ANIMAL WELL-BEING AND BEHAVIOR

A meta-analysis on the effect of environmental enrichment on feather pecking and feather damage in laying hens

Nienke van Staaveren,[*,†,1] Jennifer Ellis,[§] Christine F. Baes,[*,‡] and Alexandra Harlander-Matauschek[†]

*Centre for the Genetic Improvement of Livestock, Department of Animal Biosciences, University of Guelph, Guelph, Ontario, Canada N1G 2W1; †The Campbell Centre for the Study of Animal Welfare, Department of Animal Biosciences, University of Guelph, Guelph, Ontario, Canada N1G 2W1; ‡Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland 3001; and §Centre for Nutrition Modelling, Department of Animal Biosciences, University of Guelph, Guelph, Ontario, Canada N1G 2W1

ABSTRACT Feather pecking (**FP**) is a significant issue in laying hens, which impacts societal acceptance of poultry farming, farm productivity, and bird welfare. Environmental enrichment—modifications of the environment to stimulate biological functioning and psychological well-being of animals—is one management strategy farmers can use to mitigate FP. However, inconsistent results of environmental enrichment are reported across studies questioning its value. A meta-analysis was used to determine the effect of environmental enrichment on FP and feather damage (**FD**) in laying hens. A systematic review of published literature from 4 databases resulted in 23 publications that met inclusion criteria. Feather pecking and FD outcomes were standardized between studies using different scoring systems. Driving variables included the presence of enrichment, production period when the enrichment started, housing type, beak trimming, bird strain, and age of the birds when FP and FD was measured. Considering the experiment as a random effect, linear mixed model analysis was used in a 2-step approach, whereby variables with a $P < 0.30$ in univariate analysis were included within the subsequent multivariate analysis. Variables with $P < 0.05$ in the multivariate analysis were retained in the final models. Model selection and evaluation were based on corrected Akaike information criteria, the root mean square prediction errors, and concordance correlation coefficients. A higher frequency of FP was observed in flocks lacking enrichment ($P < 0.001$), with increased age ($P = 0.001$), and in cage housing systems ($P = 0.002$). Similarly, FD increased in flocks lacking enrichment ($P = 0.018$), with increased age ($P < 0.001$), in the absence of beak trimming ($P = 0.001$) and in cage housing systems ($P = 0.042$). This meta-analysis confirmed the effectiveness of environmental enrichment in reducing FP and FD. Nevertheless, the modest ability of enrichment to dampen FD (-0.14 ± 0.06, 1-4 scale) suggests that other management strategies must be implemented in conjunction with environmental enrichment to successfully manage FP and resulting FD.

Key words: poultry, injurious pecking, plumage, enrichment, housing

2021 Poultry Science 100:397–411
https://doi.org/10.1016/j.psj.2020.11.006

INTRODUCTION

Billions of laying hens are kept for egg production worldwide (IEC, 2015). These laying hens are descendants of the red jungle fowl native to south and east Asia which spends most of its time foraging in vegetation-rich, complex environments (Dawkins, 1989). The behavior of laying hens and jungle fowl is similar (Kruijt, 1964); however, genetic selection for production may have inadvertently introduced feather pecking behavior (Korte et al., 1997; Su et al., 2006). Feather pecking (**FP**) is a form of injurious behavior where birds peck at, pull, and potentially remove and consume each other's feathers resulting in feather damage (**FD**) (Savory, 1995; Bilčík and Keeling, 1999).

This behavior is a large problem in the laying hen sector, affecting anywhere between 8 to 65% of flocks and 15 to 95% of birds within those flocks depending on various internal and external factors (reviewed by van Staaveren and Harlander-Matauschek, 2020). The behavior has large implications for bird welfare, farm

Received September 2, 2020.

Accepted November 11, 2020.

[1]Corresponding author: nvanstaa@uoguelph.ca

productivity, and societal acceptance of poultry farming (Nicol et al., 2013; van Staaveren and Harlander-Matauschek, 2020). The main control strategy includes trimming of the sharp upper and lower mandible tips of the beak; however, this practice can cause short-term and long-term pain, the formation of neuromas, loss of normal beak function, and consequent changes in behavior (Hughes and Gentle, 1995; Kuenzel, 2007; Nicol, 2018). Furthermore, it does not eliminate FP behavior but rather reduces the consequences or symptoms of the behavior (i.e., FD) (Nicol, 2018). Therefore, the practice is under increasing scrutiny because of animal welfare and ethical and societal concerns (Nicol, 2018).

Scientists, veterinarians, farmers, and poultry sector representatives are continuously working to identify the underlying mechanisms and risk factors in an effort to prevent and reduce FP and FD (Nicol et al., 2013; Rodenburg et al., 2013). Unfortunately, the precise mechanisms for FP are still unknown. From an ethological point of view, it is suggested that FP may result from a lack of stimuli for normal species-specific behavior in their barren environment. In particular, FP is considered as a form of redirected pecking behavior stemming from frustration at the lack of foraging and feeding opportunities (Wennrich, 1974; Blokhuis, 1986; Huber-Eicher and Wechsler, 1997; Weeks and Nicol, 2006; Dixon et al., 2008). Modified or new housing systems have consequently been implemented in an attempt to reduce the barrenness of the laying hens' environment and to improve laying hen welfare (Blokhuis et al., 2007). However, this ethological viewpoint can be challenged by the fact that FP still occurs in all different types of housing systems from barren wire enclosures (i.e., cage housing systems) to complex single-tier or multi-tier barns (i.e., noncage housing systems) with outdoor access where they can perform species-specific foraging behavior (Green et al., 2000; Nicol et al., 2013; Rodenburg et al., 2013; Bestman and Wagenaar, 2014; Elkhoraibi et al., 2014).

Environmental enrichment is one of many management strategies farmers can use to reduce FP and FD (Lambton et al., 2013; Rodenburg et al., 2013; Jung and Knierim, 2018). Newberry (1995) defined environmental enrichment as "modifications within the environment that result in the improvement of biological functioning of the animal". More recently, it was stated that environmental enrichment should "enhance animal welfare by providing them sensory and motor stimulation, through structures and resources that facilitate the expression of species-specific behavior and promote psychological well-being through physical exercise, manipulative activities, and cognitive challenges according to species-specific characteristics" (National Research Council, 2011). However, diverse, inaccurate, and vague definitions have led to a large variation of possible enrichments, with some likely being biologically irrelevant (Würbel and Garner, 2007).

In laying hen flocks, enrichment can be provided in various forms such as objects or materials which are suitable for foraging or dustbathing (Campbell et al., 2019; Schreiter et al., 2019). Recommendations published by administrations, associations, universities, breeding companies, or food labels aimed at helping farmers reduce FP often encourage the provision of enrichment (Jung and Knierim, 2018). However, even in these guidelines, environmental enrichment is not always mentioned (7 out of 15 and 11 out of 15 guidelines recommended enrichment during the rearing and laying period, respectively) (Jung and Knierim, 2018). A possible reason for the variation could be the inconsistent, or even contradictory, results from epidemiological studies and experimental studies in terms of the effectiveness of enrichment material in reducing FP or FD (reviewed by Jung and Knierim, 2018; Campbell et al., 2019; Schreiter et al., 2019). Effectiveness of enrichment plays a large role in the uptake of this strategy by farmers (Newberry, 1995; Lambton et al., 2013), and inconsistent results make interpretation of findings difficult, thus undermining the strength of evidence in favor of providing enrichment to reduce FP. A quantitative meta-analysis that integrates the results of different independent studies (Phillips, 2005) to determine the effectiveness of enrichment is currently lacking. Therefore, the aim of this article was to use a meta-analysis approach to quantify the effect of enrichment on FP and FD in laying hens while considering several possible co-variates that could influence the results, including bird age, bird strain, beak trimming status, and housing system.

MATERIALS AND METHODS

Literature Search

A systematic literature review was conducted in January 2020 to identify relevant literature using Web of Science, CAB Direct, PubMed, and Agricola databases. Searches were performed using the following keywords "chicken OR poultry OR laying hen" AND "feather pecking OR feather damage OR plumage damage OR injurious pecking" AND "enrichment OR toy OR device" for all fields. No limitation was set on the language or year of publication to ensure the highest number of returned publications. A manual search of references cited in articles and reviews collected was also conducted to find any additional publications. Authors of articles for which additional information was needed were contacted.

After the removal of duplicates, publications were screened based on relevance of the title, abstract and keywords. Owing to the inherently different nature of cage and noncage housing systems, different types of environmental enrichment are observed within these systems. Therefore, within the database, environmental enrichment was defined as "any additional modifications within the environment" (Newberry, 1995), over any provisions that should be present in the housing system under investigation. Studies had to include a control treatment which was a similar housing system without the specific enrichment present. Consequently, studies that compared different housing systems (i.e., conventional cage, furnished cage, noncage systems, and outdoor systems), different litter substrates, or range quality were excluded. Housing systems were separated into cage systems (i.e., conventional and furnished cages) and noncage systems (i.e., single-tier and multi-tier systems with or without outdoor access). Items such as perches, claw abrasive devices, nest, or scratch areas were not considered as enrichments per se in cage systems, as these would be present in furnished cages (EFSA, 2005). Studies that were not specific to laying hens were excluded. The literature funnel (Moher et al., 2009) for the development of this database is presented in Figure 1.

Owing to the large variety in types of enrichment materials that were used in the studies (Table 1), we classified enrichment materials into 3 groups as per Schreiter et al. (2019). Specifically, 1) enrichment materials for foraging and consumption as feed, 2) enrichment materials for dustbathing, and 3) objects with no foraging or dustbathing opportunities. Most studies looked at one type of environmental enrichment, though some studies compared multiple types, or included a combination of the different types within 1 treatment (Table 1).

Data Extraction and Calculations

Information regarding type of studies and whether or not studies were performed on commercial farms is presented in Table 1. The total number of birds used in the different studies ranged from 40 to 900 on research facilities, whereas studies using commercial flocks reported flock sizes up to 16,000 birds. Reported means of the outcome measures (FP, FD) and potential x-variables were extracted from individual papers into a single line entry in Excel. When multiple experiments were presented within 1 paper, experiments were coded separately. Measures of variance (SD or SE) were extracted where possible, as well as the number of experimental units and the number of birds assessed per treatment. Data presented in graphs within the original publication were digitized to extract the mean, SD, and SE values of the FP or FD outcome (WebPlotDigitizer, https://automeris.io/WebPlotDigitizer/). Detailed information regarding the enrichment was extracted which was classified as 1 of the 3 types (i.e., foraging, dustbathing, objects) defined by Schreiter et al. (2019). Treatments that combined the different types were noted as such (Table 1). Additional descriptive parameters included housing system (cage vs. noncage systems), beak trimming (yes vs. no), production period at which the enrichment started (rearing vs. laying period), age at which FD and FP were measured (in weeks), and genetic strain of the birds. In experiments where the outcome was recorded at multiple time-points, the means of individual time-points were extracted where possible. When individual time-points were not provided, the overall mean of the outcome was extracted, and the average age at which the outcome was measured was used.

Both outcomes (FP, FD) were standardized to allow for comparison between experiments. Feather pecking was standardized to the average number of pecks per bird per min. Data were extracted for severe FP where possible; however, it should be noted that not all studies used the same definitions or looked at multiple types of bird-to-bird pecking (e.g., gentle FP, severe FP, aggressive pecking). Additionally, some studies reported the different types of pecking separately, whereas others only provided a value where all different types of bird-to-bird pecking were summed together. Where severe FP was not reported separately, data on the combined forms of bird-to-bird pecking were used instead. Feather damage was scored on different scales depending on the experiment (Table 2) and was either presented for individual body areas or presented as an overall sum of scores (range: 1–11 body areas). Scores were adjusted to represent 1 body area (if more than 1 area was represented, the old score was divided by the number of areas before conversion to the new score), rather than an overall sum of scores and transformed into the 1-4 scale following Tauson et al. (2005), where 4 indicates a better feather cover (Table 2) assuming a linear conversion between scales (Bedere et al., 2018). When scores were given for separate body areas, the tail/back/rump area was chosen as this is most likely to reflect FP (Wood-Gush and Rowland, 1973; Bilčík and Keeling, 1999). In some studies, the number or percentage of birds with FD was provided, and this was reconstructed to calculate the average score. The percentage of birds was

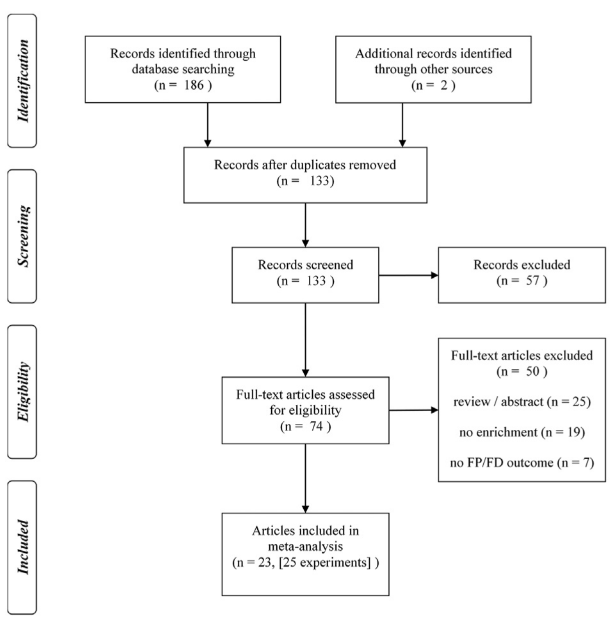

Figure 1. Literature funnel (Preferred Reporting Items for Systematic reviews and Meta-analyses diagram adapted from (Moher et al., 2009)). Abbreviations: FD, feather damage; FP, feather pecking.

Table 1. Summary description of publications included in the meta-analysis, with information on type of study, housing type, beak trimming status, enrichment type provided, the outcome measured, and the results found.

Publication	Study type	Housing	N[1]	Beak trim	Enrichment type	Outcome[2]			Result[3]	
						FD	FP	Pecking	FD	FP
Alm et al. (2017)	Factorial	Noncage	12	No	Foraging	x	-	-	=	-
Campbell et al. (2018)	Factorial	Noncage	6	Yes	Combination	x	-	-	NA[4]	-
Chow et al. (2005)	Factorial	Cage	10	NA[4]	Objects	-	x[5]	-	-	↓[6]
Cronin et al. (2018)	Factorial	Noncage	16	No	Foraging	x[5]	-	x	↓[6]	↓[6]
Daigle et al. (2014)	Factorial	Noncage	30	Yes	Objects, Foraging	x	x[5]	-	=	=
Decina et al. (2019)	Observational[7]	Noncage	38	Yes	NA[4]	x	-	-	↑	-
Dixon et al. (2010)	Factorial	Cage	48	No	Foraging, Objects, Dustbathing	-	-	x[5]	-	↓
Hartcher et al. (2015)	Factorial	Noncage	16	Both	Combination	x[5]	x	-	=	=
Ito et al. (2002)	Factorial	Cage	12	Yes	Objects	x	x[5]	-	↓[6]	=
Johannson et al. (2016) [8]	Factorial	Cage	8	Yes	Foraging	x	-	x	↓	↓
Jones et al. (2002)	Factorial	Cage	20	No	Objects	-	-	x	-	NA[4]
Klein et al. (2000)	Factorial	Noncage	16	No	Foraging	-	-	x[5]	-	↓
Liebers et al. (2019)	Factorial[5]	Noncage	18	No	Foraging	x	-	-	↓[6]	-
McAdie et al. (2005) [8]	Factorial	Cage	60	No	Objects	x	x[5]	-	↓	↓
Mielenz et al. (2010)	Factorial	Noncage	8	Yes	Combination	x	-	-	NA[4]	-
Morrissey et al. (2016)	Factorial[5]	Cage	64	Both	Objects	x	-	x	=	↓
Norgaard-Nielsen et al. (1993)	Factorial	Cage	16	Yes	Foraging	x[5]	-	-	↓[6]	-
Pettersson et al. (2017)	Intervention[7]	Noncage	12	Yes	Combination	x	x	-	=	↓
Schmidt et al. (2019)	Intervention[7]	Noncage	1	No	Foraging	x[5]	-	-	NA[4]	-
Steenfeldt et al. (2007)	Factorial	Noncage	16	No	Foraging	x	x	-	↓[6]	↓[6]
Tahamtani et al. (2016)	Factorial[7]	Noncage	23	No	Foraging	x[5]	x	-	NA[4]	=
Wechsler et al. (1998)	Factorial	Noncage	16	No	Foraging	x	x[5]	-	=	↓
Zepp et al. (2018)	Factorial[7]	Noncage	12	No	Foraging	-	x	-	-	↓

[1]Number of groups within the publication (depending on the housing system this refers to a cage, pen, or flock).
[2]Outcome measured: FD = feather damage, FP = feather pecking, pecking = combination of different forms of bird-to-bird pecking including FP.
[3]Result found: = no effect, ↓ improvement in FD/FP, ↑ worsening of FD/FP, NA not reported or not analyzed.
[4]NA = information not explicitly reported or not analyzed because of lack of birds affected.
[5]Data extracted using WebPlotDigitizer.
[6]Statistical tendency (P-value < 0.10) or results only observed for certain age group and/or body area assessed.
[7]Research performed on commercial farms.
[8]Two experiments within the publication.

reconstructed to the number of birds with a certain score, and the average score was calculated by summing all scores and dividing this by the total number of birds. Finally, a total of 210 treatment means were included based on 25 experiments from 23 studies; however, not all studies could be used for both of the outcome variables (Table 1).

Model Development

Separate analyses were performed for FD and FP as the dependent variables. Independent variables in the meta-analysis included enrichment, housing system, beak trimming, bird strain, production period at which the enrichment started, and age at which the outcome

Table 2. Scoring scale transformations for feather damage (FD). Within the original scoring scale, we indicate if the higher value represents a better or worse feather cover. The number of body areas that were assessed per bird and the formula used to transform the original score (FD_{scale}) onto the 1–4 (best) scale while adjusting for one body area (FD_{trans}) are shown.

Publication	Original scale	No. of body areas	Transformation
Alm et al. (2017)	1–4 (best)	6	$FD_{trans} = FD_{scale}/6$
Campbell et al. (2018)	0–1 (worst)	1[1]	$FD_{trans} = -3 \times (FD_{scale}/1) + 4$
Cronin et al. (2018)	0–1 (worst)	7	$FD_{trans} = -3 \times (FD_{scale}/7) + 4$
Daigle et al. (2014)	0–5 (worst)	11	$FD_{trans} = -0.6 \times (FD_{scale}/11) + 4$
Decina et al. (2019)	0–2 (worst)	1[2]	$FD_{trans} = -1.5 \times (FD_{scale}/1) + 4$
Hartcher et al. (2015)	0–4 (worst)	1[1]	$FD_{trans} = -0.75 \times (FD_{scale}/1) + 4$
Ito et al. (2002)	0–3 (worst)	5	$FD_{trans} = -1 \times (FD_{scale}/5) + 4$
Johannson et al. (2016)	1–4 (best)	5	$FD_{trans} = FD_{scale}/5$
Liebers et al. (2019)	1–4 (best)	3	$FD_{trans} = FD_{scale}/3$
McAdie et al. (2005)	0–5 (worst)	10	$FD_{trans} = -0.6 \times (FD_{scale}/10) + 4$
Mielenz et al. (2010)	0–6 (worst)	1[2]	$FD_{trans} = -0.5 \times (FD_{scale}/1) + 4$
Morrissey et al. (2016)	0–1 (worst)	1[1]	$FD_{trans} = -3 \times (FD_{scale}/1) + 4$
Norgaard-Nielsen et al. (1993)	1–4 (best)	5	$FD_{trans} = FD_{scale}/5$
Pettersson et al. (2017)	1–4 (best)	5	$FD_{trans} = FD_{scale}/5$
Schmidt et al. (2019)	0–4 (worst)	1[2]	$FD_{trans} = -0.75 \times (FD_{scale}/1) + 4$
Steenfeldt et al. (2007)	1–4 (best)	5	$FD_{trans} = FD_{scale}/5$
Tahamtani et al. (2016)	0–2 (worst)	1[1]	$FD_{trans} = -1.5 \times (FD_{scale}/1) + 4$
Wechsler et al. (1998)	1–4 (worst)	6	$FD_{trans} = -1 \times (FD_{scale}/6) + 5$

[1]Presented as presence or absence–area not specified.
[2]Back/rump/tail area value extracted.

was measured. All statistical procedures were conducted using SAS Studio (SAS Inst. Inc., Cary, NC). Initial data exploration was conducted to calculate descriptive means and SE (PROC MEANS) and visually examine normality of the data (PROC UNIVARIATE) for continuous variables. Frequency tables (PROC FREQ) were used for categorical variables to determine the number of observations within each class (Table 3). Owing to the limited and unbalanced nature of the data, the enrichment treatment was condensed into a binary (yes vs. no) variable, and strain was omitted from the analysis; a proxy in the form of feather color (brown vs. white colored birds) was explored instead.

Associations between categorical independent variables were assessed using chi-square tests (PROC FREQ). Variables associated with each other were not included within the same model. Spearman rank correlations between the outcome measures and the age of the birds were also evaluated (PROC CORR).

Linear mixed models were developed with experiment as a random effect (St-Pierre, 2001; Sauvant et al., 2008). Only variables that may have influenced the outcome measure in univariate analysis (arbitrary liberal $P < 0.30$) were considered for multivariate equation development to reduce the potential of over-fitting models to the data (Dohoo et al., 2009). The period in which the enrichment was started (rearing vs. laying period) and feather color (brown vs. white colored) did not meet the selection criteria of $P < 0.30$ for both FD and FP and were thus excluded from further analysis. Beak trimming only met the selection criteria for the FD outcome and not for the FP models. In addition, beak trimming is not thought to biologically influence the actual FP behavior, instead exerting its effects because of the potential damage to the feather cover birds with intact or trimmed beaks can do (Nicol, 2018). Therefore, beak trimming was only considered as a biologically relevant dependent variable for FD. The main factor of enrichment was included in the model, and remaining variables (i.e., age of birds, housing, or beak trimming) were added one by one using a forward selection approach. Variables that were significant at $P < 0.05$ were retained in the model(s) developed. Interactions (enrichment × housing, enrichment × beak trimming, enrichment × age of birds) were explored, but interactions were omitted from the final models because of limited or unbalanced data or nonsignificance of interactions. Models could not be weighed to account for heterogeneous errors and differences in accuracy across studies (St-Pierre, 2001) because of insufficient data (SE or SD) being available for this step.

The assumptions of normally distributed residuals and homogeneity of variance were examined graphically with the use of conditional studentized residual plots. Normality and homogeneity of random effects were evaluated visually using histograms and Q-Q plots. The optimal variance-covariance matrix structure was evaluated by comparing the corrected Akaike information criterion (AICc). Influential

Table 3. Descriptive statistics of the dependent variables (FP: feather pecking, FD: feather damage), continuous independent variables, and categorical independent variables in the database.

Variable	N[1]	Mean (SD)	SE	Median	Min	Max
Dependent variables						
FP (pecks/bird/min)	110	0.03 (0.063)	0.006	0.01	0.0	0.60
FD (1–4 scale)	114	3.1 (0.75)	0.07	3.3	1.4	4.0
Continuous independent variables						
Age at FP (week)	126	21.6 (11.44)	1.02	23.0	1.0	53.0
Age at FD (week)	118	40.8 (14.57)	1.34	40.0	10.0	72.0
Categorical independent variables						
Housing type						
Cage	60 (28.6%)					
Noncage	150 (71.4%)					
Beak trimming[2]						
No	122 (61.6%)					
Yes	74 (37.4%)					
Both[3]	2 (1.0%)					
Enrichment						
No	91 (43.3%)					
Yes	119 (56.7%)					
Enrichment period						
Laying	137 (65.2%)					
Rearing	73 (34.8%)					
Feather color						
Brown	76 (36.2%)					
White	118 (56.2%)					
Silver	8 (3.8%)					
Combination[3]	8 (3.8%)					

[1]N is total number of observations for which information was available. Presented as N (% of observations) for categorical variables.

[2]Note that not all studies specified beak trimming status and therefore N does not equal 210 observations.

[3]Data were not presented separately for the different categories.

points were investigated using the Cook's distance test within PROC MIXED and removed when necessary. Statistical significance was considered at $P < 0.05$, and tendencies are reported when $0.05 \leq P \leq 0.1$.

Model Evaluation

Models developed during the analysis were evaluated to assess the precision and accuracy of the predictions (Tedeschi, 2006). Owing to the limited size of the data set, no independent evaluation was performed, but rather the model was evaluated back on the developmental data set. The first evaluation was performed by calculating the mean square prediction error (**MSPE**) as

$$MSPE = \sum_{i=1}^{n} O_i - P_i^2 \Big/ n \quad [1]$$

where O_i is the observed value of the i^{th} observation, P_i is the predicted value of the i^{th} observation, and n is the total number of observations. The square root of the MSPE (**RMSPE**) was subsequently expressed as the percentage of the observed mean to provide a metric of the overall prediction error (Bibby and Toutenburg, 1977).

$$RMSPE = \frac{\sqrt{MSPE}}{\overline{O}} x100\% \quad [2]$$

RMSPE was decomposed into the overall bias error (**ECT**), regression slope deviation (**ER**), and error due to "random" disturbance (**ED**) (Bibby and Toutenburg, 1977).

$$ECT = (\overline{P} - \overline{O})^2 \quad [3]$$

$$ER = (s_p - R\ x\ s_o)^2 \quad [4]$$

$$ED = (1 - R^2)x\ s_o{}^2 \quad [5]$$

where $\overline{O}$ is the observed mean, $\overline{P}$ is the predicted mean, s_o and s_p are the observed and predicted standard deviations, respectively, and R is the Pearson correlation coefficient between the observed and predicted values.

A second evaluation in the form of the concordance correlation coefficient (**CCC**) was calculated as per Lin (1989). The CCC value ranges from -1 to +1, with -1 indicating that the predicted and observed values are in perfect disagreement, 0 indicating that there is no relationship, and +1 indicating that they are in perfect agreement. The formula for CCC can be expressed as

$$CCC = R\ x\ C_b \quad [6]$$

where R is the Pearson correlation coefficient as a measure of precision, and C_b is the bias correction factor as a measure of accuracy (Lin, 1989; Tedeschi, 2006). C_b is calculated as

$$C_b = \frac{2}{\left[v + \frac{1}{v} + \mu^2\right]} \quad [7]$$

$$\text{with } v = \frac{s_p}{s_o} \quad [8]$$

$$\text{and } \mu = \frac{\overline{O} - \overline{P}}{\left(s_{o\ x\ s_p}\right)^{1/2}} \quad [9]$$

where v indicates a measure of scale shift, where a value > 1 indicates that less variance is explained by the prediction model than is observed (ideal value: 1), and μ is a measure of location shift where a positive value signals under prediction and a negative value signals over prediction.

The predicted vs. observed plot and conditional residual (predicted-observed) vs. predicted plot were also visually examined for patterns. The intercept of the predicted vs. observed plot and the slope of the conditional residual vs. predicted plot were tested for significant difference from zero using PROC REG to determine mean and slope bias in the conditional residuals, respectively.

RESULTS

Feather Pecking Model and Evaluation

Four prediction equation models were developed for FP (Table 4). Model evaluation was based on AICc, RMSPE, and CCC (Table 5) and residual plots shown in Figure 2. The final model for FP (Model FP3) used a lognormal distribution and included enrichment, housing, and age as fixed effects (PROC GLIMMIX), with experiment as a random effect using an unstructured variance-covariance matrix. In general, there was a higher frequency of FP in birds without access to enrichment and in those kept in a cage housing system (Table 5, Model FP1-4), and the frequency increased with age (Table 4, Model FP2-4). A significant interaction was observed between enrichment and housing system (Model FP4), indicating that the highest level of FP was present in flocks without enrichment that were kept in cage systems. However, the residuals for the random effect did not follow a normal distribution. Furthermore, the interaction was unbalanced (8 observations and 16 observations with no enrichment and enrichment in cage housing, respectively, vs. 38 and 48 observations with no enrichment and enrichment in noncage housing, respectively), which is also reflected in the high SE, and differences observed between the 4 treatment combinations which are difficult to interpret (Table 4). As Model FP3 (no interaction included) and Model FP4

Table 4. Backtransformed least square means (LSM ± SE) for models developed to quantify the effect of environmental enrichment on feather pecking (FP) in laying hens (pecks/bird/min).

	Model FP1		Model FP2		**Model FP3**[1]		Model FP4[1]	
Variable	LSM	*P*	*LSM*	*P*	*LSM*	*P*	*LSM*	*P*
Enrichment		<0.001		<0.001		<0.001		<0.001
No	0.03 ± 0.008		0.03 ± 0.007		0.04 ± 0.009		0.05 ± 0.012	
Yes	0.01 ± 0.003		0.01 ± 0.003		0.02 ± 0.003		0.01 ± 0.003	
Age[2] (week)	NI		0.04 ± 0.012	0.001	0.04 ± 0.010	0.001	0.04 ± 0.010	0.001
Housing	NI		NI			0.002		0.001
Cage					0.05 ± 0.017		0.06 ± 0.020	
Noncage					0.01 ± 0.003		0.01 ± 0.003	
Enrichment × Housing interaction	NI		NI		NI			0.001
No—Cage							0.141 ± 0.0558[a]	
Yes—Cage							0.023 ± 0.0082[b,c]	
No—Noncage							0.020 ± 0.0043[b]	
Yes—Noncage							0.010 ± 0.0021[c]	

Final model is bolded (Model FP3). Variables with a *P*-value < 0.05 were retained in the final model.
Abbreviation: NI, not included in the model.
[1]1 outlier was removed from the model.
[2]Estimate parameters are on the untransformed lognormal scale.

(interaction included) did not differ much in the further model evaluation analysis, such as the RMSP and CCC (Table 5), Model FP3 was considered as the final model.

Feather Damage Model and Evaluation

Five prediction equation models were developed for FD (Tables 6 and 7) with different combinations of

Table 5. Evaluation of model equations for feather pecking (FP) in laying hens.

Evaluation parameter	Model FP1	Model FP2	**Model FP3**	Model FP4
N	106	106	105	105
AICc[1]	270.82	267.09	233.94	223.76
Mean ± SE[2]	−4.2 ± 0.06	−4.2 ± 0.07	−4.2 ± 0.07	−4.2 ± 0.08
SD[2]	0.65	0.69	0.74	0.77
RMSPE (%)[3]	−17.1	−16.5	−14.6	−13.7
ECT (%)[4]	0.0	0.0	0.0	0.0
ER (%)[5]	1.1	0.7	0.3	0.3
ED (%)[6]	98.9	99.3	99.7	99.7
CCC[7]	0.645	0.683	0.756	0.790
R[8]	0.712	0.736	0.787	0.814
C_b[9]	0.907	0.929	0.961	0.971
V[10]	1.567	1.476	1.331	1.278
μ[11]	0.000	0.000	0.000	0.000
Plots				
Slope[12]	0.11 ± 0.108	0.09 ± 0.098	0.05 ± 0.081	0.04 ± 0.073

Model evaluation included square root MSPE and CCC analysis as well as evaluation of predicted vs. observed and conditional residual vs. predicted plots. Independent variables included enrichment (Model FP1); enrichment and age (Model FP2); enrichment, age and housing (Model FP3—final model); enrichment, age, housing and enrichment × housing interaction (Model FP4).
Abbreviation: MSPE, mean square prediction error.
[1]Corrected akaike information criterion as a measure of goodness-of-fit.
[2]Mean, SE, and SD of predicted values on the lognormal scale. Observed mean ± SD were 0.03 ± 0.064 (FP1), 0.03 ± 0.064 (FP2), 0.03 ± 0.064 (FP3), and 0.03 ± 0.064 (FP4).
[3]Root mean square prediction error expressed as a percentage of the observed mean on the lognormal scale.
[4]Error due to bias expressed as a percentage of MSPE.
[5]Error due to regression slope deviation expressed as a percentage of MSPE.
[6]Error due to disturbance expressed as a percentage of MSPE.
[7]Concordance correlation coefficient calculated as R × C_b.
[8]Pearson correlation coefficient.
[9]Bias correction factor.
[10]Scale shift measure.
[11]Location shift measure.
[12]Slope of conditional residual vs. predicted regression as calculated in PROC REG. Values are presented as estimate ±SE and * indicates a significant difference from zero, *P* < 0.05.

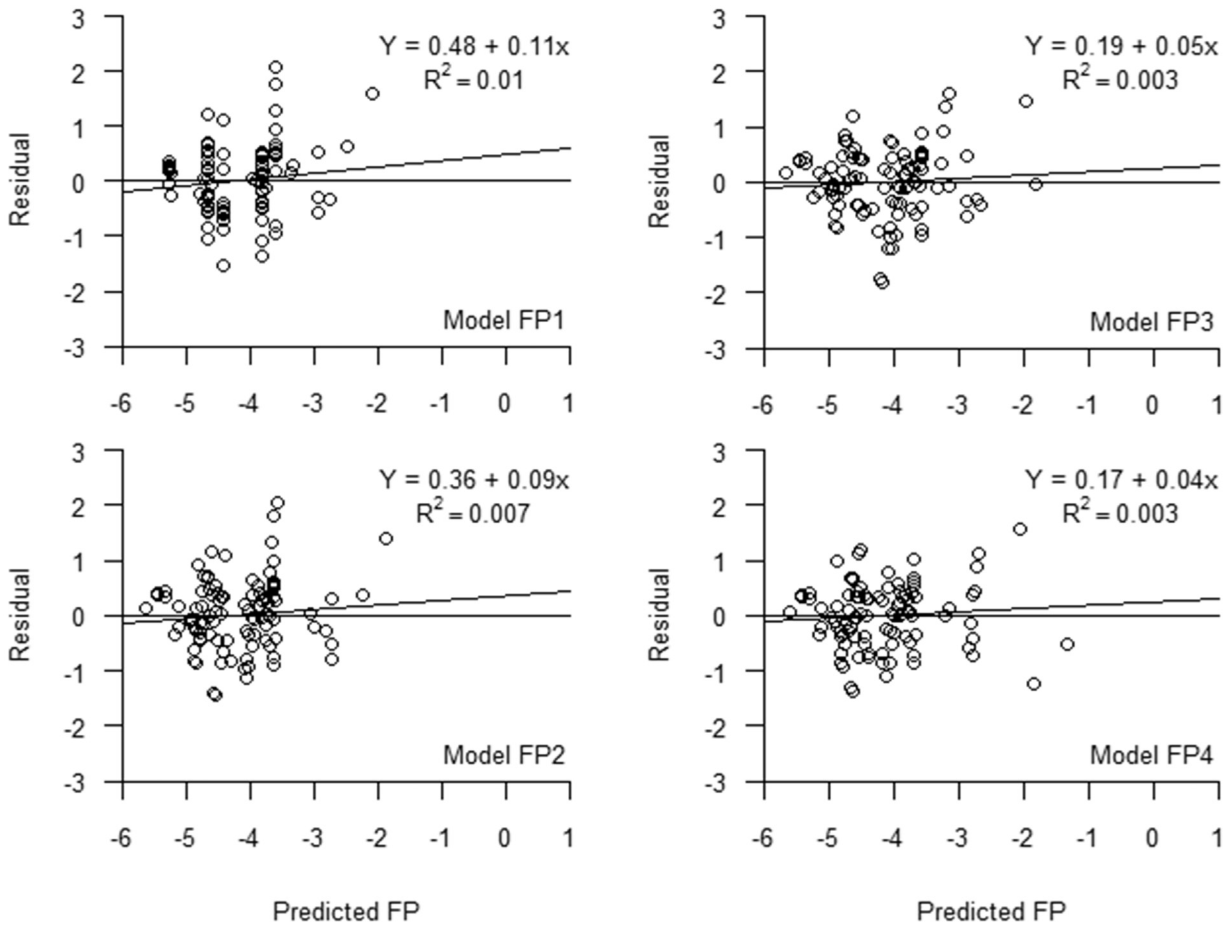

Figure 2. Conditional residual (predicted–observed) vs. predicted plots for the different model equations for feather pecking (FP, pecks/bird/min on the lognormal scale) in laying hens. Independent variables included enrichment (Model FP1); enrichment and age (Model FP2); enrichment, age, and housing (Model FP3—final model); enrichment, age, housing and enrichment × housing interaction (Model FP4).

independent variables, including environmental enrichment, age, beak trimming status, and housing type. Model evaluation was based on AICc, RMSPE, and CCC (Table 8) and residual plots shown in Figure 3. The final model for FD (Model FD5) followed a Gaussian distribution and included enrichment, housing, beak trimming, and age as fixed effects (PROC MIXED), experiment as a random effect, and used an unstructured variance-covariance matrix structure. Negative parameter estimates for the FD score were found when no enrichment was provided, birds had an increased age, birds were nonbeak trimmed,

Table 6. Parameter estimates for the model equations developed to quantify the effect of environmental enrichment on feather damage (FD).

	Model FD1		Model FD2		Model FD3		Model FD4		**Model FD5**[1]	
Variable	Estimate	*P*	Estimate	*P*	Estimate	*P*	Estimate	*P*	Estimate	*P*
Intercept	3.18 ± 0.151	<0.001	4.17 ± 0.176	<0.001	4.40 ± 0.196	<0.001	4.29 ± 0.192	<0.001	4.59 ± 0.217	<0.001
Enrichment										
No	−10.02 ± 0.08	0.786	−0.02 ± 0.08	0.057	−0.13 ± 0.06	0.046	−0.13 ± 0.07	0.059	−0.14 ± 0.06	0.018
Yes	–		–		–		–		–	
Age (week)	NI		−0.03 ± 0.003	<0.001	−0.03 ± 0.003	<0.001	−0.03 ± 0.003	<0.001	−0.02 ± 0.003	<0.001
Beak trim	NI		NI				NI			
No					−0.49 ± 0.18	0.008			−0.57 ± 0.17	0.001
Yes					–				–	
Housing	NI		NI		NI					
Cage							−0.36 ± 0.25	0.176	−0.57 ± 0.26	0.042
Noncage							–		–	

A higher feather damage score (range: 1–4) indicates a better feather cover. Final model is bolded (Model FD5). Variables with a *P*-value < 0.05 were retained in the final model.

Abbreviation: NI, not included in the model.

[1]3 outliers were removed from Model FD5 which was considered the final model.

Table 7. Least square means (LSM ± SE) for models developed to quantify the effect of environmental enrichment on feather damage (FD) in laying hens.

	Model FD1		Model FD2		Model FD3		Model FD4		**Model FD5**[1]	
Variable	LSM	P	LSM	P	LSM	P	LSM	P	LSM	P
Enrichment		0.786		0.057		0.046		0.0587		0.018
No	3.2 ± 0.15		3.0 ± 0.13		3.0 ± 0.13		2.9 ± 0.13		2.9 ± 0.13	
Yes	3.2 ± 0.15		3.1 ± 0.13		3.1 ± 0.13		3.1 ± 0.13		3.0 ± 0.13	
Age (week)	NI		−0.03 ± 0.003	<0.001	−0.03 ± 0.003	<0.001	−0.03 ± 0.003	<0.001	−0.02 ± 0.003	<0.001
Beak trim	NI		NI			0.008	NI			0.001
No					2.8 ± 0.16				2.7 ± 0.16	
Yes					3.3 ± 0.16				3.2 ± 0.15	
Housing	NI		NI		NI			0.1763		0.042
Cage							2.8 ± 0.21		2.7 ± 0.21	
Noncage							3.2 ± 0.15		3.2 ± 0.15	

A higher feather damage score (range: 1–4) indicates a better feather cover. Final model is bolded (Model FD5). Variables with a P-value < 0.05 were retained in the final model.

Abbreviation: NI, not included in the model

[1]3 outliers were removed from Model FD5 which was considered the final model.

and cage housing systems were used (Table 4). This is also represented through the lower least square mean values (Table 7), indicating a worsening of the feather cover. It should be noted that enrichment on its own did not significantly explain the variance observed in the FD score (Model FD1, P = 0.786). The remaining variables had to be added in order for enrichment to become a significant variable. Furthermore, housing system was not significant (Model FD4) until included together with beak trimming in the final model (Model FD5). The addition of all significant variables in Model FD5 resulted in slight improvements in the RMSPE and CCC (Table 8). The majority of the RMSPE was attributed to the ED in all models. The CCC indicated a close to perfect relationship between predicted and observed FD values. Model FD5 explained nearly all variance as observed by the scale shift being close to 1 and extremely low location shift indicating very little

Table 8. Evaluation of model equations for feather damage (FD) in laying hens.

Evaluation parameter	Model FD1	Model FD2	Model FD3	Model FD4	**Model FD5**
N	114	108	108	108	105
AICc[1]	181.7	127.2	121.2	126.2	104.7
Mean ± SE[2]	3.1 ± 0.056	3.0 ± 0.065	3.0 ± 0.066	3.0 ± 0.065	3.0 ± 0.068
SD[2]	0.6001	0.6757	0.6840	0.6770	0.6970
RMSPE (%)[3]	12.8	9.9	9.4	9.9	8.7
ECT (%)[4]	0.0	0.0	0.0	0.0	0.0
ER (%)[5]	1.2	0.6	0.5	0.5	0.4
ED (%)[6]	98.8	99.4	99.5	99.5	99.6
CCC[7]	0.834	0.913	0.922	0.913	0.934
R[8]	0.855	0.919	0.927	0.919	0.938
C_b[9]	0.975	0.993	0.995	0.993	0.996
V[10]	1.252	1.124	1.111	1.122	1.092
μ[11]	0.000	0.000	0.000	0.000	0.000
Plots					
Intercept[12]	0.97 ± 0.12*	0.55 ± 0.11*	0.50 ± 0.10*	0.55 ± 0.11*	0.43 ± 0.10*
Slope[13]	0.07 ± 0.06	0.03 ± 0.04	0.03 ± 0.04	0.03 ± 0.04	0.02 ± 0.04

Model evaluation included square root MSPE and CCC analysis as well as evaluation of predicted vs. observed and conditional residual vs. predicted plots. Independent variables included enrichment (Model FD1); enrichment and age (Model FD2); enrichment, age, and beak trimming (Model FD3); enrichment, age, and housing (Model FD4); enrichment, age, beak trimming, and housing (Model FD5—final model).

Abbreviation: MSPE, mean square prediction error.

[1]Akaike information criterion as a measure of goodness-of-fit.

[2]Mean, SE, and SD of predicted values.

[3]Root mean square prediction error expressed as a percentage of the observed mean.

[4]Error due to bias expressed as a percentage of MSPE.

[5]Error due to regression slope deviation expressed as a percentage of MSPE.

[6]Error due to disturbance expressed as a percentage of MSPE.

[7]Concordance correlation coefficient calculated as R × C_b.

[8]Pearson correlation coefficient.

[9]Bias correction factor.

[10]Scale shift measure.

[11]Location shift measure.

[12]Intercept of predicted vs. observed regression as calculated in PROC REG. Values are presented as estimate ±SE and * indicates a significant difference from zero.

[13]Slope of conditional residual vs. predicted regression as calculated in PROC REG. Values are presented as estimate ±SE and * indicates a significant difference from zero, P < 0.05.

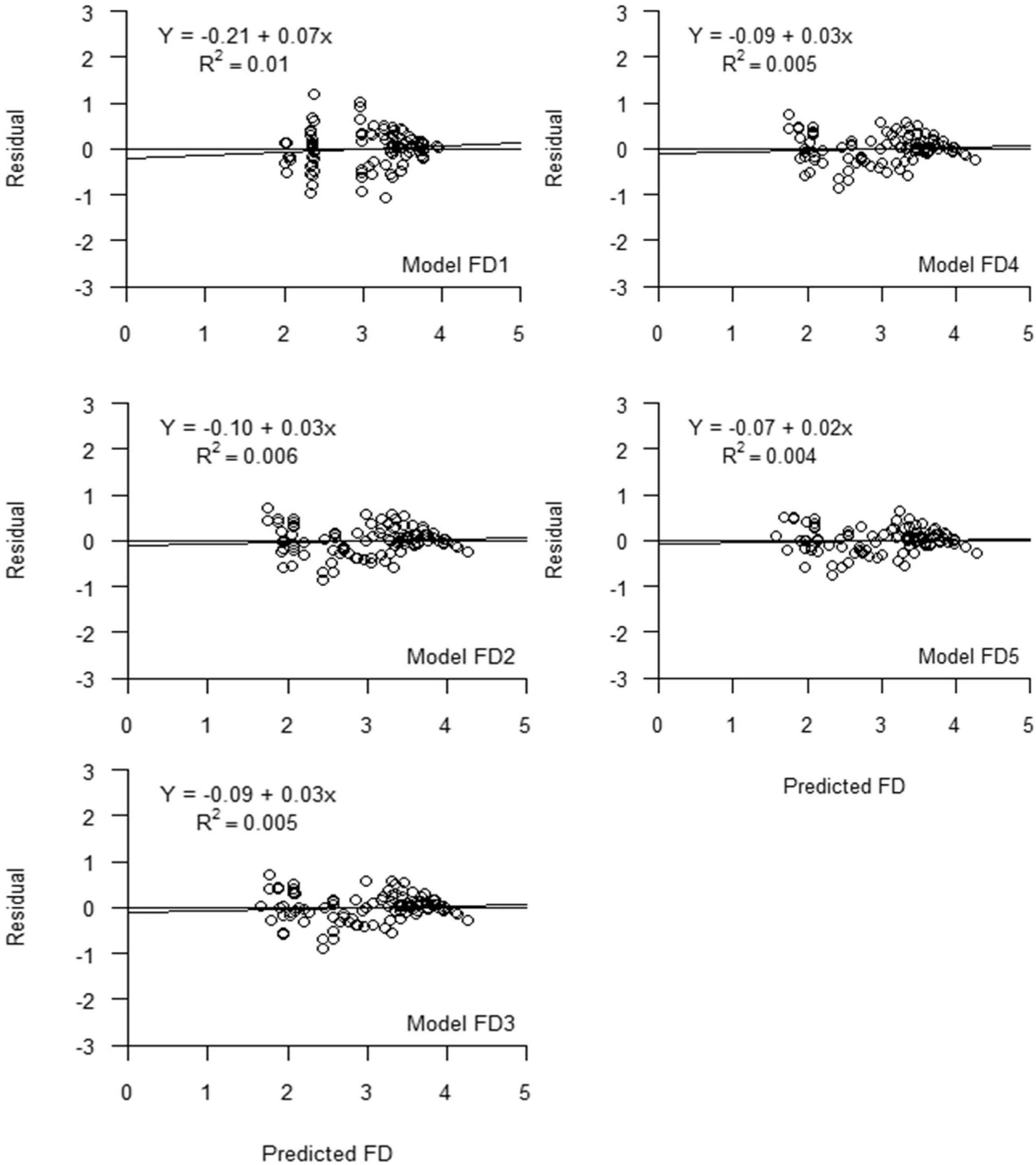

Figure 3. Conditional residual (predicted–observed) vs. predicted plots for the different model equations for feather damage (FD, range: 1–4) in laying hens. Independent variables included enrichment (Model FD1); enrichment and age (Model FD2); enrichment, age, and beak trimming (Model FD3); enrichment, age, and housing (Model FD4); enrichment, age, beak trimming, and housing (Model FD5—final model).

under prediction. Importantly, the model fit of Model FD5 was distinctly improved compared with the other models based on the AICc (Table 8). The residual vs. predicted plot for Model FD1 shows a pattern that is reduced when additional variables are included in the model (Model FD2-5). The slope of the residual vs. predicted plots was not significantly different from zero ($P > 0.05$; Figure 3 and Table 8), indicating that predicted and observed values did not differ from one another and lie close to the line of unity, as also indicated by all C_b values being close to one (Table 8).

DISCUSSION

To the authors' knowledge, only few previous meta-analyses have been conducted in regard to poultry welfare and have focused on mortality in different housing systems (Weeks et al., 2016) or the effect of feed and water deprivation posthatch in broilers (de Jong et al., 2017). This study provides the first meta-analysis on the impact of enrichments on FP and FD. The need for a meta-analysis on FP and subsequent FD had been highlighted in previous studies (Kjaer et al., 2011; Freire and Cowling, 2013).

Effects of Enrichment on Feather Pecking and Feather Damage

The provision of enrichment was significantly associated with lower FP. The frequency of FP was approx. 2 times higher in flocks without enrichment (Table 4). Furthermore, lack of enrichment was significantly associated with a lower FD score (Table 7), indicating a worse feather cover. The decrease in FD score when no enrichment was provided was relatively small (-0.14 ± 0.06, Table 6); however, this is a 4.7% change when considering that the score can range from 1 to 4. Additionally, it should be noted that by condensing the FD score (e.g., a range of 11–55 being fit into the 1–4 range), the differences between treatment means are indeed on a smaller scale. This provides strong evidence that it is an effective measure to reduce FP and subsequently FD.

It should be noted that the enrichment provided in the different studies was diverse (Table 1). Studies provided a combination of different enrichment types (i.e., objects, foraging, dustbathing materials) within 1 treatment, switched enrichment types, or confounded enrichment type with the housing system. Even within an enrichment type, often multiple forms of enrichment were provided (different objects together, e.g., plastic caps, wind chimes, strings within the object type; or different foraging substrates together, e.g., pecking stones, lucerne bales within the foraging type). This large variety in provided enrichment forced us to consider enrichment as a binary yes or no variable. While this showed a clear result that enrichment per se is capable of reducing FP and FD, it limited the possibility of sophisticated further analyses to truly elucidate if the type of enrichment plays a role. Types of enrichment that give birds opportunities to forage, and to a lesser extent dustbathe, are thought to be more effective in reducing FP (Blokhuis, 1986; Vestergaard and Lisborg, 1993; Rodenburg et al., 2013). Most of the studies provided foraging opportunities as enrichment (16 studies), wheres only 2 studies provided additional dustbathing opportunities as enrichment. In contrast, objects which are often used (10 studies) are considered less effective in reducing FP or FD, though they might have benefits in reducing fearfulness which is also linked to FP (Campbell et al., 2019). Dixon et al. (2010) indeed found that foraging material was most effective in reducing FP, though dustbathing material and objects also reduced FP compared with the control group. They suggested that the environmental enrichment could work by occupying the birds' time or having some stress-reducing effects (Dixon et al., 2010). However, this was the only study that compared all types of enrichment within the same group of birds with a Latin square design while most other studies used factorial designs. Further work is needed to investigate how the type of enrichment influences the FP and FD outcomes. Other remaining questions regarding environmental enrichment for laying hens were highlighted by Schreiter et al. (2019), for example whether enrichment should be provided preventatively or curatively. The majority of studies included in this meta-analysis did not provide enrichment preventatively (only 12% of observations), and this factor could therefore not be further analyzed.

The review by Schreiter et al. (2019) emphasized that enrichment was shown to be beneficial in cage systems, but less clear effects were observed in noncage systems. The current meta-analysis did not find a significant interaction between enrichment and housing type for FD (Tables 6 and 7), suggesting that the effect of enrichment is, in fact, similar in cage and noncage housing systems. In terms of FP analysis, there was some evidence that pointed toward a significant interaction (Table 4), which indeed showed a stronger effect of enrichment within cage systems as opposed to noncage systems. However, owing to the issue of unbalanced data (i.e., unbalanced number of observations in the treatment groups), this interaction was omitted from the final model. Nevertheless, it would be interesting to reconsider this aspect when more detailed analysis regarding enrichment provision becomes possible. Furthermore, we had intended to explore responses over time, but only 4 (FP) and 6 (FD) studies reported true repeated observations, whereas the majority either measured at one time-point or did not present values for each time-point separately. Consequently, there is room to improve this meta-analysis in the future when additional data become available, especially in terms of type of enrichment, method of enrichment provision, and short- and long-term effects.

While the positive effect of enrichment on reducing FP and FD was found in the current meta-analysis, its impact can be considered relatively small. We should, therefore, consider the possible additive effect of providing enrichment as one of multiple, effective management strategies (Bracke et al., 2004; Lambton et al., 2013). Lambton et al. (2013) found correlations indicating lower levels of FP, FD, and mortality at 40 wk of age in flocks where farmers used more of the suggested management strategies. While not providing details on the strength of these correlations, an average of 19 out of 46 possible management strategies were used in the flocks (Lambton et al., 2013), showing the need for multiple strategies. The provision of enrichment materials is a relatively simple strategy in terms of feasibility and on-going commitment which could influence uptake by farmers (Lambton et al., 2013). The results of the current meta-analysis can also sway farmers who indicate that a lack of guarantee that strategies, such as enrichment, would be effective, was a main barrier to implement strategies (Palczynski et al., 2016).

Effects of Age, Housing System, and Beak Trimming on Feather Pecking and Feather Damage

While not the main factor of interest in this meta-analysis, age, housing system, and beak trimming status were needed to explain the variance observed in FP frequency and FD score. Feather pecking frequency increased with each 1-week increase in *age* of the birds (log estimate 0.04 ± 0.001, Table 4). Similarly, a 0.02 (±0.003) reduction in FD score or worsening of the feather cover was observed as birds increased in age (Model FD5, Tables 6 and 7). This is consistent with the literature which supports the fact that FP and FD increase as flocks get older (Bilčík and Keeling, 1999; Huber-Eicher and Sebö, 2001; Lambton et al., 2013; Decina et al., 2019). If we assume a linear relationship, our results suggest that the FD score of a bird would decrease by one full point in a 50-week period. Laying hens are typically kept in production up to 70 to 72 wk of age (Pelletier, 2017), though it can be longer (Bain et al., 2016; Fernyhough et al., 2020). We should, however, caution that the relationship of worsening feather cover over time might not be linear, as assumed in this extrapolation. Additionally, it is extremely likely that other corroborating factors will further reduce the quality of the feather cover because of the multifactorial nature of FP (van Staaveren and Harlander-Matauschek, 2020). Regardless, our results suggest that birds will show more FP and have worse feather cover toward the end of production.

Beak trimming and noncage housing both reduced FD score to a larger extent than enrichment and age, and, interestingly, the extent with which they reduced FD was similar (Table 6). It also should be noted that *beak trimming* was more common in cage systems, likely because of noncage systems often involving free-range flocks, where beak trimming often is not allowed. With beak trimming being more common in cage systems, it could be expected that cage systems would actually be associated with a higher FD score and better feather cover. This, however, was not the case, suggesting that housing system and beak trimming both have their own effect on FD.

The finding that beak trimming (−0.57 ± 0.17, Table 6) was associated with a better feather cover is not unexpected, as this practice minimizes the damage birds can do to one another (Nicol, 2018). However, the practice does not eliminate the behavior (beak trimming did not significantly affect FP in the univariate analysis, data not shown). Furthermore, beak trimming is under increasing scrutiny because of ethical and societal concerns and is being banned or phased out in various countries (Nicol et al., 2013; Rollin, 2015; Nicol, 2018; van Staaveren and Harlander-Matauschek, 2020). As such, it is likely that beak trimming will no longer be available in the future. This further highlights the need for effective and socially acceptable management strategies that can counter the larger damage that is possible in flocks with intact beaks.

We found that *housing system* influenced the feather cover, with the FD score being lower in birds kept in cage systems compared with noncage systems (−0.57 ± 0.26, Table 6). A previous review of literature using a vote counting approach similarly found that feather condition was worse in cages compared with alternative systems (Freire and Cowling, 2013). However, they compared conventional cages against furnished cage, single-tier and multi-tier barn systems, and outdoor systems. No distinction was made between the different type of cage and noncage housing systems within the current analysis. Likely, the lower FD or worse feather cover in cage systems (Table 7) is a consequence of inherent differences between these systems. Noncage systems often provide some form of litter as a foraging substrate or outdoor access, which is considered the main protective factor as FP is theorized to be a form of redirected foraging (Blokhuis, 1986; Rodenburg et al., 2013). This would also explain the 5 times lower frequency of FP observed in noncage systems compared with cage systems (Table 4). However, at the same time, it is often suggested that noncage housing systems are associated with increased risk and difficulty of FP outbreaks because of larger group sizes (Nicol et al., 1999; Bilčík and Keeling, 2000). The current results seem to suggest the contrary, although the developmental data set was not completely balanced for housing systems. There were less observations within cage systems which likely also explains the slightly higher standard error for its estimate (Tables 4, 6, and 7). Differences between housing systems in terms of FP or FD are not always identified (Sherwin et al., 2010; Decina, 2018; van Staaveren and Harlander-Matauschek, 2020), suggesting that the housing system management is more influential in affecting the FP or FD within a flock rather than the housing system per se.

Feather Pecking Frequency and Feather Damage Score Within the Database

The FP frequency was expressed as the number of pecks per bird per min, which was generally low, but extrapolating it would mean that a bird pecks on average 1.8 times per hour (Table 3). Previously reported frequencies ranged between 0.5 and 30.2 pecks per hour (LayWel, 2006). However, comparing the frequency of pecks reported in different studies is difficult because of different definitions, methods, housing system, breeds, and bird ages involved. Within the current meta-analysis, FP was more often measured in young birds (average 21 wk of age), which could explain the low frequency. Furthermore, it likely was more a representation of different forms of bird-to-bird pecking, rather than FP per se. Future analysis should attempt to tease apart the different forms of FP and other forms of bird-to-bird pecking (van Staaveren and Harlander-Matauschek, 2020).

The converted FD score in the experiments was on average 3.1 (range: 1.4–4.0, Table 3). This implies that most studies had birds within the upper range of the scale with limited FD, though some observations reached the lower range with severe FD. It also highlighted that FD is not always observed or present even in studies that specifically aim to research this phenomenon (Campbell et al., 2018). It should be kept in mind that the current findings are therefore only applicable within the current observed range and area of inference (Cooper et al., 2009). The number of observations toward the lower range of the scoring scale was smaller, with the majority of observations falling within the higher range (50% of observations had FD > 3.3). Furthermore, the model was evaluated back on the developmental database, and independent evaluation is needed, preferably with a larger inclusion of low FD observations. With more standardized reporting of FD and FP in the field, this may be possible.

Standardization of Feather Pecking and Feather Damage Outcomes

The issue of standardization because of different scoring scales and limited data reported had previously prevented the use of a meta-analysis approach (Kjaer et al., 2011; Freire and Cowling, 2013). Indeed, both FP and FD data were presented in different ways in literature which required thorough standardization. While FP data were generally easier to standardize (in our case, to number of pecks per bird per min), some studies had to be omitted as they presented number or percentage of birds or the percentage of time spent on the behavior. Additionally, caution should be taking in the interpretation because of different definitions of forms of bird-to-bird pecking and presentation of these forms together rather than separate values for FP in several studies. More complicated was the standardization of at least 8 different FD scoring scales that were used on 1–11 body areas. Subsequently, FD data were presented per different area, as a total sum, or a total averaged score. To allow proper comparison between the studies, we used a similar approach to Bedere et al. (2018), assuming a linear relationship between the different scoring scales to convert them to the recognized 1–4 scale developed by Tauson et al. (2005). For discrete scores, such as for FD, the calculation of means or sums should be avoided (Botreau et al., 2007); therefore, we adjusted it to one body area to retain the meaning of the 1–4 score (Table 2). As such, the values of FD calculated for the current dataset can perhaps be best considered to more represent a continuous visual analog scale. Visual analogue scales are suggested to increase observer reliability, considered potentially more sensitive, and they allow for different processing and analyses than categorical scales (Tuyttens et al., 2009). Visual analog scales have been developed to assess for example, lameness in cows and pigs (Tuyttens et al., 2009; Nalon et al., 2014), keel bone fractures in laying hens (Rufener et al., 2018), and behavior descriptors (Welfare Quality, 2009), but to the authors' knowledge, none are developed for FD. While the conversion allowed us to standardize the FD score across studies, it could be discussed whether this adjustment was the best method to use or whether more appropriate adjustments are available. However, until consistency within research is reached and the same scoring scale is used across studies, some sort of calculation will be needed for it to be possible to perform a meta-analysis on this kind of outcome. Furthermore, measures of variance (i.e., SD or SE) were infrequently reported, further limiting the meta-analysis as also noted by Freire and Cowling (2013). Therefore, we were unable to perform any weighting of the observations according to their precision, as suggested by St Pierre (2001).

In conclusion, environmental enrichment is provided in various forms in laying hen production. Following standardization of FP and FD outcomes, this meta-analysis presents the first evidence that providing environmental enrichment per se is capable of reducing FP and limiting FD in laying hen flocks. The analyses of behavior and feather cover were in good agreement which further strengthens the argument for the benefit of environmental enrichment. The extent to which enrichment lowers FD, however, suggests that enrichment alone will not be able to eliminate FD completely. Further research is needed to determine the effect of different types of enrichment, method of enrichment provision, and short- or long-term effects. Environmental enrichment should be encouraged as part of a comprehensive approach involving multiple management strategies aimed at reducing FP and FD.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge all authors who provided clarification or additional data on their published papers that were used with this meta-analysis. This study was funded by the Natural Sciences and Engineering Research Council of Canada. The authors also would like to extend a heartfelt thank you to Neila Ben Sassi, Emily Leishman, Claire Mindus, Valerie Monckton, and Carolina Reyes Rodriguez for their valuable feedback and questions during the development of this manuscript.

DISCLOSURES

The authors declare no conflicts of interest

REFERENCES

Alm, M., R. Tauson, and H. Wall. 2017. Mussel shells as an environment enrichment and calcium source for floor-housed laying hens. J. Appl. Poult. Res. 26:159–167.

Bain, M. M., Y. Nys, and I. C. Dunn. 2016. Increasing persistency in lay and stabilising egg quality in longer laying cycles. What are the challenges? Br. Poult. Sci. 57:330–338.

Bedere, N., E. Cutullic, L. Delaby, F. Garcia-Launay, and C. Disenhaus. 2018. Meta-analysis of the relationships between reproduction, milk yield and body condition score in dairy cows. Livest. Sci. 210:73–84.

Bestman, M., and J.-P. Wagenaar. 2014. Health and welfare in Dutch Organic laying hens. Animals 4:374–390.
Bibby, J., and T. Toutenburg. 1977. Prediction and Improved Estimation in Linear Models. John Wiley & Sons, Chichester, UK.
Bilčík, B., and L. J. Keeling. 1999. Changes in feather condition in relation to feather pecking and aggressive behaviour in laying hens. Br. Poult. Sci. 40:444–451.
Bilčík, B., and L. J. Keeling. 2000. Relationship between feather pecking and ground pecking in laying hens and the effect of group size. Appl. Anim. Behav. Sci. 68:55–66.
Blokhuis, H. J. J. 1986. Feather-pecking in poultry: its relation with ground-pecking. Appl. Anim. Behav. Sci. 16:63–67.
Blokhuis, H. J., T. F. Van Niekerk, W. Bessei, A. Elson, D. Guemene, J. B. Kjaer, G. A. M. Levrino, C. J. Nicol, R. Tauson, C. A. Weeks, and H. A. de Weerd. 2007. The LayWel project: welfare implications of changes in production systems for laying hens. Worlds. Poult. Sci. J. 63:101–114.
Botreau, R., M. B. M. Bracke, P. Perny, A. Butterworth, J. Capdeville, C. G. Van Reenen, and I. Veissier. 2007. Aggregation of measures to produce an overall assessment of animal welfare. Part 2: analysis of constraints. Animal 1:1188–1197.
Bracke, M. B. M., B. Hulsegge, L. Keeling, and H. J. Blokhuis. 2004. Decision support system with semantic model to assess the risk of tail biting in pigs: 2. "Validation." Appl. Anim. Behav. Sci. 87:45–54.
Campbell, D. L. M., E. N. de Haas, and C. Lee. 2019. A review of environmental enrichment for laying hens during rearing in relation to their behavioral and physiological development. Poult. Sci. 98:9–28.
Campbell, D. L. M., G. N. Hinch, J. A. Downing, and C. Lee. 2018. Early enrichment in free-range laying hens: effects on ranging behaviour, welfare and response to stressors. Animal 12:575–584.
Chow, A., and J. A. Hogan. 2005. The development of feather pecking in Burmese red junglefowl: the influence of early experience with exploratory-rich environments. Appl. Anim. Behav. Sci. 93:283–294.
Cronin, G. M., R. L. Hopcroft, P. J. Groves, E. J. S. Hall, D. N. Phalen, and P. H. Hemsworth. 2018. Why did severe feather pecking and cannibalism outbreaks occur? An unintended case study while investigating the effects of forage and stress on pullets during rearing. Poult. Sci. 97:1484–1502.
Cooper, H. M., L. V. Hedges, and J. C. Valentine. 2009. The Handbook of Research Synthesis and Meta-Analysis. 2nd ed. Russell Sage Foundation, New York, NY.
Daigle, C. L., T. B. Rodenburg, J. E. Bolhuis, J. C. Swanson, and J. M. Siegford. 2014. Use of dynamic and rewarding environmental enrichment to alleviate feather pecking in non-cage laying hens. Appl. Anim. Behav. Sci. 161:75–85.
Dawkins, M. S. 1989. Time budgets in Red Junglefowl as a baseline for the assessment of welfare in domestic fowl. Appl. Anim. Behav. Sci. 24:77–80.
Decina C. 2018. An investigation of feather damage in Canadian laying hens. MSc Diss. Univ. Guelph, ON, Canada
Decina, C., O. Berke, N. van Staaveren, C. F. Baes, T. M. Widowski, and A. Harlander-Matauschek. 2019. A cross-sectional study on feather cover damage in Canadian laying hens in non-cage housing systems. BMC Vet. Res. 15:435.
de Jong, I. C., J. van Riel, M. B. M. Bracke, and H. van den Brand. 2017. A "meta-analysis" of effects of post-hatch food and water deprivation on development, performance and welfare of chickens. PLoS One 12:e0189350.
Dixon, L. M. M., I. J. H. J. H. Duncan, and G. Mason. 2008. What's in a peck? Using fixed action pattern morphology to identify the motivational basis of abnormal feather-pecking behaviour. Anim. Behav. 76:1035–1042.
Dixon, L. M., I. J. H. Duncan, and G. J. Mason. 2010. The effects of four types of enrichment on feather-pecking behaviour in laying hens housed in barren environments. Anim. Welf 19:429–435.
Dohoo, I., W. Martin, and H. Stryhn. 2009. Veterinary Epidemiologic Research. 2nd ed. VER Inc., Charlottetown, Prince Edward Island, Canada.
EFSA. 2005. Opinion of the Scientific Panel on Animal Health and Welfare on a request from the Commission related to the welfare aspects of various systems of keeping laying hens. EFSA J. 197:1–23.
Elkhoraibi, C., R. A. Blatchford, M. E. Pitesky, and J. A. Mench. 2014. Backyard chickens in the United States: a survey of flock owners. Poult. Sci. 93:2920–2931.
Fernyhough, M., C. J. Nicol, T. van de Braak, M. J. Toscano, and M. Tønnessen. 2020. The Ethics of laying hen genetics. J. Agric. Environ. Ethics 33:15–36.
Freire, R., and A. Cowling. 2013. The welfare of laying hens in conventional cages and alternative systems: first steps towards a quantitative comparison. Anim. Welf 22:57–65.
Green, L. E., K. Lewis, A. Kimpton, and C. J. Nicol. 2000. Cross-sectional study of the prevalence of feather pecking in laying hens in alternative systems and its associations with management and disease. Vet. Rec. 147:233–238.
Hartcher, K. M., K. T. N. Tran, S. J. Wilkinson, P. H. Hemsworth, P. C. Thomson, and G. M. Cronin. 2015. The effects of environmental enrichment and beak-trimming during the rearing period on subsequent feather damage due to feather-pecking in laying hens. Poult. Sci. 94:852–859.
Huber-Eicher, B., and F. Sebö. 2001. The prevalence of feather pecking and development in commercial flocks of laying hens. Appl. Anim. Behav. Sci. 74:223–231.
Huber-Eicher, B., and B. Wechsler. 1997. Feather pecking in domestic chicks: its relation to dustbathing and foraging. Anim. Behav. 54:757–768.
Hughes, B. O., and M. J. Gentle. 1995. Beak trimming of poultry - its implications for welfare. Worlds. Poult. Sci. J. 51:51–61.
IEC. 2015. Egg Industry Review 2015. International Egg Commission. Accessed Jan. 2019 http://www.internationalegg.com/wp-content/uploads/2015/08/AnnualReview_2015.pdf.
Ito, S., T. Tanaka, and T. Yoshimoto. 2002. Effects of an "enrichment feeder" on behavior and feather conditions of caged laying hens. Anim. Sci. J. 73:149–153.
Johannson, S. G., C. Raginski, K. Schwean-Lardner, and H. L. Classen. 2016. Providing laying hens in group-housed enriched cages with access to barley silage reduces aggressive and feather-pecking behaviour. Can. J. Anim. Sci. 96:161–171.
Jones, R. B., T. M. McAdie, C. McCorquodale, and L. J. Keeling. 2002. Pecking at other birds and at string enrichment devices by adult laying hens. Br. Poult. Sci. 43:337–343.
Jung, L., and U. Knierim. 2018. Are practice recommendations for the prevention of feather pecking in laying hens in non-cage systems in line with the results of experimental and epidemiological studies? Appl. Anim. Behav. Sci. 200:1–12.
Kjaer, J. B., H. Glawatz, B. Scholz, S. Rettenbacher, and R. Tauson. 2011. Reducing stress during welfare inspection: validation of a non-intrusive version of the LayWel plumage scoring system for laying hens. Br. Poult. Sci. 52:149–154.
Klein, T., E. Zeltner, and B. Huber-Eicher. 2000. Are genetic differences in foraging behaviour of laying hen chicks paralleled by hybrid-specific differences in feather pecking? Appl. Anim. Behav. Sci. 70:143–155.
Korte, S. M., G. Beuving, W. Ruesink, and H. J. Blokhuis. 1997. Plasma Catecholamine and Corticosterone levels during manual Restraint in chicks from a high and low feather pecking line of laying hens. Physiol. Behav 62:437–441.
Kruijt, J. P. 1964. Ontogeny of social behaviour in burmese red junglefowl (Gallus Gallus Spadiceus) bonnaterre. Behav Suppl. 12:1–201.
Kuenzel, W. J. 2007. Neurobiological basis of sensory perception: welfare implications of beak trimming. Poult. Sci. 86:1273–1282.
Lambton, S. L., C. J. Nicol, M. Friel, D. C. J. Main, J. L. McKinstry, C. M. Sherwin, J. Walton, and C. A. Weeks. 2013. A bespoke management package can reduce levels of injurious pecking in loose-housed laying hen flocks. Vet. Rec. 172:423.
LayWel. 2006. Welfare implications of changes in production systems for laying hens - Deliverable 4.2-Prevalence of feather pecking in various production systems. Accessed Jan. 2019. https://www.laywel.eu/web/pdf/deliverable%2042.pdf.
Liebers, C. J., A. Schwarzer, M. Erhard, P. Schmidt, and H. Louton. 2019. The influence of environmental enrichment and stocking density on the plumage and health conditions of laying hen pullets. Poult. Sci. 98:2474–2488.
Lin, L. I. 1989. A concordance correlation coefficient to evaluate Reproducibility. Biomatrics 45:255–268.

McAdie, T. M., L. J. Keeling, H. J. Blokhuis, and R. B. Jones. 2005. Reduction in feather pecking and improvement of feather condition with the presentation of a string device to chickens. Appl. Anim. Behav. Sci. 93:67–80.

Mielenz, N., J. Spilke, and E. von Borell. 2010. Analysis of ordered categorical data with threshold models exemplified by plumage damage scores from laying hens differing in their genotype and rearing environment. Poult. Sci. 89:2521–2534.

Moher, D., A. Liberati, J. Tetzlaff, D. G. Altman, and T. P. Group. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA Statement. PLoS Med. 6:e1000097.

Morrissey, K. L. H., S. Brocklehurst, L. Baker, T. M. Widowski, and V. Sandilands. 2016. Can non-beak treated hens be kept in commercial furnished cages? Exploring the effects of strain and Extra environmental enrichment on behaviour, feather cover, and mortality. Animals 6:17.

Nalon, E., D. Maes, S. Van Dongen, M. M. J. van Riet, G. P. J. Janssens, S. Millet, and F. A. M. Tuyttens. 2014. Comparison of the inter- and intra-observer repeatability of three gait-scoring scales for sows. Animal 8:650–659.

National Research Council. 2011. Guide for the Care and Use of Laboratory Animals. 8th ed. The National Academies Press, Washington, DC.

Newberry, R. C. 1995. Environmental enrichment: increasing the biological relevance of captive environments. Appl. Anim. Behav. Sci. 44:229–243.

Nicol, C. 2018. Feather pecking and cannibalism: Can we really stop beak trimming? In Advances in Poultry Welfare. J. A. Mench, ed. Woodhead Publishing, Duxford, UK.

Nicol, C. J., M. Bestman, A.-M. Gilani, E. N. De Haas, I. C. De Jong, S. Lambton, J. P. Wagenaar, C. A. Weeks, and T. B. Rodenburg. 2013. The prevention and control of feather pecking: application to commercial systems. Worlds. Poult. Sci. J. 69:775–788.

Nicol, C. J., N. G. Gregory, T. G. Knowles, I. D. Parkman, and L. J. Wilkins. 1999. Differential effects of increased stocking density, mediated by increased flock size, on feather pecking and aggression in laying hens. Appl. Anim. Behav. Sci. 65:137–152.

Norgaard-Nielsen, G., K. Vestergaard, and H. B. Simonsen. 1993. Effects of rearing experience and stimulus enrichment on feather damage in laying hens. Appl. Anim. Behav. Sci. 38:345–352.

Palczynski, L., H. Buller, S. Lambton, and C. Weeks. 2016. Farmer attitudes to injurious pecking in laying hens and to potential control strategies. Anim. Welf 25:29–38.

Pelletier, N. 2017. Life cycle assessment of Canadian egg products, with differentiation by hen housing system type. J. Clean. Prod. 152:167–180.

Pettersson, I. C., C. A. Weeks, and C. J. Nicol. 2017. Provision of a resource package reduces feather pecking and improves ranging distribution on free-range layer farms. Appl. Anim. Behav. Sci. 195:60–66.

Phillips, C. J. C. 2005. Meta-analysis - a systematic and quantitative review of animal experiments to maximise the information derived. Anim. Welf 14:333–338.

Rodenburg, T. B., M. M. van Krimpen, I. C. de Jong, E. N. de Haas, M. S. Kops, B. J. Riedstra, R. E. Nordquist, J. P. Wagenaar, M. Bestman, and C. J. Nicol. 2013. The prevention and control of feather pecking in laying hens: identifying the underlying principles. Worlds. Poult. Sci. J. 69:361–373.

Rollin, B. E. 2015. Telos, Conservation of welfare, and ethical issues in genetic Engineering of animals. G Lee, J Illes, and F Ohl, Eds. Curr. Top. Behav. Neurosci. 19:99–116.

Rufener, C., S. Baur, A. Stratmann, and M. J. Toscano. 2018. A reliable method to assess keel bone fractures in laying hens from Radiographs using a Tagged visual analogue scale. Front. Vet. Sci. 5:124.

Sauvant, D., P. Schmidely, J. J. Daudin, and N. R. St-Pierre. 2008. Meta-analyses of experimental data in animal nutrition. Animal 2:1203–1214.

Savory, C. J. 1995. Feather pecking and cannibalism. Worlds. Poult. Sci. J. 51:215–219.

Schmidt, M., J. Stracke, K. Kulke, N. Kemper, and B. Spindler. 2019. Case study of an Automatic enrichment device for laying hens on a free-range laying hen farm. Agriculture 9:91.

Schreiter, R., K. Damme, E. von Borell, I. Vogt, M. Klunker, and M. Freick. 2019. Effects of litter and additional enrichment elements on the occurrence of feather pecking in pullets and laying hens - a focused review. Vet. Med. Sci. 5:500–507.

Sherwin, C. M., G. J. Richards, and C. J. Nicol. 2010. Comparison of the welfare of layer hens in 4 housing systems in the UK. Br. Poult. Sci. 51:488–499.

St-Pierre, N. R. 2001. Invited review. Integrating quantitative findings from multiple studies using mixed model methodology. J. Dairy Sci. 84:741–755.

Steenfeldt, S., J. B. Kjaer, and R. M. Engberg. 2007. Effect of feeding silages or carrots as supplements to laying hens on production performance, nutrient digestibility, gut structure, gut microflora and feather pecking behaviour. Br. Poult. Sci. 48:454–468.

Su, G., J. B. Kjaer, and P. Sorensen. 2006. Divergent selection on feather pecking behavior in laying hens has caused differences between lines in egg production, egg quality, and feed efficiency. Poult. Sci. 85:191–197.

Tahamtani, F. M., M. Brantsaeter, J. Nordgreen, E. Sandberg, T. B. Hansen, A. Nodtvedt, T. B. Rodenburg, R. O. Moe, and A. M. Janczak. 2016. Effects of litter provision during early rearing and environmental enrichment during the production phase on feather pecking and feather damage in laying hens. Poult. Sci. 95:2747–2756.

Tauson, R., J. Kjaer, G. A. Maria, R. Cepero, and K.-E. Holm. 2005. Applied scoring of integument and health in laying hens. Anim. Sci. Pap. Rep. 23:153–159.

Tedeschi, L. O. 2006. Assessment of the adequacy of mathematical models. Agric. Syst. 89:225–247.

Tuyttens, F. A. M., M. Sprenger, A. van Nuffel, W. Maertens, and S. van Dongen. 2009. Reliability of categorical versus continuous scoring of welfare indicators: lameness in cows as a case study. Anim. Welf 18:399–405.

van Staaveren, N., and A. Harlander-Matauschek. 2020. Chapter 16-Cause and prevention of injurious pecking in poultry. In Understanding the Behaviour and Improving the Welfare of Chickens. C. Nicol, ed. Burleigh Dodds Science Publishing, Cambridge, UK.

Vestergaard, K. S., and L. Lisborg. 1993. A model of feather pecking development which relates to dustbathing in the fowl. Behaviour 126:291–308.

Wechsler, B., and B. Huber-Eicher. 1998. The effect of foraging material and perch height on feather pecking and feather damage in laying hens. Appl. Anim. Behav. Sci. 58:131–141.

Weeks, C. A., S. L. Lambton, and A. G. Williams. 2016. Implications for welfare, productivity and Sustainability of the variation in reported levels of mortality for laying hen flocks kept in different housing systems: a meta-analysis of ten studies. PLoS One 11:e0146394.

Weeks, C. A., and C. J. Nicol. 2006. Behavioural needs, priorities and preferences of laying hens. Worlds. Poult. Sci. J. 62:296–307.

Welfare Quality. 2009. Welfare Quality®, assessment protocol for poultry (broilers, laying hens). Welfare Quality® Consortium, Lelystad, Netherlands.

Wennrich, G. 1974. Ethological studies of domestic chickens (gallus domesticus) of different hybrid origin in floor management with special reference to aggressive behaviour as well as feather pecking and cannibalism. 1st Communication: pecking movements in feeding behaviou. Arch. für Geflügelkd 38:143–149.

Wood-Gush, D. G. M., and C. G. Rowland. 1973. Allopreening in the domestic fowl. Rev. Comport. Anim. 7:83–91.

Würbel, H., and J. P. Garner. 2007. Refinement of rodent research through environmental enrichment and systematic randomization. Natl. Cent. Replace. Refinement Reduct. Anim. Res. 9:1–9.

Zepp, M., H. Louton, M. Erhard, P. Schmidt, F. Helmer, and A. Schwarzer. 2018. The influence of stocking density and enrichment on the occurrence of feather pecking and aggressive pecking behavior in laying hen chicks. J. Vet. Behav. 24:9–18.