



HAL
open science

Reducing stress during welfare inspection - Validation of a non-intrusive version of the laywel plumage scoring system for laying hens

Joergen Kjaer, Henrike Glawatz, Britta Scholz, Sophie Rettenbacher, Ragnar Tauson

► To cite this version:

Joergen Kjaer, Henrike Glawatz, Britta Scholz, Sophie Rettenbacher, Ragnar Tauson. Reducing stress during welfare inspection - Validation of a non-intrusive version of the laywel plumage scoring system for laying hens. *British Poultry Science*, 2011, 52 (02), pp.149-154. 10.1080/00071668.2011.554799 . hal-00686679

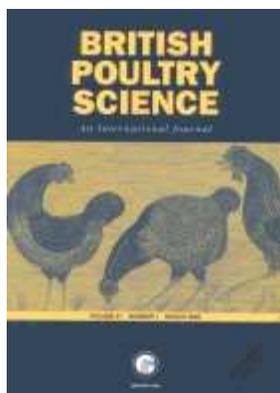
HAL Id: hal-00686679

<https://hal.science/hal-00686679>

Submitted on 11 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Reducing stress during welfare inspection -
Validation of a non-intrusive version of the laywel plumage
scoring system for laying hens**

Journal:	<i>British Poultry Science</i>
Manuscript ID:	CBPS-2010-041.R3
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	16-Aug-2010
Complete List of Authors:	Kjaer, Joergen; Friedrich-Loeffler-Institut, Institute for Animal Welfare and Animal Husbandry Glawatz, Henrike; FBN Dummerstorf, Fachbereich Genetik und Biometrie Scholz, Britta; Friedrich-Loeffler-Institut, Institute for Animal Welfare and Animal Husbandry Rettenbacher, Sophie; University of Veterinary Medicine, Biochemistry, Department of Biomedical Sciences Tauson, Ragnar; Swedish University of Agricultural Sciences, Animal Nutrition and Management
Keywords:	Behaviour, Welfare, Feather pecking, Laying hens, Physiology, Stress

SCHOLARONE™
Manuscripts

1
2 1 | **Reducing stress during welfare inspection: validation of a non-intrusive version of the**
3
4 2 | **LayWel plumage scoring system for laying hens**

5 3
6 4 | KJAER, J. B., GLAWATZ, H., SCHOLZ, B., RETTENBACHER, S. AND TAUSON, R.

7 5
8 6 | *Friedrich-Loeffler-Institut, Institute for Animal Welfare and Animal Husbandry,*
9
10 7 | *Dörnbergstrasse 25-27, D-29223 Celle, Germany*

11 8
12 9 | **Running title: Reducing stress during welfare inspection**

13
14 10 | **Accepted for publication: 24th Aug 2010**

15
16 11 |
17 12 |
18 13 |
19 14 | **Correspondence to:** Dr Joergen B. Kjaer, Friedrich-Loeffler-Institut, Institute for Animal
20
21 15 | Welfare and Animal Husbandry, Dörnbergstrasse 25-27, D-29223 Celle, Germany.
22
23 16 | Telephone +49 5141 384612; Fax +49 5141 3846117
24
25 17 | E-mail: joergen.kjaer@fli.bund.de
26
27 18
28
29 19
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2 20 **Abstract.** 1. The objectives of the present study were to validate a reduced, non-intrusive
3
4 21 version (RLS) of the LayWel plumage scoring system in domestic laying hens with reference
5
6 22 to complete, intrusive scoring (CLS) and to investigate the effect of these two scoring
7
8 23 methods on corticosterone metabolite concentrations.

9
10 24 2. A total of 312 medium-heavy laying hens from 4 commercial hybrids kept in 24 floor pens
11
12 25 were scored by two experienced teams. Another 150 hens from two hybrids kept in 6 pens
13
14 26 were used for estimating scoring treatment effects on corticosterone metabolites in droppings.

15
16 27 3. Plumage scores were in general higher using the RLS method compared to the CLS
17
18 28 method. The agreement between teams for plumage scores (CLS) were on a high (total score)
19
20 29 to an excellent (single body part except breast and cloaca) level.

21
22 30 4. Birds subjected to CLS tended to have higher concentrations of corticosterone metabolites
23
24 31 in droppings 2 h after scoring compared with birds in the control treatment (not scored). Birds
25
26 32 subjected to RLS had intermediate concentrations.

27
28 33 5. It was concluded that a reduced version of the LayWel scoring system is a valid and
29
30 34 reliable scoring method which tends to induce less stress to the subjects than the original
31
32 35 procedure.

33
34 36
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

INTRODUCTION

Animal welfare inspection of laying hen flocks often includes the evaluation of plumage condition which is considered to be an indirect measure of the amount of feather pecking in the flock (Hughes, 1982). Such scoring systems are important since recording the incidence of feather pecking in a flock by direct observation is difficult and very time consuming. Several methods of plumage scoring have been used, however, for practical reasons, subjective scoring is by far the most common method. The methods are classified as (i) application of a general score for the plumage of the whole body (Hughes and Duncan, 1972; Hill, 1980) and (2) application of a specific score for individual parts of the body (Tauson *et al.*, 1984; Gunnarsson *et al.*, 1995; Abrahamsson, 1996; Kjaer, 2000; Gunnarsson, 2000; Tauson and Holm, 2003). In the LayWel EU-project (LayWel, 2005) it was concluded that meta-analysis of plumage data was difficult due to the use of several different scoring systems. Hence a new practical system based on the scoring of several body parts was proposed as a standard to use in future research (Tauson *et al.*, 2005; 2007).

This new system involves the capture and handling of individual birds, which are scored for plumage condition on 6 parts of the body (neck, breast, back, wings, cloaca/vent and tail) along with scores for pecking damage to the rear of the body and comb, keel bone deviations and bumble foot lesions. Scores of 1-4 are awarded with higher scores indicating better condition of the integument. The system can be used to compare the incidence of scores 1-4 for individual body parts, or the scores can be pooled to give a whole body score ranging from 6-24. In the present study, we considered only the plumage condition element of the scoring system, which we refer to as the 'complete Laywel plumage scoring system' or CLS.

Scoring systems should be (1) simple to apply and allow for good repeatability (Tauson *et al.*, 1984), (2) impose as little disturbance to the birds as possible, and (3) be time-effective, enabling the user to score a large number of birds under commercial conditions in a

1
2 62 short period of time, and without stressing the birds. Capture and handling are stressful
3
4 63 procedures which can raise plasma corticosterone concentrations within minutes (Kannan and
5
6 64 Mench, 1996). In addition, capturing a representative sample in floor or free range systems
7
8 65 can be difficult and there are potential risks of sampling errors. CLS fulfils the requirements
9
10 66 of the first criterion, but not the second and third. We therefore developed a reduced version
11
12 67 of the LayWel scoring system (RLS) by introducing two major changes: (1) scoring only 4
13
14 68 body parts (neck, back, wings and tail) and (2) scoring birds on the ground without catching
15
16 69 and handling birds at any point. To our knowledge, no results from objective recording of
17
18 70 physiological stress variables after a combination of catching, handling and plumage scoring
19
20 71 have yet been published. We therefore assessed adrenocortical activity non-invasively (by
21
22 72 quantifying corticosterone metabolites in the droppings) after scoring, comparing values
23
24 73 obtained with these two scoring methods with concentrations of un-scored controls.

25
26 74 Thus, the objectives of the present study were to (1) validate the reduced, non-
27
28 75 intrusive version of the LayWel plumage scoring system in domestic laying hens (*Gallus*
29
30 76 *gallus domesticus*) with reference to the complete, intrusive scoring, and (2) to investigate the
31
32 77 effect of these two scoring methods on corticosterone metabolite concentrations.

33 34 78 35 79 MATERIALS AND METHODS

36 79 37 80 **Animals, housing and management**

38 80
39 81 A total of 1100 non-beak-trimmed chickens of 4 hybrid strains (ISA Warren (IW), Lohmann
40
41 82 Brown (LB), Lohmann Tradition (LT) and Tetra SL (TS)) were reared on litter floor and
42
43 83 housed at 16 weeks in 44 (11 per breed) partly (2/3) slatted floor pens (1.15 m wide x 4.10 m
44
45 84 long) at a stocking density of 6 hens per m² (25 hens per pen) and with a communal nest at the
46
47 85 rear of the pen. Layer mash (200 g CP/kg, 11.2 MJ ME/kg feed) and water were supplied *ad*
48
49 86 *libitum*. A 14L:10D light programme was applied with an approximate light intensity of 15
50
51
52
53
54
55
56
57
58
59
60

1
2 87 lux. The hens were part of an experiment investigating the design of 'on farm' performance
3
4 88 tests of laying hen hybrids for organic egg production.
5

6 89

7
8 90 **Data collection**

9
10 91 ***Plumage scoring***

11
12 92 Plumage condition of 312 hens was scored at the age of 73 weeks by two teams, each
13
14 93 consisting of two scorers and one assistant. Both teams scored the same 13 hens from each of
15
16 94 22 randomly chosen pens (6 per breed of LB and LT, 5 pens per breed of ISA and TB). The
17
18 95 practical scoring procedure consisted of the following steps:

19
20 96 (1) All hens were shepherded gently to one half of the pen and kept there by dividing it with a
21
22 97 catching frame made of steel and wire netting.

23
24 98 (2) Hens were separated by an assistant and individually guided through an opening to the
25
26 99 other half of the pen, while the scorers were standing nearby (2-3 m) scoring (blind to the
27
28 100 other scoring team) and recording the scores on paper (RLS method, see below).

29
30 101 (3) The assistant then caught, marked and crated the hen that had just been scored. Blue dye
31
32 102 marks on the legs enabled the second team to identify and score the same birds previously
33
34 103 scored by the first team, in order to produce independent scores by both methods on each bird.
35
36 104 The second team thus produced a CLS score blind to the RLS score given by the first team.

37
38 105 (4) After scoring, birds were immediately released back into the pen, thus avoiding the risk of
39
40 106 recapture.

41
42 107 The Complete LayWel Plumage Scoring System (CLS) is described in the
43
44 108 Introduction. Supplementary material with colour illustrations is available at the LayWel
45
46 109 homepage (LayWel, 2005) and also in Tauson *et al.* (2007). In the present experiment, only
47
48 110 that part of the integument representing plumage quality was used and validated and this
49
50 111 system will be called the 'complete LayWel plumage scoring system' (CLS).

1
2 112 The Reduced LayWel Plumage Scoring System (RLS) introduced two major changes:
3
4 113 (1) scoring only 4 body parts, neck, back, wings and tail, and (2) scoring birds on the ground
5
6 114 without catching and handling birds at all. The scorer stood outside the pen during the
7
8 115 scoring, thus minimising disturbance.

9
10 116 No scoring was carried out in the control treatment (CON) and no humans stood in
11
12 117 front of the pen. Only normal animal care, as practised in the other pens in the house, was
13
14 118 permitted.

15 16 119 17 18 120 ***Measurement of corticosterone metabolites***

19
20 121 Three pens of LB and three pens of LT hens, not previously scored for plumage condition,
21
22 122 were used to investigate the effect of scoring method and genotype on stress hormone
23
24 123 metabolites. One of three scoring methods was applied to each genotype: CLS, RLS or CON.
25
26 124 In CLS, all birds of a pen were caught and crated before scoring 30 min later. Birds were
27
28 125 scored while handled and released into the pen again. In RLS, the scorer went to the front of
29
30 126 the pen, opened the door and scored the birds individually from a distance; scoring time was
31
32 127 approximately 5 min per pen. In CON, there was no scoring or crating and no humans stood
33
34 128 in front of the pen. Excreta samples (consisting of faeces and urates together) were taken from
35
36 129 all hens in a pen beginning 120 min after initiation of the treatment. Excreta sampling lasted
37
38 130 approximately 20 min per pen. Only samples of adequate size (>0.1 g) from 86 out of a total
39
40 131 of 133 sampled birds were analysed. Samples were frozen at -21°C until later analysis for
41
42 132 3,11-dione glucocorticoid metabolites, as described by Rettenbacher *et al.* (2004) using a
43
44 133 cortisone enzyme immunoassay previously validated for chicken excreta (Rettenbacher *et al.*,
45
46 134 2009). Data were expressed as ng hormone equivalent per g excreta.

47 135

48 49 136 **Statistical analysis**

50
51
52
53
54
55
56
57
58
59
60

1
2 137 ***Plumage scoring***
3

4 138 For each combination of method (CLS, RLS) and scoring team (A, B), a total feather score
5
6 139 for each hen was calculated by summing the scores for each body region and then calculating
7
8 140 the average score per pen. This score had a minimum of 4 points (1 p per body part × 4 body
9
10 141 parts) and a maximum of 16 points (4 × 4) for the RLS method, summing scores for neck,
11
12 142 back, wings and tail. For the CLS method, two sums were calculated; a partial sum used for
13
14 143 comparing methods was calculated exactly like for the RLS method, summing neck, back,
15
16 144 wings and tail. Further, a complete sum was calculated for testing repeatability of teams
17
18 145 within this method, summing neck, back, tail, wings, breast and cloaca, thus making the
19
20 146 maximum possible sum 24 points per hen.
21

22 147 Pen average plumage score was the statistical unit. Method (CLS, RLS), team (A, B),
23
24 148 the interaction between method and team, and genotype (LT, LB, TS, IW) were included as
25
26 149 fixed factors into a mixed model procedure in SAS (SAS Inst. Inc., Cary, IL, USA). Method
27
28 150 and team were included as repeated measures (Littell *et al.*, 1996), and normal distribution of
29
30 151 residuals was accepted after visual inspection.
31

32 152

33
34 153 ***Reliability of methods and teams***
35

36 154 The between-method and between-team reliability were estimated using several measures of
37
38 155 agreement, namely kappa, weighted kappa and prevalence-adjusted bias-adjusted kappa
39
40 156 (PABAK), since there is controversy as to the best statistic. The kappa statistic alone is
41
42 157 appropriate if the marginal totals for the 2 × 2 table are relatively balanced, but if the
43
44 158 prevalence of a given response is very high or low, the value of kappa may indicate poor
45
46 159 agreement even when the observed proportion of agreement is quite high. Therefore we also
47
48 160 present the prevalence-adjusted bias-adjusted kappa (PABAK) to more fully characterise the
49
50 161 extent of the agreement between the two methods and teams respectively (Byrt *et al.*, 1993).
51
52
53
54
55
56
57
58
59
60

1
2 162 The procedure FREQ of SAS was used to calculate kappa, asymptotic standard error,
3
4 163 standardised test statistic, the P value for testing the two-sided hypothesis $Pr > |T|$, and
5
6 164 weighted kappa. A custom-written SAS-program was used to calculate PABAK, following
7
8 165 Cunningham (2009) and using the formula $PABAK = ([k + p]-1)/(k-1)$, where k is the number
9
10 166 of categories and p the proportion of agreement. According to Fleiss *et al.* (1993), a PABAK
11
12 167 score larger than 0.40 can be interpreted as good agreement and a score larger than 0.75 as
13
14 168 excellent agreement.

15
16 169

17 18 19 170 ***Hormone data***

20 171 Glucocorticoid metabolite concentrations, with pen mean as statistical units, were subjected to
21
22 172 one-way analysis of variance using the GLM procedure of SAS. *Post hoc* inspection of the
23
24 173 data revealed no effects of hybrid and the model included treatment only (CLS, RLS, CON).

25
26 174

27 28 175 RESULTS

29 30 176 **Plumage scoring**

31 32 177 ***Effects of method, team and genotype***

33
34 178 Plumage scores were significantly higher for the wings and tail with the RLS method and for
35
36 179 the neck and tail by team A. Total scores were subsequently significantly higher for the RLS
37
38 180 method and team A (see Table 1). Significant interactions were found between method and
39
40 181 team for neck, wings and tail score (Table 1). Only neck score showed a significant main
41
42 182 effect of genotype, with LB and TS having slightly lower scores even though pair-wise
43
44 183 differences were not significant (Table 1).

45
46 184

47
48 185 [Table 1 here]

49
50 186

51

52

53

54

55

56

57

58

59

60

1
2 187 ***Agreement between methods and teams***

3
4 188 In general, kappa and weighted kappa values between methods and teams for the total
5
6 189 plumage score ranged from 0.25 to 0.34 and all were significantly larger than zero (Table 2).
7
8 190 PABAK values for total scores were higher (0.35 to 0.55). Scoring the single body parts gave
9
10 191 excellent PABAK scores ranging from 0.75 to 0.97, with a few (cloaca and breast) being
11
12 192 somewhat lower (0.73 and 0.40 respectively) (Table 2). Kappa values for these body parts
13
14 193 were also low (0.38 and 0.18 respectively).

15
16 194
17
18 195 [Table 2 here]
19
20 196

21
22 197 **Corticosterone metabolites**

23
24 198 Treatment had no overall significant effect on concentration of corticosterone metabolites, but
25
26 199 birds subjected to CLS tended to excrete more corticosterone metabolites 2 h after scoring
27
28 200 compared with birds in CON (158 ± 25 ng/ml vs. 88 ± 25 ng/ml, $P=0.19$). Concentrations
29
30 201 were intermediate after the RLS method (137 ± 36 ng/ml).
31
32 202

33
34 203 **DISCUSSION**

35
36 204 The results showed that it is possible to obtain good precision by scoring birds at a distance
37
38 205 compared with catching them and scoring while handling. Furthermore, we found that
39
40 206 adrenocortical activity tended to be highest when birds were scored after catching and
41
42 207 handling, intermediate when scored at a distance, and lowest when not scored at all.

43
44 208 RLS scored significantly higher on all body parts than CLS. The same finding (RLS
45
46 209 scoring lower damage than CLS) was reported by Bright *et al.* (2006). It makes sense that
47
48 210 more damage is recorded when (1) birds are closely inspected (handled) and (2) when the
49
50 211 number of body parts scored increase. In addition, teams scored significantly differently. This

1
2 212 is a well known risk and the effect is normally balanced and included in the statistical models
3
4 213 used.

5
6 214 Agreement between teams using the CLS method, based mainly on PABAK scores,
7
8 215 were high for total plumage scores and excellent for single body regions, except for the breast
9
10 216 and cloaca. These findings are comparable to those found when validating the original version
11
12 217 of the LayWel scoring system (Tauson *et al.*, 1984) and other plumage scoring systems
13
14 218 (Adams *et al.*, 1978; Bright *et al.*, 2006). Interestingly, the somewhat lower agreement
15
16 219 between teams within CLS for the body parts breast and cloaca indicates that these body parts
17
18 220 are more difficult to score and special attention should be given here during the introduction
19
20 221 and training phase for scoring teams. Agreement was very high for the back score, which is
21
22 222 one of the most important areas for estimating the extent of feather pecking in the flock.

23
24 223 Our data support and extend those of Bright *et al.* (2006), who compared handling *vs.*
25
26 224 scoring at a distance using a scoring method (Bilcik and Keeling, 1999) only slightly different
27
28 225 to the one used in the present study. The two methods (handling *vs.* distance) used by Bright
29
30 226 *et al.* (2006) were found to be significantly correlated. However, there was a potential risk of
31
32 227 recapture in the investigation of Bright *et al.* (2006), thus possibly inflating the correlation. In
33
34 228 the present study this problem was overcome by the experimental design.

35
36 229 To our knowledge, this is the first study to investigate the potential of plumage scoring
37
38 230 to elicit physiological stress responses. The differences (45-65 ng/g) in corticosterone
39
40 231 metabolite concentrations between controls and treated birds were smaller than those reported
41
42 232 by Janczak *et al.* (2007), who found elevations of between 100 to 250 ng/g faecal
43
44 233 corticosteroid metabolites when unpredictable feeding schedules were applied to hens.
45
46 234 Baseline levels were very similar in the present experiment to those of Janczak *et al.* (2006),
47
48 235 even though, among other things, genotypes differed (White Leghorns *vs.* medium-heavy
49
50 236 strains). This is in accordance with findings of Fraisse and Cockrem (2006), who found no

1
2 237 differences in baseline concentrations of plasma corticosterone, but different stress response
3
4 238 patterns between two layer strains. Contrary to Janzak *et al.* (2006), severe feather pecking in
5
6 239 young pullets did not induce any significant change in the concentration of corticosterone
7
8 240 metabolites in excreta (Riber *et al.*, 2006).

9
10 241 In conclusion, a reduced version of the LayWel scoring system is a valid and reliable
11
12 242 scoring method which tends to induce less stress to the subjects than the original procedure.
13
14 243 The scores cannot, however, be expected to equal those obtained with the complete method.

15 244 16 245 ACKNOWLEDGMENTS

17
18 246 The authors thank Brigitte Müller and Anna-Lea Kopmann for valuable technical assistance.
19
20 247

21 248 REFERENCES

- 22 249 ABRAHAMSSON, P. (1996) Furnished cages and aviaries for laying hens. Effects on
23 250 production, health and use of facilities. *Dissertation*. Report 234. Department of Animal
24 251 Nutrition and Management, Swedish University of Agriculture, Uppsala.
- 25 252 ADAMS, A.W., CRAIG, J.V. & BHAGWAT, A.L. (1978) Effects of flock size, age at
26 253 housing, and mating experience on two strains of egg-type chickens in colony cages. *Poultry*
27 254 *Science*, **57**: 48-53.
- 28 255 BILCÍK, B. & KEELING, L. (1999) Changes in feather condition in relation to feather
29 256 pecking and aggressive behaviour in laying hens. *British Poultry Science*, **40**: 444-451.
- 30 257 BYRT, T., BISHOP, J. & CARLIN, J.B. (1993) Bias, prevalence and kappa. *Journal of*
31 258 *Clinical Epidemiology*, **46**: 423-429.
- 32 259 BRENNINKMEYER, C., DIPPEL, S., MARCH, S., BRINKMANN, J., WINKLER, C. &
33 260 KNIERIM, U. (2007) Reliability of a subjective lameness scoring system for dairy cows.
34 261 *Animal Welfare*, **16**: 127-129.
- 35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2 262 BRIGHT, A., JONES, T.A. & DAWKINS, M.S. (2006) A non-intrusive method of assessing
3
4 263 plumage condition in commercial flocks of laying hens. *Animal Welfare* **15**: 113-118.
5
6 264 CUNNINGHAM, M. (2009) More than Just the Kappa Coefficient: A Program to Fully
7
8 265 Characterize Inter-Rater Reliability between Two Raters. Paper no. 242-2009 in SAS Institute
9
10 266 Inc.: Proceedings of the SAS®Global Forum 2009 Conference (Cary, NC, SAS Institute
11
12 267 Incorporated).
13
14 268 FLEISS, J.L., LEVIN, B. & PAIK, M.C. (1993) Statistical methods for rates and proportions,
15
16 269 3rd edition (New York, John Wiley & Sons). Cit. from BRENNINKMEYER *et al.* (2007).
17
18 270 FRAISSE, F. & COCKREM, J.F. (2006) Corticosterone and fear behaviour in white and
19
20 271 brown caged laying hens. *British Poultry Science*, **47**: 110-119.
21
22 272 GUNNARSSON, S. (2000) Laying hens in loose housing systems. *Thesis*. Veterinaria 73,
23
24 273 Swedish University of Agriculture, Skara.
25
26 274 GUNNARSSON, S, ODÉN, K., ALGERS, B., SVEDBERG, J. & KEELING, L. (1995)
27
28 275 Poultry health and behaviour in a tiered system for loose housed layers. Report no. 35.
29
30 276 Swedish University of Agriculture, Skara.
31
32 277 HILL, A. (1980) Feather loss in layers. *MAFF Poultry Booklet* no.7, pp. 46-52. Gleadthorpe
33
34 278 Experimental Husbandry Farm.
35
36 279 HUGHES, B.O. (1982) Feather pecking and cannibalism in domestic fowls. *Hohenheimer*
37
38 280 *Arbeiten: Disturbed Behaviour of Farm Animals*, **121**: 138-146.
39
40 281 HUGHES, B.O. & DUNCAN, I.J.H. (1972) The influence of strain and environment upon
41
42 282 feather pecking and cannibalism in fowl. *British Poultry Science*, **13**: 525-547.
43
44 283 JANCZAK, A.M., TORJESEN, P., PALME, R. & BAKKEN, M. (2007) Effects of stress in
45
46 284 hens on the behaviour of their offspring. *Applied Animal Behaviour Science*, **107**: 66-77.
47
48 285 KANNAN, G. & MENCH, J.A. (1996) Influence of different handling methods and crating
49
50 286 periods on plasma corticosterone concentrations in broilers. *British Poultry Science*, **37**: 21-
51
52
53
54
55
56
57
58
59
60

- 1
2 287 31.
3
4 288 KJAER, J.B. (2000) Diurnal rhythm of feather pecking behaviour and condition of
5
6 289 integument in four strains of loose housed laying hens. *Applied Animal Behaviour Science*,
7
8 290 **65**: 331-347.
9
10 291 LAYWEL (2005) <http://www.laywel.eu>.
11
12 292 LITTELL, R.C., MILLIKEN, G.A., STROUP, W.W. & WOLFINGER, R.D. (1996) SAS
13
14 293 System for Mixed Models (Cary, NC, SAS Institute Inc.). ISBN 1-55544-779-1.
15
16 294 RETTENBACHER, S., MÖSTL, E., HACKL, R., GHAREEB, K. & PALME, R. (2004)
17
18 295 Measurement of corticosterone metabolites in chicken droppings. *British Poultry Science*, **45**:
19
20 296 704-711.
21
22 297 RETTENBACHER, S. & PALME, R. (2009) Biological validation of a non-invasive method
23
24 298 for stress assessment in chickens. *Berliner und Munchener Tierarztliche Wochenschrift*, **122** :
25
26 299 8-12.
27
28 300 RIBER, A.B., PALME, R. & FORKMAN, B. (2006) Effect of brooders on feather pecking
29
30 301 and cannibalism in domestic fowl (*Gallus gallus domesticus*). *Applied Animal Behaviour*
31
32 302 *Science*, **99**: 287-300.
33
34 303 TAUSON, R. & HOLM, K.-E. (2003) Evaluation of "Aviplus" - Big Dutchman - furnished
35
36 304 cage for 10 laying hens according to the §7 in the Swedish Animal welfare Ordinance and
37
38 305 according to the New-Technique Evaluation Program at the Swedish Board of Agriculture.
39
40 306 Report 251. Swedish University of Agriculture, Uppsala.
41
42 307 TAUSON, R., AMBROSEN, T. & ELWINGER, K. (1984) Evaluation of procedures for
43
44 308 scoring the integument of laying hens – Independent scoring of plumage condition. *Acta*
45
46 309 *Agricultura Scandinavica*, **34**: 400-408.
47
48 310 TAUSON, R., KJAER, J.B., MARIA LEVRINO, G. & CEPERO BRIZ, R. (2005) Applied
49
50 311 scoring of integument and health in laying hens. Proc. 7th European Symposium on Poultry
51
52
53
54
55
56
57
58
59
60

- 1
2 312 Welfare, 15-19 June, Lublin, Poland. *Polish Academy of Sciences* **23(Suppl. 1)**: 153-159.
3
4 313 TAUSON, R., KJAER, J.B., MARIA LEVRINO, G., CEPERO BRIZ, R. & HOLM, K.-E.
5
6 314 (2007) Applied scoring of integument and health in laying hens. URL:
7 315 <http://poultry.huv.slu.se/hen/scoringsystem/scoringsystem.pdf>.
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review Only

Table 1. Plumage score (least squares means) for methods (CLS = complete LayWel scoring; RLS = reduced LayWel scoring), teams and genotypes (IW = ISA Warren, LB = Lohmann Brown, LT = Lohmann Tradition and TS = Tetra SL)

Body part	Method		Team		Genotype				P values ¹			
	CLS	RLS	A	B	IW	LB	LT	TS	Meth.	Team	M*T	Genotype
Neck	2.28	2.29	2.36	2.21	2.35	2.22	2.36	2.21	ns	**	*	* ²
Back	3.68	3.77	3.71	3.74	3.62	3.68	3.76	3.84	ns	ns	ns	ns
Wings	3.58	3.78	3.71	3.65	3.74	3.62	3.71	3.65	***	ns	*	ns
Tail	3.62	3.84	3.76	3.70	3.65	3.71	3.86	3.71	***	*	**	ns
Total (CLSP)	13.2	13.7	13.5	13.3	13.4	13.2	13.7	13.4	***	*	ns	ns

¹*, ** or *** indicates a significant effect with $P < 0.05$, $P < 0.01$ or $P < 0.001$ respectively

²After Tukey-Kramer adjustment for multiple tests, no pair-wise differences between genotypes were significant

Table 2. Measures of agreement between methods (CLS = Complete LayWel Scoring; RLS = Reduced LayWel Scoring) within teams (A and B) and between-teams within methods. SE(kappa), Z and $P < |Z|$ are the asymptotic standard error, the standardised test statistic and the P value of the two sided test of the simple kappa, respectively

Comparison	Variable	Neck	Breast	Back	Wings	Tail	Cloaca	Total ¹
Method CLS vs. RLS								
<i>Within team A</i>								
	Kappa	0.59	-	0.62	0.51	0.23	-	0.28
	SE(kappa)	0.05	-	0.05	0.05	0.05	-	0.04
	Z	11	-	13	10	5	-	9
	$P < Z $	***	-	***	***	***	-	***
	Weighted Kappa	0.63	-	0.70	0.54	0.26	-	0.52
	PABAK	0.84	-	0.90	0.84	0.75	-	0.47
<i>Within team B</i>								
	Kappa	0.69	-	0.74	0.45	0.32	-	0.34
	SE(kappa)	0.06	-	0.05	0.05	0.05	-	0.04
	Z	13	-	16	9	6	-	11
	$P < Z $	***	-	***	***	***	-	***
	Weighted Kappa	0.73	-	0.79	0.53	0.40	-	0.61
	PABAK	0.94	-	0.95	0.78	0.77	-	0.53
Team A vs. B								
<i>Within method CLS</i>								
	Kappa	0.60	0.18	0.78	0.72	0.64	0.38	0.25
	SE(kappa)	0.05	0.04	0.05	0.05	0.05	0.06	0.03
	Z	11	5	15	13	12	6	11
	$P < Z $	***	*	***	***	***	***	***
	Weighted Kappa	0.64	0.41	0.83	0.76	0.69	0.42	0.61
	PABAK	0.87	0.40	0.96	0.90	0.87	0.73	0.35

Within method RLS

Kappa	0.44	-	0.78	0.59	0.34	-	0.33
SE(kappa)	0.06	-	0.05	0.06	0.07	-	0.04
Z	9	-	15	10	6	-	9
$P < Z $	***	-	***	***	***	-	***
Weighted Kappa	0.51	-	0.82	0.62	0.42	-	0.59
PABAK	0.79	-	0.97	0.90	0.88	-	0.55

¹When comparing CLS with RLS the total score includes 4 body parts only (neck+back+wings+tail)