General enquiries on this form should be made to:
Defra, Science Directorate, Management Support and Finance Team,
Telephone No. 020 7238 1612
E-mail: research.competitions@defra.gsi.gov.uk

SID 5 Research Project Final Report

- **Note**
  In line with the Freedom of Information Act 2000, Defra aims to place the results of its completed research projects in the public domain wherever possible. The SID 5 (Research Project Final Report) is designed to capture the information on the results and outputs of Defra-funded research in a format that is easily publishable through the Defra website. A SID 5 must be completed for all projects.
- This form is in Word format and the boxes may be expanded or reduced, as appropriate.
- **ACCESS TO INFORMATION**
  The information collected on this form will be stored electronically and may be sent to any part of Defra, or to individual researchers or organisations outside Defra for the purposes of reviewing the project. Defra may also disclose the information to any outside organisation acting as an agent authorised by Defra to process final research reports on its behalf. Defra intends to publish this form on its website, unless there are strong reasons not to, which fully comply with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.
  Defra may be required to release information, including personal data and commercial information, on request under the Environmental Information Regulations or the Freedom of Information Act 2000. However, Defra will not permit any unwarranted breach of confidentiality or act in contravention of its obligations under the Data Protection Act 1998. Defra or its appointed agents may use the name, address or other details on your form to contact you in connection with occasional customer research aimed at improving the processes through which Defra works with its contractors.

Project identification

1. Defra Project code: AW0235

2. Project title:
   A study to compare the health and welfare of laying hens in different types of enriched cage

3. Contractor organisation(s):
   SAC
   West Mains Road
   Edinburgh
   Midlothian
   EH9 3JG

4. Total Defra project costs (agreed fixed price): £ 364,288

5. Project: start date ..... 01 September 2005
   end date ..... 31 October 2008

6. It is Defra's intention to publish this form.
Please confirm your agreement to do so.....**YES** ☒ **NO** ☐

(a) When preparing SID 5s contractors should bear in mind that Defra intends that they be made public. They should be written in a clear and concise manner and represent a full account of the research project which someone not closely associated with the project can follow.
Defra recognises that in a small minority of cases there may be information, such as intellectual property or commercially confidential data, used in or generated by the research project, which should not be disclosed. In these cases, such information should be detailed in a separate annex (not to be published) so that the SID 5 can be placed in the public domain. Where it is impossible to complete the Final Report without including references to any sensitive or confidential data, the information should be included and section (b) completed. NB: only in exceptional circumstances will Defra expect contractors to give a "No" answer.
In all cases, reasons for withholding information must be fully in line with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.

(b) If you have answered NO, please explain why the Final report should not be released into public domain

n/a

Executive Summary

7. The executive summary must not exceed 2 sides in total of A4 and should be understandable to the intelligent non-scientist. It should cover the main objectives, methods and findings of the research, together with any other significant events and options for new work.

This study set out to assess the behaviour, health and welfare of brown (Br) and white (W) hens housed in enriched cages that differed in design and colony size. This work was of interest because enriched cages will be the only caged housing option for laying hens within the EU from 2012. Enriched cages provide hens with at least 750 cm^2 floor space per bird, a nest, a pecking and scratching area, at least 15 cm perch space per bird, and 12 cm feed trough space per bird. Currently, producers commonly use 60-bird colony cages, but it was desirable to assess if other colony sizes were beneficial, and how they compared with one another. In addition, birds used here were non-beak trimmed, because the UK is expected to ban routine beak trimming from 2011. It will be important to be able to manage such flocks in enriched cages while preserving bird welfare (i.e. low incidences of feather pecking and cannibalism).

Two flocks of hens were housed at point of lay for one year each in 72 cages (comprised of 2 cage types x 3 colony sizes x 2 bird strains). Cages were purchased from two cage manufactures (referred to as type A and type B). The cage designs differed in cage layout and amount of scratch mat and nest box space, and colony sizes tested were 20-, 40- and 80-bird colonies. Data collected included: body weights, feed intake, egg production, egg quality, behaviour, white blood cell counts, blood titres, tonic immobility, condition of the feathers, combs, claws and feet, bone strength and keel bone damage, and mortality. Data were analysed using Linear Mixed Models and Generalised Linear Mixed Models.

There was no single cage type, colony size or strain that was superior to the others in the factors assessed here. Behaviour was influenced by cage design, strain and colony size. During daylight hours, there was generally better use of enriched cage furniture (i.e. nest boxes, perches and scratch mat) with cage type A (where the nest box and scratch mat space per hen was also greater). Footpad damage was also lower with cage type A, which is possibly related to less time spent on the wire floors (however overall footpad damage was very low – less than 0.5 lesions per bird). As colony size increased, scratch mat use increased with cage type A, and decreased with cage type B, whereas time spent on the floor decreased with cage type A and increased with cage type B with increasing colony size. With strain, both Br and W birds spent similar amounts of time on all resources with cage type B, but with cage type A, W birds spent more time on enriched cage furniture and less time on the floor compared to Br birds. Br birds spent more time feeding than W (which agrees with body weight and feed intake data), and less time sitting and walking than W. At night infrared cameras were used to assess birds' perch use. About 90% of all birds observed used perches, and this was uninfluenced by strain, cage type or colony size.

Bird strains showed preferences for nest box designs and/or space per hen: Br birds preferred cage type B (which was more enclosed, and provided less space/hen), while W birds preferred cage type A (which was more open, and provided more space/hen). Br birds were better nest box users overall (based on eggs laid in the nest box). These data were not corroborated by behaviour observations (probably

because observer presence influenced W hens in particular to bolt to nest boxes), suggesting this is not a good indicator of nest box use. However, reduced nest box use does not necessarily influence egg quality *per se* (for example, if hens lay eggs on the scratch mats instead), but indicates design suitability to hens' needs.

White blood cell counts (i.e. of heterophils and lymphocytes) and titre responses to vaccination were used as physiological indicators of stress. However, there was no effect of cage type, colony size or strain on these parameters. Tonic immobility (a fear test) indicated that W birds were more fearful than Br birds (because they stayed on their backs longer, i.e. mean 202 versus 152 sec), and this was corroborated by behaviour (W birds were observed walking more than Br birds, and interrupted their behaviour more often than Br birds when people approached their cages).

Body condition (based on feather cover, comb scratches, claw length, footpad score, bone strength and bone fracture) varied. Feather damage (an indicator of feather pecking) was poorer for Br than W birds, and increased with increasing colony size, and was slightly worse for cage type A than B. Mean claw length (an indicator of how well the claw shortening device in the cage is working) was longer for W hens (but W hens' claws grow faster than Br hens'). For Br hens, claws were longer in cage type A (1.45 cm) than in cage type B (1.10 cm), probably because the claw shortening device in cage type B covers a considerably larger area than that in cage type A. The mean number of comb scratches per hen (an indicator of aggression) was higher in W (5.6) than Br birds (1.9), and higher in 20-bird cages (4.0) than 40 or 80-bird cages (3.1 and 2.8 respectively), although differences were very small, and aggressive interactions were rarely observed during behaviour observations. Footpad lesions (mentioned above) were low. Bone strength did not vary with cage design or colony size. Keel bone damage was low across all treatments (0-12.5% of birds sampled had old keel bone fracture) compared to previous work on enriched cages and gave no consistent effect with cage type, colony size or strain.

Mortality (7.7%) was higher than expected for both strains of birds (c.f. 4-5%), most likely as a result of intact beaks and a lack of stimulation with fresh feed leading to high levels of feather pecking/cannibalism. There was also a considerable proportion of mortalities attributed to aggression/bullying-related injuries and not eating (which is likely to be a symptom of a bullied bird), particularly in Br birds – these effects are likely to be worse with untrimmed beaks. For bird welfare reasons, we beak trimmed entire cages of birds where two or more birds were culled or died due to outbreaks of cannibalism – this resulted in more 80-bird than 20- or 40-bird colony cages being trimmed.

Feed intake results were unreliable due to feed spillage in both flocks, which was exacerbated by manual feeding and the design of the food trough (which should be used in conjunction with track-and-chain system with pulse feeding several times a day). Egg production (number of eggs per hen-day) was poorer than expected for Br birds due to feed quality problems in flock 1 up to 27 weeks of age. Colony size appeared to affect egg production in cage type B (20 bird colony = 83%, 40 bird = 81%, 80 bird = 79%), but this was most likely a reflection of egg eating in this cage type (there was no egg shock wire with these cages). The proportion of second class eggs (6% of all eggs) were slightly lower than seen in other studies using enriched cages (7.8%). However, cage type B had more dirty eggs than A, (probably due to broken eggs from egg eating) but A had more cracked eggs (possibly as eggs built up on the egg belt). Both of these would be exacerbated by manual egg collection. Ignoring the loss of second quality eggs as an artefact of the study, total cumulative egg mass per hen was greater with cage type B by 0.5 kg/hen.

Overall, W birds were more fearful, had overgrown claws, and used the nest boxes less than Br birds, but W birds used enriched cage furniture more, had better feather cover, and died less from aggression and bullying than Br birds. These results are perhaps counterintuitive, as it may be expected that the more fearful bird would perform less well in this type of system. The design of cage type A resulted in better furniture use and fewer feather pecking/cannibalism-related deaths but slightly worse feather damage than cage type B, and nest box designs were favoured by different bird strains. The other cage design-related factors (egg production, egg quality) would probably disappear with automatic feeding and automatic egg collection. Colony size's major effects were on feather scores (which got worse with increasing colony size) and on the requirement to beak trim (more 80-bird cages than 20- or 40-bird cages), plus some effects on use of resources and claw length (which were all better with smaller colony sizes). Thus, overall in this study, 20-bird colony sizes were considered better for bird welfare. Because birds were fed manually once a day, birds were not stimulated to peck at feed as often had they been presented with fresh feed several times a day (as is the case commercially). As a consequence, birds may have redirected the motivation to forage on one another, as indicated by poor feather scores and mortality. This damage would be exacerbated by intact beaks. With the expected ban on routine beak trimming due to take place from 2011, it would be important to carefully manage feed stimulation, and to tackle pecking problems in colony cages swiftly by beak trimming the entire cage where outbreaks (i.e. 2 or more birds removed due to pecking injury) occur. Data seen here might be considered to have been collected in a 'worst case scenario' both from the hen's and egg producer's perspectives. In a commercial setting,

producers will use pulse feeding and automatic egg collection, which should reduce feed waste, improve feather quality, reduce pecking-related mortality and reduce second-class eggs and egg losses. Future work should consider other methods of stimulating hens to peck at appropriate objects (such as by increasing the number of presentations of feed onto the scratch mat), ways to prevent aggression, and genetic differences between bird strains with regards to their appropriateness for enriched cages.

Project Report to Defra

8. As a guide this report should be no longer than 20 sides of A4. This report is to provide Defra with details of the outputs of the research project for internal purposes; to meet the terms of the contract; and to allow Defra to publish details of the outputs to meet Environmental Information Regulation or Freedom of Information obligations. This short report to Defra does not preclude contractors from also seeking to publish a full, formal scientific report/paper in an appropriate scientific or other journal/publication. Indeed, Defra actively encourages such publications as part of the contract terms. The report to Defra should include:

- the scientific objectives as set out in the contract;
- the extent to which the objectives set out in the contract have been met;
- details of methods used and the results obtained, including statistical analysis (if appropriate);
- a discussion of the results and their reliability;
- the main implications of the findings;
- possible future work; and
- any action resulting from the research (e.g. IP, Knowledge Transfer).

Scientific objectives
Objective 1: to compare the behaviour of hens (both brown and white strains) housed in enriched cage systems that differ in terms of furniture provision and colony size

Objective 2: to assess the health and welfare of hens (both brown and white strains) housed in enriched cage systems that differ in terms of furniture provision and colony size

Extent to which objectives have been met
Both objectives have been fulfilled in full.

Background
This study set out to assess the behaviour, health and welfare of brown (Br) and white (W) hens housed in enriched cages that differed in design, furniture provision, and colony size. This work was of interest because enriched cages will be the only caged housing option for laying hens within the EU from 2012. Enriched cages provide the hens with at least 750 cm^2 floor space per bird, a nest, a pecking and scratching area, at least 15 cm perch space per bird, and 12 cm feed trough space per bird. Currently, producers commonly use 60-bird colony cages, but it was desirable to assess if other colony sizes were beneficial, and how they compared with one another. Different cage designs were investigated, since how resources are laid out in the cage, and the amount of particular resources (i.e. nest box and scratch mat) will differ with cage manufacturer. In addition, birds used here were non-beak trimmed, because the UK intends to ban routine beak trimming from 2011. It will be important to be able to manage such flocks in enriched cages while preserving bird welfare (i.e. low incidences of feather pecking and cannibalism).

Materials and Methods

Housing

Two consecutive flocks were used for this project. Within both flocks, non-beak trimmed white (W) and brown (Br) pullets (from the same breeding company) were reared at one rearing farm and transferred together to the research site at point of lay. These point-of-lay hens were housed in one laying shed that contained two types of enriched cage (type A and type B), in two banks of three tiers [top (T), middle (M) and bottom B)] with three colony sizes (20, 40 and 80 bird cages) per cage type. This gave 12 combinations of strain x cage type x colony size, of which there were 6 replicates each (i.e. 72 cages total, 3360 hens per flock) (Figure 1). The first flock was placed in November 2005 at 17 weeks of age, and the second flock was placed in May 2007 at 16 weeks of age. Both flocks were housed until 72 weeks of age. The banks were swapped within the shed between the two flocks, and where cages had held Br birds in flock 1, they held W birds in flock 2, and vice versa. Temperature and relative humidity (%) were recorded daily, and light intensity was recorded every four weeks. Lights were initially on for 10 h/day, which was stepped up by 15 min/week to 16 h/day (08:00 – 00:00) by 40 weeks of age, according to the breed management guidelines. Notation throughout the rest of the report will be by strain-cage type-colony size (e.g. W-A-20) where appropriate.

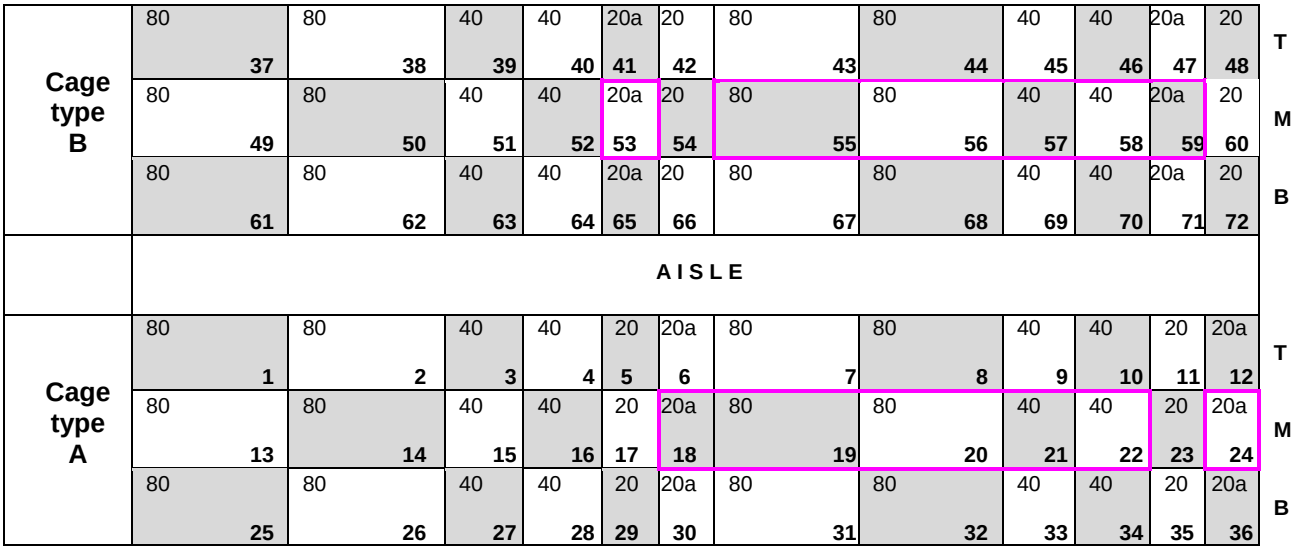


Figure 1 Cage layout of flock 1, showing cage type A or B, top (T), middle (M) and bottom (B) tiers, with colony size (i.e. numbers of birds per cage, 80, 40, or 20). Numbers in bold are the cage number. 20-bird cages are back-to-back; '20a' means 'facing the aisle' (as opposed to facing the wall). Highlighted boxes = brown strain, white boxes = white strain. Pink border = night time recording. The sides of the shed are the long walls (top and bottom), the ends of the shed are the short walls (left and right). For flock 2, the banks of cages were swapped.

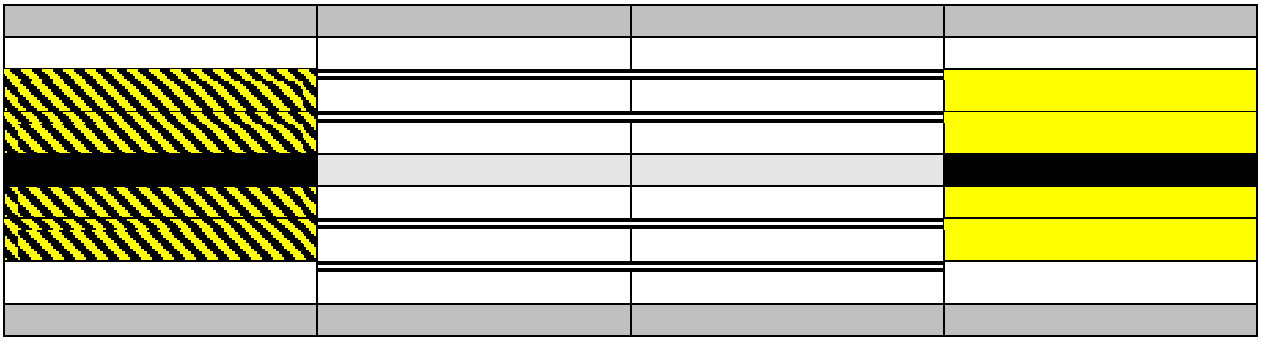
Cage design and layout

Both models of enriched cages contained a scratch mat, feed troughs, nipple drinkers, a nest box, perches, claw shortening devices and at least 750 cm^2 floor space per hen (see Table 1). In appearance, the major differences between the two cage types were that type A always positioned the nest box and scratch mat at opposite ends of the cage, and the nest box sides were made of two solid sides, one wire mesh side and one plastic 'curtain'; type B positioned the nest box nearer to the centre, with the scratch mat at the one end of the cage (20 and 40 colony) or in the centre (80 colony), and the nest box sides were made of three solid sides (see Figure 2). Additionally, the claw shortening devices differed: in cage type A it was an abrasive strip glued to the feed trough baffle; in cage type B the entire baffle was perforated (Figure 3). Cage A had an electrified wire to prevent hens damaging eggs on the egg belt, and an 'egg saver' wire that slowed eggs down before rolling onto the egg belt.

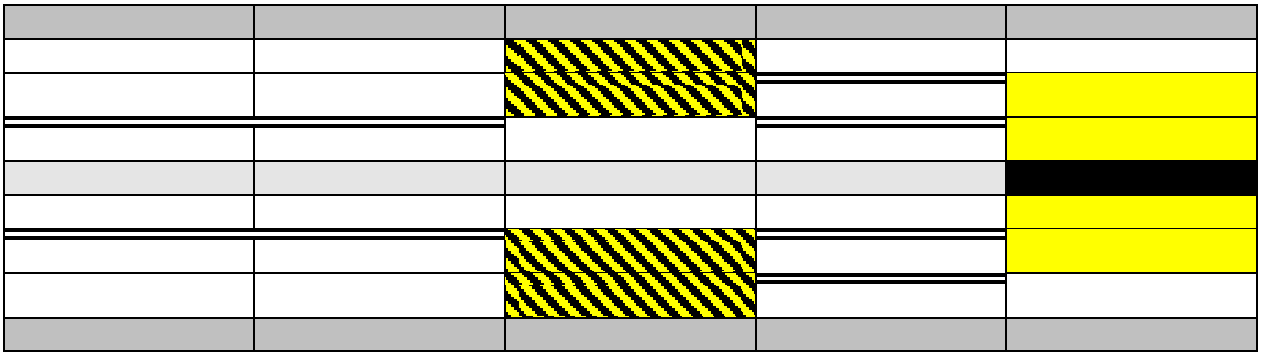
Table 1 Resource space per bird, in cm^2 unless otherwise stated, according to cage type and colony size.

	Cage type A			Cage type B		
Colony size →	**20**	**40**	**80**	**20**	**40**	**80**
Scratch mat	97	97	97	50	50	25
Nest box	97	97	97	78	78	78
Perch (length, cm)	15	15	15	16	17	18
Floor area (including nest and scratch mats)*	753	753	753	795	807	811
Food trough (length, cm)**	12	12	12	12	12	12
Claw shortening device	3.75	3.75	3.75	44	44	44

* = this includes the floor area which is inaccessible due to protruding feed trough baffle (where claw shortening devices were located) – this part of the trough reduces the floor area by 10-12%.
** = some of the food trough space in cage type B is not accessible due to the presence of the nest box – this reduces the usable space to about 11 cm per bird.



Cage type A



Cage type B

Figure 2 Diagrams of 40-bird cages of design type A and type B, showing the feed trough ( ), nest box ( ), scratch mat ( ) and perches (=). Birds could also perch on the centre auger tube (light grey). A barrier (in black) prevented birds from moving from one side of the nest box (type A only) and scratch mat (both cage types) to the other.

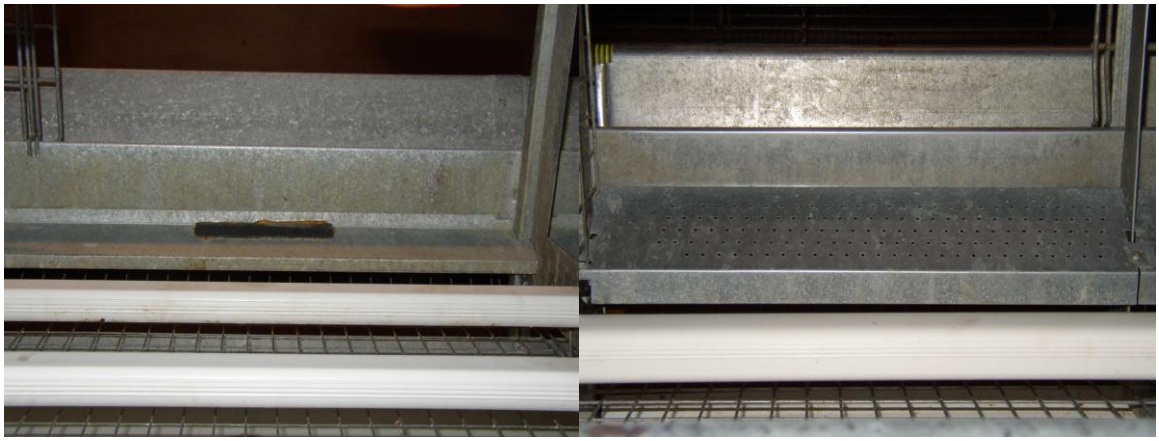


Figure 3 Claw shortening devices in cage type A (left) and cage type B (right).

Lighting for both cage types was supplied by cage type A's manufacturers, and consisted of vertical tubing designed to illuminate the top, middle and bottom tiers uniformly. Lights were hung opposite the scratch mat; light intensity was recorded (in lux) every four weeks using a light meter, with measurements taken at bird head height at the cage bars opposite the scratch mat and nest box. Where intensity was found to be uneven (for example, one bank brighter than another), the lights were adjusted and re-recorded.

Body weights, feed and water

Ten birds per cage were bulk weighed on the day of entry to the laying farm at 16-17 weeks and also prior to depopulation at 71 weeks (not necessarily the same birds). Birds were fed a layer's mash (pre- and post-peak rations, formulated based on guidelines given in the strain management handbooks) *ad libitum* by hand distribution into discrete troughs per cage, from pre-weighed feed bags. Feed was formulated by SAC and made initially by a trials feed manufacturer, but due to deficiencies in brown egg diets in flock 1 (determined: 10.9 MJ/kg ME, 15.2% CP, compared to formulated: 11.7 MJ/kg ME and 17.2% CP), feed manufacturing of all diets (Br and W) was changed to commercial feed at 27 weeks of age. Commercial feed was used throughout flock 2. Every four weeks, any remaining feed per cage was weighed in order to assess mean feed intake (i.e. feed added - remaining feed) per hen day per cage, taking into account exact bird days (i.e. accounting for mortality). Due to feed wastage seen in flock 1, grids were placed over all food troughs in flock 2 to prevent birds from raking feed out of the troughs. Water was provided *ad libitum* from nipples. Layer's mash was dispensed automatically onto scratch mats once per day (at approximately 14:00), except in flock 2 where, from 63 weeks of age until the end of the study, mash was dispensed onto scratch mats four times a day (10:00, 12:00, 14:00 and 16:00) to reduce feather pecking.

Egg production and quality

All eggs were collected by hand once a day, and the number of eggs laid and where they were located (i.e. opposite the nest box or not), as an indicator of nest box use, was noted. Eggs produced and the numbers of hen days (numbers of hens in that cage per day, taking into account daily mortality) were used to calculate %eggs per hen day. On thirteen occasions (every 4 weeks to 68 or 69 weeks), all eggs laid from each cage on two consecutive days were assessed for mean egg weight (by bulk-weighing all eggs on each day of collection, omitting any double-yolked or cracked eggs), external appearance (i.e. soft shells, double-yolks, cracks, mis-shapes such as slab-sided) and contamination of the shell with organic material (e.g. faeces, blood, referred to as 'dirty'). Weekly, cumulative, and total cumulative egg mass was also estimated (total number of eggs laid per cage x estimated egg weight, divided by hen days), with and without second-quality eggs, from 18-71 weeks. A sub-sample of 4 clean eggs per cage were assessed for shell thickness (in microns) and internal quality including albumen height (to calculate Haugh units, HU, indicative of egg freshness – the higher the figure, the fresher the egg), yolk colour (based on Roche scale of 1 (lightest) to 15 (darkest)), and presence of inclusions (blood and meat spots).

Behaviour

Once every four weeks for 14 periods, most birds (20-bird cages) or all birds between the cage front and nipple line (40- and 80-bird cage) were scan sampled once every three hours for two days during the light period (3-5 scans per day, depending on the light schedule). Daytime scans were grouped for analysis into start (first scan), end (last scan), and middle (all other scans). During lights off one quarter (40 and 80 bird colonies) to one half (20 bird colonies) of the cage front in one replicate (middle tier only, see Figure 1) per strain-cage type-colony size (i.e. 12 cages total) was digitally recorded (over several nights) using infra-red lights and cameras. Data from night-time videos was also scan sampled, with the first scan taking place within 45 min of lights off, then every three hours during the dark period, with the final scan assessed 15 min before lights on (4-5 scans per night). Night-time scans were also classed into start, middle and end scans for anlaysis. Birds were observed for time spent in an ethogram of behaviour classess (i.e. stand, sit, walk, feed, drink, non-aggressive peck, aggressive peck, peck/scratch at the mat, object peck, sham dustbathe, and preen) according to cage location (i.e. nest box, perch, scratch mat, or wire floor). Non-aggressive pecks was made up of feather pecks and pecks to the beak of another bird, object peck was pecking at any object within the cage apart from the nipple line, feed, and other birds, and sham dustbathe were dustbathing movements on either the wire or either of the plastic mats. During night time scans, some behaviour classes were pooled into a category ‘other’ because of a high proportion of zero counts in these classes making them otherwise impossible to analyse. So at night-time behaviours analysed were stand, sit, walk, preen and other. Likewise, some behaviours could only be performed in particular locations (for example, peck/scratch at the mat had to be performed on the nest box or scratch mat; drink could not occur in the nest box), or during night-time scans nest boxes and scratch mats were not recorded. Because of the enormous amount of data and subsequent computational difficulty of modelling such a large data set, the 14 scan

periods were collapsed for analysis into periods 'pre-peak' (17-24 weeks), 'peak' (25-35 weeks) and 'post-peak' (36-71 weeks) ages.

Indicators of stress and immune response

Every eight weeks from 24 weeks of age, two birds per cage were blood sampled, from which a blood smear was stained and assessed for white blood cells (heterophils and lymphocytes). One hundred blood cells were assessed and the H/L ratio calculated. At the 40-week sample, blood was also taken to assess immune response to a vaccination given at 35 weeks (to either infectious bursal disease, IBD, flock 1, or Newcastle disease, ND, flock 2). At blood sampling, birds were wing-tagged to avoid re-sampling. Just prior to depopulation (at 72 weeks for flock 1, and 70 weeks for flock 2), 1-2 birds/cage (flock 1) or 2 birds/cage (flock 2) were assessed for tonic immobility (i.e. time to right itself from lying on its back) as an indicator of fear (Jones, 1986). The test was stopped (censored) if the bird did not right itself within 10 min.

Feather, comb, claw, and foot condition

Feather damage was recorded as an indirect indicator of feather pecking (Kjaer *et al.*, 2001). Ten birds per cage were remotely feather scored (which has been shown to correlate well with feather scoring while handling, Bright *et al.*, 2006) at 35-36 and 55-56 weeks of age, and five birds per cage were feather scored with handling at 72 weeks of age, except for cage 9 (depleted at 51 weeks of age) and cage 45 (depleted at 64 weeks of age, due to feather pecking or bullying) in flock 2. The dorsal side of the neck, back, tail, wings and belly were assessed for degree of feather loss and pecking wounds, based on a scoring system of 0 – 5 by Savory and Mann (1999) where 0= perfect plumage condition, 1= slight feather loss/damage with no bare skin, 2= some damage with up to 1 cm^2 bare skin, or exposed feather vanes (wings and tail), 3= damage with up to 25 cm^2 bare skin or up to 1 cm^2 minor haemorrhage or few or bare wing or tail feathers, 4= more than 25 cm^2 bare skin or less than 25 cm^2 bare skin with up to 1 cm^2 haemorrhage or no wing or tail feathers, 5= more than 1 cm^2 haemorrhage, or more than 25 cm^2 bare skin and up to 1 cm^2 haemorrhage. The underside of the neck and breast were never scored, because these could be damaged by cage wear during feeding (and is thus not related to feather pecking). With many birds during remote scoring, some body sites were not visible, because of bird position within the cage. In flock 2, cage 9 birds were scored with handling when depopulated at 52 weeks, and this data contributed to the 55-56 weeks data set. Birds that were feather scored at 72 weeks were also assessed for claw length (middle digit of one foot per bird, measured to the nearest 0.5 cm, as an indicator of efficacy of claw shortening devices), numbers of comb scratches (an indicator of aggression), and numbers of foot pad lesions (both feet per bird, as an indicator of damage due to wire floors).

Bone strength and fractures

At 71 weeks, four birds per cage were culled and one tibia per bird was assessed for bone strength (by three-point bending test, which gives the force at which the leg bone breaks). The maximum strength measureable was 300 Newtons. The same carcasses were examined for fresh and old bone fractures (assessed by radiography) to the keel bone, furculum, wings (radius, ulna, humerus) and remaining leg (femur, tibia). The keel bone was scored from 0-3 where 0 = perfect keel; 1 = slight deformation but no fracture; 2 = pronounced deformity or minor fracture; and 3 = as with (2) but more severe fracture or twisting of the boney keel, and may involve severe displacement.

Mortality

Mortality (i.e. birds found dead or culled) was recorded daily. Where cause of death or reason for culling was unknown, birds were assessed post-mortem. Causes of mortality were categorised into aggression/bullied, found dead reason unknown, feather pecking/cannibalism, lame, not eating, or other (which consisted of symptoms such as egg peritonitis, prolapse, swollen gut, dropped crop). If two or more birds were identified with pecking wounds, that entire cage of birds was beak trimmed by hot blade to prevent further losses. Some cages were re-trimmed if wounding resumed. In flock 2, wounding or bullying persisted in two cages (cages 9 and 45, see Figure 1) thus all remaining birds were culled.

The differences between the two flocks were:

1. In flock one, cage bank A was positioned on the north-east side, and cage bank B on the south-west side, of the poultry house. In between flocks 1 and 2, the two banks were swapped, to balance for side-of-house effects.
2. Where cages had housed white hens in flock 1, they housed brown hens in flock 2, and vice versa.
3. Flock 1 was housed at 17 weeks of age; flock 2 at 16 weeks of age
4. Because of feed wastage seen in flock 1, grids were laid over the troughs to prevent birds from raking feed into the cage and thus onto the droppings belt.
5. The type of vaccines used to stimulate and asses immunity were IBD for flock 1 and ND for flock 2.
6. In flock 1, ten birds per cage were identified with spiral leg rings on day of entry to the laying house, to serve as focal birds during observing; however, these birds could not easily be seen and some birds suffered damage to their feet/legs from the rings, so the leg rings were removed. All visible birds were scanned as described above.
7. In both flocks, feed was dispensed onto the scratch mat once per day in the afternoon (approximately 14:00); however in flock 2 at 63 weeks of age until the end of the study, this was increased to four times per day (10:00, 12:00, 14:00 and 16:00) in an attempt to reduce feather pecking.
8. Tonic immobility tests were carried out on initially 2 and then 1 bird per cage at 72 weeks of age in flock 1. The number of birds assessed per cage was reduced because of the exceptionally long period of time that tests took, due to power cuts to the poultry shed (and subsequent fear effects on birds, increasing time taken in each test). For flock 2, we began the tests earlier (at 70 weeks of age), and tested 2 birds from every cage as planned.

Statistics

Linear Mixed Models (LMMs) were used to analyse continuous or near continuous variables. Log transformations were applied when required in order to satisfy the assumptions of normally distributed random errors and homogeneity of variances. In some LMMs units were weighted differently in order to account for extra variability in the data. Generalised Linear Mixed Models (GLMMs) with a logit link transformation and binomially distributed errors were used to analyse binary data or binomial count data (proportions). GLMMs with a log link transformation and poisson distributed errors were used to analyse count data and multinomial count data (proportions falling into more than 2 classes): for multinomial data a factor whose levels correspond to each table of counts was fitted as the first fixed effect in order that analysis corresponds to proportions falling into the different classes.

All LMMs and GLMMs were fitted using Residual Maximum Likelihood (REML). The natural logarithm ($\log_e$) is used throughout this report (referred to throughout as log). Tests of main effects and interactions in LMMs and GLMMs are approximate F-tests unless otherwise stated. The less accurate Wald tests are only used when computational difficulties preclude use of the approximate F-test. Stated P-values for fixed effects are those adjusted for all other fixed effects of the same order and lower (for example, a main effect tested in the model with all other main effects included, a 2 way interaction tested in the model with all 2 way interactions and main effects included, and so on) unless otherwise stated. Graphs and tables of means and standard error of the difference (SED) estimated from the models on the linear (transformed) scale and corresponding post hoc *t* tests, and of estimated means back transformed, were used to aid interpretation of effects but it should be noted that post hoc tests are approximate in the case of GLMMs. Where data were transformed, significant results in the text are shown on the back transformed scale as these are more meaningful, but in some instances they may not match the original figures shown in raw data tables (for example, mortality). Estimates of variation (SED or standard deviation, SD) are given only when estimates on the linear (transformed) scale are shown.

The effects of appropriate fixed factors were investigated, and always included strain, cage type, colony size, tier and their interactions. For longitudinal data (i.e. data collected over time) the effect of age period is also included as a fixed term. Some analyses include additional factors: for example, feather score included body site (back, tail, etc), light intensity and behaviour included position in cage (scratch mat, nest box, wire floor, perch – latter two with behaviour only), behaviour included scan class (start, middle end). It should be noted that, due to the large amount of data, often very small effects were picked up as significant, and so marginal effects should be interpreted with caution. In general marginal interactions are not reported, particularly for high order terms (e.g. 3 way interactions).

Models took into account random factors such as flock, cage position (to account for spatial position in the house), bird group (i.e. flock by cage), and lower level terms as appropriate (for example flock by cage by age for longitudinal data, flock by cage by bird when birds within cages were assessed, and so on). In GLMMs dispersion was fixed at 1 unless there was strong evidence that it was larger than 1, in which case it was estimated.

Additional analyses methods included proportional hazards models (PHM) for survival analysis of tonic immobility and time to beak trimming and exponential curve fitting to estimate birds weights over all ages in order to estimate egg mass from egg production and weight data. Genstat (Genstat, 10th edition, Lawes Agricultural Trust, VSN International Ltd, Oxford, UK) was used for all the statistical analyses. The results presented here are distilled from an extensive (>500 pages) statistical report.

Results

Light intensity

LMMs were fit to light intensity (log) including fixed effects of location (within the cage, and facing aisle or wall). As desired, light intensity was higher (P<0.001) in the scratch mat area than in the nest box (Figure 4). Light intensity was also higher in the aisle side of cages than the wall side (P<0.001), especially at the nest box, because of the greater number of lights in the aisle. Light intensity was higher for cage type B than A (P<0.001), particularly in the nest box area and (more marginally) at the wall side, and for larger colony sizes but less so in the top tier. There were no effects of strain or interactions with strain. In the first flock, light intensity was turned down from 16 weeks to 25 weeks, because of outbreaks of cannibalism, particularly with white hens.

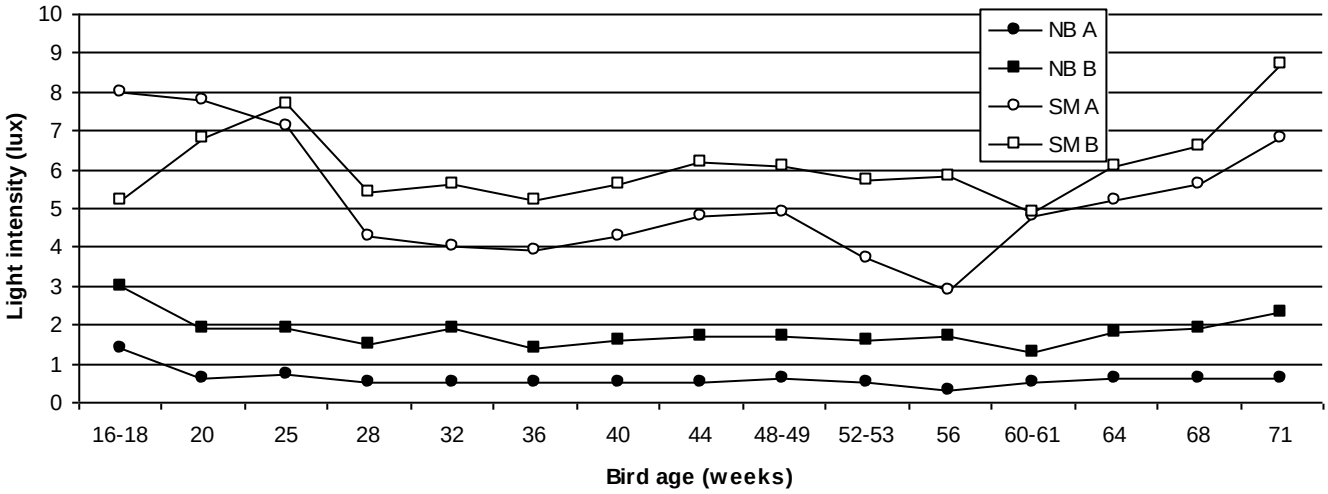


Figure 4 Back transformed means of light intensity according to cage position (nest box, NB or scratch mat, SM) and cage type (A or B), measured at approximately 4 week intervals (from LMM).

Body weight

LMMs were fit to body weights (log). Body weights were slightly behind target at point-of-lay (16-17 weeks, Table 2) but they increased significantly with bird age, and Br birds were significantly heavier than W birds, as expected according to breed standards (both effects P<0.001). Cage type had a marginal effect on body weights (P=0.014) with birds in cage type B slightly heavier (1.53 kg) than those in cage type A (1.50 kg). There were no significant effects of tier or colony size.

Table 2 Original mean bird weights (kg) according to strain at various ages, compared to target given in the strain's management handbook. The range is shown for the 6 cage types x colony sizes.

	16-17 weeks actual	16-17 week target	71 weeks actual	71 weeks target
Brown	1.33-1.37	1.43-1.47	1.93-2.08	1.94
White	1.14-1.18	1.19-1.22	1.68-1.75	1.55

Feed intake

LMMs were fit to feed intake per hen day with units weighted by hen days (since variability is likely to decrease with increasing hen days) and inversely by estimated within flock variability (to account for increased spillage in flock 1). Feed intake was higher than expected and more so for flock 1 than flock 2 (Figure 5). Feed intake was greater for Br than for W at all ages, but the difference was smaller at early ages (age by strain interaction, P<0.001). Feed intake was higher in the top tier especially as birds got older (age by tier interaction, P<0.001), but this increase for the top tier only occurred for cage type B (tier by cage type interaction, P=0.004) and for colony sizes 20 and 40 (tier by colony size interaction, P=0.022). When all 2 way interactions were dropped the main effects of age, strain, cage type and tier were all highly significant (P<0.001), with feed intake much greater for Br birds, greater in the top tier and greater for cage type B, but colony size was insignificant.

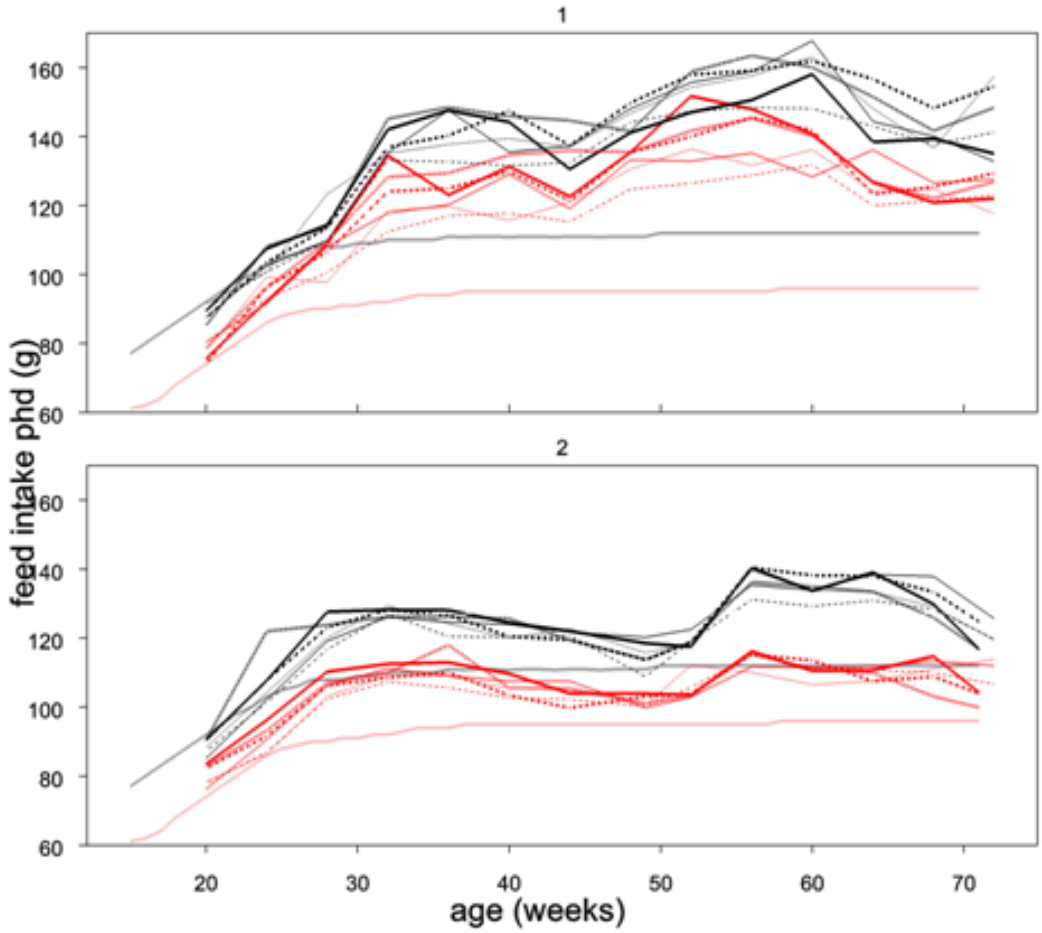


Figure 5 Mean feed intake per hen day (phd) for the different strains (Br = black lines, W = red lines); cage types (A = thin line, cage type B thick line); colony sizes (20 = solid line, 40 = small dash (……) 80 = large dash (----)) with both flocks shown (flock 1, top graph; flock 2, bottom graph). Expected intake (according to each strain's management handbook) is shown (solid grey and pink lines).

Egg production

GLMMs were fit to the weekly egg counts (poisson, link log) with log(total weekly hen days) included as an offset and dispersion (flock by cage by age) estimated. Data from age 18-71 weeks only are included, because the data at age 17 and 72 weeks was incomplete and may bias estimates. Egg production for Br birds was significantly lower than W birds at ages 18-29 and 67-71 weeks (Figure 6) (age by strain interaction, P<0.001). However, egg production was similar for Br and W birds in flock 2 and over a large range of the ages was slightly higher for Br than W, as expected according to targets. Colony size had little effect on egg production with cage type A (20 colony =81.0%, 40 =81.0%, 80 =80.2%) but for cage type B egg production decreased marginally with increasing colony size (20 colony =83.0%, 40 =80.7%, 80 =78.9%) (cage type by colony size interaction, P=0.03). However, this may have been affected by egg eating in these cages, and not a true production result *per se*. Tier was

insignificant (P=0.053, from model with significant 2 way interactions and all main effects included). When main effects were examined, there was no evidence of a cage type effect overall on egg production, but strain was significant (P=0.014) with production lower overall for Br birds, and colony size overall was significant (P=0.02) with greater production in 20-bird cages.

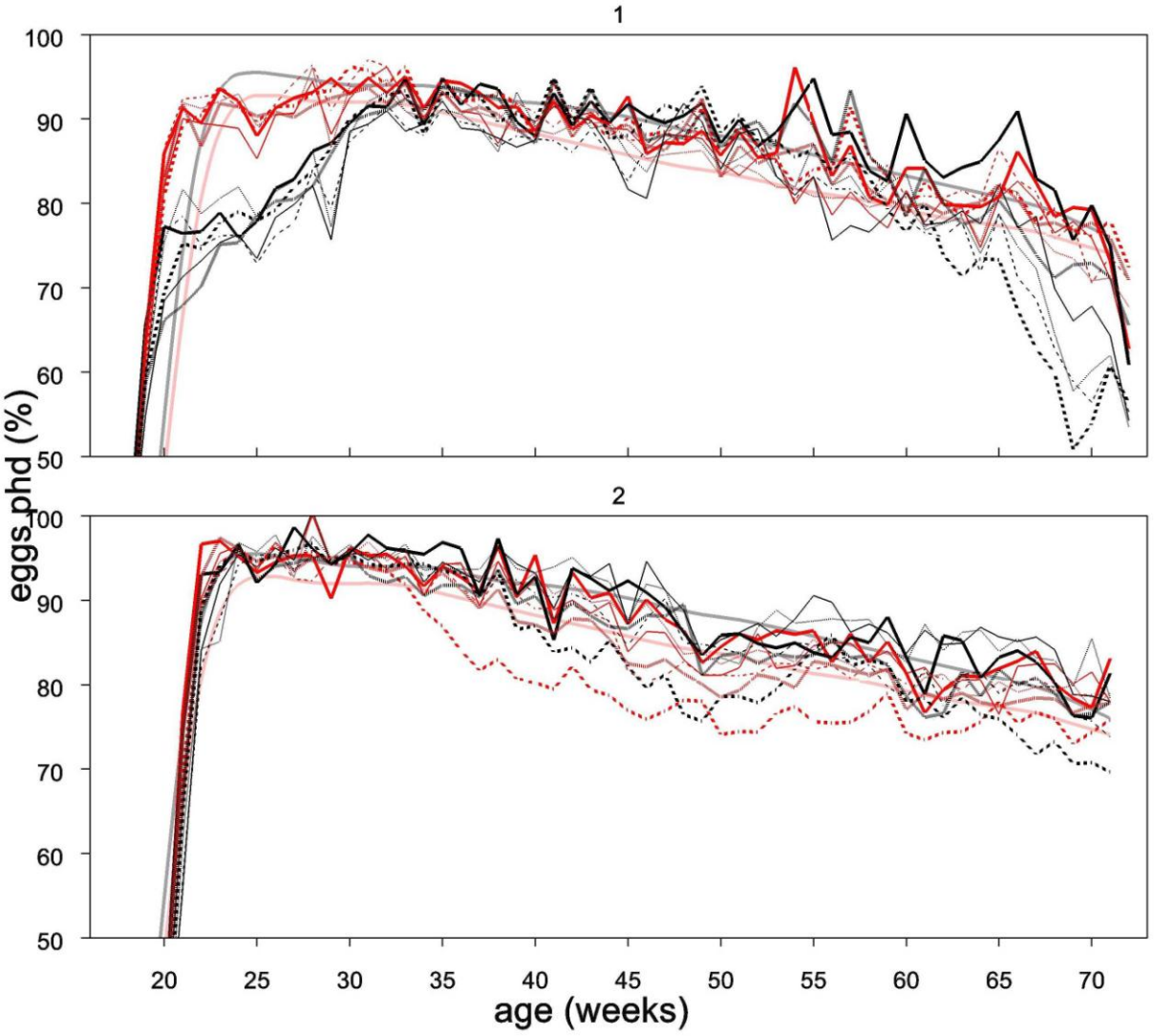


Figure 6 Mean egg production per hen day (eggs phd, shown in %) over the study for the different strains (Br = black lines, W = red lines); cage types (A = thin line, cage type B thick line); colony sizes (20 = solid line, 40 = small dash (……) 80 = large dash (----)) with both flocks shown (flock 1, top graph; flock 2, bottom graph). Data are plotted against age (at start of week) for the two flocks with expected egg production for the two strains also shown (solid grey and pink lines).

GLMMs were fit to numbers of eggs in the nest box as a proportion of those collected (binomial, link logit) with dispersion (flock by cage by age) estimated. Br birds laid a greater proportion of eggs in the nest box than W birds (83% and 64%, respectively) across both cage types (P<0.001). However, Br birds appeared to prefer cage type B nest boxes, while W birds preferred cage type A nest boxes (see Figure 7). Br birds laid significantly more nest box eggs than W at all ages in cage type B; with cage type A, Br birds exceeded W birds at 53-72 weeks only (age by strain by cage type interaction, P<0.001, Wald test). There was some evidence of a tier effect with the proportion of nest eggs higher in the middle tier. Colony size effects were not consistent for the two cage types. In both flocks some cages of birds were consistently poor nest-box users.

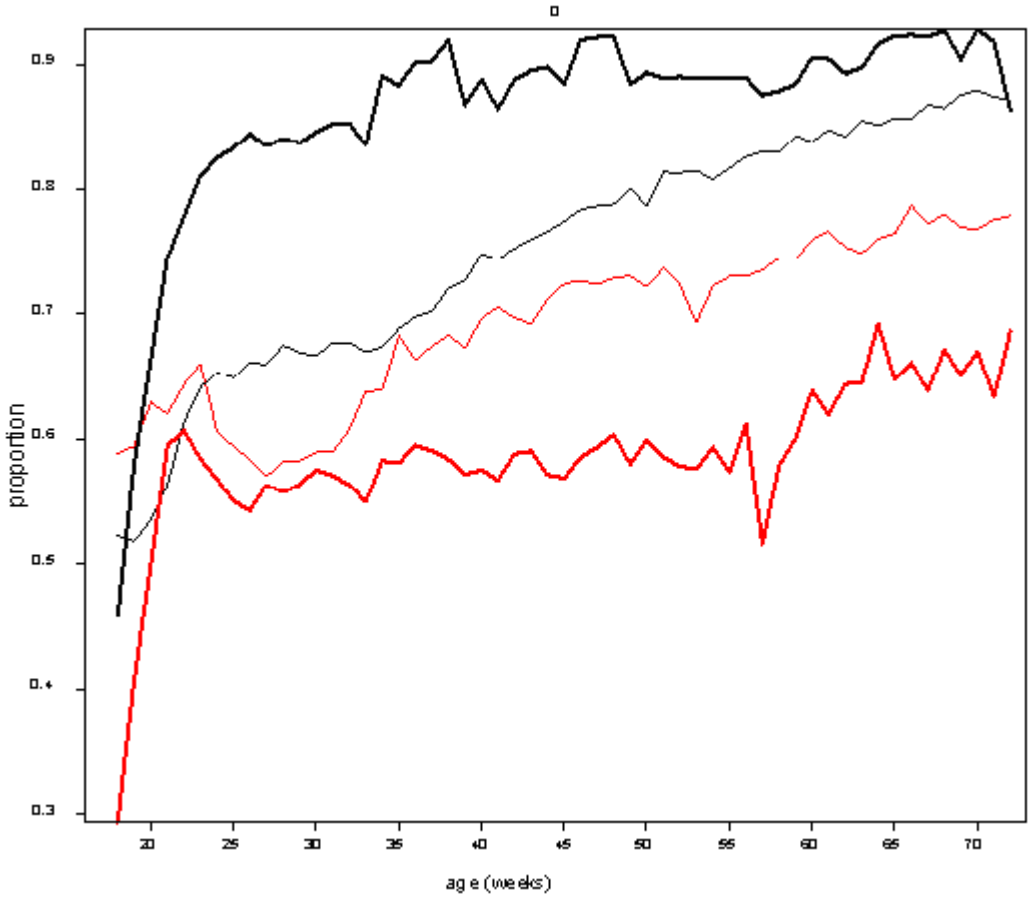


Figure 7 Weekly mean proportion of eggs laid in the nest box over the study for the different strains (Br = black line, W = red line) and cage types (A = thin line, B = thick line) over both flocks (from GLMM).

External and internal egg quality

LMMs were fit to average egg weight with units weighted by the numbers of eggs weighed (since variability is likely to increase with decreasing numbers of eggs weighed i.e. in smaller colony cages). Overall mean egg weight was higher for Br birds (64.2 g) than W birds (60.7 g) (SED=0.10, P<0.001), as expected (breed management guides, to 69 weeks: Br 63.1 g/egg, W 59.9 g/egg). There were no consistent effects of cage type on egg weight. Overall, mean egg weights increased marginally with colony size (20 =62.1 g, 40 =62.5 g, 80 =62.8 g; SED=0.14-0.15, P<0.001) but this is effect was not consistent within different cage type by tiers.

GLMMs were fit to numbers of second eggs as a proportion of eggs assessed (binomial, link logit) with dispersion (flock by cage by age by day) fixed at 1. Of all the eggs assessed (130,000 eggs), 6% were classed as seconds. Cracked (3-3.4% per flock) and dirty (1.5-2.9% per flock) were the most common classifications of seconds, and the proportion of seconds increased with bird age. The number of dirty eggs was exacerbated by egg eating, particularly in cage type B where there was no egg shock wire to prevent hens from damaging eggs. However, there were less seconds with cage type B than A (due to less cracked eggs), except in the top tier for central age ranges (due to dirty eggs in flock 2 only – see Figure 8). The proportions of seconds increased with colony size for B cages only, as a result of a large number of seconds for flock 2 colony size 80, in the top tier (age by tier by cage type interaction, P=0.002). Strain and colony size effects and overall tier effects were inconsistent and fairly marginal.

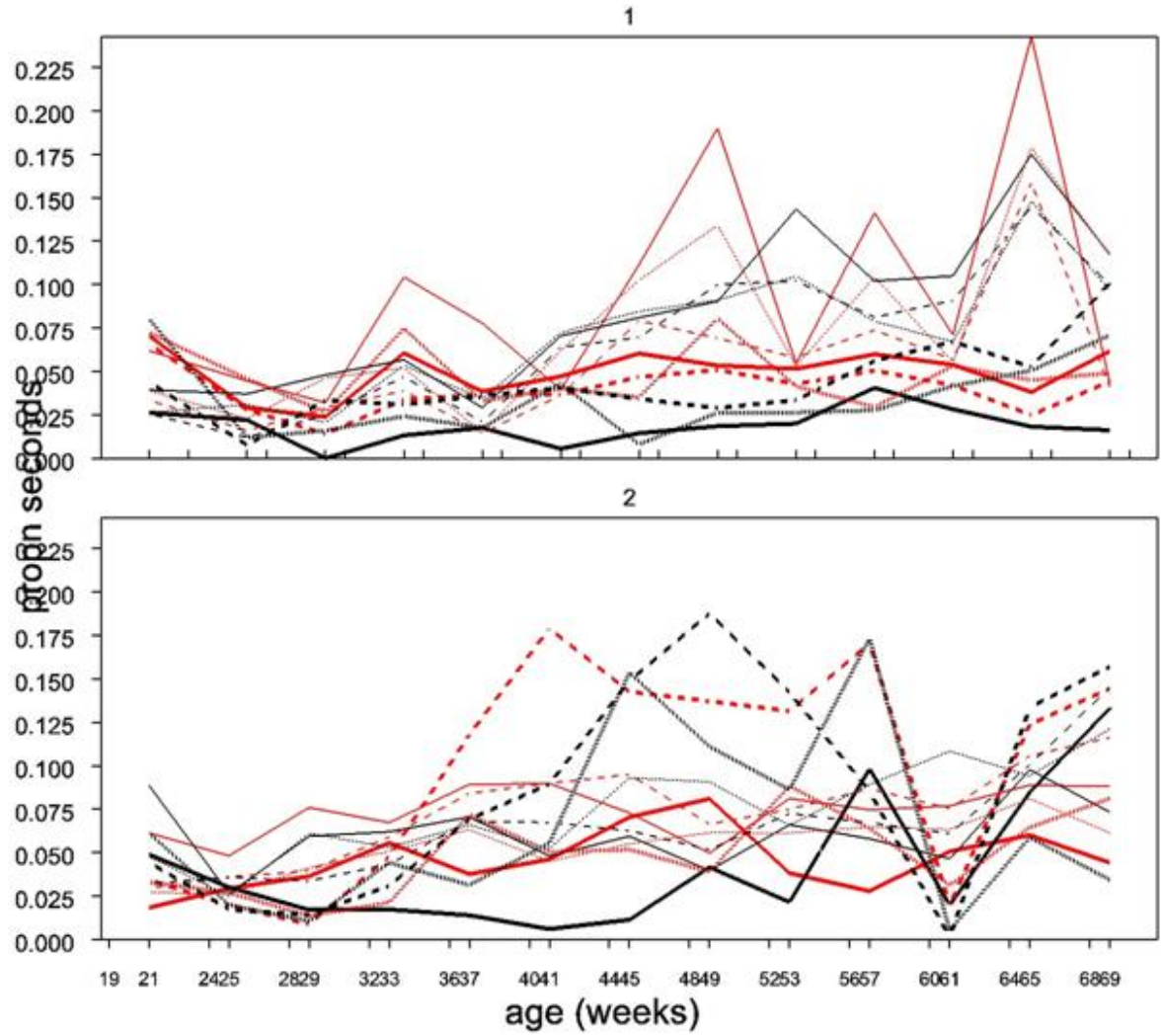


Figure 8 Mean proportion of seconds for the different strains (Br = black lines, W = red lines); cage types (A = thin line, cage type B thick line); and colony sizes (20 = solid line, 40 = small dash (……) 80 = large dash (----)) plotted against age (at start of week of sampling) for the two flocks (flock 1, top, flock 2, bottom).

LMMs were fit to shell thickness. Overall, mean shell thickness initially increased, then decreased, with bird age. As expected, shell thickness was higher in Br (345.9 microns) than W (326.4 microns) birds (SED=0.72) (P=0.001) at all ages, colony sizes and tiers, although values were lower than expected according to the strain guidelines for Br and W birds in the same age range (350 and 340 microns, respectively). There were some marginal, and often inconsistent, effects of strain by age, strain by colony size, and of tier and colony size. Cage type or its interactions had no effect on shell thickness.

Figures for estimated egg mass per hen day were generally unaffected by the presence of second-quality eggs, so all figures reported here include seconds unless stated. Weekly egg mass was higher for Br birds than W apart from at early and late ages in flock 1, while cumulative egg mass was greater for Br birds than W consistently in flock 2. Despite this, cumulative egg mass was as expected overall for Br hens (actual and expected, both 20.1 kg/hen). W birds performed better (19.2 kg/hen) than expected (17.8 kg/hen) (although not when seconds were excluded, in which case production dropped to 18.0 kg/hen). LMMs were fit to the estimated total cumulative egg mass. Despite early poor performance in flock 1, total cumulative egg mass per hen was still significantly greater with Br (20.1 kg/hen) than W (19.2 kg/hen) birds (SED=0.11, P<0.001). Total cumulative egg mass increased with decreasing colony size for cage type B only, with total cumulative egg mass being highest for B-20 (Table 3) (cage type by colony size interaction, P=0.039). Overall, there was no effect of colony size or cage type apart from for when seconds were excluded, in which total cumulative egg mass per hen was greater for cage type B (18.7 kg) than for A (18.2 kg, SED=0.16, P=0.003).

Table 3 Mean total cumulative egg mass per hen (kg) at age 71 weeks (SED=0.192-0.241) for cage type (A or B) and colony size (20, 40 and 80 birds per cage) (from LMM).

	Colony size		
	20	**40**	**80**
A	19.5	19.6	19.5
B	20.0	19.7	19.3

Internal egg contents were assessed fitting LMMs to haugh units (HU) (calculated from data on albumen height) and to yolk colour (Roche). HU decreased with age (as expected) from 105.4 to 85.5 (Br) and 104.1 to 90.1 (W) (SED=0.51-0.76) and was greater overall for W (mean 95.5) than for Br (93.3) (SED=0.19) particularly at later ages (age by strain interaction, age and strain, all P<0.001). HU were slightly greater for cage type A (overall mean 94.8) than B (94.0) at some ages (age by cage type interaction, P=0.005). HU at the final measurement (69 weeks, 85.5 Br and 90.1 W) were greater than expected values (80 Br and 86 W, both at 70 weeks, according to the strain management handbooks). With yolk colour assessments, some data were ignored (from the first two sample periods) because yolks were too small to measure properly using automotive equipment (the subsequent presence of albumen would have reduced the Roche value). From 25 to 41 weeks of age, yolk colour was higher for Br (range 7.0-8.1) than W (range 6.6-7.5, SED=0.127-0.188) (age by strain interaction, P<0.001) but this was highly influenced by flock 1. Yolk colour varied with age (P<0.001), and with tier by age (P=0.015), but in no discernible pattern. Cage type and colony size or their interactions were not significant.

GLMMs were fit to numbers of eggs with blood or meat spots separately as a proportion of eggs assessed (binomial, link logit) with dispersion fixed at 1. The (back transformed) mean proportions of eggs with blood and meat spots were much higher for Br (4% and 13%, respectively) than for W birds (both 0%) (both P<0.001, Wald test). There were some evidence of age affecting blood spots and meat spots in Br birds, with the proportion increasing with bird age, but for blood spots this was only true in flock 1 (age by strain interaction for meat spots, P=0.003, Wald test). All other effects were marginal or non-existent.

Behaviour

GLMMs were fit to the counts of birds observed in different locations or behaviour classes (poisson, link log) with dispersion fixed at 1 with a factor identifying each multinomial table of counts fitted first in the fixed effects and including the class and interactions with class in order to analyse effects on proportions in the different classes

- Daytime (location)

Overall, observed birds spent 52% of their time on the floor, 27% on perches, 13% on the scratch mat, and 7% in the nest box. There were several interaction effects on bird location (all P<0.001 Wald test) of which the most relevant ones are reported here. Location was significantly affected by cage type and colony size, with birds in cage type A spending more time on the perch, scratch mat and nest box, whereas birds in cage type B spent more time on the floor, at all comparable colony sizes (Table 4). Birds in cage type A increased time observed on the scratch mat with increasing colony size, whereas the opposite was true with cage type B. Conversely, birds in cage type B increased time on the floor with greater colony size, and the reverse was true for cage type A. For 20-bird cages, most of these differences were small, and for 40-bird cages with regards to the nest box (i.e. <2.5% difference), whereas for 80-bird cages there was a much larger difference (i.e. >10%) in time spent at the scratch mat and on the floor.

Table 4 Proportion of birds (in %) by location within the cage, according to cage type (A or B) and colony size (20, 40 or 80) (back transformed from log). Notice that some differences between cage types within colony sizes are small, i.e. <2.5% (highlighted), while others are larger i.e. >10% (bold).

	A			B		
	20	**40**	**80**	**20**	**40**	**80**
Perch	27.0	30.0	28.7	25.8	23.9	25.1
Scratch Mat	14.7	16.7	**18.4**	12.9	12.5	**7.0**
Floor	50.5	45.6	**43.5**	55.1	57.3	**61.5**
Nest box	7.8	7.7	9.5	6.2	6.3	6.4

With strain and cage type, Br birds were observed more often on the floor during the daytime than W within both cage types (Table 5). Both strains spent more time on the floor with cage type B. W birds spent more time than Br birds on the perch and scratch mat with both cage types, and in the nest box with cage type A. In cage type B, Br and W birds spent similar amounts of time on each of the resources, while in cage type A, W birds spent more time on resources and less time on the floor than Br birds. Within each strain more time was spent on the perch, scratch mat and nest box with cage type A than type B.

Table 5 Proportion of birds (in %) by location within the cage during the day, according to strain (brown; white) and cage type (A or B) (back transformed from log). Notice that some differences between strains within cage types are small, i.e. <3% (highlighted), while others are larger i.e. >10% (bold).

	Brown		White	
location	**A**	**B**	**A**	**B**
Perch	26.5	24.1	30.5	26.0
Scratch Mat	14.7	9.6	18.5	11.3
Floor	**51.7**	59.6	**41.4**	56.7
Nest box	7.1	6.7	9.6	6.0

There were significant strain and age effects on the percent of time spent in various locations in the cage during the day, but many of these differences (between strains, within age groups) were less than 5% (for example, time observed in the nest box ranged from 6.3-8.3% across all strains and ages). Time observed on the floor differed most between strains within ages, with Br birds spending more time on the floor at all ages (range 53-59.1%) than W birds (range 44.9-51.2%). Overall, during the daytime birds spent more time on the additional cage furniture with cage type A (perches 28.6%, scratch mats 16.6% and nest boxes 8.3%) than with cage type B (perches 25.1%, scratch mats 10.5% and nest boxes 6.3%), and more time on the floor with cage type B (58.2%) than cage type A (46.5%). Other effects of colony size and tier were marginal, and are thus not reported on further.

- Daytime (behaviour activity)

With behaviour activity, birds spent most observed time standing followed by feeding, with less than 1% of time spent in each of non-aggressive peck, object peck, sham dust bath or aggressive peck (Table 6).

Table 6 Overall time (in %) spent in various behaviours, back transformed from log (non agg = non-aggressive, agg = aggressive).

stand	**sit**	**walk**	**feed**	**drink**	**non agg peck**	**peck/ scratch mat**	**object peck**	**sham dust bathe**	**preen**	**agg peck**
52.9	8.1	2.8	23.5	5.6	0.8	3.6	0.5	0.1	2.0	0.1

Of the 3-way interactions (Wald test, all P<0.001), several were statistically significant and the most important ones are reported here. There were significant interactions between strain and cage type with behaviour activity (Figure 9). Of the time observed, Br birds showed more feeding, and less sitting and walking, than W birds. Birds in cage type B spent more time standing and less time drinking and pecking/scratching the mats than in A cages. Preening was greatest with W birds in cage type B, and lowest with Br birds in cage type A. Object pecking was lowest with W-A birds, and of similar levels with all other strain/cage types. Differences between time observed in non-aggressive pecking (0.7-0.9%), sham dustbathing (all 0.1%) and aggressive pecking (0.0-0.01%) were minimal.

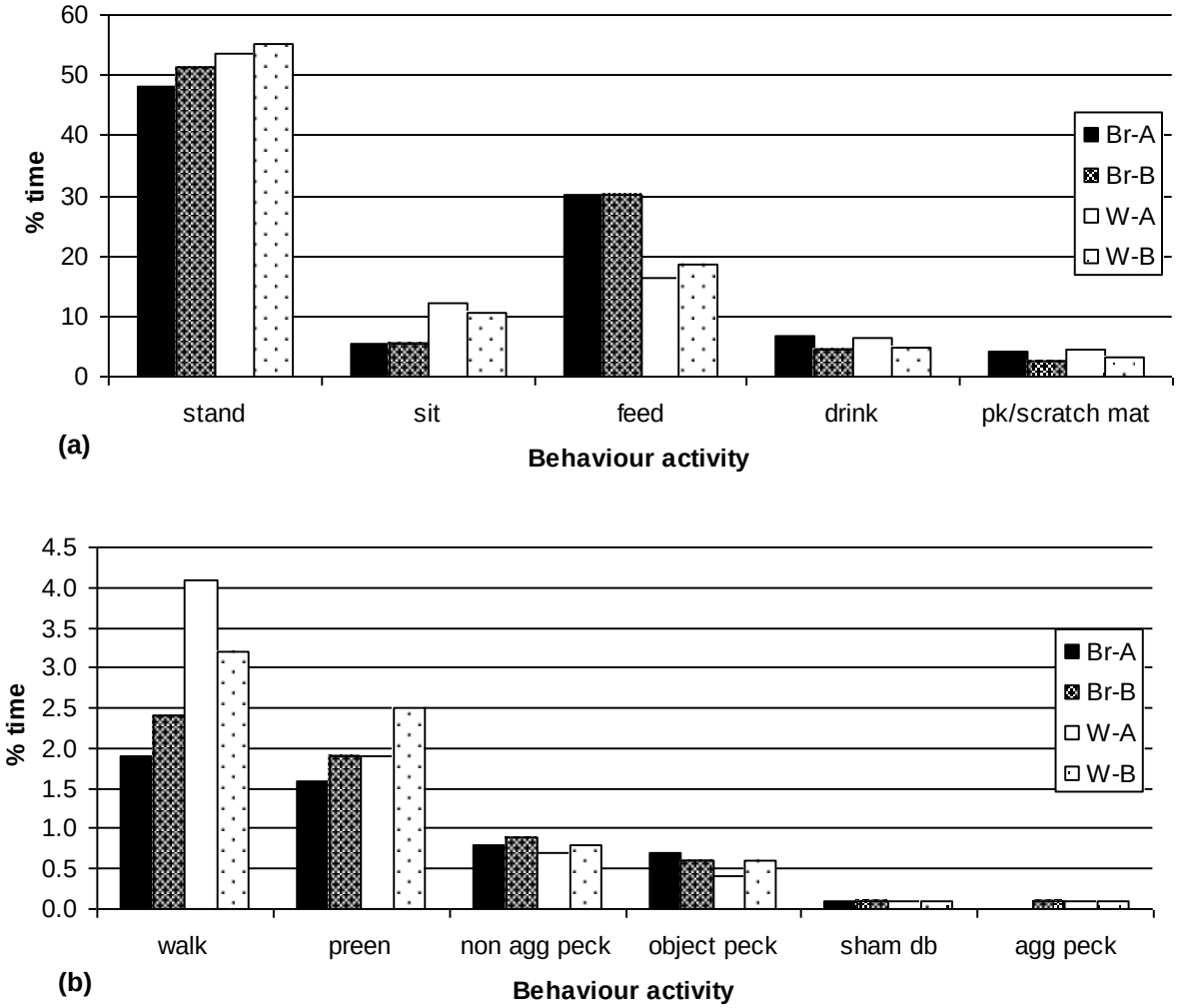


Figure 9 Percent time (back transformed from log) in various behaviour activities during the daytime, according to bird strain (Br = brown, W = white) and cage type (A or B). (a) shows more common behaviours, while (b) gives rarer behaviours (note different y-axis values).

The percent time observed in various behaviour activities also differed significantly with time of day and colony size (Table 7). Standing decreased throughout the daytime with all colony sizes, but increased overall with greater colony size. Sitting peaked during the middle of the day, but decreased overall with greater colony size. Time spent in both feeding and drinking increased slightly over the course of the day, and in both cases were highest at the end of the daytime with colony size 20. Walking behaviour was fairly consistent across both colony sizes and times of day. Pecking and scratching at the mat was lowest in 80-bird colonies at all times of day. All other behaviours occurred at very low levels. Non-aggressive and object pecking approximately doubled at the end of the day compared to other times of day, and sham dustbathing peaked during the middle of the day, with all colony sizes. Preening was observed in a greater proportion of birds in 20- and 40-bird colonies than 80-bird colonies. Aggressive pecking was rarely observed.

Table 7 Time spent (%) in various behaviours (back transformed from log) according to time of day (start, middle, and end) and colony size (20, 40, and 80-bird cages). non agg = non-aggressive, agg = aggressive.

		stand	sit	feed	drink	walk	peck/ scratch mat	non agg peck	object peck	sham db	preen	agg peck
20	start	52.9	9.2	21.8	5.4	2.7	4.2	0.6	0.4	0.0	2.7	0.0
	mid	48.5	12.2	22.7	6.7	2.7	3.5	0.7	0.4	0.3	2.5	0.0
	end	45.9	5.2	26.9	9.4	3.3	4.9	1.3	0.9	0.1	2.0	0.1
40	start	55.4	8.7	21.4	4.1	2.9	4.0	0.6	0.5	0.1	2.3	0.1
	mid	51.7	11.8	21.8	5.0	2.6	3.4	0.7	0.4	0.3	2.1	0.1
	end	50.5	5.2	24.7	6.9	3.1	4.7	1.5	1.0	0.1	2.1	0.1
80	start	57.0	8.7	22.8	3.8	2.5	2.7	0.5	0.4	0.1	1.5	0.0
	mid	55.1	10.5	22.2	4.5	2.6	2.5	0.6	0.3	0.2	1.5	0.0
	end	54.9	4.6	24.9	6.3	2.9	3.1	1.2	0.8	0.0	1.4	0.1

The largest overall effects on percent times observed in various behaviours were time of day and strain (both P<0.001, Wald test), which have been discussed above. Behaviours also changed with bird age (P<0.001, Wald test), with drinking and walking generally decreasing with age, sitting and preening generally increasing with age, and pecking/scratching at the mats at its highest at peak production (i.e. 25-35 weeks of age).

- Night time (location)

Birds were observed during the night-time in order to assess if perch use differed with cage type, colony size or bird strain. Overall, approximately 90% of birds observed were on the perches, with 10% on the floor. When comparing bird location according to start, middle and end scans and bird age periods (i.e. pre-peak, peak, and post-peak) (Table 8), the proportion of birds on the perches increased (compared to the floor) with age in the middle and end of the dark period, but with the first scan after lights off (i.e. 'start') the proportion of birds initially increased, then decreased with bird age (age by scan interaction, P<0.001, Wald test). Within each age period, proportion of birds perching was highest at the start and middle of the night, and lowest at the end of the night. All other interactions were insignificant. Overall, the proportion of birds perching increased with bird age (P<0.001, Wald test) and was higher in the middle, compared to beginning and end, scans (P=0.006, Wald test). All main other effects were insignificant (proportion of birds perching according to strain: Br =89%, W =91%; cage type: A =92%, B =88%; colony size 20 =92%, 40 =90%, 80 =88%).

Table 8 Proportion of birds (in %) on the perches and on the floor (back transformed from log) during pre-peak (17-24 weeks), peak (25-35 weeks) and post-peak (37-71 weeks) ages, according to night-time scans that took place during the start, middle, or end of the dark period.

	Pre-peak			Peak			Post-peak		
	start	**middle**	**end**	**start**	**middle**	**end**	**start**	**middle**	**end**
Perch	86.8	88.9	79.9	91.9	92.2	84.4	88.6	97.2	91.2
Floor	13.2	11.1	20.1	8.1	7.8	15.6	11.4	2.8	8.8

- Night time (behaviour activity)

There were significant differences in the behaviour of birds observed during the night according to scan and age periods (Table 9) (P<0.001, Wald test). Within all ages, the proportion of birds standing increased across night-time scan. Conversely, sitting decreased across scans, apart from during the post-peak period when sitting was highest in the middle of the night. Birds observed spent the most time preening at the end of the night, followed by the start of the night, apart from during peak when birds spent the most time preening at the end and middle of the night. Proportions of time spent walking were greatest at the start and end of the night. Overall, there was a significant effect of strain on behaviour (P=0.007, Wald test) with W birds doing more preening (4.5%) than Br birds (2.8%), but all other strain-behaviour effects were small. Other effects of cage type or colony size were not significant, although the proportion of birds observed standing tended to decrease (from 46% to 37%), and sitting tended to increase (from 50% to 59%), with increasing colony size.

Table 9 Proportion of birds (in %) performing various behaviours (back transformed from log) during pre-peak (17-24 weeks), peak (25-35 weeks) and post-peak (37-71 weeks) ages, according to night-time scans that took place during the start, middle, or end of the dark period.

	Pre-peak			Peak			Post-peak		
	start	**middle**	**end**	**start**	**middle**	**end**	**start**	**middle**	**end**
stand	56.9	61.8	60.6	33.8	37.8	40.6	23.4	23.8	32.5
sit	36.5	35.6	30.5	63.0	58.6	47.0	67.0	74.8	61.0
walk	1.4	0.1	1.0	0.9	0.2	1.2	3.8	0.0	1.0
preen	4.9	2.4	7.2	1.9	3.1	10.4	1.8	1.3	4.4
other	0.4	0.2	0.7	0.4	0.3	0.8	4.0	0.0	1.1

Indicators of stress and immune response

LMMs were fit to H/L ratios (log+1). Flock 2 had higher heterophils (%) and consequently higher H/L ratios than flock 1, and H/L ratios differed with age (P<0.001), however the pattern of change with age was slightly different between the two flocks (Figure 10). There were no clear or consistent effects of strain, cage type, tier or colony size.

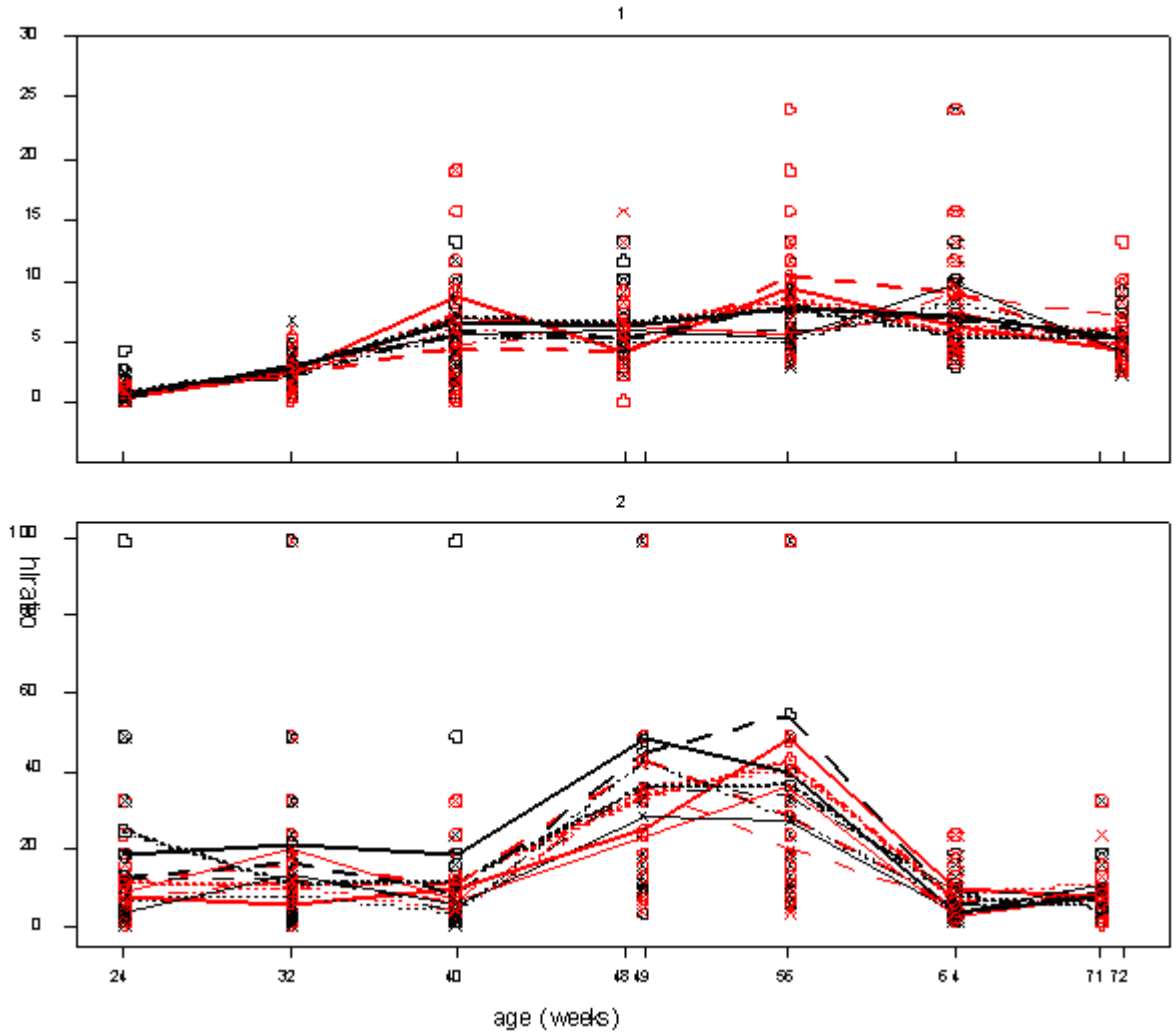


Figure 10 Heterophil/lymphocyte (H/L) ratio from blood smears taken at 8-weekly intervals in flock 1 (top) and flock 2 (bottom). Br birds = black lines, W birds = red lines; cage type A = thin line x, cage type B thick line o; colony size 20 = solid line, 40 = small dash (……) 80 = large dash (----).

LMMs were fit to IBD titre (log+1000) and ND titre (log–6). GLMMs were fit to numbers positive as a proportion of birds tested (binomial, link logit) with dispersion fixed at 1. In assessing blood samples for titres to the relevant vaccine, neither vaccine type in either flock revealed significant differences (Table 10). There were more negative titre results with IBD (up to 42% of samples per treatment), compared to no negative titre results with ND.

Table 10 Mean titre values and standard deviation (SD) from blood samples taken at 40 weeks of age to IBD (flock 1) and ND (flock 2) vaccination at 35 weeks of age.

strain	**Brown**						**White**					
cage type	**A**			**B**			**A**			**B**		
colony size	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**
IBD	1325	1388	1075	899	1214	1054	2026	1328	1469	1670	1252	1283
SD	846	956	1061	893	1028	1138	1245	1323	771	912	1321	1040
ND	9.2	8.9	8.8	8.2	8.9	8.5	8.6	8.8	8.7	8.5	8.5	8.9
SD	1.4	0.7	0.9	0.4	1.5	0.8	1.1	0.6	1.0	0.5	0.7	0.8

With tonic immobility tests, 18% of birds sampled stayed down for the maximum 600 seconds. Mean time spent in tonic immobility by treatment ranged from 209–343 sec (SD 191-243) and was higher for W birds (range 252-343 sec) than for brown birds (range 209-281 sec). GLMMs were fit to those birds censored at 600 seconds as a proportion of birds tested (binomial, link logit) with dispersion fixed at 1 and there were no significant effects of tier, strain, cage type or colony size. Survival analysis was applied to tonic immobility to allow for censoring;

PHMs were fit with fixed effects flock + tier*strain*cagetype*colonysize including interactions up to 3 way. Only flock was significant (P=0.007, approximate χ^2 test of the deviance), with tonic immobility higher for flock 1 (Figure 11).

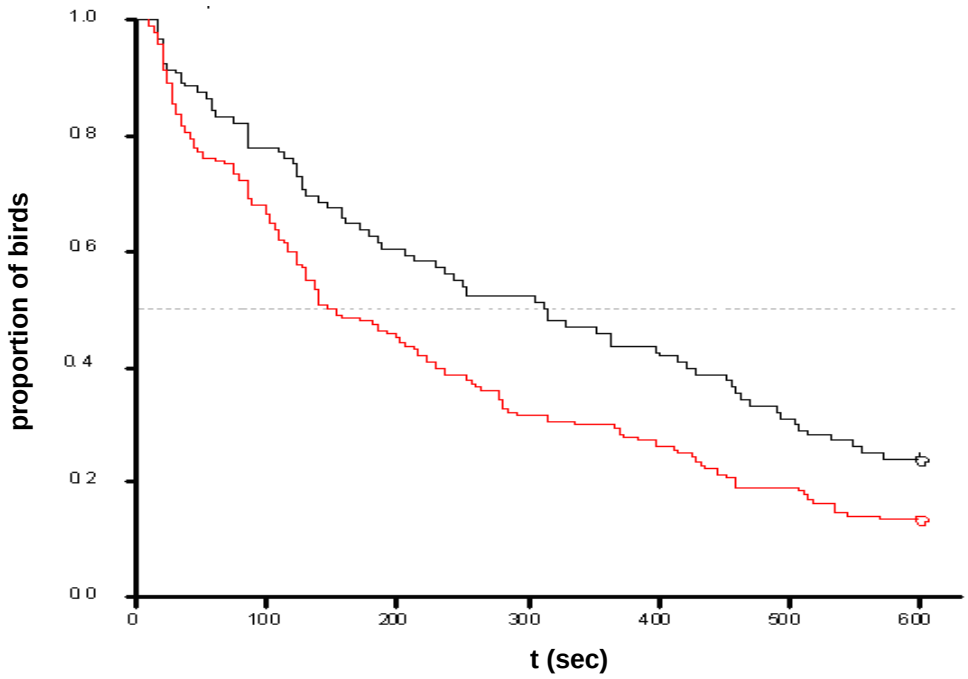


Figure 11 Kaplan-Meier plot, showing proportion of birds that were in tonic immobility (i.e. still on their backs) at time t according to flock (black = flock 1; red = flock 2). The test was stopped (censored) at 600 seconds (i.e. 10 min).

Feather, comb, claw and foot condition

LMMs were fit to feather scores including the effect of body site along with the other fixed effects. Interactions with body site were significant but since patterns were essentially similar for the different sites only main effects are reported. Feather scores increased (i.e. got worse) with age (Table 11), and were higher for Br (1.7) than for W birds (1.0) (SED=0.038-0.039 age, 0.065 strain, both P<0.001). Feather scores were higher with increasing colony size (20 =1.0, 40 =1.3, 80 =1.7, SED=0.082) and moderately higher with cage type A (1.5) than B (1.2) (SED=0.065) (both P<0.001). Feather scores were highest on the back and belly, followed by the tail, compared to other body sites (P=0.006), and this was consistent with both strains. There were also effects of tier, with feather scores being generally best from birds in the bottom tier (1.2 versus 1.5 in M and 1.4 in T tiers, SED=0.082) (P=0.004).

Table 11 Mean feather scores (scores 0-5) and standard deviation (SD) averaged over 6 body sites (neck, back, tail, wings, belly) at three ages, according to bird strain, cage type and colony size. Ten birds per cage were sampled remotely at 35-36 and 55-56 weeks, and 5 birds per cage were sampled with handling at 72 weeks of age*.

	strain	**Brown**						**White**					
	cage type	**A**			**B**			**A**			**B**		
	colony size	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**
35-36 weeks	mean	0.4	0.5	0.7	0.2	0.8	0.8	0.2	0.3	0.4	0.2	0.1	0.4
	SD	0.5	0.6	0.7	0.4	0.7	0.8	0.4	0.4	0.4	0.3	0.2	0.5
55-56 weeks	mean	1.7	2.2	2.8	0.7	2.1	2.5	1.1	1.3	1.7	0.6	0.5	1.4
	SD	1.1	1.1	0.8	0.9	0.9	0.9	0.7	0.8	0.7	0.8	0.4	0.8
72 weeks	mean	2.6	2.7	3.2	1.8	2.6	3.1	1.8	2.2	2.5	1.3	1.3	1.9
	SD	0.9	0.9	0.7	0.9	0.8	0.7	0.6	0.6	0.6	0.6	0.5	0.6

* For the 55-56 week sample, one cage (Br-40-A) was depleted at 52 weeks of age, and was scored then, and in one cage (W-40-B) only five birds were scored instead of 10; at 72 weeks of age both cage types A and B had one missing replicate of Br-40 due to earlier culling.

LMMs were fit to claw length. There was a significant difference with strain regardless of cage type, with W hens having longer claws than Br hens (P<0.001, see Figure 12) and Br-A had longer claws (1.45 cm) than Br-B (1.10 cm), whereas W birds did not differ with cage type (2.10 and 2.12 cm, SED=0.06) (strain by cage type interaction, P<0.001). For 20-bird colony cages, claws were longer in cage type A (1.85 cm) than B (1.52 cm) whereas for the other colony sizes there were no differences in claw length for the two cage types (SED=0.07) (cage type by colony size interaction, P=0.03). There were no flock or tier effects.

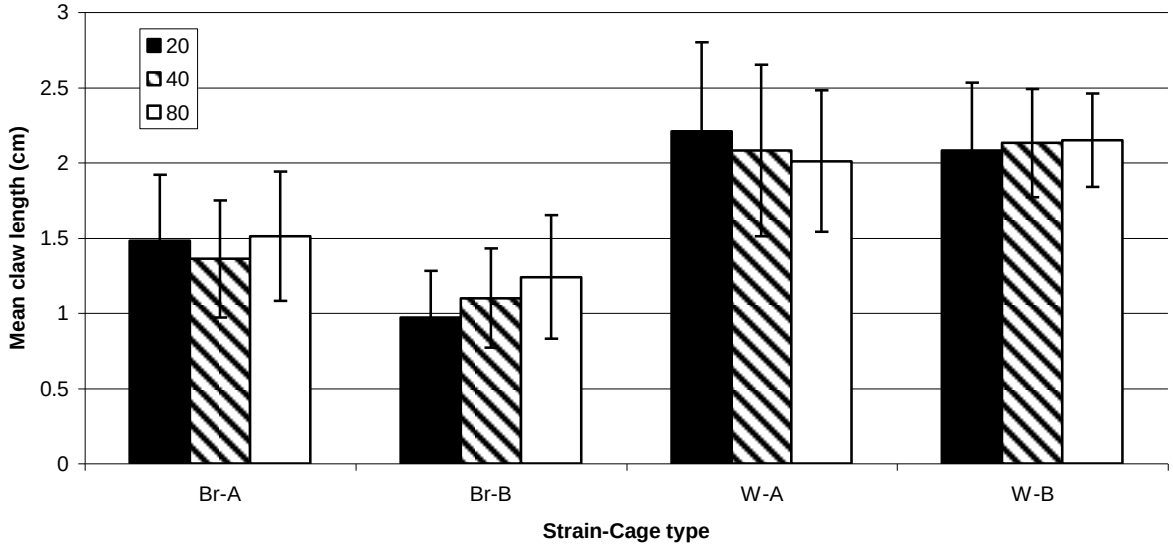


Figure 12
Mean middle digit claw length (cm) and standard deviation (SD) according to strain (Br = brown, W = white), cage type (A or B) and colony size (20, 40, or 80 birds/cage).

GLMMs were fit to the counts of comb scratches (poisson, link log) (Table 12). Scores ranged from 0 to 47 scratches. Birds from flock 2 had more scratches than flock 1. W birds had significantly more mean comb scratches/bird (5.6) than Br birds (1.9) (P<0.001). Colony size had a significant effect on comb scratches/bird (P=0.003) with more mean comb scratches for colony size 20 (4.0) than for colony size 80 (2.8) or 40 (3.1) but no difference between 40 and 80 colonies. There were no significant interactions or effects of tier or cage type.

Table 12 Mean number of comb scratches and standard deviation (SD), measured at 72 weeks of age, according to bird strain, cage type and colony size, over two flocks.

strain	**Brown**						**White**					
cage type	**A**			**B**			**A**			**B**		
colony size	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**
mean	2.8	2.2	2.1	2.7	1.9	1.5	8.5	6.8	6.5	8.1	6.3	6.1
SD	2.7	2.5	2.2	2.7	1.8	1.4	9.7	6.6	4.8	6.7	5.2	5.4

GLMMs were fit to foot pad scores (poisson, link log) (i.e. numbers of lesions per bird, see Table 13) which ranged from 0-4, with 63-87% of birds assessed per treatment having a foot score of 0, except for W-B-20 in which only 48% of birds assessed had footpad score 0. Mean foot pad scores were higher for W-20 (0.5) and Br-80 (0.4) than for Br-20 (0.2) while W-20 had higher foot pad scores (0.5) than W-40 (0.2) (strain by colony size interaction, P=0.003). There was a large effect of cage type (P<0.001, from model with significant 2-way interactions and all main effects included) with higher foot pad scores for cage type B than A (0.4 versus 0.2, respectively). There were no significant effects of tier. However, in all cases, foot pad damage was low.

Table 13 Mean number of foot pad lesions per bird and standard deviation (SD), measured at 72 weeks of age, according to bird strain, cage type and colony size, over two flocks.

strain	**Brown**						**White**					
cage type	**A**			**B**			**A**			**B**		
colony size	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**
mean	0.2	0.3	0.5	0.5	0.5	0.6	0.4	0.2	0.3	0.8	0.4	0.5
SD	0.5	0.7	0.9	0.8	0.8	0.9	0.6	0.5	0.6	0.9	0.7	0.6

Bone strength and fractures

LMMs were fit to leg bone breaking strength (log) (Table 14), for which 4.4% of tibia bones tested reached the censor point of 300 N. Leg bones were stronger in flock 1 than flock 2. Strain was highly significant (P<0.001) with substantially higher breaking strength for W birds (193.8 N) than for Br birds (171.4 N) (Table 14).

Table 14 Mean leg bone breaking strength (N) and standard deviation (SD), measured at 71 weeks of age, according to bird strain, cage type and colony size, over two flocks.

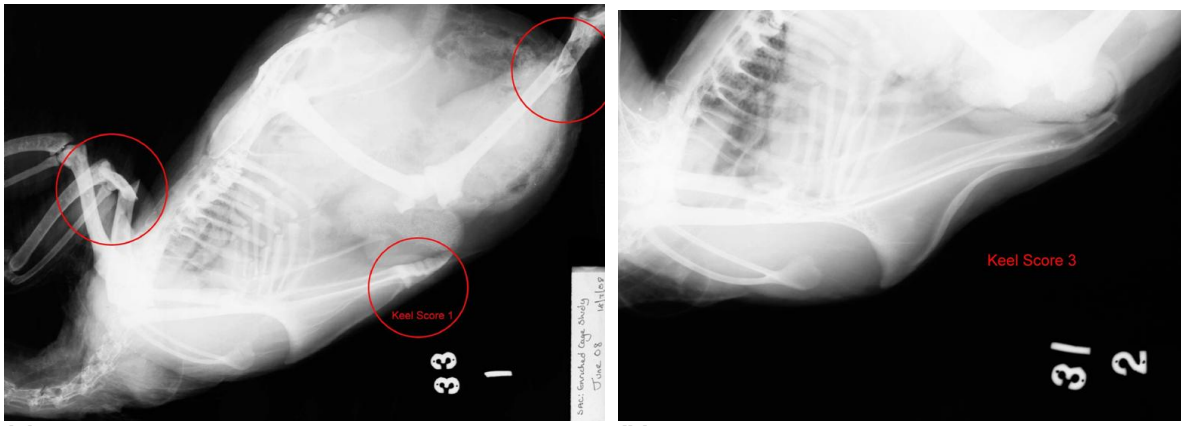
strain	Brown						White					
cage type	A			B			A			B		
colony size	20	40	80	20	40	80	20	40	80	20	40	80
Mean leg strength (N)	173.1	182.0	185.7	172.1	180.8	180.9	192.4	210.2	215.3	193.6	191.3	195.1
SD	48.7	54.7	59.2	48.1	51.5	49.8	42.5	44.4	53.9	53.4	43.3	45.3

Mean keel condition score was low (i.e. good) across all treatments, but in the interaction between strain, cage type, and colony size, 15-38% of birds sampled had some level of keel bone damage (i.e. condition score >0, see Table 15). All keel bone damage was old. GLMMs were fit to birds with positive keel condition scores (score >0) and to birds with keel fractures (score >1), as a proportion of those assessed (binomial, link logit) with dispersion fixed at 1. There were no significant effects of strain or tier on positive scores and other effects were inconsistent. Keel bone condition improved with increasing colony size for cage type A (20=36%, 40=35%, 80=23% scores >0) but got worse for cage type B (20=12%, 40=25%, 80=27% scores >0) (cage type by colony size interaction, P=0.024); B-20 had better keel bone condition than A-20 (12% compared to 36%) and B-20 was better than B-40 and B-80 (12% compared to 25% and 27%, respectively). There were no significant effects on the percent of birds with keel fracture (i.e. scores >1).

Table 15 Mean keel bone condition score (based on scores of 0-3) and standard deviation (SD). Also, percentage of birds with: keel bone condition score >0 (indicating some level of damage); keel fracture (i.e. score >1), fracture to the furculum, wings, legs, or any fracture (keel, furculum, wings, legs).

strain	Brown						White					
cage type	A			B			A			B		
colony size	20	40	80	20	40	80	20	40	80	20	40	80
Mean keel bone condition score	0.5	0.3	0.4	0.2	0.3	0.3	0.5	0.6	0.3	0.1	0.3	0.4
SD keel condition score	0.8	0.6	0.7	0.6	0.6	0.6	0.7	0.8	0.7	0.4	0.6	0.7
% KB condition score > 0	37.5	27.3	29.2	16.7	25.0	22.9	35.4	45.8	18.8	14.6	27.1	33.3
% keel fracture	12.5	2.3	8.3	6.3	9.1	6.3	10.4	10.4	8.3	0	6.3	10.4
% furculum	2.1	0.0	4.2	0.0	4.5	0.0	0.0	2.1	0.0	0.0	0.0	0.0
% wing	2.1	9.1	0.0	4.2	6.8	6.3	2.1	2.1	2.1	2.1	4.2	0.0
% leg	0.0	4.5	0.0	0.0	2.3	0.0	0.0	4.2	0.0	2.1	0.0	4.2
% any fracture	16.7	13.6	12.5	10.4	18.2	10.4	12.5	18.8	10.4	4.2	8.3	14.6

GLMMs were fit to birds with any fracture (keel, furculum, wing and leg) as a proportion of those assessed (binomial, link logit) with dispersion fixed at 1. For Br birds there were more fractures the lower the tier (T = 9%, M = 12%, B = 18%) whereas for W there were more fractures in the higher tiers (T = 15%, M = 10%, B = 6%) (tier by strain interaction, P=0.032); there were more fractures for Br (18%) than W (6%) birds in the bottom tier. For samples of bone damage, see Figure 13.



(a) **(b)**

Figure 13 Radiographs of hens examined at 71 weeks. (a) Hen with new humerus (left circle) and tibia (right circle) fractures, and keel bone score 1. (b) Hen with keel bone score 3.

Mortality

Of 6720 hens over two flocks, 518 birds (7.7%) died or were culled, 223 (6.6%) in flock 1 and 295 (8.8%) in flock 2. This is higher than the levels suggested by the strain management handbooks (i.e. 4% Br, 5% W to 80 weeks). In flock 2, two cages (one of each cage type, both Br-40) were depleted early due to feather pecking/cannibalism (cage 9 at 52 weeks of age, by which time 9 birds had already been culled), or aggression/bullying (cage 45 type B, at 64 weeks of age, by which time 13 birds had been culled). Interestingly, both were top-tier cages directly opposite one another (see Figure 1b). The 22 birds culled one at a time from these cages are included in the analyses; the remaining 31 and 27 birds that were culled at once are not. Ignoring these bulk culls, 460 birds (6.8%) died/were culled overall.

GLMMs were fit to the number of mortalities as a proportion of colony size (binomial errors, link logit) with fixed and random factors included. There were significant effects on mortality counts but these were mainly due to the large mortality (i.e. 22 birds) in the two subsequently culled cages in the second flock, which had an overriding influence on the results of the analyses, rather than to any consistent trend over all. Other than that, there was some suggestion that mortality tended to be higher for Br (5.8% of all Br birds) than for W (5.5% of all W birds) hens (P=0.048), particularly with classes 'aggression/bullying' and 'not eating' (which may in fact be due to bullying) and for cage type B (6.5 % of all birds housed in B) than A (4.9% of all birds housed in A) (P=0.023) (Table 16).

Table 16 Mean percent mortality per treatment (i.e. strain-cage-colony size) across two flocks, excluding mortality due to culling of entire cages (one Br-A-40, one Br-B-40), and relative proportion of mortality type. 'Other' included symptoms such as prolapse, egg peritonitis, dropped crop, and swollen gut.

strain	**Brown**						**White**					
cage type	**A**			**B**			**A**			**B**		
colony size	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**	**20**	**40**	**80**
% mortality (of all birds in that treatment):												
	2.5	5.8	6.7	5.0	11.9	9.9	9.6	5.0	5.0	4.6	5.8	6.7
Mortality type (by proportion of % mortality):												
aggression/bullying	0.33	0.18	0.20	0.08	0.12	0.23	0.00	0.00	0.02	0.00	0.04	0.00
found dead, not sure	0.33	0.29	0.30	0.42	0.35	0.45	0.30	0.42	0.35	0.09	0.18	0.30
feather pecking/cannibalism	0.00	0.25	0.19	0.08	0.39	0.13	0.30	0.13	0.19	0.36	0.54	0.44
lame	0.00	0.04	0.06	0.17	0.05	0.07	0.13	0.04	0.15	0.27	0.07	0.11
not eating	0.00	0.21	0.11	0.17	0.05	0.05	0.09	0.21	0.04	0.00	0.04	0.00
other	0.33	0.04	0.14	0.08	0.04	0.06	0.17	0.21	0.25	0.27	0.14	0.16

GLMMs were fit to the multinomial counts of the six mortality types (poisson, link log, adjusted for total mortality counts) with dispersion fixed at 1 in order to analyse effects on proportions of the different types (relative to total mortality). Additional GLMMs were fit to each mortality type as a proportion of total mortality and of colony size (binomial, link logit) with dispersion fixed at 1. About 60% of all mortalities fell into the classes 'found dead/not sure' or 'feather pecking/cannibalism'. Mortality due to feather pecking/cannibalism occurred earlier, and mortality due to aggression/bullying occurred later, in the flocks' lifetimes compared to other types of mortality. More Br birds died from aggression/bullying (18.0% of all mortality for Br birds, P<0.001 both relative to total mortality and relative to total numbers that could have died) and from not eating (10.4% of all mortality, P=0.035 relative to total mortality, P=0.043 relative to total numbers that could have died), compared to W birds (1% aggression/bullying and 4.4% not eating, of all mortality for W birds). A larger percentage of W bird mortalities were due to 'other' causes (22.1%) compared to Br birds (8.7%) when assessed out of total mortality (P=0.004) and also compared to total possible numbers of birds (P=0.04), whereas 'feather pecking/cannibalism' was significantly higher in W birds (32.9%) compared to Br birds (18.4%) when compared to total mortality only (P=0.008). Therefore, there were not more feather pecking/cannibalism deaths in W birds overall. There was a cage type effect on the different mortality types (mortality type by cage type interaction, P=0.008) with a higher proportion of deaths due to feather pecking/cannibalism in B cages (34% of all mortality) than A cages (20% of all mortality) relative to total mortality (P=0.017) and also to total possible numbers of birds (P=0.024). There was no evidence of colony size or tier effects on mortality.

Beak trimming

Birds in 26 cages (18% of all cages, but 23% of all birds) were beak trimmed during the study, with 10 cages in flock 1 and 16 cages in flock 2. GLMMs were fit to the whether a cage was beak trimmed or not (binomial errors, link logit) with fixed and random factors included. Survival analysis was applied to the age at first beak trim with censored ages for non beak trimmed cages set to age at end of flock. PHMs were fit with fixed effects flock + tier + strain + cage type + colony size + (tier by cage type).

Both the GLMM and the survival analyses implied more (27% of cages) and earlier beak trimming for colony size 80 than for 40 (9% of cages) or 20 (7% of cages) colony sizes (P=0.022, Wald-test, from model with significant two-way interactions and all main effects included; P=0.018 approximate χ^2 test of the deviance, respectively) (Figure 14, right-hand graph). The survival analysis suggested earlier beak trimming for W than Br strain (P=0.004, approximate χ^2 test of the deviance) (Figure 14, left-hand graph) but the GLMM did not suggest a significantly higher proportion of W bird cages beak trimmed (17% W, 9% Br, but back-transformed values).

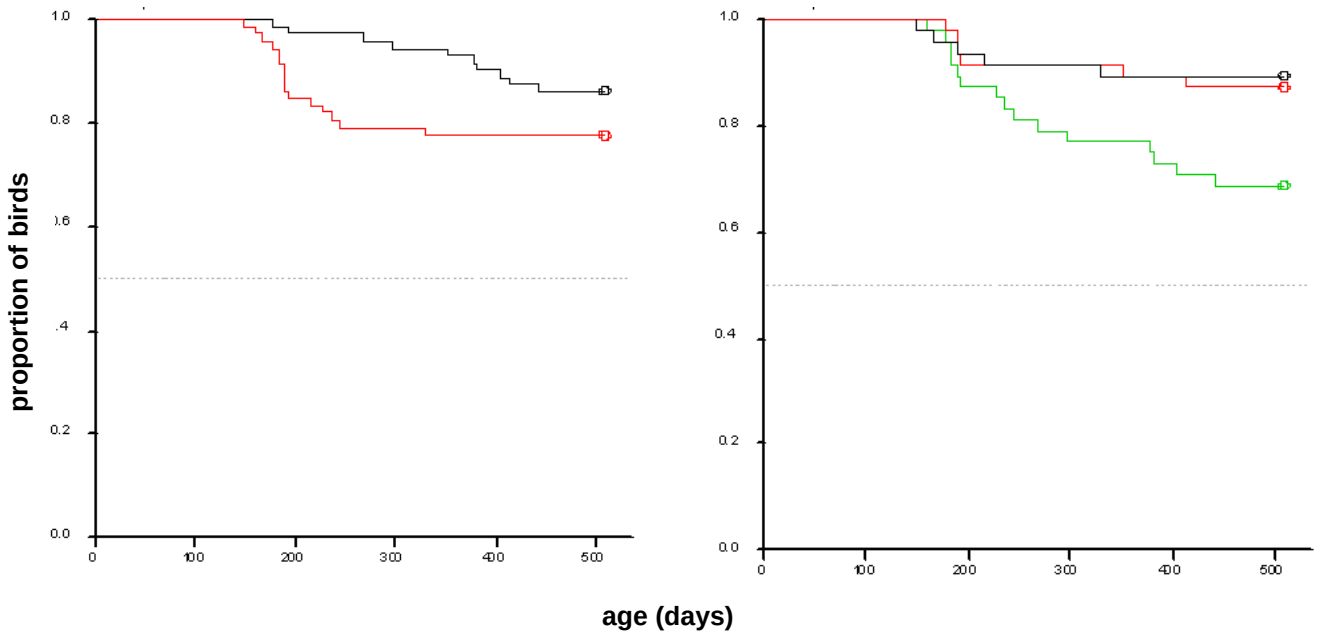


Figure 14 Kaplan-Meier survival plots, showing the proportion of birds that were not beak trimmed at any given age (in days) according to strain (left, black = Br, red = W) and colony size (right, black =20, red =40, green =80 bird cages).

Discussion of results

Before embarking on this project, considerable time was spent in discussion with the cage manufacturers regarding cage layout. Ideally, a bank would have contained all replicates (in other words, containing both cage types), but joining up two different cage makes was impossible. Per flock, there were six replicates of every treatment (strain x cage type x colony size), but in order to control for flock and cage type position (i.e. bank) effects, we ran the second flock with the banks of cages in alternate positions.

Cage type

Cage type alone had a few effects on parameters measured here. Body weights were higher in cage type B compared to A, despite there being slightly less available food trough space available in B due to the nest box position, but the effect was marginal. Cage type alone did not result in different times observed feeding. Cage type B had less scratch mat space per bird (Table 1) compared to A, and also did not necessarily increase scratch mat space per bird with colony size (50 cm^2/bird for 20- and 40-bird cages, 25 cm^2/bird for 80-bird cages). A lower proportion of birds were observed on the scratch mats with cage type B compared to cage type A particularly in 80-bird cages (Table 4 and 5), and pecking/scratching at the mats was lowest in 80 bird colonies (both logical given the lower space/bird with cage type B), yet there was less feather damage overall (as an indicator of feather pecking) in cage type B. However, there were more feather pecking/cannibalism deaths with cage type B than A. Since feather pecking is thought to be related to redirected foraging behaviour (Blokhuis, 1986), enabling more birds to express foraging behaviour on the scratch mat (by providing more space) may have been beneficial to behaviour and reduced feather pecking/cannibalism-related deaths in cage type A. However, time observed at the pecking/scratching mat did not appear to differ with cage type overall (Figure 9a).

Observed feather pecking behaviour (incorporated into the behaviour class 'non-aggressive peck) was very low in this study (less than 1%) and did not appear to differ with cage type. There was no effect of cage type overall on egg production, although there was tendency for egg production to decrease with increasing colony size with cage type B: however, this may have been unduly influenced by egg eating. Footpad lesions were lower with cage type A, where birds also spent less time on the wire floors and more time on the perches, scratch mat and nest boxes compared to cage type B (see Tables 4 and 5), and higher with Br-80 than Br-20 cages. Footpad lesions are more likely to develop with more time standing on wire floors (Appleby *et al.*, 1993). Behaviour indicated that birds spent more time standing with increased colony size (Table 7), but not necessarily on the wire. The greater proportion of birds observed seen on the furniture (i.e. nest boxes, scratch mats, and perches) with cage type A is seen as beneficial to the hen, because it indicates that the furniture are fulfilling hen's needs,

and over a long period of time. However, no comparable advantages were seen with keel bone damage or bone strength with cage type.

Claw lengths were shorter in cage type B with Br hens, and in cage type B with 20-bird colony sizes. Previous work has indicted that claws are shorter where the shortener is larger (Elson, 2003), as is the case with cage type B compared to A. Mortality was higher in cage type B, but the reasons why are not clear. There was a striking difference in nest box usage with cage type, with W birds preferring type A and Br birds preferring B. Nest box design can affect their use (Appleby *et al.*, 1993), and in particular if nest boxes are similar in design to the scratch mat (as with cage type A), this may reduce nest box use (because birds also use the scratch mat for laying). Why this should affect one strain (Br) in particular is unclear. Conversely, W birds appeared to prefer the nest box design in A, which was noted to be less enclosed and provided more space per hen than design B. It may be that W hens have an aversion to the enclosed space in cage type B nest boxes, making them less likely to use them. Reduced nest box use may not necessarily influence egg quality *per se* (for example, if hens lay eggs on the scratch mats instead), but indicates design suitability to hens' needs. Overall, however, Br birds were better nest box users, which may indicate their greater malleability in considering what makes a suitable nest site.

Bird strain

Both W and Br strains of non-beak trimmed hens were used to assess if one strain coped better with the enriched cage environment than another. The immediate observation was that white hens were much more flighty to human presence throughout the lifetimes of the two flocks, which was corroborated by more time observed in walking with this strain. Data collected at end of lay showed that they were more fearful, based on longer tonic immobility times. White birds have been shown previously to exhibit a higher fear state (Roll *et al.*, 2005). White birds also had longer claw lengths, but white birds' claws are stronger and grow faster than brown birds', plus brown birds have been shown to exhibit more scratching behaviour (Van Emous, 2003). Here, there did not appear to be a strain effect on time observed in pecking/scratching at the mat (Figure 9a) however white birds were seen on the scratch mat more frequently than brown birds (see Table 5). It may be that white birds, which were more susceptible to operator disturbance than brown birds, interrupted pecking and scratching behaviour when the observer approached. Data also showed that white birds were less prone to feather damage (despite some initial mortality losses due to cannibalism in flock 1, which was probably related to manual feeding rather than due to the birds *per se*) and had lower mortality. Lower mortality with white birds is in keeping with previous work (for example, Tauson and Holm, 2005; Blokhuis *et al.*, 2007), however it should be noted that mortality may have been higher here due to birds not being beak trimmed (particularly given problem with bullying/aggression/not eating with brown birds). This is also in alignment with the LayWel study results (for an overview, see Blokhuis *et al.*, 2007). White birds also had greater bone strength than brown birds, but similar studies have found the opposite (Riczu *et al.*, 2004; Vits *et al.*, 2005). White hens showed better egg production than brown birds here, however this was largely due to behind-target production in brown hens in particular periods of flock one, which was most likely caused be poor feed quality early on in that flock. Brown hens also had a greater proportion of blood and meat spots than white birds, but levels have been seen to vary between strains, and fell within normal levels (i.e. up to 10% for blood spots, and 3-30% for meat spots, www.thepoultrysite.com, 2008). Despite these, egg production was generally good and second quality eggs were low, compared to other work on enriched cages (mean 7.8%, ADAS 2006) and compared to overall levels of seconds seen at UK packing stations for all egg production systems in 2008 (mean 7.3%, Defra 2009). Clearly, improving the egg collection method to reflect commercial practice would reduce the level of seconds seen here. Based on egg production results, brown hens used the nest boxes more than white hens, and there was a significant difference in strain preferences for nest box design (discussed above). However, time observed in the nest boxes (Table 5) did not reflect the proportions of eggs laid there (Figure 7). Despite this, white birds were observed to spend the most time in nest boxes of design A (which is also where they laid the most eggs), but this is thought to be due to their propensity to scatter to both the nest boxes and the scratch mat when the observer approached. Comb scratches (an indicator of aggression) were higher in white birds, but observed aggression was low (about 0.1%) and could not be related reliably to strain or colony size.

There was no evidence that immune status (which can be compromised when animals experience chronic stress) differed with cage types, colony size, or strain. This may indicate that factors measured here were all equally stressful, or not stressful, to laying hens. The IBD challenge used in flock 1 was discarded for flock 2 due to the poor immune response to the vaccine. We would recommend using ND vaccine for similar studies in future. Feather scores, unsurprisingly, increased with bird age (as damaged feathers are not replaced in most adult hens). Feather damage also increased with bigger colony sizes, as has been found previously (Allen and Perry, 1975; Bilcík and Keeling, 2000), and time spent in non-aggressive pecking was highest in 80-bird colonies. These results may be because perpetrators have more victims in bigger groups. Feather damage was also noted to be quite high overall (by the cage manufacturer representatives, by a commercial egg producer) which is likely to be related to their intact beak status (Croxall *et al.*, 2005). With night time behaviour, the proportion of birds using the perches was high (~90%) with all combinations of strain, cage type and colony size, suggesting that birds make good use of perches at night. Total bone fracture seen here (range 4.2-18.8%) was low compared to previous work assessing damage after depopulation in commercial enriched cages (c.f. 13% new fractures, 30%

old fractures, Sandilands and Sparks, 2006), and there was some indication of less damage with smaller colony sizes.

Reliability
Compared to commercially-run enriched cages, this study had two major drawbacks: in order to measure feed intake, we hand-fed each cage of birds once a day from weighed feed sacks, whereas commercial birds would be fed automatically several times a day. Pulse feeding with fresh feed will help to direct foraging behaviour onto the feed trough and its contents, rather than birds directing pecking at one another. This, along with using birds with intact beaks, could be a contributing factor to the high feather loss seen here. Feeding by hand also meant that food toughs were fuller than they are designed to be, enabling hens to rake feed forward into the cage and onto the droppings belt, thus accounting for a high degree of food wastage. Installing grids over the troughs helped somewhat, based on feed use in flock 2, but was still much higher than expected (Figure 5). The second major drawback was that we collected eggs by hand in order to assess where eggs had been laid, but this resulted in a build-up of eggs in front of the nest box which led to some egg eating, particularly in cage type B (where there was no egg shock wire) and thus egg loss, and more cracked eggs with cage type A. This will have affected the egg production and egg quality figures to a degree.

Main implications
Overall, W birds were more fearful, had overgrown claws, and used the nest boxes less than Br birds, but W birds used enriched cage furniture more, had better feather cover, and died less from aggression and bullying than Br birds. These results are perhaps counterintuitive, as it may be expected that the more fearful bird would perform less well in this type of system. The design of cage type A resulted in better furniture use and fewer feather pecking/cannibalism-related deaths but slightly worse feather damage than cage type B, and nest box designs were favoured by different bird strains. The other cage design-related factors (egg production, egg quality) would probably disappear with automatic feeding and automatic egg collection. Colony size's major effects were on feather scores (which got worse with increasing colony size) and on the requirement to beak trim (more 80-bird cages than 20- or 40-bird cages), plus some effects on use of resources and claw length (which were all better with smaller colony sizes). Thus, overall in this study, 20-bird colony sizes were considered better for bird welfare. Because birds were fed manually once a day, birds were not stimulated to peck at feed as often had they been presented with fresh feed several times a day (as is the case commercially). As a consequence, birds may have redirected the motivation to forage on one another, as indicated by poor feather scores and mortality. This damage would be exacerbated by intact beaks. With the expected ban on routine beak trimming due to take place from 2011, it would be important to carefully manage feed stimulation, and to tackle pecking problems in colony cages swiftly by beak trimming the entire cage where outbreaks (i.e. 2 or more birds removed due to pecking injury) occur. Data seen here might be considered to have been collected in a 'worst case scenario' both from the hen's and egg producer's perspectives. In a commercial setting, producers will use pulse feeding and automatic egg collection, which should reduce feed waste, improve feather quality, reduce pecking-related mortality and reduce second-class eggs and egg losses. Future work should consider other methods of stimulating hens to peck at appropriate objects (such as by increasing the number of presentations of feed onto the scratch mat), ways to prevent aggression, and genetic differences between bird strains with regards to their appropriateness for enriched cages.

Possible future work
A study focussing on ways to reduce damage and injury in intact brown hens in enriched cages is warranted, including the effects of genetics (c.f. white strains) and behaviour. The study should include various ways encouraging use of the scratch mat (i.e. depositing small amounts of feed onto the mat several times a day, greater amounts of feed onto the mat fewer times per day, and greater amounts of feed several times per day), the benefits of frequent presentation of fresh feed (using automatic feeding system), and the addition of other appropriate pecking substrates (i.e. bundles of string, peck-a-blocks, mineral blocks etc.).

Any action resulting from the research (IP, KT)
Two presentations of the results at UK conferences have taken place. With acceptance of this report, a publication will be submitted in a relevant peer-reviewed journal.

References
ADAS 2006. Report on production and egg quality (WP6). The LayWel project. Welfare implications of changes in production systems for laying hens. FP6 European Research Programme.

Allen, J., Perry, G.C., 1975. Feather pecking and cannibalism in a caged layer flock. Br. Poult. Sci. 16, 441–451.

Appleby, M. C., Smith, S. F., Hughes, B. O., 1993. Nesting, dustbathing and perching by laying hens in cages: effects of design on behaviour and welfare. Brit. Poult. Sci. 34, 835-847.

Bilcík, B., Keeling, L.J., 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., 1986. Feather-pecking in poultry: its relation with ground-pecking. App. An. Behav. Sci. 16, 63-67.

Blokhuis, H. J., van Niekerk, T. F., Bessei, W., Elson, A., Guémené, D., Kjaer, J. B., Levrino, G. A. M., Nicol, C. J., Tauson, R., Weeks, C. A., van de Weerd, H. A., 2007. The LayWel project: welfare implications of changes in production systems for laying hens. World. Poult. Sci. J. 63, 101-114.

Bright, A., Jones, T. A., Dawkins, M. S., 2006. A non-intrusive method of assessing plumage condition in commercial flocks of laying hens. An. Welf. 15, 113-118.

Croxall, R., Elson, A., Walker, A., 2005. Effects of beak trimming on laying hens in small group furnished cages. Animal Science Papers and Reports 23, 71-76 (7th European Symposium on Poultry Welfare, Poland).

Defra, 2009 Egg statistics notice (29 January) http://www.defra.gov.uk/foodrin/poultry/statistics/

Elson, H.A., 2003. Positive results from claw shortener monitoring. Poult. World 157 (8):14 and 157 (9):26

Jones, R. B., 1986. The tonic immobility reaction of domestic fowl: a review. World. Poult. Sci. J. 42, 82-96.

Kjaer, J. B., Sørensen, P., Su, G., 2001. Divergent selection on feather pecking behaviour in laying hens (*Gallus gallus domesticus*). App. An. Behav. Sci. 71, 229-239.

Poultry Site (2008) http://www.thepoultrysite.com/publications/1/egg-quality-handbook/29/meat-spots

http://www.thepoultrysite.com/ourbooks/1/egg-quality-handbook/28/blood-spots Both accessed 03-Dec-2008.

Riczu, C. M., Saunders-Blades, J. L., Yngvesson, A. K., Robinson, F. E., Korver, D. R., 2004. End-of-cycle bone quality in white- and brown-egg laying hens. Poult. Sci. 83, 375-383.

Roll, V. F. B., Maria, G. A., Cepero, R., Buil, T., 2005. Ethological parametres and performance of HY Line White and ISA Brown hens when housed in furnished cages. Animal Science Papers and Reports 23, 77-84 (7th European Symposium on Poultry Welfare, Poland).

Savory, C. J., Mann, J. S., 1999. Feather pecking in groups of growing bantams in relation to floor litter substrate and plumage colour. Brit. Poult. Sci. 40, 565-572.

Sandilands and Sparks, 2006, The welfare effects of different methods of depopulation on laying hens. Defra report AW0231.

Tauson, R., Holm, K.-H., 2005. Mortality, production and use of facilities in furnished small group cages for layers in commercial egg production in Sweden 1998-2003. Animal Science Papers and Reports 23, 95-102 (7th European Symposium on Poultry Welfare, Poland). .

Van Emous, R. 2003. Claw length of layers in alternative systems – to measure or not to measure. World Poult. 19, 34-35

Vits, A., Weitzenburger, D., Hamann, H., Distl, O., 2005. Production, egg quality, bone strength, claw length, and keel bone deformities of laying hens housed in furnished cages with different group sizes. Poult. Sci. 84, 1511-1519.

References to published material

9. This section should be used to record links (hypertext links where possible) or references to other published material generated by, or relating to this project.

Full paper publications were not considered appropriate until all replicates, i.e. both flocks, were complete. We published one technical article and two abstracts, with Defra's permission, to inform people of the work that was ongoing, during the lifetime of the project:

Sandilands, V., 2006. From enriched cages to organic pullets – poultry welfare research. The Scottish Farmer, 13 May 2006, p10.

Sandilands, V., McDevitt, R. M., Sparks, N. H. C., 2007. Effects of enriched cage design and colony size on production, health and welfare in two strains of laying hens. Brit. Poult. Sci. 3, 16-17. Presented at WPSA UK branch meeting, April 2007.

Sandilands, V., Baker, L., and Brocklehurst, B., 2009. The reaction of brown and white strains of hens to enriched cages. Brit. Poult. Sci. (in press). To be presented at WPSA UK branch meeting, April 2009.