75	.88	10.96
4	.74	9.23	.74	9.27
5	.53	6.67	.54	6.71
6	.24	3.03	.23	2.88
7	.04	.47	.03	.42
8	.01	.17	.01	.14

Table 5 shows that Factor 1 described final score, feed and legs, udder traits, mammary system and dairy

Table 5. Factor pattern (*100) of two genetic factors and the communality estimates of scorecard type traits

Trait	Cows			Sires		
	F1	F2	Communality	F1	F2	Communality
FS	84		96.47	88		96.54
F/Cap		74	62.89		72	60.68
Rump		83	69.56		76	61.68
FL	64		46.31	61		59.12
FU	74		70.68	77		72.54
RU	86		81.79	88		82.61
MS	87		89.56	89		90.34
DC	53		32.11	57		33.48

FS = Final Score, F/Cap = Frame/Capacity, FL = Foot and Legs, FU = Fore Udder, RU = Rear Udder, MS = Mammary System, and DC = Dairy Character.

character. However, frame and capacity and rump were described by Factor 2. The distribution of the eight traits within the two factors was similar for cows and sires. Communality estimates, for cows and bulls, were greater than 80% for final score, rear udder and mammary system. On the other hand, dairy character had the lowest communality estimate for both sexes (32% for cows and 33% for bulls) indicating that it has less impact on final conformation scores than other major type components.

CONCLUSIONS

Conformation systems should describe all aspects of the cow's form that relate to utility. In cases where conformation traits are highly related, one may be enough. This study shows a very high communality among traditional major score card traits, and establishes that dairy character is not closely related to final type score in the era under study, 1982 to 1994. The study also indicates shifts in emphasis from final score card and udder traits to size and stature traits in certain five year periods. Although Final Classification was always associated with udder traits throughout this twelve-year period, size and stature traits had the greatest emphasis in the latest period (1991-1994), and in 1985-1988. A new system in which final score is computed from the various type components with pre-assigned weight may partially solve this problem. This study also indicates that the number of traits can be represented by fewer factors without reduction in accuracy in describing the cow's and sires' genotypes for conformation. Finally, udder traits should be given the the sight priority and the most emphasis in judging and evaluation.

REFERENCES

- Ali, A. K. A., M. AL-Enazi, E-Hayes and M. AL-Saidy. 1995. Phenotypic Factor and Image analysis of type traits of Holstein Cows in the Masstock Saudi Herds. *J. King Saud University*. 7:143.
- Boldman, K. G., A. E. Freeman, B. L. Harris and A. L. Kuck. 1992. Prediction of sire transmitting abilities for herd life from transmitting abilities of linear type traits. *J. Dairy Sci.* 75:SS2.
- De Hann, M. H. A., B. G. Cassell, R. E. Reardon and B. B. Smith. 1992. Relationship between net income, days of productive life, production and linear type traits in grade and registered Holsteins. *J. Dairy Sci.* 75:3553.
- Foster, W. W., A. E. Freeman, P. G. Berger and A. Kuck. 1988. Linear type trait analysis with genetic parameter estimation. *J. Dairy Sci.* 71:223.
- Hill, W. G. and R. Thompson. 1978. Probabilities of non-positive definite between group or genetic covariance matrices. *Biometrics* 34:429.

- Honnette, J. E., We. E. Vinson, J. M. White and R. H. Kliever. 1980. Predictor of herd life and life-time production from first lactation and individual type traits in Holstein Cows. *J. Dairy Sci.* 63:816.
- Kaiser, H. F. 1958. The Varimax Criterion for Analytic Rotation in Factor Analysis. *Psychometrika*, 23:187.
- Keller, D. S. and F. R. Allaire. 1987. Relationship of first lactation milk and type traits to cow survival and a sire index for discounted total milk. *J. Dairy Sci.* 70:2116.
- Koots, K. R., K. M. Wade, B. W. Kennedy, J. C. M. Dekkers, G. C. Smith and E. B. Burnside. 1994. Method and effect of adjustment for heterogeneous variance of Holstein conformation traits. *J. Dairy Sci.* 77:294.
- Lin, C. Y., A. J. Lee, A. J. McAllister, T. R. Batra, G. L. Roy, J. A. Veseley, J. M. Wauthy and K. A. Winter. 1987. Intercorrelations among milk production traits and body and udder measurements in Holstein heifers. *J. Dairy Sci.* 70:2385.
- Misztal, I., K. Weigel and T. J. Lawlor. 1995. Approximation of estimates of (Co) variance components with multiple-trait restricted maximum likelihood by multiple diagonalization for more than one random effect. *J. Dairy Sci.* 78:1862.
- Moore, R. K., S. Higgins, B. W. Kennedy and E. B. Burnside. 1981. Relationships of teat conformation and udder height to milk flow rate and milk production in Holsteins. *Canadian J. Anim Sci.* 61:493.
- Mulaik, S. A. 1972. *The Foundation of Factor Analysis*. New York: McGraw-Hill Book Company.
- Norman, H. D., B. G. Cassell, R. E. Pearson and G. R. Wiggans. 1981. Relation of first lactation production and conformation to life-time performance and profitability in Jerseys. *J. Dairy Sci.* 64:104.
- SAS User's Guide. 1986. *Statistical Analysis System*. SAS Inst. Inc., Box 8000, Cary NC. 27511-8000, U.S.A.
- Schaeffer, G. B., W. E. Vinson, R. E. Pearson and H. D. Norman. 1985. Genetic and phenotypic relationships among type traits scored linearly in Holsteins. *J. Dairy Sci.* 68:2984.
- Seiber, M., A. E. Freeman and P. N. Hinz. 1987. Factor analysis for evaluating relationships between first lactation type scores and production data of Holstein dairy cows. *Journal of Dairy Science*, 70:1018.
- Seiber, M., A. E. Freeman and P. N. Hinz. 1988. Comparison between factor analysis form phenotypic and genetic correlation matrix using linear type traits of Holstein dairy cows. *Journal of Dairy Science*, 71:477.
- Short, T. H. and T. J. Lawlor. 1992. Genetic parameters of Conformation traits, milk yield and herd life for Holsteins. *J. Dairy Science*, 75:1987.
- Snell, E. J. 1964. A scaling procedure for ordered categorical data. *Biometrics* 20:592.
- Sullivan, B. P., B. J. Van Doormal and E. B. Burnside. 1986. An analysis of the linearized type characteristics as predictors of stayability in Canadian Holstein. *J. Dairy Sci.* 69(Suppl) :267(Abstr).
- Van Doormaal, E. B. Burnside and L. R. Schaeffer. 1986. An analysis of the relationships among Stayability, Production, and Type in Canadian Milk-Recording Programs. *J. Dairy Sci.* 69:510.
- Van Raden, P. M., E. L. Jensen, T. J. Lawlor and D. A. Funk. 1990. Prediction of Transmitting abilities for Holstein Type Traits. *J. Dairy Sci.* 73:191.
- Vinson, W. E., R. E. Pearson and L. P. Johnson. 1982. Relationships between linear descriptive type traits and body measurements. *J. Dairy Sci.* 65:995.
- Visscher, P. M. 1994. Bias in Genetic R^2 from Half-sib Designs. *Proceedings of the 5th World Congress on Genetic Applied to Livestock Production*. Ontario, Canada. 18:394.
- Weigel, D. J., B. G. Cassell, I. Hoechele and R. E. Pearson. 1992. Multiple-trait prediction of Transmitting abilities for herd life and estimate of economic weights using relative net income and income adjusted for opportunity cost. *J. Dairy Sci.* 78:639.