Differences in evaluators and genetic parameter estimations using subjective measurements of beef cattle temperament

Elfren F. Celestino Jr,[†] Jordan K. Hieber,[†,||] Carl R. Dahlen,[†] David G. Riley,[‡] Sarah A. Wagner,[†] and Lauren L. Hulsman Hanna[†,1]

[†]Department of Animal Sciences, North Dakota State University, Fargo, ND 58108; [‡]Department of Animal Science, Texas A&M University, College Station, TX 77843; and ||Present address: Department of Animal and Range Sciences, Montana State University, Bozeman, MT 59717

Transl. Anim. Sci. 2019.3:1769–1773
doi: 10.1093/tas/txz084

INTRODUCTION

When handled, cattle that are calm have greater average daily gain, increased feed efficiency, and pregnancy rates (Voisinet et al., 1997). Cattle with excitable temperament have impaired feedlot performance, poor carcass, and meat quality traits (Francisco et al., 2015). Temperament is associated with stress experienced during husbandry procedures (Sebastian et al., 2011). It is therefore, important to select cattle with calm temperament to improve production and reproductive traits.

Objective temperament methods include exit velocity (Burrow et al., 1988) and movement measuring devices (Sebastian et al., 2011; Yu, 2016). Subjective methods include chute score (Grandin, 1993), pen score (Hearnshaw and Morris, 1984), and qualitative behavior attributes (QBA; Sant'Anna and Paranhos da Costa, 2013). There are concerns of evaluator bias due to evaluator experience and interpretation, where limited literature exists investigating if it affects genetic evaluations. The objective of this study was to 1) determine evaluator effect on subjective measures of temperament and 2) compare genetic parameter estimates when evaluator was included in the model. We hypothesize that evaluator has a significant effect for subjective measures of temperament and will affect genetic parameter estimations (heritability and breeding value ranking).

MATERIAL AND METHODS

Animals

All procedures were reviewed and approved by the Institutional Animal Care and Use Committee of North Dakota State University. Weaning age calves (n = 1,542) were used over a 4-yr period (2014: n = 420, 2015: n = 382, 2016: n= 337, and 2017: n = 403). Calves were produced by the North Dakota State University Central Grasslands Research Extension Center cow herd. This herd consists of roughly 425 Angus-based (AN) and Hereford-based (HP) females (mature cows and heifers) that are bred to either AN or HP bulls.

Breed Composition

Calf dams had unknown pedigree and breed composition if born prior to 2012. Some heifers born from 2012 to 2015 were retained for use in breeding, leading to a better estimation of breed composition since sire breed was known. Over the 4-yr period, dams were mated to either AN or HP bulls, except in the first year (AN only).

These matings resulted in calves of eight breed types (%): 75 AN 25 Unknown (UN; n = 362), 50

[1]Corresponding author: lauren.hanna@ndsu.edu
Received April 4, 2019.
Accepted May 26, 2019.

AN 50 UN (n = 951), 50 HP 25 AN 25 UN (n = 147), 50 HP 50 UN (n = 31), 87.5 AN 12.5 UN (n = 30), 62.5 AN 25 HP 12.5 UN (n = 13), 50 AN 25 HP 25 UN (n = 4), and 50 HP 37.5 AN 12.5 UN (n = 4). On the basis of primary breed (50% or greater), this resulted in 1,360 AN- and 182 HP-influenced calves.

Temperament Scoring

Collection of docility score (DS), temperament score (TS), and QBA followed Hieber (2016). Briefly, DS was a scale of 1 to 6 with the head caught in the chute and TS ranged from 1 to 5, where the intermediate score (3) was removed to avoid choosing the median value (Sant'Anna and Paranhos da Costa, 2013). For QBA, evaluators scored each attribute (n = 12) on a 136-mm line to indicate the level of expression. The QBA score was the distance of the mark from the far left side measured with a digital fractional caliper (General Tools & Instruments, New York, NY). Evaluators per year (n = 6) were assigned to two of three methods to avoid fatigue in scoring (i.e., 4 evaluators per method).

Calves entered the handling facility based on management group (young vs. old dams) and first encountered the Silencer chute (Moly Manufacturing Inc., Lorraine, KS), where weaning weight and DS were recorded. Calves then entered a four-platform standing scale that measured weight distribution 8 to 10 times per second before moving into a working pen where TS and QBA were evaluated. A handler was present so that evaluators could score specific attributes of TS and QBA.

Statistical Analysis

Principal component analysis (PCA) on QBA in SAS, v.9.4 (SAS Institute, Inc., Cary, NC) per evaluator produced the temperament index (TI) using the first principal component (PC; Sant'Anna and Paranhos da Costa, 2013). Each trait (n = 15) was evaluated in SAS for fixed effects of evaluator (n = 4 per trait; 11 evaluators total), sex (n = 2), breed composition (n = 8), interactions of evaluator by sex and breed composition by sex, as well as a fixed covariate based on year, day, and sequence of evaluation using a repeated measures design. The final model across traits was then applied using pedigree in ASReml 4.2 (Gilmour et al., 2015) to calculate estimates of additive genetic variance, permanent environmental variance, its ratio over phenotypic variance ($\hat{R}$) and heritability.

RESULTS AND DISCUSSION

Statistical Modeling

Breed composition was significant in 7 of 15 models. Upon review, smaller sample sizes were

Table 1. Record summary per evaluator for docility score (DS), temperament score (TS), qualitative behavior attributes (QBA), and temperament index (TI)

	Evaluator										
Method	1	2	3	4	5	6	7	8	9	10	11
No. of years[1]	1	3	1	4	1	2	4	2	1	4	1
DS	418	—	382	1,541	419	702	1,534	740	398	—	—
TS	—	1,181	382	1,542	420	—	—	739	—	1,532	336
QBA[2]											
Apathetic	420	1,203	—	—	—	719	1,542	—	402	1,541	337
Calm	420	1,204	—	—	—	719	1,539	—	402	1,542	337
Curious	420	1,205	—	—	—	719	1,538	—	402	1,541	337
Happy	419	1,205	—	—	—	719	1,542	—	402	1,542	337
Pos. occupied	418	1,202	—	—	—	719	1,534	—	402	1,541	337
Relaxed	419	1,205	—	—	—	719	1,542	—	402	1,542	337
Active	420	1,205	—	—	—	719	1,542	—	402	1,542	337
Agitated	419	1,201	—	—	—	719	1,527	—	402	1,542	337
Attentive	419	1,202	—	—	—	718	1,539	—	402	1,539	337
Distressed	419	1,205	—	—	—	719	1,542	—	401	1,542	337
Fearful	420	1,204	—	—	—	718	1,539	—	402	1,542	337
Irritated	419	1,204	—	—	—	719	1,537	—	402	1,540	337
TI	420	1,205	—	—	—	719	1,542	—	402	1,542	337

[1]Number of years the evaluator scored as part of the project.

[2]QBA are grouped by positive-like (apathetic, calm, curious, happy, positively [pos.] occupied, and relaxed) and negative-like (active, agitated, attentive, distressed, fearful, and irritated) behavior.

driving significance rather than finding true breed composition differences. Primary breed ($n = 2$) was used as a fixed effect instead. Interactions of evaluator by sex and primary breed by sex were not included in the final model. Majority of models ($n = 8$ or 9 of 15, respectively) indicated these interactions were not significant. The final model across traits included fixed effects of evaluator, primary breed, sex, and covariate of year–date sequence.

Evaluator Effect

Evaluators vary across years as some could not return, where attempts were made to keep experience level equivalent. Evaluator summaries are presented in Table 1. Least square means for evaluator effect are presented in Tables 2 and 3. Evaluators scored differently in all evaluation methods ($P < 0.001$), except for TI ($P = 0.48$).

For DS, differences ranged from 0.18 (3% of the scale) to 1.2 (20%; Table 2). Similarly, differences ranged from 0.19 (4.75%) to 0.65 (16.25%) for TS (Table 2). Active (95.07, 69.9%), attentive (80.37, 59.10%), apathetic (74.93, 55.10%), curious (62.24, 45.76%), and relaxed (49.35, 36.28%) QBA had large differences observed (Table 3). Attributes of positively occupied, happy, calm, and fearful had moderate differences of 42.74 (31.43%), 39.34 (28.93%), 38.84 (28.56%), and 34.31 (25.23%), respectively. Agitated, distressed, and irritated QBA had small differences of 17.77 (13.07%), 15.31 (11.26%), and 7.06 (5.19%), respectively. Four of six negative-like behaviors had low differences seen between evaluators (less than 26%), meaning evaluators scored similarly for these types of behaviors. Negative QBA included active, agitated, attentive, distressed, fearful, and irritated.

For TI, evaluators' scores were not different from each other (Table 3). The first PC accounted for 39.64% to 45.90% variation across evaluators, which means TI captured the majority of variation for the 12 QBA (Figure 1). There is still concern if this truly explains temperament sufficiently for selection purposes. According to Kaiser (1960), PC with eigenvalues greater than 1 should be retained for interpretation. However, Ledesma and

Table 2. Least squares means and standard errors for evaluator on docility score (DS) and temperament score (TS)

Evaluator	Method	
	DS	TS
1	1.59 ± 0.04^{d}	—
2	—	1.63 ± 0.06^{d}
3	2.52 ± 0.04^{a}	2.28 ± 0.07^{a}
4	1.76 ± 0.04^{c}	1.89 ± 0.06^{c}
5	2.01 ± 0.05^{b}	1.88 ± 0.07^{c}
6	2.14 ± 0.04^{b}	—
7	1.51 ± 0.03^{d}	—
8	1.33 ± 0.04^{e}	1.84 ± 0.07^{c}
9	1.32 ± 0.04^{e}	—
10	—	2.09 ± 0.06^{b}
11	—	$1.93 \pm 0.07^{b,c}$

[a,b,c,d,e]Within a column, different superscript letters differ ($P < 0.05$).

Table 3. Least squares means and standard errors for evaluator on qualitative behavior attributes (QBA) and temperament index (TI)

Method	Evaluator						
	1	2	6	7	9	10	11
QBA[1]							
Apathetic	56.37 ± 2.55^{b}	22.68 ± 2.20^{d}	10.00 ± 2.32^{f}	19.97 ± 2.15^{e}	36.65 ± 2.56^{c}	84.93 ± 2.15^{a}	21.89 ± 2.65^{d}
Calm	85.76 ± 2.95^{a}	75.51 ± 2.71^{a}	51.40 ± 2.79^{c}	65.15 ± 2.68^{b}	81.30 ± 2.96^{a}	69.20 ± 2.68b	90.24 ± 3.03^{a}
Curious	40.29 ± 1.90^{b}	36.53 ± 1.46^{b}	7.81 ± 1.61^{e}	23.08 ± 1.41^{d}	$35.25 \pm 1.92^{b,c}$	28.22 ± 1.41^{c}	70.05 ± 2.00^{a}
Happy	30.84 ± 2.25^{c}	25.11 ± 1.92^{c}	11.14 ± 2.03^{d}	10.90 ± 1.88^{d}	30.19 ± 2.26^{c}	41.96 ± 1.88^{b}	50.24 ± 2.34^{a}
Pos. occupied	$49.82 \pm 1.68^{a,c}$	34.79 ± 1.31^{a}	7.08 ± 1.44^{c}	10.04 ± 1.27^{c}	12.58 ± 1.70^{c}	46.66 ± 1.27^{a}	28.13 ± 1.76^{b}
Relaxed	90.85 ± 2.78^{a}	72.00 ± 2.53^{a}	41.50 ± 2.62^{c}	59.06 ± 2.50^{b}	$75.84 \pm 2.79^{a,b}$	68.91 ± 2.50^{a}	$83.43 \pm 2.86^{a,b}$
Active	20.96 ± 2.12^{f}	37.33 ± 1.86^{e}	55.69 ± 1.95^{d}	38.57 ± 1.83^{e}	65.22 ± 2.13^{c}	79.01 ± 1.83^{b}	116.03 ± 2.20^{a}
Agitated	20.36 ± 1.87^{c}	20.12 ± 1.66^{c}	31.06 ± 1.73^{b}	$24.45 \pm 1.63^{c,d}$	32.31 ± 1.88^{b}	37.89 ± 1.63^{a}	23.93 ± 1.93^{c}
Attentive	55.31 ± 1.83^{b}	$44.86 \pm 1.46^{c,d}$	39.12 ± 1.58^{c}	40.48 ± 1.41^{c}	58.10 ± 1.85^{b}	61.35 ± 1.41^{b}	119.49 ± 1.92^{a}
Distressed	10.29 ± 1.07^{c}	12.77 ± 0.89^{c}	21.26 ± 0.95^{a}	13.56 ± 0.86^{c}	12.72 ± 1.08^{c}	5.95 ± 0.86^{d}	16.40 ± 1.12^{b}
Fearful	11.69 ± 1.64^{c}	$14.78 \pm 1.43^{c,e}$	39.74 ± 1.50^{b}	23.85 ± 1.40^{c}	46.00 ± 1.65^{a}	23.11 ± 1.40^{c}	$18.85 \pm 1.70^{c,d}$
Irritated	$19.71 \pm 1.65^{a,c}$	$22.94 \pm 1.46^{a,b,c}$	$21.03 \pm 1.53^{b,c}$	18.35 ± 1.44^{c}	$23.19 \pm 1.66^{a,b,c}$	$25.33 \pm 1.44^{a,b}$	25.41 ± 1.71^{a}
TI	-0.32 ± 0.13^{a}	-0.01 ± 0.08^{a}	0.03 ± 0.09^{a}	0.02 ± 0.07^{a}	0.37 ± 0.13^{a}	0.026 ± 0.07^{a}	0.14 ± 0.13^{a}

[a,b,c,d,e,f]Within a row, different superscript letters differ ($P < 0.05$).

[1]QBA are grouped by positive-like (apathetic, calm, curious, happy, positively [pos.] occupied, and relaxed) and negative-like (active, agitated, attentive, distressed, fearful, and irritated) behavior.

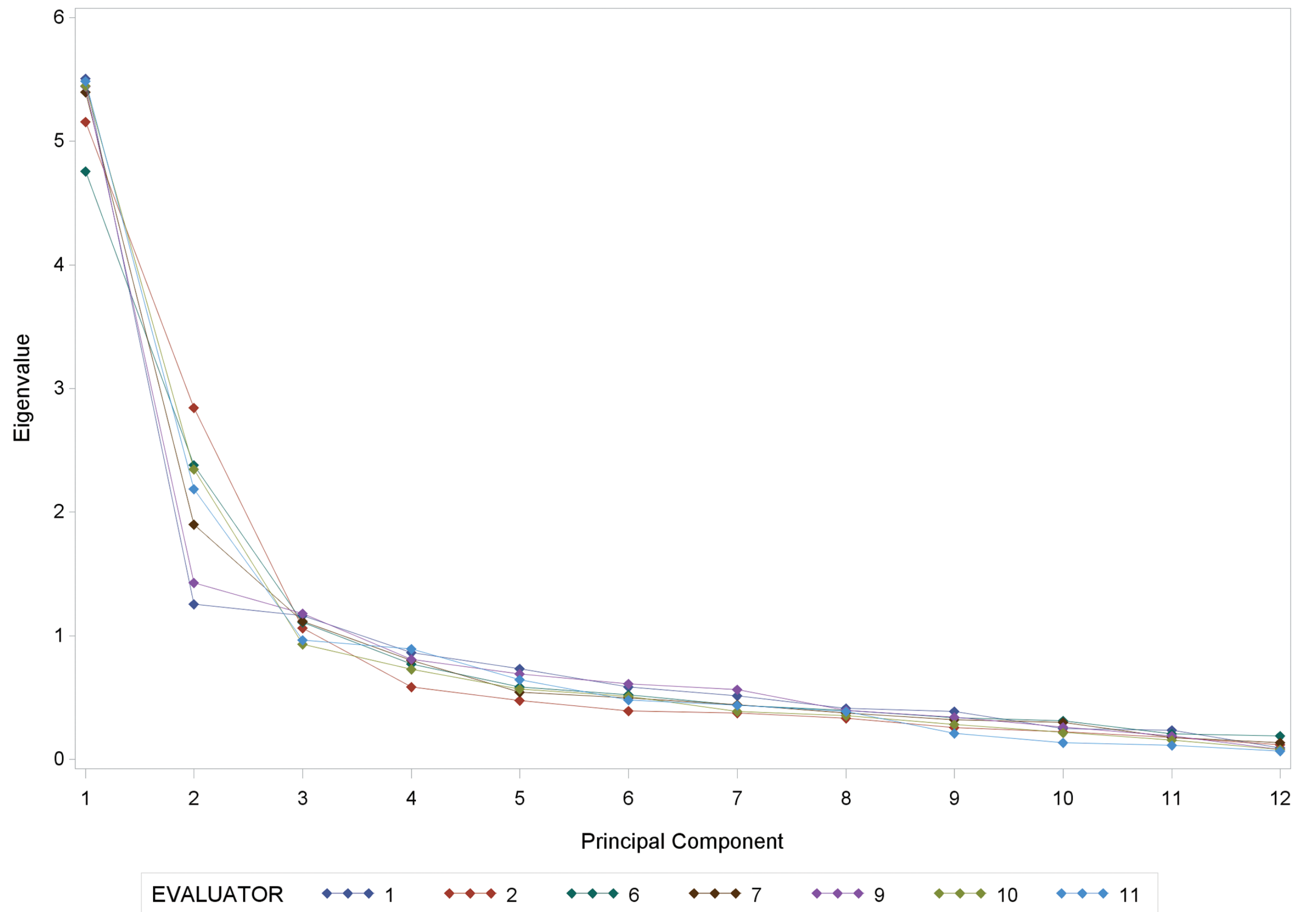

Figure 1. Principal component analysis scree plot across evaluators (n = 7). Eigenvalues greater than 1 contribute to significant variation in the data (Kaiser criterion, 1960).

Table 4. Genetic parameter estimates across evaluators for docility score (DS), temperament score (TS), qualitative behavior attributes (QBA), and temperament index (TI)[1]

Method	N[1]	$\hat{\sigma}_a^2$	$\hat{\sigma}_{pe}^2$	$\hat{\sigma}_p^2$	$\hat{h}^2$	$\hat{R}$
DS	6,134	0.06 ± 0.02	0.14 ± 0.02	0.44 ± 0.01	0.13 ± 0.04	0.45 ± 0.01
TS	6,132	0.22 ± 0.05	0.31 ± 0.04	0.87 ± 0.03	0.25 ± 0.05	0.61 ± 0.01
QBA[2]						
Apathetic	6,164	294.88 ± 44.54	28.59 ± 34.45	1,141.70 ± 25.17	0.26 ± 0.04	0.28 ± 0.02
Calm	6,163	422.50 ± 73.74	362.04 ± 60.54	1,436.00 ± 38.83	0.29 ± 0.05	0.55 ± 0.01
Curious	6,162	77.96 ± 18.65	18.60 ± 17.55	840.82 ± 15.83	0.09 ± 0.02	0.11 ± 0.01
Happy	6,166	201.12 ± 32.03	67.87 ± 26.31	936.52 ± 20.20	0.22 ± 0.03	0.29 ± 0.02
Pos. occupied	6,153	72.34 ± 8.60	0.00 ± 0.00	621.57 ± 11.72	0.12 ± 0.01	0.12 ± 0.01
Relaxed	6,166	415.76 ± 70.00	285.73 ± 57.15	1,504.50 ± 37.65	0.28 ± 0.04	0.47 ± 0.01
Active	6,167	176.64 ± 35.19	156.00 ± 29.89	835.02 ± 19.53	0.21 ± 0.04	0.40 ± 0.01
Agitated	6,147	135.07 ± 28.20	163.56 ± 24.52	655.75 ± 15.97	0.21 ± 0.04	0.46 ± 0.01
Attentive	6,156	88.62 ± 19.39	26.61 ± 17.39	732.78 ± 14.19	0.12 ± 0.03	0.16 ± 0.01
Distressed	6,165	28.04 ± 8.37	59.30 ± 8.07	264.73 ± 5.66	0.11 ± 0.03	0.33 ± 0.01
Fearful	6,162	95.55 ± 19.69	80.29 ± 17.23	571.60 ± 12.21	0.17 ± 0.03	0.31 ± 0.01
Irritated	6,158	100.45 ± 22.78	149.20 ± 20.24	525.81 ± 12.94	0.19 ± 0.04	0.47 ± 0.01
TI	6,167	0.00 ± 0.00	0.00 ± 0.00	5.26 ± 0.10	0.00 ± 0.00	0.00 ± 0.00

[1]N = number of records across evaluators, $\hat{\sigma}_a^2$ = estimated additive genetic variance, $\hat{\sigma}_{pe}^2$ = estimated maternal permanent environment variance, $\hat{\sigma}_p^2$ = estimated phenotypic variance, $\hat{h}^2$ = estimated heritability, and $\hat{R}$ = the sum of additive genetic and permanent environmental variances divided by phenotypic variance, which represents repeatability of evaluators. Least square means and standard errors were calculated using ASReml 4.2 (Gilmour et al., 2015) using fixed effects of evaluator, primary breed, sex, and fixed covariate of year–date sequence and random effect of animal (with and without pedigree).

[2]QBA are grouped by positive-like (apathetic, calm, curious, happy, positively [pos.] occupied, and relaxed) and negative-like (active, agitated, attentive, distressed, fearful, and irritated) behavior.

Valero-Mora (2007) suggest Kaiser's method is problematic. Various simulation studies demonstrated that PCA substantially overestimate or underestimate the number of factors retained (Zwick and Velicer, 1986), which could explain findings in this study.

Genetic Parameter Estimates

Differences in heritability ($\hat{h}^2$) were observed for the DS, TS, and QBA given the fixed effect of evaluator in the model (Table 4). The QBA scale used a 136-mm line, whereas DS and TS used discrete scales. Fordyce et al. (1996) reported $\hat{h}^2 = 0.14 \pm 0.11$ for weaned *Bos indicus* crosses using a chute test similar to DS. Phocas et al. (2006) reported 0.18 ± 0.01 for Limousin heifers using a scale similar to TS. Hoppe et al. (2010) reported that chute score $\hat{h}^2$ ranged from 0.11 to 0.33 in five different *Bos taurus* breeds. Heritability estimate of DS was lower than most traits captured in the working pen, whereas the permanent environmental effect ($\hat{R}$) was much higher. This coincides with observations of animals being scored differently in DS compared to TS or QBA (data not shown). Investigations are ongoing related to impact on genetic merit estimation.

ACKNOWLEDGMENTS

The authors express their gratitude to the North Dakota State University (NDSU) Central Grasslands Research Extension Center director and staff, the many NDSU staff and students involved in data collection and processing, and funding assistance from the North Dakota Agricultural Experiment Station.*Conflict of interest statement:* The authors confirm that there are no known conflicts of interest associated with this work.

LITERATURE CITED

Burrow, H. M., G. W. Seifert, and N. J. Corbet. 1988. A new technique for measuring temperament in cattle. Proc. Aust. Soc. Anim. Prod. 17:154–157.

Fordyce, G., C. J. Howitt, R. G. Holroyd, P. K. O'Rourke, and K. W. Entwistle. 1996. The performance of Brahman-Shorthorn and Sahiwal-Shorthorn beef cattle in the dry tropics of northern Queensland. 5. Scrotal circumference, temperament, ectoparasite resistance, and the genetics of growth and other traits in bulls. Aust. J. Exp. Agr. 36:9–17. doi:10.1071/ea9960009

Francisco, C. L., F. D. Resende, J. M. Benatti, A. M. Castilhos, R. F. Cooke, and A. M. Jorge. 2015. Impacts of temperament on Nellore cattle: physiological responses, feedlot performance, and carcass characteristics. J. Anim. Sci. 93:5419–5429. doi:10.2527/jas.2015-9411.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, S. J. Welham, and R. Thompson. 2015. ASReml User Guide Release 4.1. [Accessed April 4, 2019]. https://asreml.kb.vsni.co.uk/knowledge-base/asreml_documentation/.

Grandin, T. 1993. Behavioral agitation during handling of cattle is persistent over time. Appl. Anim. Behav. Sci. 36:1–9. doi:10.1016/0168-1591(93)90094-6

Hearnshaw, H., and C. A. Morris. 1984. Genetic and environmental effects on temperament score in beef cattle. Aus. J. of Ag. Res. 35(5):723–733. doi:10.1071/AR9840723

Hieber, J. K. 2016. Temperament evaluation in beef cattle: understanding evaluator bias within subjective measurements of docility score, temperament score, and qualitative behavior assessment. [master's thesis]. Fargo, ND: Department of Animal Sciences, North Dakota State University.

Hoppe, S., H. R. Brandt, S. König, G. Erhardt, and M. Gauly. 2010. Temperament traits of beef calves measured under field conditions and their relationships to performance. J. Anim. Sci. 88:1982–1989. doi:10.2527/jas.2008-1557.

Kaiser, H. F. 1960. The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20:141–51. doi:10.1177/001316446002000116

Ledesma, R. D., and P. Valero-Mora. 2007. Determining the number of factors to retain in EFA: an easy to use computer program for carrying out parallel analysis. Pract. Assess. Res. & Eval. 12(2):1–11.

Phocas, F., X. Boivin, J. Sapa, G. Trillat, A. Boissy, and P. Le Neindre. 2006. Genetic correlations between temperament and breeding traits in Limousin heifers. Anim. Sci. 82:805–811. doi:10.1017/asc200696

Sant'Anna, A. C., and M. J. R. Paranhos da Costa. 2013. Validity and feasibility of qualitative behavior assessment for the evaluation of Nellore cattle temperament. Livest. Sci. 157:254–262. doi:10.1016/j.livsci.2013.08.004

Sebastian, T., J. M. Watts, J. M. Stookey, F. Buchanan, and C. Waldner. 2011. Temperament in beef cattle: methods of measurement and their relationship to production. Can. J. Anim. Sci. 91:557–565. doi:10.4141/cjas2010-041

Voisinet, B. D., T. Grandin, J. D. Tatum, S. F. O'Connor, and J. J. Struthers. 1997. Feedlot cattle with calm temperaments have higher average daily gains than cattle with excitable temperaments. J. Anim. Sci. 75:892–896. doi:10.2527/1997.754892x.

Yu, H. 2016. The exploration of a four-platform standing scale in the application of measuring temperament in beef cattle. [master's thesis]. Fargo, ND: Department of Animal Sciences, North Dakota State University.

Zwick, W. R., and Velicer, W. F. 1986. Comparison of five rules for determining the number of components to retain. Psych. Bulletin 99:432–442.