Graduate School for Cellular and Biomedical Sciences

University of Bern

Examining spatial personalities of commercial laying hens and behavioural responses to keel bone fractures and management practices

PhD Thesis submitted by

Camille Marie Montalcini

for the degree of

PhD in Biomedical Sciences

Supervisor
PD Dr. Michael J. Toscano
Institute/Department of Animal Welfare
Vetsuisse Faculty of the University of Bern

Co-advisor
Prof. Dr. Stephan Morgenthaler
Institute/Department of Mathematics
École Polytechnique Fédérale de Lausanne (EPFL)

Accepted by the Faculty of Medicine, the Faculty of Science and the Vetsuisse Faculty of the University of Bern at the request of the Graduate School for Cellular and Biomedical Sciences

Bern,	Dean of the Faculty of Medicine
Bern,	Dean of the Faculty of Science
Bern,	Dean of the Vetsuisse Faculty Bern

SUPERVISOR

Dr. Michael J. Toscano
Centre for Proper Housing of Poultry and Rabbits (ZTHZ)
Animal Welfare Division, VPH Institute
University of Bern

CO-SUPERVISOR

Dr. Matthew B. Petelle
Centre for Proper Housing of Poultry and Rabbits (ZTHZ)
Animal Welfare Division, VPH Institute
University of Bern

CO-ADVISOR

Prof. Dr. Stephan Morgenthaler
Mathematics Department
École Polytechnique Fédérale de Lausanne (EPFL)

EXTERNAL CO-REFEREE

Prof. Dr. Niels Dingemanse
Department of Biology
Ludwig Maximilians University
Germany

MENTOR GRADUATE SCHOOL

Prof. Dr. med. Jan P. Kucera
Integrative Cardiac Bioelectricity Group
Department of Physiology
University of Bern

Table of contents

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

Abstract

This thesis aimed to explore space-use and movement behaviours of cage-free commercial laying hens to quantify consistent individual differences and to explain behavioural variation in response to keel bone fractures and various management practices. We monitored the movements of 403 hens distributed across two flocks, as they transitioned between five zones: the three stacked tiers of an aviary system, the littered floor, and the winter garden, using a herein validated low-frequency tracking system. We tracked hens from the day of transfer to the laying barn until nearly the end of the production cycle (17-60 weeks of age) and extracted daily behaviours related to various aspects of their daily routine, including the sleeping, feeding, nesting, indoor movements, and outdoor usage. We found consistent individual differences during the onset of lay (the first two months in the laying barn) and during adulthood where consistent differences between individuals explained between 24% and 66% of the behavioural variation. These long-term consistencies - together with the identified syndrome comprising all behaviours except the one related to the nesting - revealed the potential applicability of these behaviours as personality traits and indicated two axes of spatial personalities that may be driven by different mechanisms. Alongside consistent inter-individual differences we exposed intra-individual variation underscoring the relevance of studying spatial behaviour to better understand how animals respond to external and internal changes. Specifically, we found that an increase in the severity of keel bone fractures led to a drop in vertical travelled distance and tended to be followed by more tiers crossed within a transition. In addition, we exemplified how tracking systems could be used to assess long- and short-term influences of different management practices on hens' behaviour, including the commercial hatchery process, the transfer to the laying barn, and the daily automated delivery of fresh feed. To assess relative benefits further research should evaluate how these behaviours correlated with animals’ physiological stress responses and affective state.

Summary

Every year, humans raise over 78 billion terrestrial animals for consumption around the world, including 70 billion chickens (*Gallus gallus domesticus*) [1]. There is a general rising social awareness of the consumer on farm animal welfare, including an increased welfare concern towards laying hens [2]. As a result, the European Citizens’ Initiative *End the Cage Age* of 15th of April 2021 aimed at abolishing the use of cages in Europe animal farming by 2027. Therefore, within Europe, battery cages in the laying hen industry should slowly be replaced by cage-free systems, such as free-range and multi-tier aviaries [3]. These alternative systems are thought to enhance animal welfare by providing greater opportunities for more natural and motivated behaviours [4]. However, despite these clear benefits for animal welfare, the health issues that these modern systems have, in particular aviaries, remain considerable and can be more challenging to control than in cage-systems [5]. In addition to

the practical need of improving husbandry practices for cage-free commercial hens, there is a more basic need to change our representation of farm animals as individuals with complex characteristics rather than as commodities [6]. Laying hens are often seen as unfamiliar [6] and perceived as cognitively simpler [7] with fewer emotions and less differentiated individual personalities [8] compared to other farm animals, despite their cognitive, emotional, and social complexity [6].

Technologies for tracking animal locations have emerged as a promising tool to uncover individuality of commercial hens within aviaries [9] but also to study behavioural variation in relation to management practices and health issues [10–12]. Yet, compared to other systems (such as free-range [13–18]) there are a limited number of studies that have monitored individual movements within multi-tier aviaries (but see, [9–11]), conceivably resulting from the challenges associated with tracking individuals in the presence of metal equipment and high animal density. This thesis aimed to examine space-use and movement behaviours of cage-free commercial laying hens to quantify consistent individual differences and to explain behavioural variation in response to keel bone fractures and various management practices. In this thesis, we will use the term **'spatial behaviour'** to generally refer to how animals use and move in their housing systems (e.g., space-use and movement behaviours).

To reliably track individual movements within multi-tier aviaries, we customised and validated a tracking system (against video observations, 337 hours; Chapter I). We used a system comparable to ultra-wideband tracking systems but with lower frequencies, to reduce possible interactions with metallic materials and animals [19]. To further facilitate the process of tracking individual in multi-tier aviaries, the system was designed to register hens' transitions between five defined zones (the three stacked tiers of the aviary, littered floor, and winter garden) instead of precise individual locations. We found that the system accurately determined hens' presence in each zone but overestimated the number of transitions, explaining only 23% of the actual variation. Therefore, we evaluated two processing methods (i.e., a deterministic rule and a classifier) to filter false registrations. The two resulting processed datasets were representative of hens' transitions, accounting for 91% and 99% of the actual variation, respectively. We concluded that this tracking system is suitable to track individual transitions between the stacked tiers of aviary systems, given appropriate data processing.

To explore the significance of spatial behaviours to uncover individuality of commercial hens, we first quantified the extent of consistent individual differences in averages of a daily composite movement behaviour during the onset of lay (first two month in the laying barn; Chapter II). Then, during adulthood, we quantified consistent individual differences over time and across contexts, of five behaviours related to the sleeping, feeding, nesting, indoor movements, and outdoor usage (Chapter

III). We found consistent individual differences from the onset of lay and throughout adulthood in all behaviours, with up to 66% of the behavioural variation attributed to individual differences, providing a first quantification of the previously identified differences in spatial behaviours among hens in multi-tier aviaries [9,11]. These results, not only indicate the presence of animal personalities in our flock, but also underscore the complex nature of commercial laying hens.

To explore the significance of spatial behaviours for animal health, we evaluated the bidirectional relationship (i.e., the dynamic) between behaviours and keel bone fractures (KBF), considered as one of the greatest welfare issues in the egg production industry [20–25] (Chapter IV). By understanding such dynamics, we could for instance identify when behavioural variation may lead to feedback loops that would exacerbate health issues. Despite their importance, there is a dearth of studies that evaluated state-behaviour feedbacks within commercially relevant settings (for other settings see: [26]). Keel bone fractures are an excellent candidate to evaluate such feedbacks as there is strong evidence that they impair mobility [10,27–29], whereas effects of spatial behaviours on bone fractures likely exist [3,20,30–33] but remains poorly understood. We found that an increase in the severity of KBF led to a drop in vertical travelled distance and tended to be followed by more tiers crossed within a transition. In contrast, we did not find any evidence that spatial behaviours altered the severity of KBF. However, similar efforts accounting for the location of fractures could unveil the potential influence of spatial behaviours in the formation and maintenance of KBF.

To evaluate the relevance of spatial behaviours in assessing animal responses to management practices, we used the standardized commercial hatchery process [34] (Chapter V), the transfer to the laying barn (Chapter II), and a daily stimulus in the laying barn (delivery of fresh feed; Chapter III). We found that for the first three months post-transfer to the laying barn (~17-30 weeks of age), hens hatched in commercial hatchery, on average, transitioned more between the aviary tiers and spent more time in the littered floor, suggesting a faster use of all available resources compared to hens hatched on-farm in less stressful conditions. We also identified individual differences in temporal plasticity of daily movements following the transfer to the laying barn, with hens increasing their movements on average for 39 days. Additionally, we identified weak, but long-term, consistent individual differences in response to the delivery of fresh feed (repeatability = 0.33 over a period of 6 months). Overall, these results demonstrated the potential of tracking systems in assessing long-term and daily animal responses to management practices.

General introduction

Behavioural variation can facilitate effective responses to a wide range of changes, including internal changes like bone fractures and external changes like fluctuations in feed availability. Two important aspects of behavioural variation include the notions of consistent individual differences [35] and of behavioural plasticity [36]. This general introduction will start by providing a background on these two notions before discussing causes and consequences of behavioural variation. Then, I will narrow the scope of this introduction to concentrate solely on spatial behaviours, focusing on their significance for farm animal welfare. Finally, the commercial laying hen housed in cage-free systems will be introduced as an interesting animal model to study behavioural variation and its association with animal welfare, particularly to KBF, considered as one of the greatest welfare concerns to over 8 billion individuals [20–25,37].

Behavioural consistency and plasticity

An animal's sensory receptors, including those in its ears, olfactory system, and eyes, receive inputs from its surroundings, which are then processed by the animal and may be translated into internal (e.g., physiological) and external responses (e.g., behavioural). **Animal behaviour** can be broadly categorized into several classes, including behaviour used to communicate (e.g., vocalizations, body language), reproduce (e.g., nest building, courtship), feed (e.g., foraging, scratching), socialise (e.g., grooming), defend (e.g., fighting, hiding), acquire new information (e.g., tool use, observational learning), and move (e.g., dispersal, space-use behaviour). Essentially, behaviour serves as an interface between an animal and its surroundings [38].

Phenotypic variability (including behavioural variability) when expressed in response to environmental variation within a single genotype is called **phenotypic plasticity** [39,40]. This ability of a genotype (or an individual) to modify its phenotype, such as behaviour, physiology, and morphology, in response to environmental variation can result in a selective advantage [40–42], such as increasing survival rate [43,44] and reproductive success [45,46]. Two of the most common categories of phenotypic plasticity are the developmental plasticity [47] and the phenotypic flexibility [39]. **Developmental plasticity** refers to an organism’s capacity to develop different phenotypes in response to different environmental conditions during its development [48]. These permanent differences in the developmental trajectory of the phenotype contrast with the reversible phenotypic transformations in response to environmental changes, called **phenotypic flexibility** [39,49].

Behaviour is considered as one of the most plastic phenotypic traits [48]. However, this plasticity may be limited by **consistent individual differences** in behaviour (but not necessarily [50]), which are prevalent across a wide range of animal taxa [35,51] and explain on average 37% of the behavioural variation. Accumulating evidence suggests that behavioural differences between individuals can be maintained across time and contexts [52], which is generally referred to as animal personality [53,54]. Personality is a multidimensional concept and can be summarized into five axes, namely aggressiveness, activity, exploration, boldness, and sociability [53,55]. Correlations among individual's average behavioural expressions, called **behavioural syndrome** [56], are common between these five axes. Typically, more aggressive individuals tend to be more explorative, active, bold, and less social [55]. These individuals would commonly be classified as proactive and contrast to more reactive animals [55,57,58]. Correlated behaviours play a crucial role in the study of animal behaviour, as they can constrain evolution and provide insights into life history strategies (e.g., activity - risk-taking syndrome affecting survival [59]) and key ecological processes (e.g., 'fast'/proactive personalities [60] tend to disperse over longer distance [61]).

Causes and consequences of behavioural variation

By understanding the causes and consequences of individual variation in behaviour, we can understand when variation enters a positive feedback loop, leading to increased variation, or when variation feedbacks negatively and dampens successively. Positive feedbacks between two processes X and Y occur when a change in X leads to a change in Y that reinforces the initial change in X, while negative **feedback** occurs when a change in X leads to a change Y that counteracts the initial change in X. By studying the dynamics between animal health and behaviour, we could identify which behaviours may lead to feedback mechanisms that amplify or stabilize changes in the health process. For example, by studying risky behaviours (such as hazardous landings in multi-tier aviary systems) we could assess whether such behaviours exacerbate certain health problems (such as KBF) which in turn could increase the propensity to engage in these risky behaviours, resulting in an undesirable positive feedback loop.

A wide range of factors has been shown to **cause** behavioural variation, regardless of whether they lead to consistent individual differences. For instance, behavioural variation can stem from variation in a range of external factors related to the abiotic environment [43,62], including temperature [63] and many other habitat characteristics [64], or related to the biotic environment, including conspecific (dominance hierarchies) [65] or heterospecific interactions (predator risk [66] or mixed-species resource availability [38]). Furthermore, the salience and importance of an environmental stimulus may differ between individuals and depend on motivational factors, such as hunger and the need to engage in enrichment activities or interact with conspecifics. The way animals experience their environment is crucial in shaping their behaviour as it can alter their perception of stimuli and provide new opportunities to learn. Therefore, individual variation in behaviour can also stem from differences in cognition processes including perception, learning, memory, and decision making [67], as well as a range of other internal factors [26,68], such as physiology [60], life historic traits [60], health status [29], and life cycle stage [69].

Consistent individual differences in behaviour are generally attributed to differences between individuals in genes and environmental conditions. Recent research has shed new light on potential mechanisms underlying consistent individual differences in behaviour. For example, a study on Amazon mollies (*Poecilia Formosa*) has found consistent differences in behaviour among genetically identical individuals reared in near-identical conditions [70]. The authors suggested amongst other explanations, that positive state-behaviour feedback [26] and maternal bet-hedging could explain individuality in absence of variations in genetic or environmental backgrounds. They concluded that individuality may be an unpredictable and inevitable result of development. Fisher et al. (2018) [71] suggested that dynamics that are sensitive to initial conditions and governed by deterministic rules (often called chaotic dynamics) may also have contributed to behavioural development in a way that could result in long-term individual differences in animals with very similar environmental backgrounds and genetics. In other words, this latter explanation proposes that deterministic chaos could explain how small initial behavioural differences would cause long-term differences.

Understanding the causes of individual variation in behaviour can provide valuable insights into the potential **consequences** of such variation. For instance, as consistent individual differences in behavioural traits can result from genetic differences, these traits could be heritable. Accumulating evidence suggests that personality traits are heritable [72–75] and have higher heritability than those of other behaviours [74]. Because these traits also relate to animal productivity [76–78], welfare [55,79,80], and fitness [81,82], they can have important implications for animals. For instance, natural selection of specific behaviours could shape adaptations by helping wild animals to survive and reproduce under environmental change, while human selection for specific personality traits could be used to breed for more robust farm animals (e.g., in laying hens [83], pigs [84], cows [85]). Another example of the potential consequences arising from consistent individual differences in behaviour, particularly at the individual-level, is the fact that such personality traits could limit behavioural plasticity and thereby prevent optimal changes of a behaviour in certain contexts.

The importance of spatial behaviours for farm animal welfare

Animal welfare is a multidimensional concept that comprises animal health, normal behaviour, and affective state [86]. Therefore, in addition to minimal suffering and negative experiences, animals should experience positive emotions [86]. Consequently, to assure optimal welfare, animals must be able to move freely in their environment and have the choice to perform natural behaviours that matter to them [87,88]. However, these conditions are not always met, given that farm animals are typically housed in human-designed artificial environments, where the emphasis on productivity was often prioritized over animal welfare. Therefore, studying movement and space-use behaviours could help identifying features of their environment that could hinder their freedom of movement or the expression of natural behaviours at specific times in their lives, such as when dealing with health issues or specific management practices.

Studying behavioural variation in response to **management practices**, such as those related to the transfer to new environments, the distribution of fresh feed (e.g., laying hens [89], dairy cattle [90]), and the social environments could provide insights into farm animal welfare. For example, given the disparities between natural and commercial environments, it is probable that farm animals need special guidance and time to effectively use all available resources needed to express all natural behaviours upon the transfer to a new environment. Another example comes from a prior study which illustrated how the combination of behavioural observations, including space-use, and production measurements could inform decisions regarding feeder space for laying hens to reduce agonistic behaviour and facilitate access to the feed while maintaining a similar feed conversion [89]. Furthermore, considering that the group size of farm animals cannot always self-regulate through cost-benefit balance and that individuals cannot join or leave a group as in natural populations [91], management practices related to the social environment should also be carefully designed. Previous literature showed negative effects arising from social instability, isolation, and crowding [92], highlighting the relevance of studying individual spatial behaviours in relation to conspecifics' spatial distributions.

The majority of farm animals, including chickens, cattle, sheep, ducks, goats, pigs, and rabbits, are considered prey species. Some prey species may find advantageous to hide any signs of vulnerability under **health issues**, including sickness behaviour, defined as a stereotypical response to injury or infection [93]. This tendency to hide vulnerability may weaken the influence of health issues on animal behaviour, making it more challenging to study [93]. This underscores the importance of sensor technologies that can collect large datasets and models that can account for the potential dynamic between an animal's state and its behaviour. These tools may facilitate the detection of subtle changes in behaviour that animals cannot mask. By understanding the interplay between spatial behaviours and health indicators, it can, for instance enable the early detection of health issues [94–97]. Given that animals can modify their spatial behaviours in response to the presence of illness [97], these behaviours could help detecting health issues at an early stage. This has been observed in various species such as broilers [98], dairy cows [99,100], pigs [101], and laying hens [102]. Such behaviours may also precede the onset of a disease and thereby predict disease susceptibility prior manifestation [98,103]. Correlations between spatial behaviours and health indicators can also inform potential modification of housing systems to decrease prevalence of health problems [30,104]. For example, Stratmann et al. (2015) [30] showed that adding ramps in aviary systems could enable hens to move between the stacked areas by walking instead of jumping or flying, which was shown to decrease incidence of falls, collisions, and keel bone damage in laying hens.

Studying spatial behaviours of farm animals through more fundamental approaches, such as those studying consistent individual differences, could help shift our representation of farm animals as individuals with complex characteristics rather than commodities. Traditionally, individual traits are assessed in laboratory settings or with standardized assays [105] on a limited number of individuals or over short periods of time, and therefore may not reflect all important aspects of behaviours expressed within commercially relevant settings [106,107]. Technologies allowing to track animal locations are among the few tools capable of automatically registering individual behaviour over long periods of time and with minimal disturbances, enabling to expose consistent individual differences within commercially relevant settings. Results from two meta-analysis, showed that on average 67% of the variation in spatial behaviours within populations could be explained by differences between individuals [51], compared to 37% when considering a broad variety of behaviours [35]. These findings highlight the general remarkable individual consistency in spatial behaviours across many taxa, also referred to as "**spatial personality**" [51]. This relatively high consistency may be partially explained by the fundamental role of spatial behaviour in common personality traits (e.g., activity and exploration behaviours) [61] and their association to other personality traits. For instance, bolder animals may be more likely to take risks and explore new environments compared to shyer animals which may be more likely to avoid risky behaviours and stay in familiar environments.

The relatively high repeatability estimates of spatial behaviours [51] sets a relatively high upper bound on their heritability that supports the use of tracking systems to study behavioural traits to breed more robust farm animals. Breeding practices of farm animals have long been mostly focused on optimising productivity. Research has now proposed the integration of personality traits as phenotypes into the breeding process to breed animals that exhibit increased resilience to challenges, such as diseases and stressors [55,79,80]. This approach has already been discussed in various farm animals, including pigs

[84], laying hens [83,108,109], and cows [85]. Therefore, tracking systems could also have the potential to benefit farm animal welfare at a global scale, but future research is needed to assess the feasibility.

In conclusion, studying spatial behaviours in relation to management practices and health conditions can yield practical benefits for animal welfare, including more appropriate housing design and management practices and the development of early warning systems. Additionally, exploring more fundamental aspects of behaviours (e.g., related to animal sentience or animal personality) could change our attitude towards animals to better account for their needs and to shed light on novel breeding practices.

Cage-free commercial laying hen as animal model

Commercial settings offer researchers a unique environment that combines the real-world complexity relevant to farm animals with a certain level of control. For example, factors that are generally partially regulated by humans includes those related to the rearing processes, lighting, timing of feed, and indoor temperature. Furthermore, these settings are sometimes globally standardized and have automated processes, such as for feeding or providing access to areas, which could also be beneficial in research. Last but not least, commercial settings provide an opportunity for animals to serve a dual purpose by serving commercial and research objectives, aligning with the ethical framework for the design of animal experimentation based on the 3R principles (Replacement, Reduction, and Refinement) [110].

Despite the benefit for researchers to study farm animals directly within their relevant settings, there are significant expenses and potential challenges associated with behavioural studies within commercial settings. It is therefore important to emphasize that conducting research in these environments is crucial when the findings are meant to assess relative benefits for the welfare of animals. Indeed, not only the intricate design of modern commercial systems increases the behavioural repertoire of animals but the social dynamics in larger groups can differ from those in smaller groups. For instance, in large groups, hens can't recognize all conspecifics which could limit their ability to form social groups [111]. Consequently, behaviours observed in laboratory settings may not always extrapolate to commercial settings and conducting research within the latter environment is important to understand how farm animals behave.

Every year, humans raise over 78 billion terrestrial animals for consumption globally, including 70 billion chickens (*Gallus gallus domesticus*) [1]. As the most common food-producing animals on land, chickens are widespread and could be of great interest to researchers studying animal behaviour due to their cognitive, emotional, and social complexity. For example, in a recent review, Marino (2017) [6] reported that chickens experience positive and negative emotions, perceive time intervals, perform basic transitive inference, and demonstrate a basic comprehension of numerical quantity, among other abilities. **Laying hens**, in particular, also exhibit a fascinating range of behaviours, including many vocalisations (20-25 discrete calls already documented [112–116]) and various highly motivated behaviours, such as roosting on elevated perches, using nest boxes for egg-laying, and dust bathing [4]. Hens' diverse modes of locomotion, including flying and running, further enhance their appeal as subjects in scientific studies related to spatial behaviours, allowing to study individual differences in the modes of locomotion, in addition to their locations.

There is a rising social awareness of the consumer on farm animal welfare with increased welfare concern towards laying hens [2]. The European Citizens' Initiative *End the Cage Age* of 15th of April 2021 aimed at abolishing the use of cages in European animal farming by 2027. Therefore, within Europe, battery cages in the laying hen industry should slowly be replaced by **cage-free systems**, such as barn, free-range, and multi-tier aviaries [3]. Multi-tier systems, the system used in this thesis, are designed so that the usable area is not limited to ground level [117]. In these systems, hens may have access to a maximum of four tiers, though the number of levels the birds can access may be higher than the number of tiers (i.e., when tiers have more than one level) [117]. To access all resources in multi-tier aviaries, hens must move horizontally and vertically, as resources are distributed throughout the stacked tiers reaching up to 3.5 meters in height. These cage-free systems are thought to enhance animal welfare mainly by providing greater opportunities for more natural and motivated behaviours. However, despite the considerable benefits for animals, the welfare problems that these modern systems have, in particular aviaries, remain considerable and can be harder to control than in battery cages [118,119].

Welfare issues in cage-free laying hens are diverse, including bacterial infection (bumblefoot), parasites (e.g., red mites), feather pecking, and **keel bone fractures**, considered as one of the greatest welfare issues in the egg production industry [20–25], potentially affecting egg-production [28,120,121] and animal welfare. Indeed, compelling evidence indicates that hens with KBF feel pain at least for a few weeks [22,122], show behavioural differences in highly motivated behaviours, including nestbox use and perching, which could indicate negative affective states [122,123], and face restricted mobility. More specifically, increased KBF have been associated fewer vertical movements [29], a change in resting location from litter to perches [124], increased latency to fly up or down from perches [27,28], more time spent on the top tier [10], and less time spent on the litter floor [10].

While the exact cause of KBF remains unclear [125,126], human selection for early sexual maturity and high egg productivity (~320 eggs/year compared to 12 eggs/year in the red junglefowl) has resulted in an increased demand for calcium for eggshell formation, which in turn increased bone fragility and susceptibility to fractures [3,125,127]. Previous literature also suggested that specific aspects of the aviary's internal structures, such as perches or more generally structures increasing risks of falls and collisions, are contributing factors to KBF prevalence [3,30,31]. In other words, the complex design of multi-tier aviaries could exacerbate KBF’s prevalence [20,33]. Additionally, impaired mobility within such housings may increase the risk of dehydration, emaciation, and floor eggs [128], as hens must move horizontally and vertically to access all resources, distributed throughout several stacked tiers reaching up to 3.5 meters in height. Consequently, these systems could also exacerbate the consequences of KBF, making multi-tier aviaries particularly interesting when studying KBF, as their complex design may intensify both its prevalence and consequences.

In conclusion, commercial laying hens are an interesting animal model to explore behavioural variation and its association with animal welfare given their diverse behavioural repertoire, numerous health issues, and worldwide presence. Additionally, cage-free housings provide researchers unique environments with real-world complexity relevant to farm animals and partially controllable settings.

Thesis objectives, hypotheses, and experiments

This thesis aimed to explore spatial behaviours of commercial laying hens with three specific objectives: (1) quantifying consistent individual differences in spatial behaviours (spatial personality), (2) examining the dynamics between behavioural responses and the amount of the keel bone affected by fractures, and (3) assessing influence of management practices on behaviours. These objectives were defined based on the relevance of the existing knowledge gap, introduced above and summarized below, that need to be addressed to ultimately enhance the welfare of commercial hens.

To shift our perception of commercial hens as individuals, it is important to underscore their intricate nature and better understand their individual need. In 2018, consistent individual differences in spatial behaviours of commercial hens within multi-tier aviaries was first identified [9], but the extent of individual consistency in spatial behaviour is not yet quantified. As a result, we also do not know how individuality persists over time and across context and differs between production phases or across different behaviours. Therefore, the first goal of this thesis was to quantify consistent individual differences in different spatial behaviours over time and across context, during the onset of lay and adulthood.

In addition to the need to change our perception of farm animals, there is the practical need of reducing health issues in commercial laying hens, where KBF are considered as one of the greatest welfare concerns. In particular, it is important to understand whether KBF may be caused or exacerbated by specific behaviours and how such fractures may impair behaviour, particularly in multi-tier aviary systems since their spatial complexity could exacerbate both KBF's prevalence and consequences. In 2019, Rufener et al. [10] provided the first evidence of an association between KBF and mobility in hens housed within multi-tier aviaries, but the causality has not yet been investigated. More generally, while there are clear associations between KBF and mobility with evidence suggesting that KBF impair mobility, the effect (whether causing or maintaining fractures) of spatial behaviours on KBF remains poorly understood. Therefore, the second goal of this thesis was to evaluate the dynamics between KBF and spatial behaviours.

Lastly, there is the practical need of improving husbandry practices for modern cage-free commercial hens. Despite the considerable benefits for animals, the welfare problems that these modern housing systems have, specifically aviaries, remain considerable and can be harder to control than in battery cages. However, there is a dearth of studies that used spatial behaviours to assess individual responses to management practices within commercial settings. Therefore, the third goal of this thesis was to study behavioural variation in relation to three types of management practices: the standard commercial hatchery practices, the transfer from the rearing to the laying barn, and recurring environmental stimulus such as the daily automated delivery of fresh feed.

Objectives

Previous literature has underscored the significance of spatial behaviours in commercial hens within multi-tier aviary systems as a means to study individuality, and to evaluate underlying causes and consequences of KBF. However, the extent of individual consistency in spatial behaviour has not been quantified, and the two halves of the KBF-mobility association was not investigated in conjunction. This thesis primarily aimed to address these two knowledge gaps and, in pursuit of further practical implications, also aimed to evaluate behavioural variation induced by three common management practices. Therefore, the objectives of this thesis were to explore the significance of spatial behaviours in commercial laying hens within multi-tier aviaries to study:

I. **Individuality**, by quantifying consistent individual differences in different spatial behaviours over time and across context, during the onset of lay and adulthood,
II. **Keel bone fractures**, by evaluating the dynamics between the amount of the keel bone affected by fractures and spatial behaviours, and
III. **Responses to management practices**, by evaluating how the standard commercial hatchery practices, the transfer from the rearing to the laying barn, and recurring environmental stimulus resulting from management practices can cause behavioural variation.

Hypotheses

The hypotheses of the thesis were formulated for each objective separately:

I. **Individuality**: We hypothesised that commercial laying hens have individual personalities that is reflected in their spatial behaviours. Therefore, we predicted that while controlling for some variation in health and the environment, a significant part of the remaining variance in spatial behaviours would be attributed to repeatable inter-individual variation.
II. **Keel bone fractures**:
    - We hypothesized that hens with more severe KBF would reduce their activity and spend more time on higher tiers with vital resources. Therefore, we predicted that an increase in KBF would lead to a drop in vertical travelled distance, greater time spent in the top tier, and more generally to a more uneven usage of the five zones (aviary tiers, littered floor and winter garden).

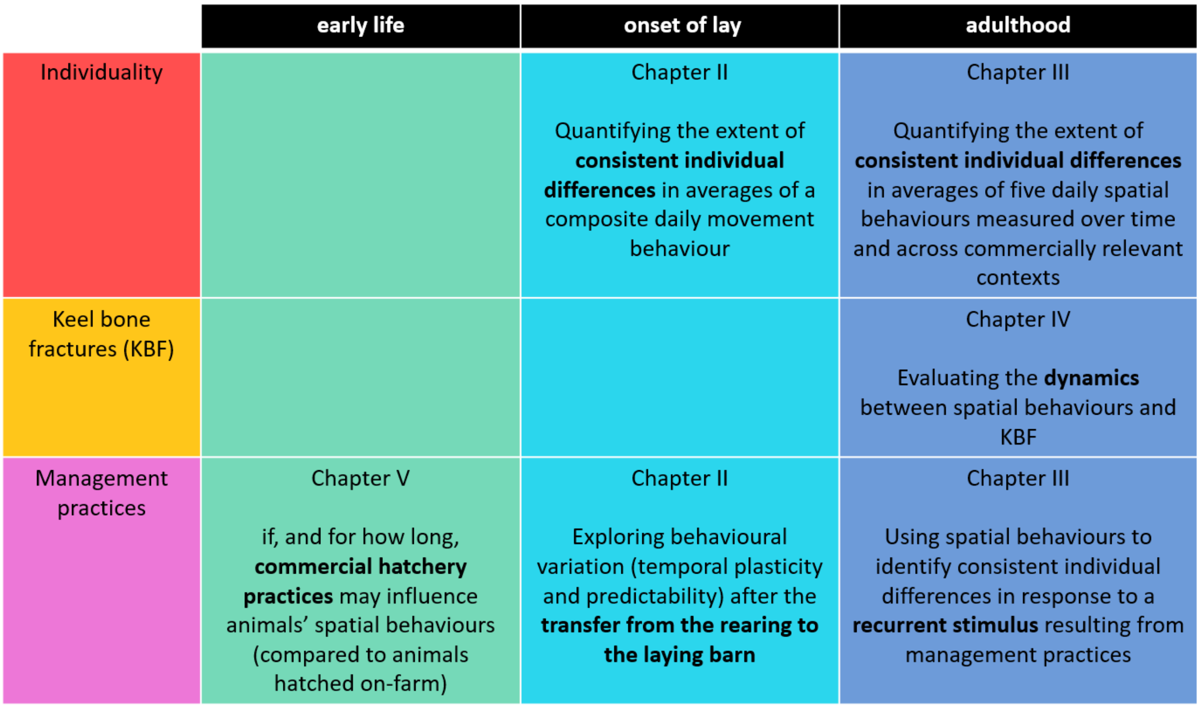
- We hypothesized that more transitions between the aviary stacked tiers would lead to a higher number of landings and, consequently, increased occurrence of falls and collisions. Therefore, we predicted that greater vertical travelled distance and greater number of tiers crossed within one transition (leading to longer and potentially more hazardous landings), would lead to increased KBF severity.

III. **Responses to management practices**: We hypothesized that spatial behaviours play an important role in responding to environmental variation in commercial aviaries. We predicted long-term influences of early-life stress induced by the standard commercial hatchery practices on hens' spatial behaviours at the population-level (when compared to hen hatched on-farm). We also predicted individual differences in temporal plasticity after the transfer from the rearing to the laying barn and consistent individual differences in response to recurring environmental stimulus resulting from management practices during the laying phase (specifically the delivery of fresh feed).

Experiments

This thesis involved two animal experiments, both conducted with Dekalb white laying hens (*Gallus gallus domesticus*) in the same aviary system over two consecutive years. In brief, our primary objective with the **first experiment** was to track hens' transitions between five zones in a multi-tier aviary from the day of transfer to the laying barn (17 weeks of age) until nearly the end of the production cycle (60 weeks of age). In total, we tracked 227 hens for varying durations, primarily dependent upon tissue collection for other experiments not pertaining to this thesis. For the purpose of objective II, we also conducted five health assessments during the laying phase including the assessment of the total amount of the keel bone affected by any fracture (KBF severity) using radiography. For the purpose of objective III, we had a treatment group, where half of the hens hatched on farm, while the other half hatched in a standard commercial hatchery. The **second experiment** was specifically conducted to augment the sample size to evaluate the dynamics between the severity of KBF and spatial behaviours (objective II). Therefore, in this second experiment we tracked hens during each week prior to radiography only (7 days of tracking/radiographs, 4 radiographs, 169 hens). **Based on these two experiments, we validated the tracking system (Chapter I) and conducted four other studies, i.e., chapters**. The first experiment was used for the five chapters, while the second experiment was used solely in Chapter IV. In brief, in Chapter II and Chapter III we studied spatial behaviour during the onset of lay and the adulthood, respectively. In Chapter IV we evaluated the dynamics between spatial behaviours and KBF, and in Chapter V we evaluated the influence of standard commercial hatchery practices on spatial behaviour during the laying phase.

More precisely, Figure 1 illustrates how the three thesis objectives were explored during hens' early life, onset of lay, and adulthood. Figure 2 illustrates the observational timeframe of each chapter.



	early life	**onset of lay**	**adulthood**
Individuality		Chapter II Quantifying the extent of **consistent individual differences** in averages of a composite daily movement behaviour	Chapter III Quantifying the extent of **consistent individual differences** in averages of five daily spatial behaviours measured over time and across commercially relevant contexts
Keel bone fractures (KBF)			Chapter IV Evaluating the **dynamics** between spatial behaviours and KBF
Management practices	Chapter V if, and for how long, **commercial hatchery practices** may influence animals' spatial behaviours (compared to animals hatched on-farm)	Chapter II Exploring behavioural variation (temporal plasticity and predictability) after the **transfer from the rearing to the laying barn**	Chapter III Using spatial behaviours to identify consistent individual differences in response to a **recurrent stimulus** resulting from management practices

Figure 1 - Simplified schematic of the three thesis objectives (rows) across the different life stage of the hens (columns) and the corresponding chapters. Chapter I is not included in the figure as it was dedicated to the validation of the tracking system.

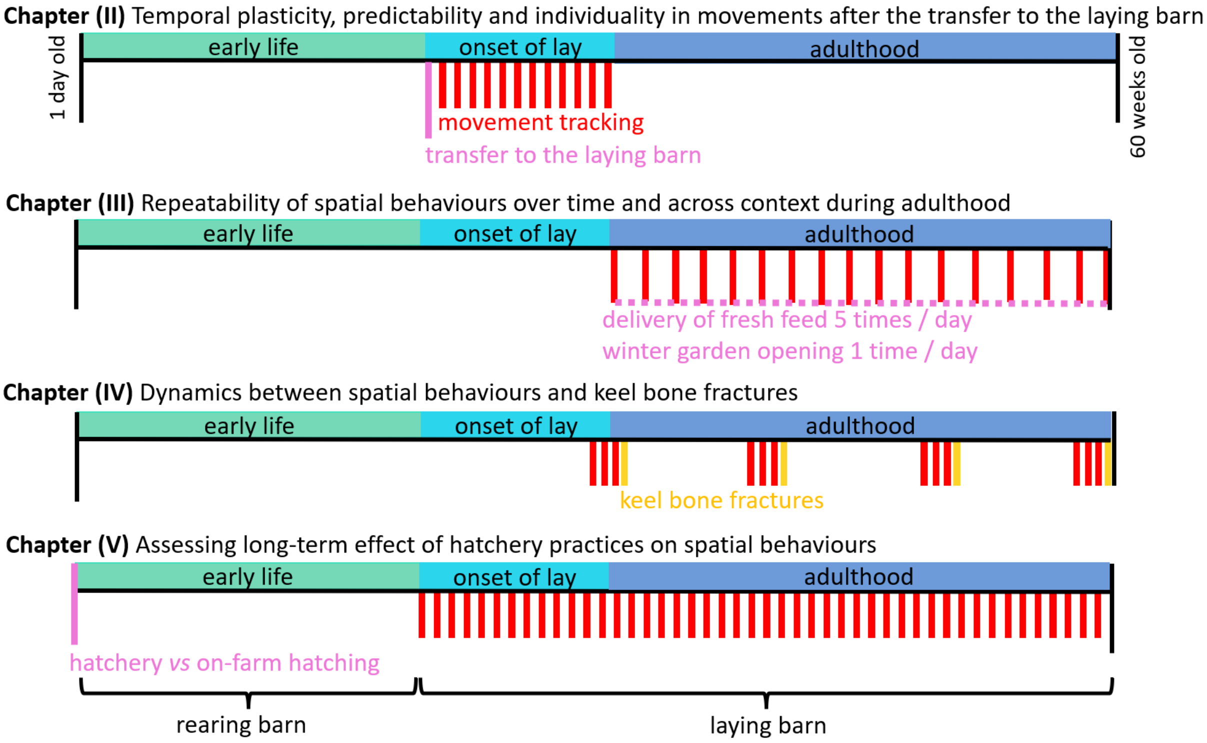


Figure 2 – Simplified schematic of the observational timeframe for each chapter. Red vertical bars indicate tracking data collection used to assess daily spatial behaviours. Pink lines represent the three different management practices studied in the thesis. Yellow lines denote time when radiographs were taken.

CHAPTER I

Based on:

Evaluation of an active LF tracking system and data processing methods for livestock precision farming in the poultry sector

Camille M. Montalcini [1,2], Bernhard Voelkl [3], Yamenah Gómez [1], Michael Gantner [4] and Michael J. Toscano [1]

1 Center for Proper Housing: Poultry and Rabbits (ZTHZ), Division of Animal Welfare, VPH Institute, University of Bern, Switzerland
2 Graduate School of Cellular and Biomedical Sciences, University of Bern, Switzerland
3 Division of Animal Welfare, VPH Institute, University of Bern, Switzerland
4 Gantner Pigeon Systems GmbH, 6780 Schruns, Austria

Published in: Sensors (2022), https://doi.org/10.3390/s22020659

Camille Montalcini participated in the installation of the tracking system, conducted the analysis and wrote the manuscript in agreement with her supervisor Dr. Michael J. Toscano and co-authors Dr. Bernhard Voelkl, Dr. Yamenah Gómez, and Michael Gantner.

Abstract

Tracking technologies offer a way to monitor movement of many individuals over long time periods with minimal disturbances and could become a helpful tool for a variety of uses in animal agriculture, including health monitoring or selection of breeding traits that benefit welfare within intensive cage-free poultry farming. Herein, we present an active low-frequency tracking system that distinguishes between five predefined zones within a commercial aviary. We aimed to evaluate both the processed and unprocessed datasets against a "ground truth" based on video observations. The two data processing methods aimed to filter false registrations, one with a simple deterministic approach and one with a tree-based classifier. We found the unprocessed data accurately determined birds' presence/absence in each zone with an accuracy of 99% but overestimated the number of transitions taken by birds per zone, explaining only 23% of the actual variation. However, the two processed datasets were found to be suitable to monitor the number of transitions per individual, accounting for 91% and 99% of the actual variation, respectively. To further evaluate the tracking system, we estimated the error rate of registrations (by applying the classifier) in relation to three factors, which suggested a higher number of false registrations towards specific areas, periods with reduced humidity, and periods with reduced temperature. We concluded that the presented tracking system is well suited for commercial aviaries to measure individuals' transitions and individuals' presence/absence in predefined zones. Nonetheless, under these settings, data processing remains a necessary step in obtaining reliable data. For future work, we recommend the use of automatic calibration to improve the system's performance and to envision finer movements.

Introduction

Tracking technologies generate sequences of chronologically ordered location data and offer a way to monitor movement of many individuals over long time periods with minimal disturbances. Tracking technologies have become valuable for detecting health issues in farm animals at an early stage [94–97] and in cage-free poultry farming, for their potential to select breeding traits that benefit welfare within cage-free systems [108,109] as well as to provide scientific information for optimal management [129]. However, cage-free housings are uniquely complex and may introduce numerous challenges for tracking technologies. For instance, cage-free housings of laying hens often contain a relatively high concentration of material that can interfere with tracking signals, including metal hardware (e.g., perches, floor, feeding lines) and multiple stacked horizontal levels that prevent direct lines of sight require by some automated tracking technologies (e.g., video tracking, infrared). Furthermore, compared to most other livestock, laying hens are relatively small animals that can be housed in large groups at very high densities, which would likely alter ultra-high frequency (UHF) radio signals [130]. Compared to most other commonly tracked livestock (e.g., swine, cattle), laying hens move differently (e.g., flying, jumping between horizontal tiers) and often faster. These challenges might induce measurement errors (as defined by the difference between a measured quantity and its true value), both of random and systematic natures [131]. Random errors are often inevitable and unpredictable, but their effects can be minimized, for example, by increasing the sample size. On the other hand, systematic errors are often predictable with consistent causes (e.g., environmental interference, improper calibration), but their effects are harder to compensate for and can lead to biases if not appropriately addressed during analysis.

Tracking systems have already been used to examine laying hens within the interior of a commercial system [9,10,132]; these tracking systems had to overcome the housing complexities described above. However, measurement errors were primarily evaluated within less complex settings (e.g., in small interior or outdoor settings) than commercial aviaries but focusing on movements of greater precision (i.e., individual location) than the current effort (transitions between predefined zones). For instance, using an ultra-wide band (UWB) system, Rodenburg et al. [108] reported an accuracy of 85% in detecting individuals' location, and Stadig et al. [16] reported an error of less than 50 cm in 80% of measurements. These results present great potential for tracking systems to represent individual positions within free-range areas, as well as a margin to refine the data. Systematic errors were also investigated, although only within settings less complex than commercial aviaries. For instance, comparing registrations generated by a UWB system against video observations, Sluis et al. [133] observed an average overestimation of 40% of in the distance of broilers moving less than 15 m and an average underestimation of 15% in the distance of broilers moving more than 30 m. Furthermore, Stadig et al. [16] observed a larger error in certain areas of the experimental field and a negative influence of rain on the percentage of successful registrations. These results suggest that various factors, such as the individual level of activity, specific areas, and weather conditions, could cause errors in measurement. Although tracking systems within cage-free housing systems are becoming more popular, they still have challenges to overcome. We therefore studied long-term tracking in commercial aviaries at the level of visited zones (with five zones) instead of precise individual locations. In the current study, we used active tags with low-frequency (LF) tracking and UHF communication that distinguished five zones with key resources, including the three stacked tiers of a commercial aviary (top floor, nest box, lower floor), the littered floor underneath, and an outside covered winter garden. This tracking system is comparable to UWB tracking systems with lower frequencies, with the aim of reducing possible interactions with the environment, such as liquid and metallic materials [19].

To overcome measurement errors, some studies have mentioned novel placement of tracking system components [16,134], filtering of registrations that are not possible [16], or filtering of individual positions that do not move more than the 95% confidence interval of the system's positioning errors [63]. When modifying the configuration of the tracking system is not an option, data processing may be the only alternative to refine and, in some cases, obtain validated data. Furthermore, tracking data often contain metadata associated to each registration, which could be used to detect false registrations and increase accuracy. Due to a potentially large number of available features and interaction effects, manually defining a rule-based algorithm can be time-consuming and suboptimal, whereas machine learning may offer a valuable solution for filtering false registrations. Despite potential for data refinement, there are only a few studies on UWB systems and related technologies that scrutinize data-processing methods, particularly within the unique settings of housings of cage-free laying hens. In the current study, we aimed to contribute to the collective effort of evaluating tracking systems for laying-hen farming, with a focus on the interior of a commercial aviary system. To achieve this aim, two analysis steps were involved. First, two data-processing approaches were applied to filter false registrations, including a simple deterministic approach that filters stays of short durations (SD method) and a machine learning approach (ML method) based on a tree-based classifier. The two processed datasets and the unprocessed dataset were compared against video-observation results (our gold standard). This evaluation was conducted in terms of the number of transitions per individual within the predefined zones and individuals' presence/absence in each zone every second. Secondly, to better evaluate the tracking system, we studied the effect of filtering false registrations based on the ML method over a two-month period on 144 tracked animals under three potential influencing factors: different areas of the aviary, external temperature, and external humidity. We selected these factors because they have already shown to be associated, to some extent, with tracking-system performance and could introduce biases in our own work and that of others using comparable technology if associated with false-registrations.

Materials and Methods

Ethical Statement

The study was conducted according to the cantonal and federal regulations for the ethical treatment of experimentally used animals and approved by the Bern Cantonal Veterinary Office (BE-45/20).

Animals and Housing

As part of a larger study examining effects of on-farm hatching, approximately 4800 chicks were reared in an Inauen Natura rearing barn previously described by Stratmann et al. [30] and located at the Aviforum facility in Zollikofen, Switzerland. At seven days of age, focal animals were selected, and at approximately 16 weeks of age, all animals were transferred to an on-site commercial laying barn containing a Bolegg Vencomatic Terrace aviary. The aviary system is split into 20 identical pens separated by a vertical grid, with each pen containing 225 animals and an outside, covered winter garden that can be accessed through a pop hole (illustrated in Figure 1). Eight of the 20 pens were used for the current study, with 18 focal animals per pen (a total of 144). On the same day as the transfer to the laying barn, we mounted a tracking tag enclosed within a cloth backpack (mass: 15.6 g; height: 14.5 cm; width: 13 cm) on the back of each focal hen. These backpacks were identifiable from video cameras based on their unique colour combination.

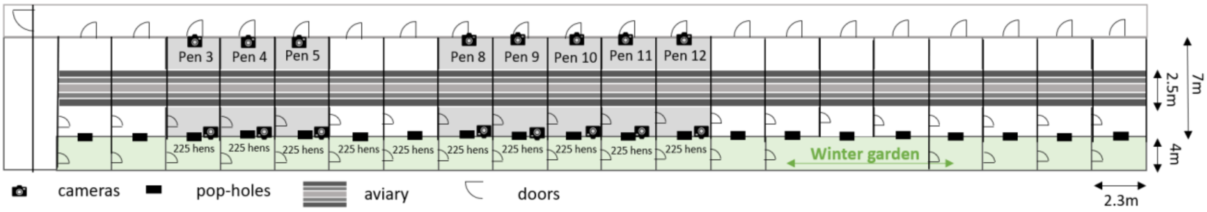


Figure 1. Housing setup, including the pens and aviary location in the barn, winter-garden zone, pop holes, and cameras.

Tracking System

To track hens across different areas within a pen, we distinguished five zones with key resources, including the three stacked tiers of a commercial aviary (top, nest box, lower), the littered floor underneath, and the winter garden, as illustrated in Figure 2A. During the laying phase, transitions between the five zones were assessed continuously for each focal hen by means of a customized tracking system. For this, three identical stations of a low-power, active tracking system (®Gantner Solutions GmbH, Schruns, Austria) were installed within the laying barn, each covering either two or three pens (Stations 3–5: pens 3, 4, 5; Stations 8–9: pens 8, 9; Stations 10–12: pens 10, 11, 12). Each station involved several components, including five markers (1 per zone) emitting signals through a cable (creating separately enclosed fields for each zone; Figure 2B); active tags (mass: 28.1 g) that can receive signals; and lastly, a reader that communicates through UHF (868 MHz), with the tags and a dedicated computer.

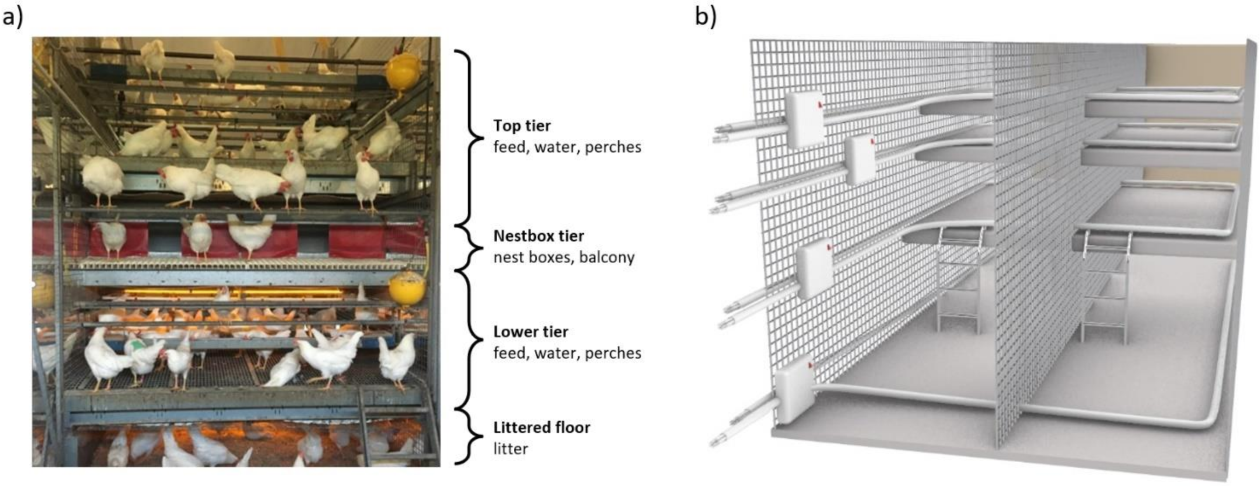


Figure 1 – a) Side view of the aviary including a single pen with its three aviary zones (top tier, nestbox, lower tier) and the littered floor. **b)** Simplified representation of the tracking system covering two pens, including the four indoor markers (square boxed) and their cables through which the signals are emitted

The receiving strength of the LF signal (RSS) is used to determine theoretical distance to the antenna loop. At almost every position in any zone, the tag can receive the signal of multiple markers. This signal last for 50 ms. It is important that during that time, the tag only receives the signal of one marker; otherwise, signals would overlap and might not be valid. Therefore, the markers send at different transmission intervals (varying from 1.6 to 2.1 s depending on the zone) a fixed low carrier frequency signal of 0.125 MHz (LF-signal) that is modulated to allow markers to be differentiated. Within a 10-s interval, a tag could theoretically receive between five and six signals per marker, but this number will often be lower, as every marker has a maximum range of only two to three metres. Every time a tag receives an LF-signal, an algorithm (tag-algorithm) is applied to the registered LF signals received within the past 10 s to evaluate whether the tagged hen has transitioned to a new zone. The tag algorithm reports a new transition when a tag receives the absolute strongest signal value from the same marker twice within 10 s and if the associated zone differs from the last registered zone (pseudo-code in the Supplementary Text S1). Following the installation of the tracking-system stations, each pen was calibrated under field conditions to ensure a correct interpretation of information obtained by the devices. More specifically, a tracking tag was positioned in each of the 44 predefined critical locations per pen (e.g., where two zones border one another) to evaluate RSS against observed distance to the antenna loop and to adjust the LF signal of specific markers as necessary.

Individual transitions to a zone registered by the tracking system are hereafter called registrations. More specifically, we will refer to correct registrations (CR) for registered zones where the animal is located (i.e., true zone as determined by video) and to false registrations (FR) for registered zones not consistent with the true zone for the bird (FR). Among CRs, we distinguish two types of registrations: (1) registrations that are not associated with a true transition (corrected registrations) and (2) registrations associated with a true transition (transitional registrations). Our goal was to obtain only transitional registrations, and data processing was used towards this objective.

Video Observations to Detect False Registrations

Two cameras per pen were placed within the indoor portion of each pen in such a manner that each location where an animal could transition between any of the three indoor zones was visible. The view did not cover the interior of the pop hole nor the winter garden and thus did not allow transitions to the winter garden to be filmed. For the generation of the video-based tracking data as a gold standard, video data were collected over the third and fourth weeks for an 11-day period simultaneously with the collection of the tracking data. Single animals were visually tracked by two trained observers independent of one another in order to classify each registration as FR or CR. An inter-rater reliability test between the two observers for 137 registrations, including four random hens and four different days, resulted in an inter-rater reliability of perfect agreement, with all recordings classified correctly by both observers.

For the evaluation of the two processing methods (SD and ML) and the unprocessed tracking data against video-based tracking data as the gold standard, two sets of registrations were analysed through video, generating two datasets: (1) the training dataset used to develop the ML method; and (2) the test dataset used to evaluate the two processing methods, as well as the unprocessed tracking data, against video-based tracking data. As described in the next section (2.5), the training dataset was used in a cross-validation process to split the data into validation and training sets and select for the optimal models.

The training dataset was composed of 4274 registrations classified as FR or CR by means of 241 h of video observations divided into 79 batches, varying from 0.5 h to 7 h, involving 44 tracking tags over 11 days. The batches were selected based on the visual representations of individuals' movement across all days to ensure a broad variation of movement sequences and a reasonable number of observations across zones, stations, and tracking tags. To avoid introducing noise in model training, the training dataset did not contain registrations from the winter-garden zone due to the limited camera view in the pop holes described earlier. The training dataset comprised 13% FR and 87% CR.

The test dataset was composed of 865 registrations classified as FR or CR by means of 96 h of video observation. More specifically, 48 batches (six/pen) of 2-hour video (including 47 randomly selected tracking tags) were randomly chosen over six days and reduced to 42 batches due to technical issues (e.g., backpacks not visible from the cameras). As the test dataset was used to evaluate two processing methods, including one that did not require training, the test dataset contained registrations from each of the five zones, including the winter garden. However, as the classifiers can only be tested on classes included in the training process, all registrations from the winter garden were processed solely by the SD method. Registrations in the winter garden were retained in the evaluation of both processing approaches for two main reasons: first, to avoid any bias towards poorer/greater performance of the SD method, if that zone would be more easily/laboriously detected by the tracking system compared to the other zones; second, even if the winter-garden zone is processed by the SD method when evaluating the ML method, its performance is still influenced by the ML method, typically when the ML method filters a registration to the litter zone reported between registrations in the winter-garden zone (as there would be one less transition to the winter garden). When a registration to the winter-garden zone could not be clearly classified through video observation (i.e., animal could be either in the pop hole or the winter garden), CR was used for biological relevance. We decided to define the pop-hole area (illustrated in Figure 1) as part of the winter-garden zone (and not the litter zone), as exposure to natural light in the pop hole is more similar to the winter-garden zone than the litter zone. To better evaluate the tracking system, in addition to the tracking system's registrations, the test dataset contained all true transitions observed during video observations that were not reported by the tracking system (missed transitions). Missed transitions represented 0.6% of the test dataset. The test dataset comprised 5% FR and 95% CR.

Evaluation of the Two Data Processing Methods

As the tracking system used in this study evaluated the location of a tag every time the tag received an LF-signal, longer records have more opportunities for self-correction and therefore are more likely to be accurately record the location. Therefore, an intuitive and simple way to process the data is to filter all registrations that last for less than a certain threshold (SD method). We used a one-minute threshold with the objective of minimizing loss of actual transitions while maintaining a good representation of the true data.

To account for more of the available information during data processing, we used a machine learning approach (ML method) based on decision-trees, which, in addition to the registration duration used by the SD method, employed 13 features of the registrations (detailed in Table 1), including the RSS, the zone, and the station identities. The zone identities of the previous and next registrations (of the same tag) were also included to account for the movement sequence. The durations of the previous and next registrations were also included, as we expected the duration to be the most important feature for detecting FR. Our goal was to build a model to process (clean) the data rather than generate predictions about hen movement patterns. We aimed to isolate the true signal of hen movement, which can be used in future research to evaluate the drivers of hen behaviour. As such, our model is independent of external factors that could be of potential interest for future investigation (e.g., weather). Three classifiers (random forest, gradient boosting, CatBoost) based on decision-trees [135] were used to account for potential non-linearity and interaction effects [136]. The gradient-boosting classifier is a greedy algorithm that sequentially trains a shallow decision tree in order to correct the errors of the previously trained tree [137], and the CatBoost model is a recently developed gradient-boosting algorithm [138,139] that was selected in this study for its ability to process categorical features during training (algorithms of the classifiers further detailed in Supplementary Text S2). Following hyperparameter selection through a 3-fold cross-validated grid search and model training on the training dataset, the performances of the classifiers were evaluated on the held-out test dataset using three common classifier performance measures [140]: (1) accuracy, defined as the fraction of predictions correctly classified by the model; (2) precision of class X, defined as the proportion of the predicted class X that is correctly classified by the model; and (3) recall of class X, defined as the proportion of the observed (true) class X, that is correctly classified by the model. To better contrast predictions of the three tree-based classifiers on the test dataset in order to select one for the ML method, we used McNemar's non-parametric test for pairwise binary classifier comparison [141] to test the null hypothesis that two models have similar proportions of errors. The normalized importance of features was generated for the selected model to understand the model's reliance on each feature when producing its predictions. Finally, the ML method used the selected classifier to classify registrations as FR and CR and then filtered FR from the unprocessed data. However, due to the limitations of video in covering the pop-hole area, the SD method was applied here to filter registrations in the winter-garden zone.

Table 1. Record features used to train the model and the normalized importance of features in the final CatBoost model.

Feature Name	Description
previous zone; zone; next zone	zone identity of the previous/considered/next registered record with the strongest LF signal, indicating the zone where the individual has transitioned/is transitioning/will transition to
RSS	a measurement of the power present in the strongest received LF signal (dB)
tracking system ID	identity of the tracking-system copy
previous duration; duration; next duration	reported time of stay in the zone from the previous/considered/next registered record
zone2	second zone identity with the strongest LF signal
RSS of zone2	a measurement of the power present in the second strongest received LF signal (dB)
zone3exist	binary feature that equals 1 if the tag registers a signal of at least three different zones during the last 10 s, and otherwise equals 0
next2zone = zone; previous2zone = zone	binary feature that equals 1 if the registered second zone from the next/previous record is the same as the occurring zone, and otherwise equals 0

We contrasted the two data-processing approaches by applying them to the unprocessed test dataset (i.e., including CR and FR). The resulting two processed datasets (ML and SD datasets), as well as the unprocessed test dataset, were then compared against the respective gold-standard dataset (i.e., registrations identified as CR through video observation). In each case, we evaluated two things: (1) the animal's location (or more specifically, their presence/absence in each zone) and (2) the animal's movement. To evaluate how well these datasets represented individuals' presence/absence in each zone at each second, we compared their associated categorical time series (containing five categories, one for each zone). The performance was evaluated in terms of accuracy, macro-averaged recall, and macro-averaged precision (where the macro-averaged recall/precision is the average of the recall/precision across each zone). To evaluate how well these datasets represented individuals' movement, we compared the total number of individual transitions per batch, per zone in each case.

Performance was evaluated with the explained variance score (EV) and the mean absolute error (MAE), defined as:

$$EV = 1 - \frac{variance\{y_{GS} - \hat{y}\}}{variance\{y_{GS}\}},\ MAE = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} |y_{GS_i} - \hat{y}_i|$$

where $\hat{\boldsymbol{y}}$ contains information from a processed dataset and $\boldsymbol{y_{GS}}$ contains the respective gold-standard information. The $\boldsymbol{EV}$ is used to measure the magnitude of the expected effect on the number of transitions [142]. The $\boldsymbol{MAE}$ is used to measure, in an unambiguous and natural manner, the magnitude of the expected average error [143] in terms of the number of transitions (for a two-hour batch). This analysis was performed with Python version 3.8.5 using the SciKit Learn package [144] for the performance measures and the CatBoost package [138] for the CatBoost classifier.

Investigation of Influencing Factors

When comparing large datasets with thousands of hours of tracking per animal, comparison with video recordings as a gold standard becomes impractical. Therefore, to further evaluate the tracking system, we used the tree-based classifier from the ML method to identify FR (IFR) and studied the estimated error rate, defined as the number of IFRs against the total number of registrations, in relation to specific factors. The estimated error rate had a value of one when all records were filtered by the ML method and a value of zero when none was filtered. This approach has some limitations due to probable FRs not being detected or some being falsely detected. However, by removing the limitation on the number of days and individuals used, a broader investigation of the systems' performance can be conducted. Data processing with the ML method is shown in the Results section to filter most of the true FRs (recall of class FR: 93%) and to filter mostly true FRs (precision of class FR: 84%). Therefore, IFRs should highlight most of the FRs from the unprocessed data and should be composed mainly of FRs. We applied the ML method over a two-month period, involving 144 animals, during which the hens were kept under similar management conditions every day, including 15 h of artificial light and six hours with access to the winter garden. To avoid biasing the data towards a greater error rate when the winter garden was closed, we excluded all registrations of transitions to the winter garden for periods when it was closed. We evaluated the estimated error rate in relation to different areas by reporting the mean ± SD of the estimated error rate across individuals for each of the five zones in each of the eight pens (40 pen-zone areas). We evaluated the estimated error rate in relation to external weather variables by fitting a mixed-effects logistic regression (link function: logit, R package "lme4" [145]) on the ratio of IFR to the total registrations minus IFR (per hour), with pen identity nested in station identity as a random term and hourly external humidity (%) and temperature (°C) as explanatory variables. External humidity was rescaled by dividing its values by 10. To control for variations barn management and animal behaviour throughout the day, the hour of the day was also added as a fixed effect. External humidity and temperature were obtained from the LSZB weather station (~12 km from the barn) and accessed via the Wolfram alpha API in Python.

Results

Evaluation of the Two Data Processing Methods

On the test dataset, the three classifiers showed stable (over 100 random seeds) accuracy, recall, and precision (Supplementary Figure S1), and the McNemar's test showed a similar proportion of errors between each classifier ($p > 0.05$). Thus, with an accuracy of 99%, we selected the CatBoost algorithm for the ML method because of its ability to handle categorical variables in Python. Additionally, 84% of the time that the model identified an FR, the model prediction was correct (precision of class FR). and 100% of the time that the model identified a CR, the model prediction was correct (precision of class CR). Additionally, 93% of the FR observations were classified by the model as FR (recall of class FR), and 99% of the CR observations were classified by the model as CR (recall of class CR). The zone identity, RSS, and the previous registrations' zone identity were the three most important features, accounting for 21%, 19%, and 13% of the overall importance of the features, respectively, while duration accounted for 7% (Figure 3A). To further illustrate the importance of the features, Figure 3B show the RSS and duration of the test dataset's registrations, split into CR and FR (from video observations). The receiving strength of the LF signal was generally higher for the correct registrations of all indoor zones. We also observed longer duration of stay to be more frequent among the correct registrations, with the exception of registrations in the lower perch zone, where no difference in the duration of stay was observed between correct and false registrations.

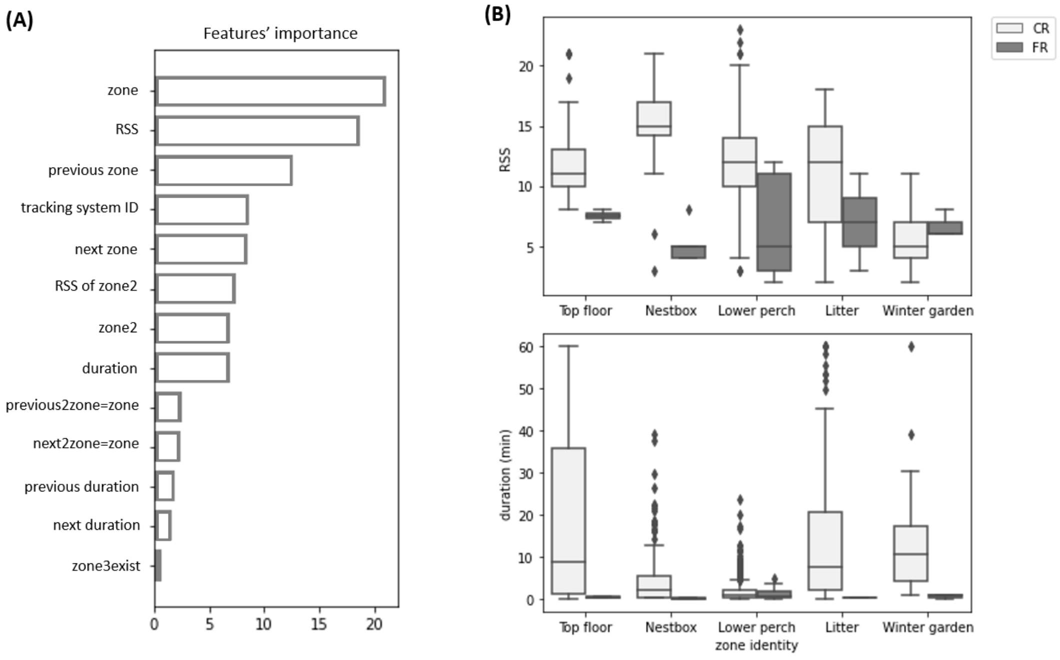


Figure 3. (**A**) Normalized importance of features for the selected CatBoost model. (**B**) Box plots of the RSS (top) and duration of stay (bottom) of the test dataset registrations, split into CR (light grey) and FR (grey), as produced by video observations, are displayed for each zone.

The unprocessed, SD and ML datasets all determined an individual's zone (every second), with an accuracy of 99%, 98%, and 100%, respectively, and displayed the same values (99%, 98%, and 100%, respectively) for the macro-averaged precision and macro-averaged recall. We found the ML and the SD datasets to underestimate the number of transitions by an average 0.27 and 0.06 transitions per zone, respectively, for a two-hour batch, in contrast to the unprocessed dataset, which overestimated the number of transitions by approximately 0.5 transitions per zone, on average, for a two-hour batch (average number of transitions per batch, per zone by video observation was 1.8). The percentage of variance of the ground-truth data recovered by the unprocessed, SD and ML datasets was 23%, 91%, and 99%, respectively, which is further illustrated in Figure 4.

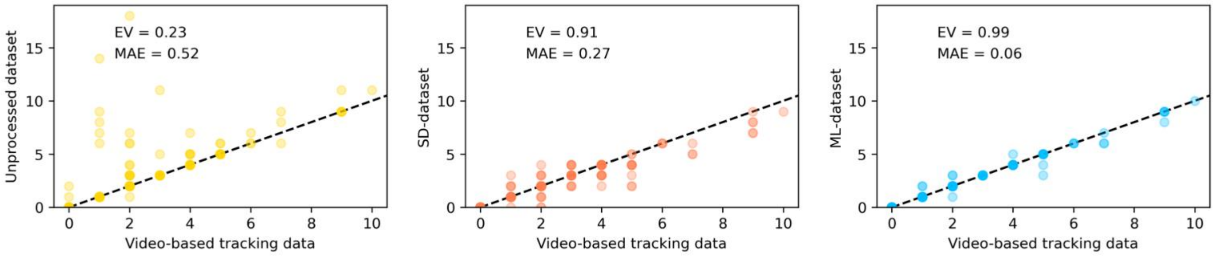


Figure 4. Number of transitions per individual (per batch, per zone) for the unprocessed, SD and ML datasets against video-based tracking data and associated EV and MAE scores. Overlapping data points are represented by darker shading.

Investigation of Influencing Factors

The estimated error rate across pen-zone areas varies from 0.0 ± 0.0 (e.g., litter area within each pen of Stations 10–12) to 0.5 ± 0.19 for Pen 8, suggesting that half of the registrations in the winter garden from Pen 8 were filtered by the ML method. The estimated error rate per pen-zone area is further detailed in Figure 5. Furthermore, we found a negative effect of humidity ($p = 0.003$) on the estimated error rate, with an odds ratio of 0.96 (95%-CI [0.94, 0.99]), indicating a 4% lower likelihood of obtaining a false registration with an increase in humidity of 10%. Additionally, we found a negative effect of temperature ($p < 0.001$) on the estimated error rate, with an odds ratio of 0.97 (95%-CI [0.96–0.98]), indicating a 3% lower likelihood of obtaining a false registration with an increase in temperature of 1 °C. The difference between the unprocessed and the processed data (by the ML-method) is further illustrated in Figure 6 through a visual representation of an animal’s transitions over eight consecutive days. For instance, observed several transitions filtered by the ML method between the lower-perch and top-floor zones.

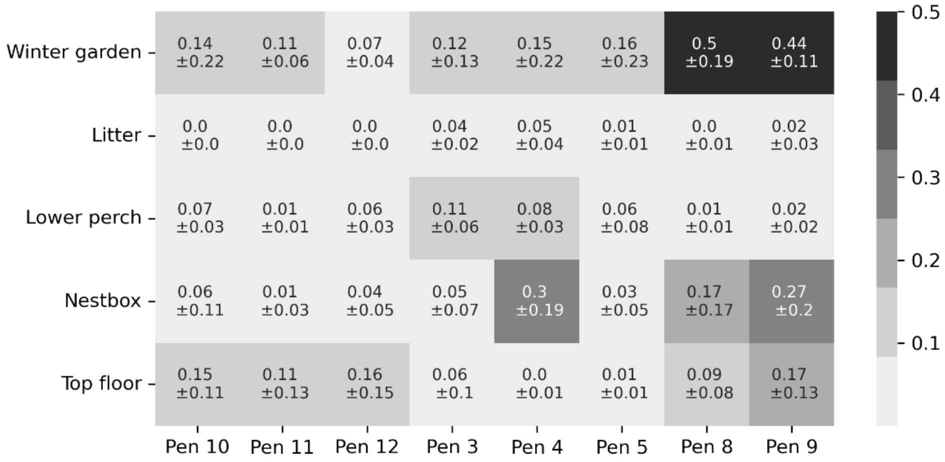


Figure 5. Estimated error rate (mean ± SD) for pen-zone areas.

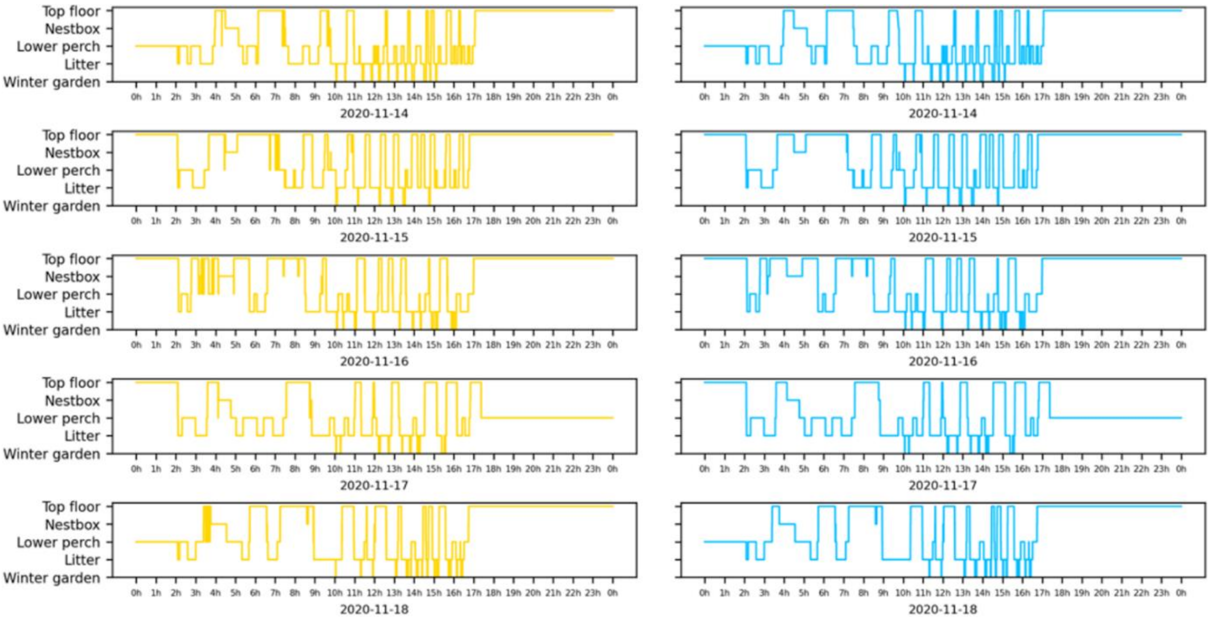


Figure 6. Unprocessed data (yellow) and processed data by the ML method (blue) of a single animal over five consecutive days. Each row represents a single 24-hour day, with the zone identities represented on the y-axis (the indoor zones are ordered following the aviary order, and the winter garden is represented below the indoor zones).

Discussion

We found the presented LF tracking system accurately determined the presence of animals in a given zone (at the second level), with macro-averaged precision, and macro-averaged recall of 99% when compared against video observations of the test dataset. This good performance might be explained by the tag algorithm, which searches for new transitions, on average, every 0.5 s (i.e., each time a tag receives an LF signal), thus regularly providing opportunities for correctional records. However, the number of transitions in a zone generated by the tracking system was overestimated and only explained 23% of the true variance (as observed by video). Therefore, the unprocessed tracking data did not constitute a good representation of individual transitions between the five zones, which could be emphasized by the observed differences in the estimated error rate within specific pen-zone areas.

On the one hand, we observed clear differences in the estimated error rate of a given zone across different stations (e.g., winter-garden zones in Stations 10–12, Stations 3–5, and Stations 8–9 had a mean estimated error rate varying, across their respective pens, between 0.07 and 0.14, 0.12 and 0.16, and 0.44 and 0.5, respectively). On the other hand, we observed differences within pens of the same station (e.g., nest-box zone in Stations 3–4 had an estimated error rate of 0.05 ± 0.07 in Pen 3, 0.3 ± 0.19 in Pen 4, and 0.03 ± 0.05 in Pen 5). The observed differences in the estimated error rate across different pen-zone areas aligned well with locations described through anecdotal notes made during video observations describing precise locations where a tracking tag generated a high amount of FR (by repeatedly switching between two, sometimes non-neighbouring zones) while the animal was immobile (weak spots). An explanation for the existence of weak spots may be the pen furnishing blocking the line of sight between tags and signal cables, which is known to cause signal interference in UWB systems [134]. More specifically, metallic materials can absorb the signal and distort the electromagnetic field, which could either block or enhance the signal, rendering RSS a poor representation of the distance to the signal cable, possibly explain errors between non-neighbouring zones. Our tracking system was designed to use a lower frequency than a common UWB system in order to avoid possible interactions with metallic materials, although signals may still be affected.

Furthermore, the existence of weak spots may be attributed to the calibration process, a manual and time-consuming step performed independently for each station and iteratively through each pen. When a tracking tag was detected in an incorrect zone during the calibration process, the LF values of specific markers were adjusted. As the pens are steel cages, the LF field generated by each marker can be slightly inhomogeneous. As a result, when an LF signal value is adjusted, all measurements must be repeated to ensure that the change in the LF value did not lead to further detection errors. The difficulty lies in setting the LF values of the markers in such a way that the correct zone is detected in all locations of the tracking system. In particular, the nest-box zone is a small zone located between two zones (Figure 2), and a change in the LF value of the marker had a greater effect on the neighbouring zones because it quickly led to the tracking tag being detected in the incorrect zone. Therefore, we recommend the use of an automatic calibration process to improve the system's performance. To achieve this, within each zone, several tags would be placed at predefined locations of critical measurement points. Each signal strength received by any tag from any marker would be registered. An algorithm would be executed every 10 s (ensuring enough time for adjustment of LF signal values to take place in the field) on all RSS registered within the past 10 s. This algorithm would identify the most problematic zone, defined, for example, by the zone with the smallest dB difference in relation to another zone (across all tags in that zone). If this difference does not exceed the limit of 1 dB in relation to another zone, the LF signal value of the associated marker is automatically adjusted. As soon as 100 consecutive runs induce no adjustment of an LF signal, the calibration is complete. An automatic calibration would save time as only one person would work on the calibration. This would also offer new opportunities, such as smaller zones, allowing for registration of finer movements. For instance, in our settings, it might be possible to differentiate between the nest boxes and the balcony in front the nest boxes (currently, both are registered as the nest-box zone). Furthermore, automatic calibration would ensure more homogeneous LF values across markers from the same zone across all pens, and consequently, more comparable datasets across different stations and pens would be generated.

Our tracking system's poor performance in representing individual transitions highlights the importance of processing automatically generated datasets. Relevant data-processing studies are lacking, although they could help to standardize this process to generate comparable datasets across different studies. The benefit of this work is most essential in light of rapid development in technology in order to manage and improve the welfare of animals within commercial livestock systems [94,95,146–151]. We showed that the data processed by a simple filtering of registrations associated with short durations (<1 min) of stays was suitable for monitoring the number of transitions per individual per zone, accounting for 91% of the actual variance (as observed by video). We further reported a gain in performance using a tree-based classifier to filter false registrations, accounting for 99% of the true variance in the number of transitions per individuals, which could partly be explained by the additional information provided to the ML method. Indeed, zone identity and RSS were the two main features upon which the tree-based classifier based its predictions, while the SD method was based solely on the duration of the stay. Interestingly, this also suggests that our expectation of the record's duration being the most important feature to detect FR was incorrect when other features are included. The current study did not allow for this comparison when a single feature is used; however, further studies using a simple rule-based approach should consider the RSS addition to the records' duration. The importance of features further suggests that the zone identities of the previous and next registered record are of greater importance than the duration of stay from the previous and next registrations. Results concerning the importance of can offer direction on how to improve similar tracking systems, for instance, by including a threshold of RSS values for each zone based, for example, on the result of an automatic calibration. Another possibility would be to include the SD method as part of the tag algorithm, although this would eliminate the possibility of registering fast transitions between two zones (<1 min).

Compared to the SD method, the ML method required additional efforts such as more video observations and statistical modelling and the choice between both methods relies on a compromise between time accorded in data processing and precision of the data. In the current study, the SD method recovered 8% less of the true variance in the number of transitions per individuals compared to the ML method. To put this value in context, we used simulated sampling to estimate the impact of a comparable loss on the effect size (measured by the Pearson correlation) of a simulated movement variable, M, on a simulated health variable, H. The two simulated variables (M and H) followed a standard normal distribution, with a Pearson correlation coefficient varying from 0.15 to 0.40 to cover potentially interesting ranges of effect sizes when studying movements in relation to health[93,152] and heritability of behaviour[54,72,153]. By adding noise to M (and calling it M'), we estimated (over 10,000 simulations per sample size) the percentage of cases where significance would be lost ($p > 0.05$), depending on the initial effect size and sample size. Our estimations suggest that a change in percentage of the initial variance explained by M' from 0.99 to 0.91 would change the significance of a critical test in 26% (or 25%) of cases when applying a sample size of 80 (or 120) and an initial effect size of 0.25 (or 0.2), respectively (see details in Supplementary Figure S2). Therefore, using a tree-based classifier to filter false registrations can be greater value for studies with small sample and effect sizes (e.g., n = 120, effect size of 0.2) than the filtering approach using stays of short duration as threshold. For large sample sizes or samples with strong correlations between the measured movement and the trait of interest, the SD method might produce equally reliable results as the ML method.

Our results further reported a marginal effect of periods of time characterized by higher humidity or higher temperature, associated with a lower estimated error rate of transitions to the winter garden. Because air is our medium of signal transmission, when humidity is changes, the magnetic field is also expected to change. As calibration was conducted in August 2020, the performance of the tracking system may be optimized for a period with higher temperature than average. Additionally, as Richards et al. [154] reported, associations between daily weather conditions and mean pop-hole usage in laying hens, including an increase in mean pop-hole usage associated with an increase in temperature, and the influence of the weather conditions on animal behaviour may be explanatory. In spite of these results, external environmental factors cannot be controlled for and are part of the experiment. However, these results can aid in interpretation and awareness of possible limitations for subsequent analyses of these or similar tracking data.

Conclusions

The active LF tracking system evaluated in this study determined the presence/absence of birds in a zone with an accuracy of 99% but overestimated the number of transitions by birds per zone, explaining only 23% of the true variation (as observed by videos). However, we showed that filtering stays of short durations rendered the data suitable for monitoring the number of transitions per individual, explaining 91% of the true variation, and that the use of a tree-based classifier to filter false registrations recovered an additional 8% of the true variation. Simulations further suggested that a machine learning approach for data processing could be of greater value than a simple deterministic approach in studies with small sample and small effect sizes. Results also suggest that filtering false registrations may reduce the effect of systematic errors towards certain pen-zone areas and towards periods of time characterized by lower humidity or temperature values. However, results also suggest that these factors might, to some extent, remain in the processed data and should be considered properly in subsequent analyses. In conclusion, this tracking system is well suited for complex indoor housing (similar to commercial aviaries) to measure the transitions of individuals and the presence/absence of birds in predefined zones (thus, duration of stays in zones). Nonetheless, under these settings, data processing remains a necessary step in obtaining reliable tracking data. For future work, we recommend the use of automatic calibration to improve the system's performance and to envision finer movements.

CHAPTER II

Based on:

Intra-individual variation of hen movements is associated with later keel bone fractures in a quasi-commercial aviary

Camille M. Montalcini[1,2], Michael J. Toscano[1], Sabine G. Gebhardt-Henrich[1] and Matthew Petelle[1]

1. Center for Proper Housing: Poultry and Rabbits (ZTHZ), Division of Animal Welfare, VPH Institute, University of Bern, Switzerland
2. Graduate School of Cellular and Biomedical Sciences, University of Bern, Switzerland

Published in: Scientific Reports (2023), https://doi.org/10.1038/s41598-023-29587-9

Camille Montalcini participated to the experimental work needed to collect the movement and health data, conducted the analysis, and wrote the manuscript in agreement with her co-supervisor Dr. Matthew B. Petelle and co-authors Dr. Micheal J. Toscano and Dr. Sabine G. Gebhardt-Henrich.

Abstract

Measuring intra- and inter-individual variation in movement can bring important insights into the fundamental ecology of animals and their welfare. Although previous studies identified consistent differences in movements of laying hens within commercial aviaries, the level of consistency was not quantified, limiting our capacity to understand the importance of individual movements for welfare. We aimed to quantify the scope of intra- and inter-individual differences in movements of commercial laying hens and examined their associations with indicators of welfare at the end of production. We quantified individual differences in one composite daily movement score for 80 hens over 54 days post-transfer to a quasi-commercial aviary. Results showed consistent inter-individual differences in movement averages, explaining 44% of the variation, as well as individual variation in predictability and temporal plasticity (at the population-level, hens increased their movements for 39 days). Hens that were more predictable in their daily movements had more severe keel bone fractures at the end of production while we found no such correlation between daily movement averages (individual intercept) and welfare indicators. Our findings highlight the importance of inter-individual difference in intra-individual variation of movements to improve poultry welfare.

Introduction

Intra-individual variation in movements of animals result from a dynamic interplay between factors such as health [10,29,155], spatial memory [156,157], need for resource acquisition [38], social interactions [65] and predation risk [158]. The temporal dynamics of these relationships illustrate the complexity of individual variation in movements as well as their importance for the fundamental ecology of animals [159] and animal welfare. For instance, intra-individual variation in movements could be used as an early warning of health issues in animals [160] but also to protect endangered species (e.g. when used as a complementary tool by wildlife managers for anti-poaching efforts [158] and for improving reintroduction success [161]). Quantifying biologically relevant behavioural variation is achieved using methods that decompose phenotypic variation into intra- and inter-individual components. The latter component provides information that is particularly relevant at the group level. For instance, differences between individuals in movements and space use behaviours may facilitate population-level adaptation in animals [51] as well as access to resources of group-housed animals where not all resources and areas can be accessed by all individuals simultaneously [89,162,163].

Until recently, measuring inter-individual variation in movements has been focused on consistent individual differences in averages (personality) or plasticity (i.e. average change in behaviour across a context) [164–166]. Focusing solely on phenotypic averages, however, restricts our understanding of important biological variation. By including residual intra-individual variation in behaviour (predictability), we obtain a more nuanced and comprehensive understanding of individuals [167–169]. To the authors' knowledge, this component of behavioural variation has never been studied in hen movements and its underlying mechanisms remain generally unclear and understudied [170]. Additionally, we lack quantification of consistent inter-individual differences in both average and variability of hen movements. Previous studies have identified consistent individual differences in daily movement patterns [9] or space use behaviours [11] within the interior of aviaries, but repeatability, a population-level measure quantifying the extent of individual differences, was not directly quantified. This gap likely extends from the challenge of monitoring individual movements in a densely populated housing containing multiple stacked horizontal levels (such as aviaries) and the various types of possible movements (e.g. flying, walking) between areas.

Within Europe, battery cages in the poultry industry are being replaced by cage-free systems (e.g. aviaries and free range) that are believed to benefit animal welfare mainly through greater ability to move freely and perform more natural behaviours [4]. Despite benefits, these modern systems, in particular aviaries, have a higher incidence of severe feather pecking, bacterial infections, and keel bone fractures compared to cage-systems [5]. These increased welfare issues in cage-free systems are due to a variety of reasons including the complexity of the structure which increases the risk of collision, and the large group size where bacterial infections or severe feather pecking are harder to detect and control. A better understanding of how individuals acclimate to and behave in this complex environment is a first, but important step to tackle these welfare issues. Because individual movements are indicative of how individuals interact with their environment, tracking systems are a promising tool to identify potential issues in cage-free systems, and improve poultry welfare (e.g. by modifying husbandry practices or housing systems to better allow expression of important behaviours [171] or reduce prevalence of potentially harmful behaviour [30]).

The transfer from a rearing to a laying barn is a standard practise in poultry farming that may be stressful to the animals. In the following weeks, hens will experience substantial environmental and internal variability (e.g. new husbandry practices, social settings; bone maturation, onset of lay), likely contributing to intra-individual variation in behaviours. Furthermore, these first weeks in the laying barn are important for hens' welfare and producers, due to an increased mortality risk at the onset of lay [172]. In this study, we quantified the scope of inter-individual differences in averages and in intra-individual variability of a composite score reflecting daily movements within a quasi-commercial aviary, over 54 days post transfer to a laying barn. We evaluated associations between the hen's welfare assessed at the end of production and both the hen's average and variability of the movement score. We expected to find individual variation in averages, predictability, and temporal plasticity of the daily movement score. We hypothesized that hens showing less behavioural variability would be behaviourally more constrained and thus less able to behaviourally cope with environmental changes (such as alterations in management practices or social structure). Therefore, we predicted that a high level of predictability or a low level of plasticity in movements would be associated with generally poorer welfare, including more severe keel bone fractures and feather damage.

Materials and Methods

Experimental design

This study involved initially four rearing pens each containing 630 Dekalb White chicks (*Gallus domesticus*, a widespread hybrid across the world within commercial egg production) from the same parent flock, housed in a standard rearing facility, containing an Inauen Natura aviary described previously [10]. All chicks came from the same commercial hatchery, but, for the purpose of a larger study, half of the chicks (housed in two of the four rearing pens) were hatched on-site within the mentioned rearing barn (on-farm hatch treatment; were transported 3 days before hatching). The other half of the flock underwent standard hatchery processing in the commercial hatchery (standard treatment) with transportation at one day of age. We classified all chicks into a more/less explorers' class (Supplementary Text S1). At seven days of age, we selected 96 focal chicks (24 / rearing pen; random selection of 10 animals amongst the more exploring class, 10 animals amongst less exploring class and four animals amongst the entire population) which we assigned to one of four identical laying pens associated with the rearing treatment (eight laying pens in total). We assigned focal hens to have equal representation of an individual's class and rearing pen throughout the laying phase. Chicks were individually identified with a leg band characterized by a unique color-number combination. At 17 weeks of age (WOA), we transferred all animals to an onsite quasi-commercial laying barn containing a Bolegg Terrace aviary split into 20 identical pens (previously described in Rufener et al. [10]), and an outside covered winter garden (WG) accessible by pop holes (barn schedule detailed in Supplementary Figure S2). On the day of transfer, we transported focal hens to their pre-selected laying pen with additional non-focal hens randomly selected across the same treatment, for a total of 225 hens/pen. All barns are located at the Aviforum facility in Switzerland where standard animal husbandry practices are used and from which we received pen-level production data. We extracted the average daily number of eggs per live hen (laid inside and outside the nest boxes) and the daily number of dead hens per pen (illustrated in Figures S3-S4). The WG and the litter area under the aviary are closed during the first eight days to encourage hens to use the nest boxes for egg-laying. To control for the external temperature (hereafter temperature) we extracted the temperature from the nearest regional weather station (LSZB, at Belp Airport, Bern, 10.2km from the barn) via the Wolfram alpha

API in python. The study was conducted according to the cantonal and federal regulations for the ethical treatment of experimentally used animals and all procedures were approved by the Bern Cantonal Veterinary Office (BE-45/20).

Tracking system

To track individuals across different areas within a laying pen, we distinguished five zones with key resources including the three stacked tiers of the aviary (top tier, nest box tier and lower tier), the littered floor and the WG. The top and lower tiers contained feed, water and perches, the nest box tier contained nest boxes and perches, and the litter floor contained litter. Due to the animal density, not all hens can be simultaneously in the nest box tier, top tier or WG. We used a low-frequency tracking system with active tags (mass: 28.1 g) enclosed in a backpack mounted on the back of the focal hens to register any movement across zones (transitions). The tracking system is composed of markers that emits low frequency signals (0.125 MHz) through a cable (creating separately enclosed fields for each zone), active tags (mass: 28.1 g) that can receive signals, and readers that communicates through ultra-high frequency (868 MHz) with the tags and a computer. The receiving strength of the signal is then used to determine theoretical distance to each cable, and in turn to each zone (see Montalcini et al. [173] for detailed description and validation). We collected tracking data from the first day in the laying barn, although subsequent analysis excluded the day of the transfer to keep only fully tracked days (i.e. the first 17h hours of tracking [10]). We stopped collecting data on day 54 because all focal birds were handled as part of another study which created a stopping point for data collection. We excluded 16 of the 96 individuals from our analysis due to equipment malfunction and death. Due to low battery level or equipment malfunction, some days were excluded for certain hens so that our analysis included 3,750 hen-days and 80 hens (average of 47 days/hen with a minimum of 40 days/hen and a maximum of 49 days/hen).

Movement Data

To quantify individual differences in average and variability of movements over time, it is important to use comparable observations throughout. Therefore, we extracted daily movement variables for each hen from when WG access was provided (i.e. from the second week onward). While the artificial light was on, we extracted the percentage of time spent in each of the five zones, the average number of stays in each zone (per hour to account for varying day length), the average travelled vertical distance (total number of the indoor zones crossed, e.g. a hen jumping from the top tier to litter crosses three zones) per hour and whether the hen entered the WG within 15 min after access was provided (scored 0-no/1-yes). While the artificial light was off (i.e., during the hens' night-time) we extracted the sleeping height, measured by the number of stacked tiers underneath the zone where an animal was for most of the night (value from 0 to 3, e.g. hens sleeping: on the top tier get a 3, on the litter get a 0). Because these movement variables may be intrinsically correlated and we aimed to assess how animals differed in their general daily movement without prior assumptions, we reduced these daily variables into one aggregate by extracting a linear composite variable from a correlation-based principal component analysis (PCA) using the psych package [174] in R. Due to all hens and all weeks not having the same amount of observations as a result of technical issues, we only included the first day of each week in the PCA for each individual (Supplementary Text S2) to ensure the same mass across weeks and individuals while accounting for the variation in movements across time and individuals. Three of the PCA's principal components had an eigenvalue > 1 [175], and explained respectively 41%, 20% and 14% of the total variation. Among the variable loadings on the first principal component, nine had an absolute value > 0.4 (Supplementary Table S1 a). On the first principal component, the percentage of time spent in the top tier loaded strongly in the opposite direction as the travelled vertical distance and the number of stays in both litter and lower perch. The loadings suggested that this first component reflected general movement throughout the indoor area, where a higher score is associated with animals that spent more time on the floor and lower tier but also transitioned more between these zones. We projected all observations onto the subspace spanned by the first principal component to obtain daily movement scores for each hen (PC1). The loadings of the other two principal components (not used in subsequent analysis) suggested two other behavioural axes, one mostly determined by the nest box tier usage and the other axis mostly determined by the WG usage (both in terms of the number of stays and the percentage of time spent; all loadings are detailed in the Supplementary Table S1 a).

We then used this daily movement score (PC1) to extract individual-level estimates (intercept, temporal plasticity, and predictability in movements) by gradually increasing the complexity of a linear mixed-effects model and using "best linear unbiased prediction" (BLUPS) to estimate random effects. Furthermore, as we could not extract the daily movement score (PC1) for the first week (due to the WG being closed for management purposes), we extracted an additional individual movement score based on the first week after the transfer to the laying barn, to further investigate the relevance of movements during the early laying phase for animal welfare. Because initial pen-level observations showed a surprisingly high number of birds not transitioning over entire days (34% of hens with no transitions during at least one of the first three days, a value reduced to 1% after 30 days), we extracted individuals' number of days without transitions between any zones (no-transitions-day). However, the first unintended interruption of the tracking system happened over several days from the 4th day onward, therefore we only used the first three days.

Welfare indicators

Two welfare indicators were assessed on each animal near the end of the production cycle (60 WOA). During the assessment, the two observers were blind to the treatment group (standard versus on-farm hatch), laying pen ID, and hen class. Keel bone fracture severity (KBF) scores (continuous, 0-100) were based on latero-lateral radiographs with a scoring methodology described in Rufener et al. [176]. The KBF score is an indicator of the total amount of bone affected by any fracture in the keel bone. To evaluate inter- and intra-observer reliability, we used the intraclass correlation coefficient with its 95% Confidence Interval (CI). Intra- and inter-observer reliability based on 40 radiographs were high, (ICC = 0.89, 95% CI = 0.74 – 0.95 and ICC = 0.92, 95% CI = 0.832 – 0.96, respectively). A feather damage score (continuous, 0-100) was assigned using the photographs of white laying hens which we rescaled to 0-100 and took the complement to 100 so that higher scores are indicative of poorer welfare (score 1: approx. 100 – 76 depending on the extent of damage; score 2: approx. 75 – 51 ; etc.) for each body part[128]. Lin's concordance correlation coefficient was used to assess intra- and inter-observer reliability (ICC = 0.92, 95% CI =89 – 94 and ICC = 0.86, 95% CI = 0.82 – 0.89, respectively).

Statistical Analysis

Quantifying inter-individual differences in movements

We quantified the extent of inter-individual differences in daily movements (PC1) average response, temporal plasticity, and predictability, by gradually increasing the complexity of a linear mixed-effects model [177]. To quantify inter-individual differences in averages we fitted a random intercept (RI) model to the movement score (PC1) with Hen ID as a random effect to allow the mean response to vary among individuals and to extract individual intercepts. We included as fixed effects: treatment, explorer class, laying pen identity, temperature (°C), time (number of days post-transfer to the barn with starting value 0), and the initial body mass assessed on the day of transfer to the laying barn with a digital scale in grams. Because hens may change their movement behaviour non-linearly across time, we added a quadratic effect of time in the fixed effects. We also accounted for individual differences caused by malfunctioning equipment by adding as a fixed effect hens' average number of tracked days post-transfer (21-24 days after access to the WG was provided). All continuous variables were scaled and centred to a mean of 0, except for time which was not centred, so that the individual intercept would reflect individual differences in their initial movements (i.e. on the first fully tracked day). To avoid convergence issues in more complex models we excluded the class identity (based on a likelihood-ratio test, LRT) and the laying pen identity random effect (explaining < 1% of the variance not accounted by fixed effects) in subsequent analysis unless otherwise mentioned. To evaluate the magnitude and effect of inter-individual difference in averages, we estimated the adjusted repeatability [178] by dividing the variance explained by the hen ID by the total phenotypic variance and its 95% credible intervals using 1,000 simulations of the posterior distribution of all variance components. For visual purposes, we estimated individual intercept (mean ± sd) with BLUPS by generating 1000 repeated samples from the posterior distributions of the RI model.

To analyse whether individuals also differed in their temporal plasticity, we extended the RI model to a random slope (RS) model, where hen ID is also able to vary across time [36] (hen ID x time; model named RS1). The inter-individual differences in their temporal plasticity were statistically tested by performing an LRT with the previous RI model. We extended this model including another random slope term allowing individuals to also vary in the curvature of their slope (hen ID x time^2; model named RS2) and evaluated its significance with an LRT. For RI, RS1 and RS2 models we reported the conditional and marginal R-squared and the Akaike Information Criteria (AIC). We checked model assumptions (i.e., normality of error and homoscedasticity) by sight. For visual purposes, we estimated individual slope (linear, quadratic) (mean ± sd) with BLUPS by generating 1000 repeated samples from the posterior distributions of the RS2 model. We fitted RI, RS and RS2 models with the lme4 package [145].

To evaluate individual movement differences in predictability we extended the RS2 model to allow estimations of residual intra-individual variation using a double hierarchical model (DHGLM) [179]. We included only significant fixed effects from the RS2 model for both the mean and the dispersion parts of the model. We used a Bayesian Markov Chain Monte Carlo (MCMC) approach using the brms package [180] in R. We ran the model with uninformative priors, 10 Markov chains, 50,000 MCMC iterations (including 25,000 for burnin), a thinning rate of three and a Gaussian distribution for the response variable (PC1). We extracted the mean and standard deviation from the posterior distribution for each individual's intercept. Individual estimates were multiplied by -1 so that individuals with higher estimates have a lower estimated residual variance and are considered as more predictable in their daily movements than individuals with lower estimates. The significance of the variance structure on the dispersion part was statistically tested with an approximate leave-one-out cross-validation (Loo-cv) [181], comparing the model with one excluding the dispersion part of the model. We quantified the strength of the inter-individual differences in predictability with the coefficient of variation in predictability (CV_p), a standardized population-level measure comparable across studies [177,179]. Specification of the model was assessed with a posterior predictive check, trace plots, the Gelman-Rubin's convergence diagnostic [182] and Loo-cv.

Association with welfare indicators

We fitted two bivariate Bayesian models to evaluate if the identified inter-individual movement differences associated with the KBF severity and feather damage scores, assessed at the end of production. We fitted two bivariate models, one with movement and KBF score as response variable and one with movement and feather damage score as response variables. As fixed effects for movement, we included the time, time^2, treatment, and temperature. As fixed effects for the welfare indicator, we included the individual predictability estimates and the number of no-transitions-day to estimate their associations with the two welfare indicators, as well as the treatment, the class, and the initial body mass to account for individual differences. We included hen ID as a random effect for both response variables, and individual random slopes (hens ID x time, hens ID x time^2) for the movement response only. We included laying pen ID as random effect for the welfare response variable. Model parameters were estimated using a Bayesian MCMC sampling method with the MCMCglmm package [183] for 700'000 iterations with a burn-in phase of 105,000 and a thinning interval of 100. As welfare scores do not have repeated measures at the individual-level we did not allow the variances of the residuals to covary, and we further constrained the residual variance of welfare scores to be fixed and close to zero. Note that to avoid a stats-on-stats issue [184] when inferring on individuals' welfare measured on a single instance, with the individual predictability estimate we would need to implement a DHGLM bivariate random regression model with fixed variance structure, which is not supported by the brms package, and so beyond the scope of this paper. We specified priors for the variance structure of the residuals (R) with scale equals to a diagonal matrix with entry 1 for the movement variance and 0.0001 for the KBF variance, and with degree of belief equal to 1.002. We specified parameter expanded priors for the variance structures of the random effects ($G1$ for the penID ; $G2$ for the HenID) with scale (V) equals to a diagonal matrix with 1 on the diagonal (I), degree of belief equal to the dimension of V (i.e. 1 for $G1$, 4 for $G2$) and a multivariate normal prior specification for the redundant working parameters with null mean vector and covariance $I \cdot 25^2$. Further details on the linear model for the latent variable can be found in the Supplementary Equation S1. Model diagnostics were checked with the trace plots, autocorrelations (< 0.05 for all parameters), and the Gelman and Rubin's convergence diagnostic with 3 chains having over dispersed starting values ($\hat{R} \leq 0.03$ for all parameter) [182]. We deemed a factor or correlation significant if the 95% credible intervals excluded 0.

Results

Quantifying inter-individual differences in movements

Hens differed in their average movement (repeatability = 0.44, 95% CI = 0.43 – 0.48), meaning that, on average, 44% of the remaining variance (after controlling for fixed effects) in movement can be attributed to differences between individuals. Hens also exhibited differences in temporal plasticity (both linearly and quadratically). Random effects variance estimates for both RI, RS, and RS2 are further detailed in Supplementary Table S2. In addition to inter-individual differences in average and temporal plasticity, results showed that hens differed in their predictability, exhibiting differences in their within-individual variability of behaviour around the mean, with a coefficient of variation in predictability among individuals (CV_p= 0.25, 95% CI = 0.20 – 0.30). Modelling the dispersion part in addition to the mean part of movement improved the model. We found a negative correlation ($r_{model\,name}$) between individual intercepts and linear random slopes (r_{RS1} = -0.79, bootstrap 95% CI = -0.82 – -0.78; r_{RS2} = -0.69, bootstrap 95% CI =-0.76 – -0.66 ; r_{DHMM} = -0.69, 95% CI = -0.80 – -0.55), a positive correlation between random intercept and quadratic random slope ($r_{RS2} =$ 0.41, bootstrap 95% CI = 0.37 – 0.54 ; r_{DHMM} = 0.42, 95% CI = 0.21 – 0.60) and a negative correlation between linear and quadratic random slopes (r_{RS2} = -0.89, bootstrap 95% CI = -0.93 – 0.89 ; r_{DHMM}= -0.90, 95% CI = -0.94 – -0.84). These results suggest that hens with initially lower movement increased their movement more rapidly than hens with higher initial movement. We found no correlation between predictability and the other individual-level metrics (intercept: r_{DHMM} = -0.02, 95% CI = -0.26 – 0.23, linear slope: r_{DHMM} = 0.10, 95% CI = -0.16 – 0.34 and quadratic slope: r_{DHMM} = -0.08, 95% CI = -0.33 – 0.17). Figure 1 illustrates these results by highlighting four daily time series of hens' initial movements (a), daily PC1 scores with the RI, RS1 and RS2 models predictions, and (b) individual intercepts, slopes and predictability estimates relatively to all other hens (c).

In addition to individual-level variation, we found both linear and quadratic effects of time at the population level. The population, overall, increased their movement until day 39 in the laying barn, at which point, their activity started to decrease (illustrated by the black line in Figure 1b; Supplementary Table 3). We also found individuals hatched on-farm moved on average less than individuals that were transported at one day of age to the rearing barn, and hens moved generally less within the aviary as the temperature outside increased. Results from fixed effects from RI, RS1 and RS2 models are detailed in Supplementary Table S3.

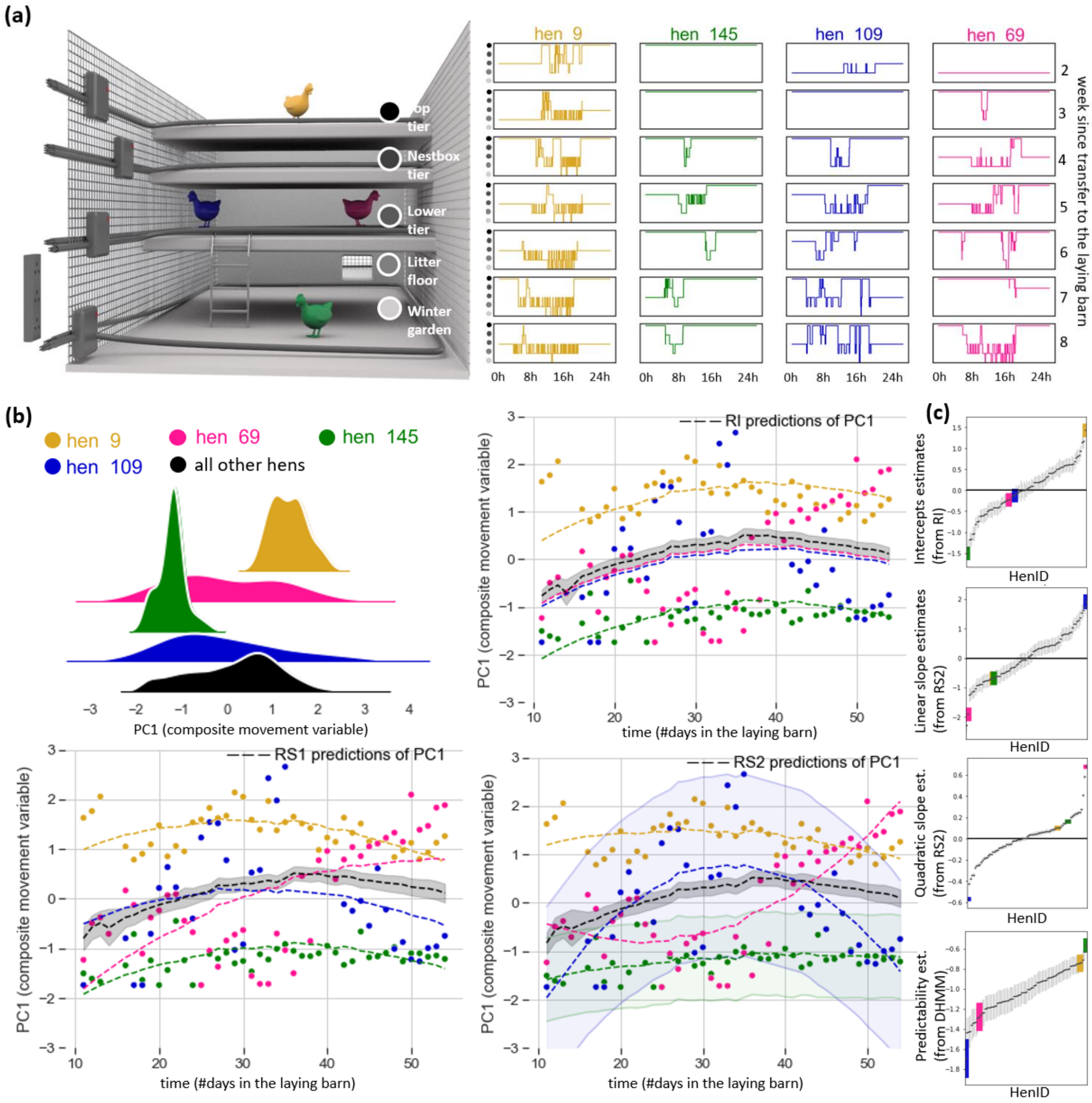


Figure 1. Visual representation of four hens' transitions across the five aviary zones, daily PC1 scores and individual estimates. Hen 145 and hen 9 have the smallest and largest intercept estimates, respectively. Hen 109 and hen 69 have the smallest and largest quadratic slope estimates, respectively. **(a)** On the left side we showed a simplified representation of the laying barn and the tracking system equipment (dark grey) and on the right side we showed seven (one per week) daily time series of the hens' transitions across the five aviary zones. Each column represents one hen and each row represents the first tracked day of a particular week. **(b)** PC1 daily scores of four hens including a kernel density estimate (top left); the daily PC1 score over time (dots); the PC1 predictions from studied models (dashed lines); shaded areas around the prediction of two hens to illustrate predictability estimates (represented by a constant value = $\max_{i \in day}(|prediction_i - observed_i|)$). **(c)** Individual estimates (mean ± sd) of studied hens (intercept, linear slope, quadratic slope, and predictability) sorted by the smallest to the highest estimates.

Association with welfare indicators

Descriptive statistics of the KBF severity and the feather damage scores are given in Table 1.

welfare indicator	mean	standard deviation	min	0.25 percentile	median	0.75 percentile	max
KBF severity	36	15	6	24	36	45	84
Feather damage	34	11	13	27	34	42	65

Table 1. Descriptive statistics of the KBF severity and the feather damage scores assessed at 60 WOA on the 80 hens used in the models.

Individuals' predictability of movements had a positive association with individuals' KBF severity (posterior mean = 23.54, 95%CI = 6.57 – 38.70, Supplementary Table S4), individuals that were more predictable had greater KBF scores (Figure 2). We did not find correlations between KBF severity and individual intercepts (r = 0.06, 95%CI = -0.22 – 0.26) or individual linear slopes (r = 0.18, 95%CI = -0.05 – 0.44), but did find a weak negative correlation between KBF severity and individual quadratic slope (r = -0.24, 95%CI = -0.51 – -0.04). Furthermore, hens that had a higher number of no-transitions-day during the first three days had higher KBF severity at the end of production (posterior mean = 5.12, 95%CI = 1.24 – 8.69, Supplementary Table S4, Figure 2).

We did not find an association between individuals' feather damage and individuals' predictability nor the number of no-transitions-day during the first week. We also found no correlations between feather damage score and individual intercepts, linear, and quadratic slopes (r = -0.01, 95%CI = -0.27 – 0.24, r = -0.22, 95% CI = -0.42 – 0.11, and r = -0.26, 95%CI = -0.06 – 0.45, respectively). Estimates from fixed and random effects are detailed in Table S4 and Equations S2-S3 from the supplementary materials.

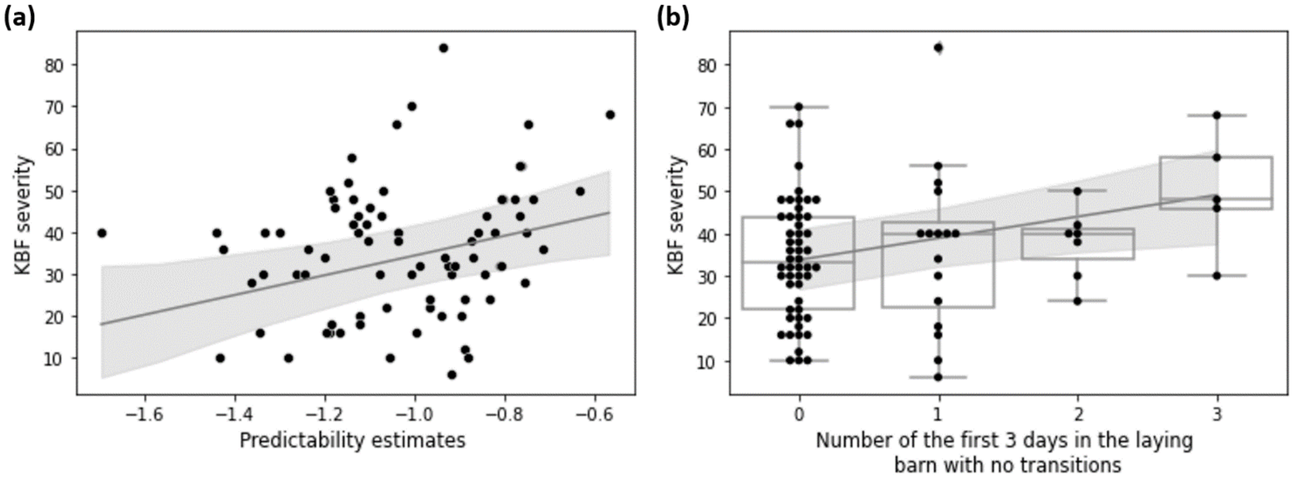


Figure 2. Raw data points of KBF severity in relation to individual predictability **(a)** and the number of days with no transitions during the first three days in the laying barn **(b)**, together with the predictions line and its 95% credible intervals from the bivariate model.

Discussion

Consistent individual differences in movement behaviour in both wild and domestic species have only recently received attention [61,159,177,185]. Previous studies have identified consistent differences in daily movement patterns [9] or space use behaviours [11] of laying hens within aviaries. In the current study, we extended that understanding by quantifying the extent of inter-individual differences in averages (with repeatability) and variability (with the coefficient of variation in predictability) of daily movements using two population-level measures that are, to some extent, comparable across traits and studies. We believe the current effort is the first to reveal intra-individual variability of laying hen movements across time and around their behavioural mean within an aviary. We also found associations between these individual movement scores, assessed during the first two months in an aviary, and the severity of keel bone fractures at the end of production. Collectively, our results highlight the importance of movements during the early laying phase and of the intra-individual variation in movements to explain occurrence of keel bone fractures at end of lay.

We found consistent inter-individual differences of hen daily movements within the first two months in a laying barn, with 44% of the variation attributed to individual differences (repeatability (R) = 0.44). Our study's repeatability is slightly higher than the average reported repeatability of behaviours (across eight taxa: R = 0.37) [35] and lower than the average repeatability based solely on spatial behaviours (across five taxa: R = 0.67) [51]. Various covariates may explain difference in repeatabilities, including individuals' life stage [51]. For instance, a meta-analysis conducted by Stuber et al. [51] found evidence that repeatability increased with increased age, so that difference between individuals during adulthood explained on average about 30% more of the behaviours' variance compared to juveniles. In this study, we used a transitional life stage between juvenile and adulthood, where individuals are probably still developing cognitively and physiologically as well as gaining spatial experience in their new housing and thus adjusting their behaviours. Therefore, individuals in this study likely exhibited a higher within-individual variation and thus lower repeatability compared to later stages.

After controlling for individual differences in average and temporal plasticity, hens still differed in how predictable they were in daily movements with a relatively low degree of variation within our population ($CV_p = 0.25$) but similar to previous results on movement distance (total distance travelled of calves *Bos taurus*: $CV_p = 0.18$ [186] and mean daily distance of wild African elephants *Loxodonta africanus*: $CV_p = 0.27$ [177]). Quantifying the degree to which individuals vary in their movements within commercial aviaries may be particularly important in light of the restricted space and the high density of animals. Further research is required to understand if a high degree of variation in specific spatial behaviours among individuals housed together would facilitate access to resources and affects animals' welfare and productivity (such as the timing of nest box usage if it reflects mean oviposition time where maintaining variability could limit competition for nest boxes [163]).

In addition to the two population-level measures (R and CV_p) estimating the degree of individual variation, we found both linear and quadratic effects of time at the population level. Overall, individuals increased their movement until day 39 in the laying barn, that is five days before the last change in the light schedule, near the end of the mortality peak (on average 2.6% of hens died in the studied pens during the first 54 days in the laying barn) and around the moment hens reached maximum production level. It is likely that the egg laying behaviour was an important factor driving the temporal plasticity in daily movements (encouraging them to move to the nest box tier) as well as stress (suggested by a mortality peak and by the high number of hens not transitioning over entire days from the top tier). We also identified differences between individuals in the rate at which they adjusted movement in the laying barn (i.e. temporal plasticity), suggesting that the effect of time at the population level will not capture all relevant information for each individual. The negative correlation between the random intercept and linear slope indicates that individuals with fewer initial movements had a more rapid increase in movement over time as compared to individuals with initially more movements, which may be explained by a regression towards the mean.

The absence of correlations between individuals' predictability and the other individual-level effects (intercept, linear or quadratic slopes) highlights the relevance of a multidimensional approach to study inter-individual differences in movement that would account for intra-individual variability in addition to individual average response. We also found that individuals exhibiting higher predictability in daily movements had more severe KBF at 60 WOA, supporting our hypothesis that hens with less variability in daily movements may be more constrained in their ability to behaviourally adapt to the commercial environment and in turn have a reduced welfare, compared to less predictable animals. A previous study suggested that space use in the laying barn may be related to differences in spatial cognitive abilities, with hens that never went outside showing lower spatial abilities [157]. Animals with less predictable daily movements may need greater abilities to maneuver in changing conditions (social and environmental) and thus would have greater spatial cognitive ability. Greater spatial cognitive ability would in turn allow individuals to navigate better, leading to lower collision rates with internal structures, a potential cause of KBF [30,187].

We found no association between temporal plasticity (linear or quadratic) or predictability in daily movements and feather damage. These results suggest that greater variability in movements and possibly more behavioural flexibility to adapt to new social constraints, did not allow hens to better avoid feather pecking, perform maintenance behaviours, or more generally, avoid feather damage. Alternatively, the absence of an association could be explained by the little variation in feather damage between individuals (i.e. most animals had low feather damage). The few studies that have reported associations between individual movements and feather damage (so that hens with greater feather damage used the outside area less frequently [155]) where based on free-range housings which have important differences with our study that contained a WG but no free-range area (i.e., no access to grass and uncovered areas). Also, the WG usage is not well represented by our composite movement score (PC1). More studies are needed to confirm that an animal's predictability in its daily movement routine associates positively with KBF, but not with feather damage, and whether spatial cognitive ability could explain the association.

Individuals with greater number of no-transitions-day had more severe KBF at the end of production. Every day a hen did not transition between tiers, the hen remained on the top tier, a zone with all essential resources, many perching locations but that lacks important resources (litter, nest boxes and direct UV sunlight). Although reduced accessibility to all resources may compromise an animal's welfare, it is unlikely that this would explain the increased KBF observed at the end of production. Baur et al. [21] observed almost no KBF at 22 WOA in hens housed in the same barn, which suggest that these days spent on the top tier were not a consequence of KBF but rather expressed in response to the new environment. Because Rufener et al. (2019) [10] found an increased duration of stay in the top tier with increasing KBF severity, it is possible that the top tier is used to offset stress in response to aversive situations such as the transfer to a barn or severe KBF (e.g. by using the perches to avoid more dominant bird [188]). Other locations, such as the nest box, are known to be used in daytime by hens to hide and escape aggression [188], and the top tier may offer similar refuge as it is accessible at any time (while nest boxes are typically closed in the afternoon).

Altogether, our findings support the relevance of hen movements during the early laying phase, in terms of predictability in daily movements and the number of days spent without transitioning between any aviary's tiers, to explain later KBF (both movement scores associated positively with KBF score), but further research is required to disentangle underlying mechanisms. Previous literature suggested that the prevalence of KBF increases with the spatial complexity and height of the housing system as well as the presence of perches [23], and reported KBF to be associated with decreased mobility [27,28]. Although, hens that did not transition for an entire day remained on the top tier (the highest tier with many perching locations) we believe it is unlikely that these movements affect the keel bone directly since there is a gap of 250 days between data collections.

Instead, we propose that these movement scores are expressed as part of a proactive/reactive coping style [80] that remain consistent through an animal's life, and which would be associated with both hens' behaviours and welfare. From a neuroendocrinological perspective, more proactive animals may have less inhibitory control and in turn may be more predictable behaviourally compared to reactive animals [189], which would be more flexible and perform better under unpredictable environmental conditions [190]. Therefore, hens with less predictable movements may have a reactive coping style and as such be more able to adjust their behaviour to the new environment compared to hens with more predictable movements, which could explain these long-term association to KBF. Further research is required to understand how coping style relate to farm animal welfare, such as KBF, though our methods offer a novel approach to aid welfare assessments.

Exploring coping behaviour in farm animals can provide valuable information to improve animal welfare [55,80] by optimising husbandry practices and allow individuals to perform effective coping behaviour [55] in the laying barn. Our study is in line with previous research showing an increased mortality risk at the onset of lay [172], and further suggest that during the early laying phase some areas may be overcrowded (here, the top tier) or not fully utilized. These findings provide new information that may help to better design and properly prepare hens for inhabiting three-dimensional systems in the future. Furthermore, exploring proxies of coping style (such as potentially predictability in movements) may in the future help to tentatively breed for more resilient farm animals [55,80]. However, the heritability of plasticity and predictability of behaviour or stress responsiveness is still relatively unstudied [191,192] and will need to be investigated to determine the relative benefits.

Conclusion

The present study revealed the presence of consistent inter-individual differences in average movements of hens as well as individual variation in predictability and temporal plasticity within the first two months in a quasi-commercial aviary. We found associations of intra-individual variability in daily movement with the severity of KBF and observed a mortality peak and a high number of hens not transitioning over entire days early on. Altogether, these findings highlight the importance of the early laying phase for animal welfare and revealed considerable individual differences providing new information that may help to better design and properly prepare hens for inhabiting three-dimensional systems in the future.

CHAPTER III

Based on:

Commercial laying hens exhibit long-term consistent individual differences and behavioural syndromes in spatial traits

Camille M. Montalcini[1,2], Matthew Petelle[1], and Michael J. Toscano[1]

1 Center for Proper Housing: Poultry and Rabbits (ZTHZ), Division of Animal Welfare, VPH Institute, University of Bern, Switzerland
2 Graduate School of Cellular and Biomedical Sciences, University of Bern, Switzerland

Published in: Royal Society Open Science (2023), https://doi.org/10.1098/rsos.230043

Camille Montalcini participated to the experimental work needed to collect the movement and health data, conducted the analysis, and wrote the manuscript in agreement with her supervisor Dr. Michael J. Toscano and co-author Dr. Matthew B. Petelle.

Abstract

Past research has supported the importance of animal personalities for the productivity and welfare of farm animals. However, current assessments of personality traits are commonly conducted over short periods using standardized assays and may not reflect all important aspects of behaviours in commercial settings throughout the production period. This study aimed to evaluate consistent behavioural differences between 194 commercial laying hens within an aviary across most of the production period (eight months). We used five spatial behaviours related to various aspects of commercial hens' daily routine, including the sleeping, feeding, nesting, indoor movements, and outdoor usage. All behaviours were repeatable over time and across contexts, with consistent differences between individuals explaining between 24% and 66% of the variation. These long-term consistencies revealed the potential applicability of the behaviours as personality traits of commercial hens. Moreover, we identified behavioural syndromes comprising all behaviours except the nesting related behaviour, indicating two axes of spatial personalities that may be driven by different mechanisms. We discussed the significance of such individual differences in using personality traits to breed more resilient farm animals. Future research should evaluate associations of these behaviours with animal welfare and productivity to inform breeding efforts.

Introduction

Animal personality is defined as repeatable individual differences in behaviour over time and across contexts [53,54]. Personality traits can limit behavioural plasticity and hinder individuals from behaving optimally in all situations. That is, if a behaviour is not plastic enough, it will be suboptimal in some contexts. For example, high levels of feeding activity will be optimal when predator abundance is low but suboptimal when predator abundance is high due to higher predation risk. Therefore, studying correlations of a behaviour across contexts could explain suboptimal behaviours and identify important trade-offs driving individuals' fitness [81]. Personality is a multidimensional concept and can be summarized into five axes: aggressiveness, activity, exploration, boldness, and sociability [53,55]. These axes are often correlated into a behavioural syndrome [56]. More aggressive individuals tend to be more active, explorative, bold, as well as less social, and would commonly be classified as proactive animals (in contrast to reactive animals) [55,57,58]. Correlated traits can constrain evolution and help to understand key ecological processes and life history strategies, such as population dynamics and survival (e.g. 'fast'/proactive personalities [60] tend to disperse over longer distance [61] and activity - risk-taking syndrome affecting survival [59]). Importantly, behavioural traits can also be correlated to morphological and physiological traits, which is of particular importance in farm animals.

Indeed, there is a growing interest in farm animal personalities owing to its association with individual welfare and productivity. For instance, previous literature suggested that less nervous cows produced more milk [76], more exploratory-active calves associated with greater average daily gain [77], and calmer temperament in cows were related with greater first lactation milk yield [78]. A greater understanding of farm animal personality could help to improve management practices and the design of housing systems for increased productivity and welfare. Research suggested another critical role of animal personalities for the welfare of farm animals by proposing the integration of personality traits as phenotypes into the breeding process to breed for more robust farm animals [55,79,80] (e.g., in pigs [84], laying hens [83], cows [85]). Yet, personality traits are commonly assessed in laboratory settings or with standardized assays [105] on a limited number of individuals and over short periods of time. As a result, these traits may not reflect all important aspects of behaviours expressed within commercially relevant settings [106,107], which could lead to important misinterpretations that could impact how animals are housed, managed, and bred. Therefore, technologies that allow automatic monitoring within commercial settings will be key in assessing farm animal personalities.

Following the incorporation of monitoring technology in the study of animal behaviour, how animals use their space is garnering increased attention [193]. Movement and space-use are fundamental behaviours for common personality traits [61] such as boldness, activity, or exploration behaviour. Spatial behaviours are of particular importance in farm settings where animals have to live within complex human-made housings and where the high animal density and bounded environment constrain freedom of movement. The importance of spatial behaviours is especially true for commercial laying hens, where cage-free housing systems are becoming more common. Although thought to provide improved welfare over cage housing, cage-free housing is associated with certain welfare issues (e.g. severe feather pecking, bacterial infections, and keel bone fractures [5]) that are more difficult to resolve with traditional management and genetic interventions. We suggest that by using tracking technologies that allow high resolution, individual-level observations over long periods of time, we can establish new spatial behavioural traits in laying hens that can in turn be interpreted as personality traits, give insights into animal needs and preferences, and offer a promising tool to breed more resilient farm animals.

This study aimed to evaluate consistent behavioural variation between commercial laying hens within an aviary system across most of the production period. Our objectives were to: (1) characterize five spatial behaviours related to a hen’s daily routine including: sleeping, feeding, nesting, indoor movements, and outdoor usage, (2) quantify the extent of consistent inter-individual differences (repeatability) over time and within four contexts for each behaviour, (3) evaluate the maintenance of these individual differences across contexts, and (4) explore the existence of behavioural syndromes. We hypothesised that such behaviours are indicative of personality traits and therefore predicted that all five behaviours would be repeatable over time and across contexts. We further hypothesized that the underlying personality traits are indicative of proactive and reactive strategies, and therefore predicted the behaviours to be correlated into behavioural syndromes.

Materials and Methods

Study design

The study was conducted according to the cantonal and federal regulations for the ethical treatment of experimentally used animals and approved by the Bern Cantonal Veterinary Office (BE-45/20). As part of a larger study, 2’520 chicks (*Gallus gallus domesticus*) were reared in one of four rearing pens (630 chicks/pen), where half of the chicks hatched on farm (OFH treatment) while the other half arrived at one day of age from a commercial hatchery (TRAN treatment). Chicks originated from a single parent flock from a standard commercial hybrid (DeKalb White). At seven days of age (DOA), all chicks were classified into a more/less explorer class. We will not use the class as an exploratory behaviour as the measurement could not be validated (Supplementary text 1), but we will control for the class in subsequent analysis. At 7 DOA, we selected 160 focal chicks (40/rearing pen; arbitrary selection of 10 animals amongst each class and 20 animals amongst the entire population). On the same day, we arbitrarily assigned (but not yet transferred) each focal chicks to one of four identical laying pens from the same treatment (4 pens/treatment for a total of 8 laying pens, with 20 focal hens each), maintaining an equal representation of an individual’s class and rearing pen throughout the laying pens. At 17 weeks of age (WOA; September 2020), we transferred all hens, including focals, to an onsite laying barn containing a Bolegg Terrace aviary (separated into 20 pens by grids, pens’ indoor area: 7 m length, 2.3 m width, 2.69 m height until the top tier grid floor) and an outside covered winter garden (WG). Animal density was 8.1 hens per square-meter of permanent accessible area (225 hens/pen of 27.92 m^2). At five timepoints (127, 173, 243, 313, 417 DOA), we randomly selected 16 focal hens (2 hens/pen) to be killed as part of the larger study and replaced them with 16 arbitrarily selected hens to continuously track the same number of hens (see below for descriptive statistics that accounts for technical issues). At five slightly different timepoints (173, 215, 243, 313, 417 DOA), all focal hens were weighed and radiographed to produce a latero-lateral image that was then used to generate a keel bone fracture (KBF) severity score (continuous, 0-100) based on a tagged visual analogue scale [176]. The KBF score is an indicator of the total amount of the keel bone affected by any fracture (intra-observer reliability: ICC = 0.89, 95% CI = 0.74 – 0.95, inter-observer reliability: ICC = 0.92, 95% CI = 0.832 – 0.96). Both rearing and laying barns are located at the Aviforum facilities in

Switzerland where standard animal husbandry practices are used. Hens were kept for commercial and experimental purposes until July 2021.

Tracking system

We tracked focal hens across five zones in the laying pen: the top and lower tiers (both containing water, feed and perches), the nestbox tier (with a balcony and nestboxes), the littered floor and the outside WG (containing water and litter). The four indoor zones and tracking system are illustrated in figure 1 from Chapter I, where a detailed description and validation is described. We used the movement data from the point at which the daily number of eggs laid by the flock peaked and the management schedule (i.e. timing of lighting and wintergarden access) became stable (i.e. at 25 weeks of age), until the end of production (i.e. at 60 weeks of age). During the tracking period, artificial light was turned on at 02:00 h and off at 17:00 h. The dataset included 194 hens tracked over 242 days (i.e. 8 months) resulting in a total of 30,780 observations (after removal of non-functional tags or days), amongst which hens had on average 159 days tracked (min = 12, max = 200, sd = 52, 25th percentile = 120, 50th percentile = 193, 75th percentile= 198 (unit: day)).

Spatial behaviours

We selected five daily spatial behaviours based on five aspects of commercial hens' daily routine, so that behaviours may be functionally different. We attempted to avoid behaviours that would be intrinsically correlated, so that any behavioural syndromes that we identified should arise from individual hens' preferences. We hypothesized that hens with a more proactive personality would: travel greater vertical distances, use the WG on more days, have earlier nesting behaviours, spend more nights on the highest tier, and express a stronger reaction to the feed being delivered by using the tier with feed more when fresh feed is available. Therefore, we predicted positive among-individual correlations (i.e., behavioural syndromes) between these behaviours. Figure 2 illustrates the percentage of focal hens in a zone at specific time of the day, along with the timing of the main husbandry practices.

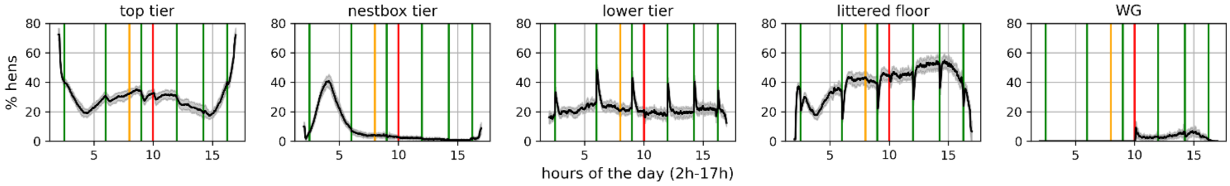


Figure 2- Black lines correspond to the mean (± SD) percentage of focal hens in a zone at specific time of the day, computed over all available days in January (arbitrarily chosen), where focal hens represent ~9% of the hens in each of the 8 pens. The main daily husbandry practices are highlighted in coloured vertical lines (orange: collecting eggs on the floor; red: access to WG opened; green: fresh feed delivery).

Vertical travelled distance

Because the four indoor zones are stacked on top of each other, we used the number of indoor zones crossed as a representation of a hen daily vertical travelled distance. More specifically, we used the mean vertical travelled distance per hour excluding the dusk, night, and dawn phases and the time when the hen was in the WG to prevent overlap with the sleeping and outdoor usage behaviours. We tracked movements between zones but not within them, and thus did not measure hens' activity. However, we hypothesized that hens that travel greater vertical distances would on average be more active and thus have a more proactive personality.

Winter garden presence

The outside covered WG (9.32 m^2) could be accessed via the littered floor through a pop-hole (length: 60cm, height: 38 cm, width: 28 cm) from 10:00 h until approximately 16:00 h on most days. Most of its area cannot be seen from the inside and individuals must go to the edge of the pop-hole to see the entire WG. Although the WG is screened, it is subjected to more variable environmental conditions than the indoor area, which may cause some animals to perceive it with greater uncertainty. Previous research in free-range housing found negative association between range-use and fearfulness [12,17,194], although free-range area differs from our WG (i.e., no uncovered areas or access to grass). We used whether the hen entered the WG on the day (scored 0-no/1-yes) as the daily outdoor usage behaviour. We hypothesized that hens using the WG on more days would be less fearful and thus would have a more proactive personality.

Nestbox tier timing

Nesting behaviour is a highly ritualized and internally motivated process [195,196]. In cage-free systems, hens should lay their eggs in shared nestboxes that cannot hold all hens simultaneously (here: 2.3 m^2 of nestbox surface per pen of 225 hens). Therefore, dominance hierarchies may affect laying behaviour, with subordinate hens laying their eggs (and potentially also performing nesting behaviour) slightly later than dominant hens [197]. Therefore, to account for nesting behaviour, we used the period when hens are expected to lay (i.e., 02:00h – 08:00h) and extracted the point in time (h) when a hen reached half of its time spent in the nestbox tier (hereafter referred to as nestbox tier timing). Because the nestbox tier contains the nestboxes but also a balcony (figure 1a), we used several other lines of evidence to show that this likely reflects nesting behaviour. We believe the measure reflects nesting behaviour as 1) the high density of hens in the nestbox tier between 02:00h and 08:00h (figure 2) likely prevents them from roosting on the balcony and 2) the distribution of the interval time between consecutive nestbox tier timing has a narrow spread around 24h that seems specific to the morning (95% of the values falling within 23.3-24.9h, compared to 20.1-27.8h for the behaviour computed over the following 6h hours (8h - 14h); illustrated in figure S1). We hypothesis that more dominant, aggressive, proactive individuals, would lay earlier in the morning than subordinate, docile, reactive individuals would.

Sleeping tier

Previous literature suggested that hens are highly motivated [198] to roost at night on the highest area [199,200], but in commercial aviaries, not all hens can be on the highest tier simultaneously. On average across all available days, 69.4 ± 5.1% (mean ± SD) of the hens spent the night on the highest tier, 14.1 ± 4.7% on the nestbox tier, 16.2 ± 3.4% on the lower tier, and 0.3 ± 0.7% on the littered floor. Therefore, we used a binary variable whether the hen spent most of the night-time on the highest tier (yes/no) as a behaviour related to the night-time routine which is hereafter referred to as sleeping tier. The night phase is preceded by 15 minute dusk phase, where light is slowly reduced. We hypothesized that hens with a more proactive personality would have some traits (e.g. more bold, risk-taking, aggressive, or faster and more active behavioural response [57] such as to the reduced light) that would enhance opportunities to roost at night on the highest tier.

Feed delivery response

Feed is delivered automatically six times throughout the day (02:30 h, 06:00 h, 09:00 h, 12:00 h, 14:15 h, 16:15 h ; figure 2) via an automatic chain feeding system that runs for three minutes at each delivery and produces an elevated noise level. Although hens had *ad libitum* access to feed, the feed delivery brings fresh particles of larger size that are generally preferred by chickens [201–203]. Hereafter, we will refer to the two tiers where feed is accessible (the lower and highest tiers) as the "feed-tiers". We defined a feed delivery response as a descriptor of the tendency of being in a feed-tier more frequently while the fresh feed is delivered than while the feed is not delivered:

$$\text{Feed delivery response} = \frac{1}{|P|}\sum_{p\in P}\left(\frac{1}{50}\sum_{r=0}^{50}\left(\frac{T_p - T_{pr}}{T_p + T_{pr}}\right)\right).$$

where $\boldsymbol{P}$ corresponds to the set of periods when the feed is delivered, $\boldsymbol{T_p}$ the time spent in the feed-tiers during the period $\boldsymbol{p}$, and $\boldsymbol{T_{pr}}$ the time spent in the feed-tiers during a random period ($\boldsymbol{pr}$) of same duration and around the same period of the day as $\boldsymbol{p}$, but where feed was not delivered. Intuitively, the behaviour is similar to:

$(T_{fd} - T_{fnd})/(T_{fd} + T_{fnd})$, where T_{fd} (and T_{fnd}) represent the daily time spent in the feed-tiers while fresh feed is delivered (and is not delivered), but where we controlled for the difference in duration between periods with feed delivery and without feed delivery. To do so, we compared each feed-delivery period (p) with 50 random periods without feed being delivered, but of the same duration and occurring within a 1 h window surrounding p, but disregarding the 15 minute period immediately before or after p. We believe 50 periods gave a reasonable representation of the overall behavioural patterns around the time of feed delivery (see electronic supplementary material, text S2 for further explanation). The feed delivery response varies from −1 to 1, where negative numbers indicate a tendency of being in the feed-tiers more frequently while the feed is not delivered, and positive numbers indicate a tendency of being in the feed-tiers more frequently while the feed is delivered. Because periods with reduced feeder space (as expected after arrival of fresh feed and suggested by figure 2) may be associated with increased aggression [89], we hypothesized that hens with higher values would generally be bolder as they risk greater aggression to get fresh feed compared to others and thus we would deem them to be more proactive. To verify that this behaviour relates to the feed becoming available in the feed-tiers, we compared the behavioural responses computed with true and false feed delivery timing (defined as 20 min forward and 20 min backward in time). When computed with false timing, we found that on average 54±5% or 53±6% (forward or backward push, respectively) of individuals per day had a strictly positive response. When computed with true timing, we found substantially higher values, with on average 77±4% of individuals per day had a strictly positive feed delivery response (illustrated in figure S2 b, d, and f), indicating that the defined feed delivery response reflects hens' reaction to the feed being delivered.

Contexts

To understand whether individuals varied their spatial behaviour across contexts, we selected two commonly occurring situations in commercial settings and two different production stages (for a total of four contexts). We chose an early and late production stage defined by the first and last three weeks of the tracking period to compare the extent of consistent individual differences from the onset of adulthood to the end of production, when animals are more likely to have poor welfare conditions. To include a common commercial perturbation, we used three vaccination events spread over 120 days, with each involving two hours of water deprivation (beginning at 08:00 h), followed by vaccination delivery through water, and a three-hour postponement of the WG opening. External environmental conditions, such as the temperature, may also influence behaviour, particularly when temperatures fall outside of the optimum thermoneutral zone, which in laying hens lies between 19-22°C [204]. Therefore, we defined a "cold external temperature" context, characterized by days with solely negative mean hourly external temperature during hours where access to the WG was provided (varying from -1 °C to -4 °C). Throughout the experiment, there were several periods characterized by consecutive days with negative temperatures. We chose the first day within each of the first three periods to limit habituation effects. Any day that would fit into more than one of the four contexts were excluded to avoid overlap between contexts. Days that fit into none of these contexts were considered to estimate repeatability over time but, due to possible autocorrelations between days close in time and uncontrolled human disturbances, we only used Saturdays (a day with limited human non-staff visits). Similarly, Saturdays were used for both the early and late production stage contexts. Thus, each context involved three distinct days, with at least 118 hens tracked during all three days.

To evaluate the repeatability of behavioural differences across contexts, we used the observation in each of the four contexts that we believed was the most representative of the context, while avoiding observations close in time (detailed in figure 3). Subsequent analysis involved 194 hens and a total of 5,047 observations, amongst which hens were tracked on average for 26 days (min = 1, max = 33, sd = 1, 25th percentile = 21, 50th percentile = 31, 75th percentile = 33 (unit: day)).

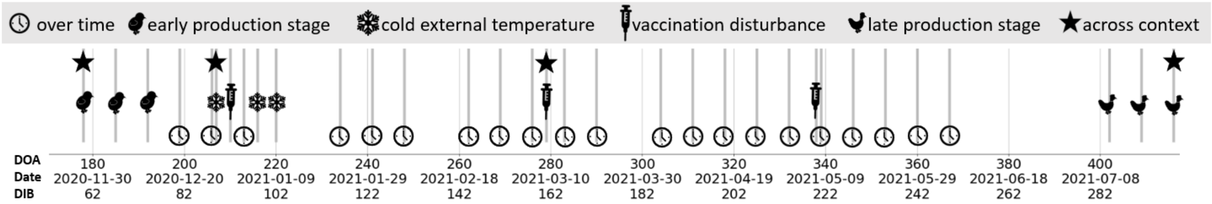


Figure 3. The timeline of the contexts in relation to the DOA, the date, and the DIB. Each vertical bold grey line represents a day used in subsequent analysis. All days selected within a context, over time or across contexts are represented here.

Statistical analysis

Analysis was conducted with Python for data processing and visualisations, and R for the statistics (code is given in the supplement).

Repeatability of behaviours

First, we evaluated individual consistency in each movement behaviours over time and across contexts. To evaluate individual consistency over time, we estimated the adjusted repeatability (defined as the proportion of the total variance that is accounted by differences among individuals, while accounting for known individual differences), hereafter referred as repeatability, of each behaviour separately. Before addressing the across contexts consistency, we investigated if individuals exhibited consistent behavioural differences within each context by estimating repeatability for each behaviour and context separately. Then, we computed the repeatability of behavioural differences across contexts. We used the rptR package [178] to calculate repeatability and fit a Gaussian distribution for both the vertical travelled distance behaviour and feed delivery response, and a binary distribution with logit-link function for both the WG presence and sleeping tier (i.e. went in the WG: yes/no; slept in one of the down tiers: yes/no), for which we reported the link-scale repeatabilities. We reported the repeatability estimate with its 95% confidence interval (CI) based on 1000 bootstraps. Because the nestbox tier timing is positively skewed, we used the glmmTMB package [205] with a gamma family and log-link function to extract the repeatability as explained in Stoffel et al. [206] (with the trigamma function to derive the observation-level variance) and used bootstrapping to get a mean estimate with confidence intervals, based on 1000 bootstrap replicates. We checked model assumptions (i.e., normality of error and homoscedasticity) by sight. Because consistent behavioural differences among individuals could arise from consistent external (environment) and internal (health) differences, also known as pseudo-repeatability [207], we accounted for KBF severity (which was already shown related to movement within the aviary [10] and is a prevalent health issues in commercial hens within aviaries [20,21]), body mass, class (MEXP/LEXP), and treatment (OFH/TRAN). Because there is a general upward trend over time for both the body mass and the KBF severity, we interpolated linearly (with monotonically increasing) both the KBF severity and body mass for each hen separately to better control for hens' health status between consecutive health assessments. To account for variation in the timing of feed delivery between pens due to sets of four pens being linked to a common feed delivery time, we controlled for a feed-chain identity in the feed delivery response. To facilitate model fit, we added a time effect (number of days since the transfer to the laying barn). All continuous fixed effects were scaled and centred to a mean of 0, so that intercepts reflect average values. The identity of the pen was previously tested for each behaviour separately and removed as the more complex models had greater Akaike information criterion ($\Delta AIC > 2$) and a same conditional explained variance ($\Delta R^2_{conditional} < 0.01$). For model convergence issues, the class was removed for all models used to compute the repeatability of the WG presence ($\Delta R^2_{conditional} = 0$). In addition to the repeatabilities, we report the trait means as well as the within- and the among-individual variance components (see supplemental material) [206].

Behavioural syndromes

We evaluated behavioural syndromes, that is, correlations among individual's average behavioural expressions hereafter referred as 'behavioural type'. We used the observations from all contexts in one multivariate model based on a Bayesian Markov Chain Monte Carlo (MCMC) approach using the brms package [180] in R. We ran the model with uninformative priors, four Markov chains with 300,000 MCMC iterations (including 200,000 for burnin), a thinning rate of 50, and with a similar distribution for the behaviour responses as in the models used to estimate the repeatabilities. To facilitate model fitting we scaled and centred (mean = 0, standard deviation = 1) the two gaussian responses (vertical travelled distance and feed delivery response) and scaled the gamma response (nestbox tier timing) (standard deviation = 1). All continuous fixed effects were scaled and centred to a mean of 0 and standard deviation of 1. Specification of the model was assessed with posterior predictive check for each response variable, trace plots, Gelman-Rubin's convergence diagnostic [182], Geweke's convergence diagnostic [208] and Leave One Out Cross-Validation (Loo-cv). We reported the mean and credible interval of each among-individual correlation between any two behaviours. Correlations were deemed significant if the credible interval did not include zero [177] after rounding to three decimal places. To illustrate the use of these multivariate model to extract a behavioural axis that accounts for most of the among-individual variation in correlated behaviours, we performed an eigendecomposition of their among-individual correlation (so that all variables are given the same weighting even though measured on a different scale) matrix. We used the eigenvalues to report what proportion of variation is explained by each eigenvector (principal component) and reported the behavioural trait loadings on the eigenvectors with an eigenvalue > 1 [175]. We used the posterior samples to compute the standard deviation of each trait loading as uncertainty measure.

Results

Repeatability of behaviours

All behaviours were repeatable over time, within, and across contexts as none of our confidence interval approached zero, indicating consistent inter-individual differences in our hen population. The most repeatable behaviour over time and across contexts was the vertical travelled distance (R = 0.66 [0.61, 0.70] and R = 0.48 [0.40, 0.56], respectively). The highest repeatability was for the sleeping tier within the late production stage context (R = 0.81 [0.65, 0.98]); the lowest repeatability was for the feed delivery response within the vaccination disturbances context (R = 0.23 [0.13, 0.35]). All estimates are provided in table 1 and a heatmap of their normalized (row wise and column wise) estimates are shown in figure 4a. Normalising each row, i.e. each behaviour separately, highlights for each behaviour which context is the most or least repeatable and therefore allows easy comparison of repeatability between context (e.g. if a context appeared as generally the least repeatable for all behaviours). Similarly, normalising each column, i.e. each context separately, allows visualizing for each context which behaviour is the most or least repeatable, and thus allows easy comparison of repeatability scores between behaviours (figure 4b). Finally, the general trend of an association between the between- and within- individual variances with the mean number of days between any two observations (as provided by the last row of table 1) is provided in figure 4c, where we generally observe a lower between-individual variance and a higher within-individual variance over longer time periods. The trait means, the between- and the within- individual variance components are reported in table S1 from the supplementary material.

	over time	early production stage	cold external temperature	vaccination disturbance	late production stage	across context
vertical travelled distance	0.66 [0.61, 0.70]	0.79 [0.74, 0.83]	0.73 [0.67, 0.79]	0.52 [0.43, 0.61]	0.78 [0.73, 0.83]	0.48 [0.40, 0.56]
nestbox tier timing	0.52 [0.48, 0.56]	0.69 [0.60, 0.78]	0.71 [0.63, 0.79]	0.57 [0.44, 0.70]	0.58 [0.41, 0.75]	0.47 [0.36, 0.59]
sleeping tier	0.48 [0.38, 0.54]	0.56 [0.34, 0.70]	0.54 [0.31, 0.68]	0.50 [0.27, 0.72]	0.81 [0.65, 0.98]	0.41 [0.22, 0.51]
WG presence	0.50 [0.41, 0.56]	0.60 [0.39, 0.69]	0.55 [0.33, 0.64]	0.30 [0.10, 0.38]	0.59 [0.35, 0.77]	0.25 [0.10, 0.31]
Feed delivery response	0.33 [0.28, 0.39]	0.36 [0.27, 0.46]	0.42 [0.33, 0.52]	0.23 [0.13, 0.35]	0.37 [0.27, 0.48]	0.24 [0.16, 0.33]
#observations	3,196	468	471	468	444	613
#individuals; #days	194; 21	157; 3	159; 3	193; 3	153; 3	190; 4
#hens with > 95% of obs.	103	155	155	118	142	109

mean and max #days between obs.	62; 168	9; 14	9; 13	85; 128	9; 14	131; 238

Table 1. Adjusted repeatability of the five behaviours over time, within each context and across contexts (in the first five rows). In the following rows, we added information on the data used to compute the repeatabilities: the number of observations, individuals, days and hens with more than 95% of observations included (in the following three rows); the mean and maximum number of days between any two observations (in the last row).

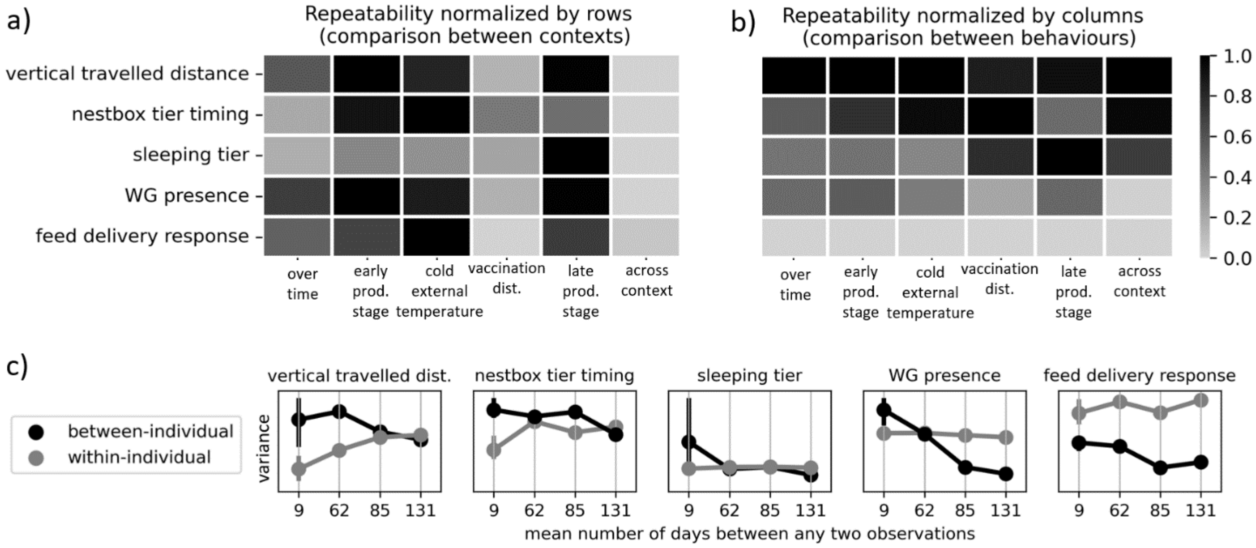


Figure 4. **(a)** Heatmap of the adjusted repeatability estimates from table 1, where each row was normalized (between 0-1) to highlight the highest (black) and lowest (light grey) estimates for each behaviour and allow easy comparison between contexts. **(b)** Heatmap of the adjusted repeatability estimates normalized by columns to highlight the highest (black) and lowest (light grey) estimates in each of the categories (over time, early and late life stage, cold external temperature, vaccination disturbance, and across context) and allow easy comparison between behaviours. **(c)** Illustration of the among- and within- individual variances underlying the repeatabilities, sorted by the "mean number of days between any two observations" (from last row of table 1). This visual is intended to highlight the general trend of the individual variances when estimated on short or long interval time, the variances are further detailed in Supplementary table S1.

Behavioural syndromes

Results revealed existence of behavioural syndromes, involving four of the five behaviours. The vertical travelled distance was positively correlated with the WG presence (r = 0.50 [0.38, 0.61]) and feed delivery response (r = 0.47 [0.35, 0.59]), as well as negatively correlated with sleeping tier behaviour (r = -0.23 [-0.39, -0.08]). The feed delivery response and the WG presence were positively correlated (r = 0.27 [0.11, 0.41]). The feed delivery response and the sleeping tier were negatively correlated (r = -0.21 [-0.36, -0.05]). Correlations between each pair of behaviours and statistically significant correlations are illustrated in figure 5.

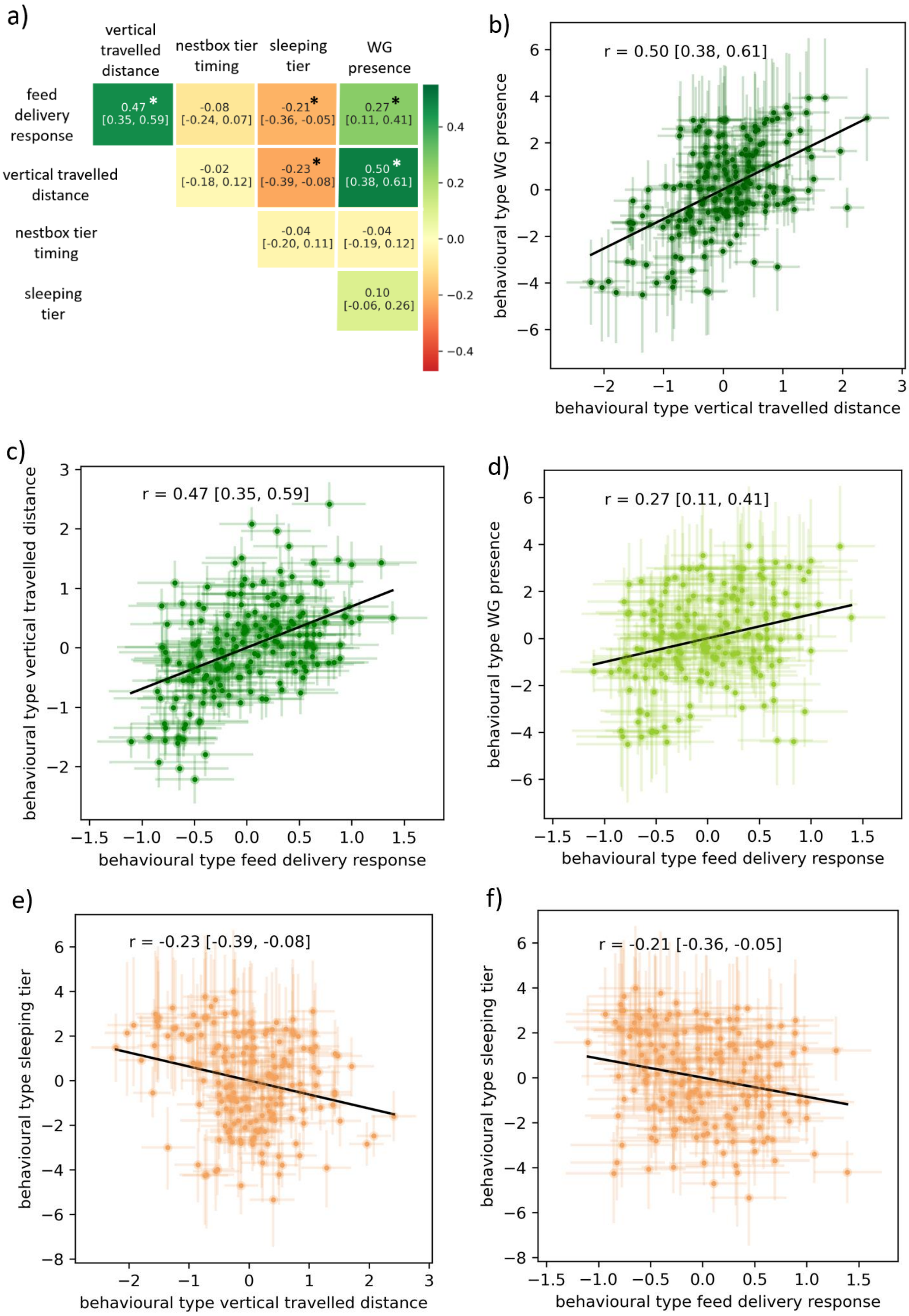


Figure 5. Correlations (±95% credible interval) of each movement behaviours from the multivariate mixed model where stars (*) highlights significant correlations (i.e. when 0 is not included in the confidence interval) in **a)**, illustrated by individual intercept estimates (±95% credible interval) of each correlated pairs of behaviours: in **b)** the WG presence and the vertical travelled distance, in **c)** vertical travelled distance and the feed delivery response, in **d)** the WG presence and the feed delivery response, in **e)** sleeping tier and the vertical travelled distance, and in **f)** sleeping tier and the feed delivery response. The slope from each regression line between these pair of behaviours is represented by a black line and calculated by dividing the covariance between both behaviours with the variance of the behaviour displayed on the x-axis. Negative correlations are represented by red colours and positive correlations by green colours.

At the population-level, the number of days spent in the barn had the largest effect on all five behaviours. With increasing days in the barn, hens on average visited the WG on more days, spent fewer nights on the highest tier, reduced their vertical travelled distance per hour spent indoor, used the nestbox tier later (as measured by a greater response of the nestbox tier timing behaviour), and increased their tendency to be in the feed-tiers upon feed delivery feed (the observed behaviours over time are illustrated in figure S3). Also, heavier birds used the nestbox tier earlier and hens that travelled greater vertical distances had more severe keel bone fracture scores and greater mass. Treatment (OFH/TRAN) had no effect. The more/less explorer class effect was statistically significant for only one behaviour where more explorative hens were found to spend fewer nights on the highest tier compared to hens randomly selected among the entire population. Estimates of population-level effects are further detailed in supplementary figure S4.

The eigendecomposition of the among-individual correlation matrix based on the correlated behaviours (all behaviours except nestbox tier timing) resulted in two principal components with an eigenvalue > 1, and explained 47% and 28% of the total variation, respectively. On the first principal component, the sleeping tier weakly loaded in the opposite direction as the feed delivery response, the WG presence, and the vertical travelled distance. Behavioural traits loadings (mean ± SD) on the first and second eigenvectors of the among-individual correlation matrix were: feed delivery response: 0.54 ± 0.03 and -0.16 ± 0.09, vertical travelled distance: 0.63 ± 0.02 and 0.03 ± 0.05, sleeping tier: -0.24 ± 0.09 and 0.83 ± 0.07, WG presence: 0.50 ± 0.04 and 0.54 ± 0.11, respectively. The first principal component suggests a behavioural axis that accounts for most of the among-individual variation in correlated behaviours. Although individuals varied along a continuum on the axis, the extremes can illustrate different "behavioural profiles" within the flock. To illustrate these extremes, we projected observed behaviours of each selected day onto the subspace spanned by the first component, and, randomly selected three hens among the 15% highest and lowest mean score to visualise their mean observed behaviours (figure 6).

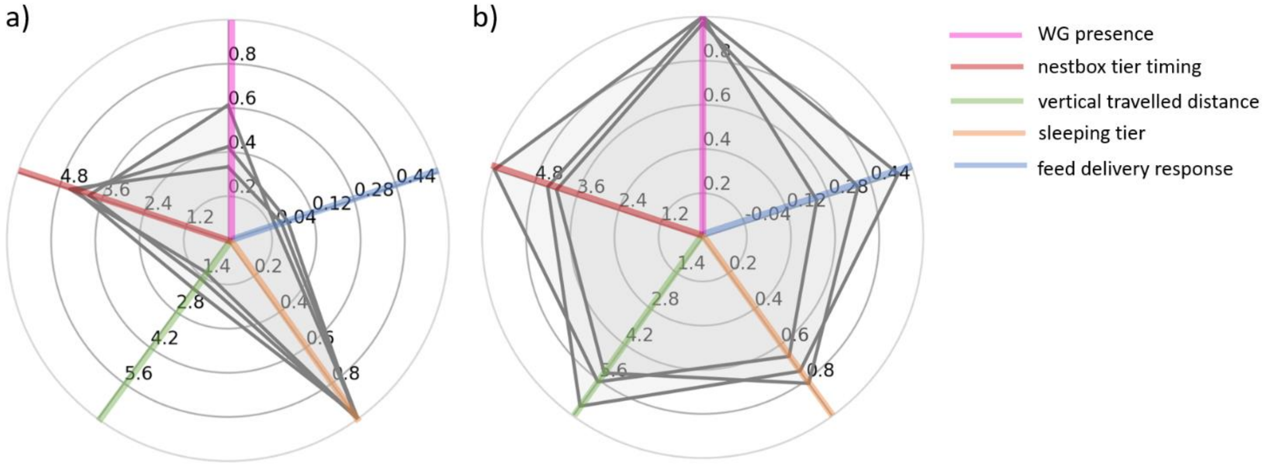


Figure 6. Illustration of two "behavioural profile" as suggested by the first eigenvector of the among-individual correlation matrix. We randomly selected three hens among the 15% lowest (a) and highest (b) mean score to visualise their mean observed behaviours, including the nestbox tier timing behaviour, showed to be independent of the other behaviours.

Discussion

This is the first study to quantify long-term individual consistency in various spatial behaviours of commercial hens within an aviary system. The behaviours were selected to represent five aspects of commercial hens' daily routine, and therefore may be functionally different. We found consistent individual differences over time and across contexts in the five behaviours. The daily vertical travelled distance and the nestbox tier timing behaviours were most repeatable. We also found behavioural syndromes involving the sleeping, feeding, indoor movements, and outdoor usage behaviours, but interestingly not the nesting related behaviour, suggesting two behavioural axes driven by different mechanisms. Altogether, these long-term individual consistencies and behavioural correlations revealed the potential applicability of such behaviours as personality traits in commercial hens and suggest two main axes of spatial personalities.

Repeatability of behaviours

Consistent inter-individual differences in behaviour may arise from intrinsic individual (genetic or epigenetic) differences, but also from consistent external and internal differences (e.g. bone fractures) between individuals. In commercial settings, numerous sources of variation are under strict human control (e.g. rearing process, light, and indoor temperature) which may reduce consistent external differences between individuals and in turn avoid spuriously high repeatabilities. In addition to offering unique conditions for controlling and standardizing environmental conditions, commercial settings often have automated processes (e.g., to provide feed or access to areas) that could be manipulated and incorporated into the design of experiments. We highlight that animals with known genetics, standardized routines, and housed in commercial settings offer an excellent system to study behavioural variation at different hierarchical levels.

For the design of future experiments, it is important to understand the effect of interval time between observations on the repeatability of behaviour. Our repeatability estimates decreased with increasing mean interval time between observations, as expected [35]. The underlying individual variances suggest that hens have more variable behaviours (i.e., higher within-individual variance) and are behaviourally more alike (i.e., lower between-individual variance) when studied over longer periods. The observed trends from figure 4c highlight the two underlying mechanisms by which repeatability of behaviours generally reduced over longer observational period, though the factors driving the mechanisms remain unknown and likely include both internal and external changes.

Repeatability may also vary with respect to an individual's life stage (e.g. due to ontogeny [69]), yet we found similar repeatabilities during our early and late production stages. However, we previously found that similar behaviours expressed previous to our early production stage on the same hens (during the first two months in the laying barn) were considerably less repeatable (R varying from 0.38 for the vertical travelled distance to 0.17 for the sleeping tier) [209]. These results show that hens were already behaviourally consistent prior to our early production stage, but that repeatability increased and stabilized during the first two months in the laying barn and seemed to be maintained until the end of production.

We found consistent individual differences in all behaviours over time and across contexts, revealing their potential applicability as personality traits of commercial laying hens. Especially the daily vertical travelled distance and the nestbox tier timing, for which consistent differences between individuals explained more than half of their variation (66% and 52%, respectively) and were maintained across contexts (48% and 47%, respectively). In contrast to these moderately high long-term repeatabilities, consistent differences in both the WG presence and the feed delivery response were only weakly maintained across contexts (R = 0.25 and R = 0.24, respectively), suggesting these behaviours may be of lesser importance to the hens or harder to maintain (e.g. due to behavioural plasticity in response to the outside temperature; or during the vaccination disturbances context). Also, because not all hens can be on the feed-tiers simultaneously due to the limited space, it may limit hens to express consistent feed delivery responses, which could explain the general low repeatabilities of the behaviour.

Although the feed delivery response did not show drastic behavioural differences within our population, it may be of particular relevance in commercial settings where animals have to regularly respond to external stimuli. It is important to note that as we are unable to control for all internal or external drivers, this behaviour may not directly reflect animals' motivation to feed. However, it is a first step towards assessing a potential proxy of the animal's affective state in response to a recurring external stimulus based on movement data. Further research should assess how such behaviours are indicative of animal wellbeing, for example by testing if hens with higher feed delivery responses also have a more optimistic attitude in a judgment bias test [210]. If supported, the behaviour could then be used as a daily proxy of the hens' affective state.

Altogether, these long-term repeatabilities demonstrated that tracking technology can be used to quantify individual differences in behaviours related to different aspects of commercial hens' routine and for long periods that are rarely accounted for in personality studies. These results demonstrated that hens from a single parent flock can differ consistently in spatial behaviours over most of the production period, suggesting that individuals have different preferences, or needs. Such individual preferences may lead to overcrowding or under-utilized areas, which could affect animal welfare (e.g., smothered hens in overcrowded areas [211]) and production (e.g., eggs on the floor due to preferred nestbox being occupied [212]). Therefore, by understanding behavioural differences in commercial settings we can design appropriate management tools and breeding practices to improve animal welfare, such as sensors to detect pilling behaviours [211], practices to encourage earlier transitions between tiers after transfer to an aviary [213], and breeding hens for specific behaviour [83,212].

Behavioural syndromes

In accordance with our predictions, we found that hens that travelled greater vertical distances, on average also went in the WG on more days and tended to use the feed-tiers more upon delivery of fresh feed. These results support our hypothesis that these behaviours are associated with a pro-reactive personality axis. We proposed that these behavioural expressions are indicative of a proactive personality. Contrary to our predictions, these hens on average also use the highest tier slightly less at night (correlation with vertical travelled distance: r = -0.23). Because hens are highly motivated [198] to roost on the highest tier [199,200], this result could suggest that more proactive hens are less successful in accessing the highest tier at night. For instance, hens that travel greater vertical distances may also stay active until later and, therefore, be less able to access the highest tier due to higher animal densities. Further research is required to evaluate associations between these behaviours and common personality traits to understand the role of pro-reactive personality axis in explaining these syndromes.

An alternative, non-mutually exclusive, mechanism to explain these behavioural differences and syndromes could be the existence of subgroups, where hens' location would reflect those of their group. In large groups such as ours, hens can't recognize all conspecifics which likely limits their ability to form social groups based on individual recognition [111]. Therefore, we believe it is unlikely that hens repeatedly associate their locations with the same individuals due to such social groups. However, hens may recognize conspecific or relevant traits (e.g. comb size) based on their status [214]. Therefore, subgroups could reflect dominance rank and hens' location would reflect their rank. Future research should investigate the social dynamics of these large groups and whether this could explain spatial behavioural differences.

Interestingly, we found no syndrome involving the nesting behaviour, which suggests that its main mechanism is independent of those involved in the other behaviours and could be, for instance, the physiological rhythm. Because commercial laying hens are under strong human selection for high egg production, we may speculate that the behaviour is more resilient to environmental change and could be expressed independently to other needs and their associated behaviours. Altogether, these correlations indicate two main axes of spatial personalities: the nesting related behaviour and the behavioural axis based on the other behaviours. We used the extreme values from the latter axis to illustrate two, probably most dissimilar, behavioural profiles in our flock (figure 6). This figure illustrates how this one axis is accounting for most of the among-individual variation and therefore may be used to extract relevant phenotypes.

To assess relative benefits of these phenotypes, future research should evaluate not only their associations with common personality traits but also with animal welfare and productivity. According to Koolhaas [215], proactive animals perform better under highly predictable conditions, or when feed is abundant [216], compared to reactive animals. However, in commercial settings, animals often have a predictable management daily routine, but unpredictable events such as diseases and vaccinations frequently occur. Therefore, it is unclear which of the proactive or reactive personalities would perform better in such settings. A greater understanding of how personality traits relate to welfare and production could inform on potential phenotypes to integrate into the breeding process for more robust farm animals [55,79,80,83]. Our repeatability estimates set a promising upper bound on the heritability of these behaviours and the behavioural syndromes highlight some potential constraints for selection. In conclusion, this study supports the use of tracking technology to assess behavioural traits and potentially breed more robust farm animals.

CHAPTER IV

Based on:

Keel bone fractures affect laying hens' mobility, but no evidence for reciprocal effects

Camille M. Montalcini[1,2], Michael J. Toscano[1], Lucy Asher[3], and Matthew Petelle[1]

1. Center for Proper Housing: Poultry and Rabbits (ZTHZ), Division of Animal Welfare, VPH Institute, University of Bern, Switzerland
2. Graduate School of Cellular and Biomedical Sciences, University of Bern, Switzerland
3. School of Natural and Environmental Sciences, Newcastle University, Agriculture Building, NE1 7RU, UK

In preparation for submission

Camille Montalcini participated to the experimental work needed to collect the movement and health data from two of the three datasets used in this study, conducted the analysis, and wrote the manuscript in agreement with her co-supervisor Dr. Matthew B. Petelle and co-author Dr. Michael J. Toscano and Lucy Asher.

Abstract

Keel bone fractures (KBF) are prevalent in commercial laying hens and are considered one of the greatest welfare concerns in the egg-production industry. While clear associations exist between KBF and animal mobility, suggesting that KBF impair mobility, the effect of mobility on KBF remains poorly understood. We combined data from three studies that scored the amount of the keel bone affected by fractures (KBF severity) and tracked hens' transitions between different zones (tiers of the aviary and the wintergarden) of a multi-tier aviary the week prior to radiograph. Using data from the tracking systems, two movement and two space-use behaviours were extracted for each hen. These spatial behaviours were the vertical distance travelled, the mean number of zones crossed within one transition, the time spent in the top tier, and the unevenness of time spent across zones. We used hierarchical Bayesian continuous time dynamic models to estimate how a change in a behaviour predicted a later change in KBF severity, and vice versa. The severity of KBF did not affect space-use behaviours, but it did alter movement behaviours. Specifically, increased KBF severity led to decreased vertical travelled distance and a tendency to cross more zones within one transition. In contrast, we found no evidence that movement or space-use behaviours may predict later change in the severity of KBF; however, similar efforts accounting for the location of fractures could unveil the potential influence of spatial behaviours in the formation and maintenance of KBF and increase our ability to mitigate their effects.

Introduction

Keel bone fractures (KBF) are recognized as one of the greatest welfare concerns in the egg production industry [20–22,24,25]. The concern for the welfare of animals with KBF has global implications, given their high prevalence across countries and commercial strains, averaging between 24% and 63% depending on the housing system [23]. Keel bone fractures may have a detrimental impact on both egg-production [28,120,121] and animal welfare, with strong evidence indicating that hens with KBF feel pain for at least several weeks [22,122] and show behavioural differences in highly motivated behaviours, including perching and nestbox use, which could indicate negative affective states [122,123]. Furthermore, increased KBF associated with impaired mobility, including reduced vertical locomotion [29], longer latency to fly from perches [27,28], and increased time spent on the aviary's top tier [10], an area with vital resource and where hens may receive less agonistic behaviours than in lower tiers [200,217]. Impaired mobility in hens with KBF is possibly due to pain and physical impairment, as the keel is the site of muscle attachment [218] and involved in breathing [219,220].

Although there are clear associations between KBF and spatial behaviours, with evidence suggesting that KBF impair mobility [27,29], the effect of spatial behaviours on KBF remains poorly understood [126]. Previous literature generally suggested movement throughout cage-free housing as a potential causal factor for KBF, suggesting that collisions with the housing structures [3,30,31] and the complex design of the system [20,33] (including the height and presence of perches [3,32]) are contributory factors to KBF. However, recent pathological evidence suggested that collisions may not be responsible for the fractures located at the caudal tip of the bone, which account for the majority of KBF [126]. Instead, these fractures may be attributed to the internal pressure exerted during egg-laying [126,221]. Also, in a recent review the highest average prevalence of KBF was reported in single-tier systems (63%), not in the more complex aviary systems (38.3%), though the latter relied on palpation which probably underestimate the prevalence of KBF [222] (out of the 27 observations only two did not rely on palpation and both observations were above 90%) [23]. Even if spatial behaviours were not causing new fractures, some behaviours may promote healing while others may exacerbate existing fractures (such as walking vs flying). Therefore, while the severity of KBF likely affect spatial behaviours, it is also possible that a change in spatial behaviours can predict a change in the severity of KBF. Yet, the two halves of this KBF-mobility dynamic are rarely studied in conjunction.

Despite the limited understanding of the influence of spatial behaviours on KBF, modifications of housing to provide safer mode of locomotion have already been shown to reduce prevalence of KBF in cage-free systems. For instance, adding ramps in multi-tier aviaries to enable hens to move between the stacked areas by walking instead of jumping or flying, was shown to decrease incidence of falls, collisions, and KBF in laying hens [30,104]. Multi-tier aviaries are especially relevant when studying KBF, as they are increasingly prevalent in commercial production, and the complex design of the system could exacerbate both KBF prevalence [20,33] and their consequences. Indeed, impaired mobility in this type of housing could increase the risk of dehydration, emaciation, and floor eggs, as individuals may be unable to access all resources across the aviary [128], which may also explain why hens with fractures spent more time on the top tier [10], a tier with feed and water (in Switzerland). In 2019, Rufener et al. [10] provided the first evidence of an association between KBF and mobility in hens housed within multi-tier aviaries, however, whether differences in mobility precede KBF or if KBF altered the behaviour, remains to be determined. Thus, in order to address KBF more comprehensively, it is essential to better understand the KBF-mobility dynamics in these complex cage-free systems.

In this study, we evaluated the potential bidirectional relationships between KBF and spatial behaviours of 376 commercial laying hens housed in a multi-tier aviary. To increase robustness and sample size, we used two published datasets in addition to a new dataset that also scored the amount of the keel bone affected by fractures (KBF severity) at 3-11 time points per hen and monitored hens' transitions between the aviary tiers during each week prior to radiograph. As spatial behaviours we extracted two movement and two space-use behaviours and estimated how a change in behaviour predicted a later change in KBF severity, and vice versa. Similar to Rufener et al. (2019) [10], we used the vertical travelled distance and the mean number of zone crossed within one transition as two movement behaviours, and the proportion of time spent on the top tier as a first space-use behaviour. Additionally, to account for the possibility that individuals may select locations other than the top tier to spend time over the day in response to KBF, and that the location may also vary across days, we also used a measure representing the evenness with which hens utilized the five zones over the day.

We hypothesized that hens with increased KBF severity would reduce their activity and spend more time on higher tiers with vital resources. Therefore, we predicted that an increase in the severity of KBF would lead to a decrease in vertical travelled distance, greater time spent in the top tier, and a more uneven usage of the zones. Also, we hypothesized that more transitions between the aviary stacked tiers would lead to a higher number of landings and, consequently, increased occurrence of falls and collisions. Therefore, we predicted that increased vertical travelled distance and increased number of tiers crossed within one transition (indicating longer and potentially more hazardous landings) would lead to increased KBF severity.

Materials and methods

Study design

We combined datasets from three experiments with similar spatial behaviours and score of KBF severity assessed repeatedly on individual hens (*Gallus gallus domesticus*) throughout the laying period. Dataset1 was published by Rufener et al. in 2019 [10] where they used an infrared tracking system. Dataset2 and Dataset3 collected location data using the same low frequency tracking system, which was validated and described in [173]). Dataset2 was collected for a previous study [223], while Dataset3 was collected for the current study. All hens were housed over different years but in the same laying barn containing an aviary system (Bolegg Terrace from Vencomatic, described in [30],separated into 20 pens by grids, pen indoor area: 7 m length, 2.3 m width, 2.69 m height until the top tier grid floor) and an outside covered winter garden (WG; $9.32\ m^2$, accessible by pop holes for 6h on most days). The barn is located at Aviforum facilities in Switzerland where standard animal husbandry practices are used. Hens were distributed among three pens for Dataset1 (pens 4-6) and eight pens for both Datasets2 and 3 (pens 3-5, pens 8-12). Stocking density was 8.1 hens per square-meter of permanently accessible area (in each pen: 225 hens / $27.92\ m^2$). In each experiment, hens' movements were tracked continuously for approximately a week before being radiographed for detection of KBF.

Movement data represented individuals' transitions between any two of the five zones in the pen: the top and lower tiers (both containing feed, water, and perches), the nestbox tier (with nestboxes and a balcony), the littered floor, and the winter garden (containing litter and water). We extracted two daily movement and two daily space-use behaviours described below. We averaged each daily spatial behaviours over the week previous to radiograph collection to have one value for each hen per radiograph (number of days tracked before being radiographed in Dataset1: 5.80 ± 0.56, Dataset2: 5.5 ± 1.77, Dataset3: 6.55 ± 1.02). We excluded hens that only had fewer than three observations (74 hens from Dataset1, 6 hens from Dataset2, more details in supplementary materials Text S1) as the analysis focus on studying variation over time, including individual variation in trends over time (for each behaviour and the KBF severity, with random slope and intercept). In subsequent analysis, we had a total of 376 hens (Dataset1: 60 Lohmann Brown hens, Dataset2 and Dataset3: 153 and 163 Dekalb White hens, respectively) with a total of 1,889 observations of the behaviour and corresponding KBF severity score (Dataset1: 593 with 10 ± 1.73 observations/hen; Dataset2: 658 with 4 ± 0.76 obs./hen; Dataset3: 638 with 4 ± 0.28 obs./hen). The youngest and oldest hens were 148 and 437 DOA respectively, at the time of being radiographed. The mean interval time between two consecutives observations on the same hen is 29.04 ±9.57 for Dataset1, 63.10 ±31.09 for Dataset2, and 72.81 ±10.33 for Dataset3. Datasets are further detailed in the supplementary Text S1.

Ethical note

The research was conducted in accordance with the cantonal and federal regulations for the ethical treatment of experimentally used animals. All procedures from the newly collected dataset were conducted in accordance with the cantonal and federal regulations for the ethical treatment of experimentally used animals, and all procedures were approved by the Bern Cantonal Veterinary Office (BE-57/21).

Keel bone fractures

The same radiograph procedure was performed in each experiment. On the first day after each tracking period, the hens were radiographed to detect fractures on the keel bone, using a mobile X-ray unit previously described by Rufener et al [176]. The hens were hung upside down from a custom built shackle for approximately 15-30 seconds to induce immobility during the radiograph procedure. Based on the latero-lateral radiographs (examples in supplementary S1 Fig), a KBF score, hereafter referred to as KBF severity (continuous, 0-100) was assessed using a scoring methodology described by Rufener et al [176]. The score is described as an indicator of the total amount of keel bone affected by any fracture. During assessment, the observers were blind to the hens' age. The supporting information (S2 Table and S2 Fig) provide descriptive statistics of the KBF severity scores for each dataset.

Movement behaviours

Because the four indoor zones are stacked on top of each other, we used the total number of indoor zones crossed (per hour to account for different day length over experiments: 14.91 ± 0.85 h) as a measure of vertical movements and hereafter referred as "vertical travelled distance". We used the mean number of zones (tiers of the aviary and the wintergarden) crossed per transition as defined by [10], hereafter referred as "mean-zone-crossed". A mean-zone-crossed value equal to one means that the hen did not skip zones over the day, while values higher than one indicate the hen skipped zones while transitioning from one zone to another. Supplementary table S2 provide the mean and standard deviation of the movement behaviours for each dataset. We illustrated the two movement behaviours using four hen-days of the raw tracking data (i.e., transitions from one zone to another zone) chosen specifically to exemplify low/high vertical travelled distance and low/high mean-zone-crossed (Fig 1a). We illustrated the relationship between the two behaviours (Fig 1b) and showed descriptive statistics of the behaviours for each timestamp and dataset in chronological order (Fig 1c), where Dataset1 is represented in the lightest grey, Dataset2 in darker grey, and Dataset3 in middle grey.

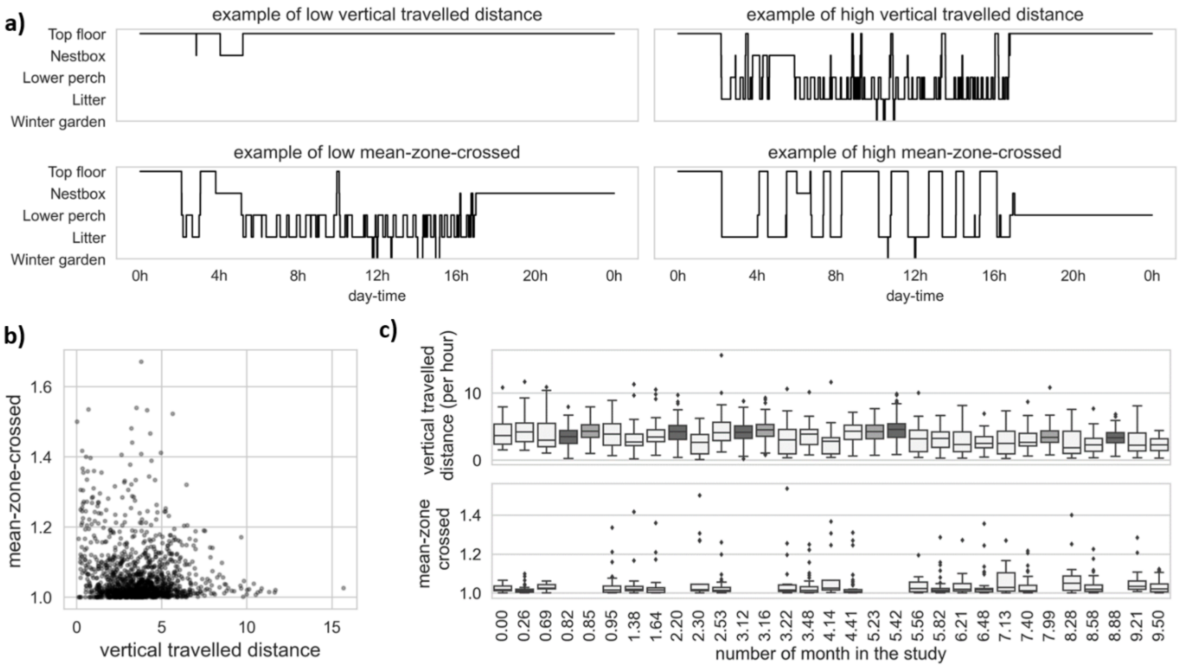


Fig 1. Illustration of the two movement behaviours. (a) Four days of the raw tracking data (transitions from one zone to another zone) chosen specifically to illustrate a low/high vertical travelled distance and low/high mean-zone-crossed (described by the title). **(b)** Scatter plot of the two behaviours with respect to each other, where overlapping data points are represented by darker shading. **(c)** Boxplots of the two behaviours for each timestamp and available dataset (Dataset1: lightest grey, Dataset2: darker grey, Dataset3: middle grey).

Space-use behaviours

As space-use behaviours we used: 1) the proportion of the daily time spent on the top tier and 2) a score reflecting the evenness with which hens spent their daily time across zones to account for the possibility that individuals may select alternative locations to spend time in response to KBF and that the location may also vary across days. For unevenness, we contrasted the hens' proportion of time spent in each zone (represented by a vector x where O_i is the proportion of time spent in zone *i*) to what would be expected from an equal usage of all surface area (E_i). Thus, we accounted for uneven daily usage of the five zones, defined by:

$$unevenness(x) = \frac{1}{5}\sum_{i=1}^{5}\frac{|E_i - O_i|}{E_i},$$

where E_i is equal to the surface area (m^2) of zone *i* normalized by the proportion of hours available to access that zone during the day (i.e., WG: 0.11, littered floor: 0.48, lower tier: 0.17, nestbox tier: 0.11, and top tier 0.13; detailed in the supplementary table S1). Supplementary table S2 provide the mean and standard deviation of the space-use behaviours for each dataset. We illustrated the two space-use behaviours providing 100 random examples of the proportion of time spent in each zone that corresponds to a low and high unevenness (darker colours represent greater proportion of time) in Fig 2a. The relationship between the two behaviours shows that hens with a high daily unevenness score, typically spent most of the day on the top tier (Fig 2b). However, we can also observe that some hens with relatively high unevenness scores have spent the majority of the day on the lower tier without visiting the top tier (Fig 1a). Given that the top and lower tiers are the only two tiers equipped with feed and water, these observations are not surprising but underscore the distinctiveness of the two space-use behaviours. We also present the descriptive statistics of the behaviours for each timestamp and dataset in chronological order, with Dataset1 represented in the lightest grey, Dataset2 in darker grey, and Dataset3 in middle grey (Fig 2c).

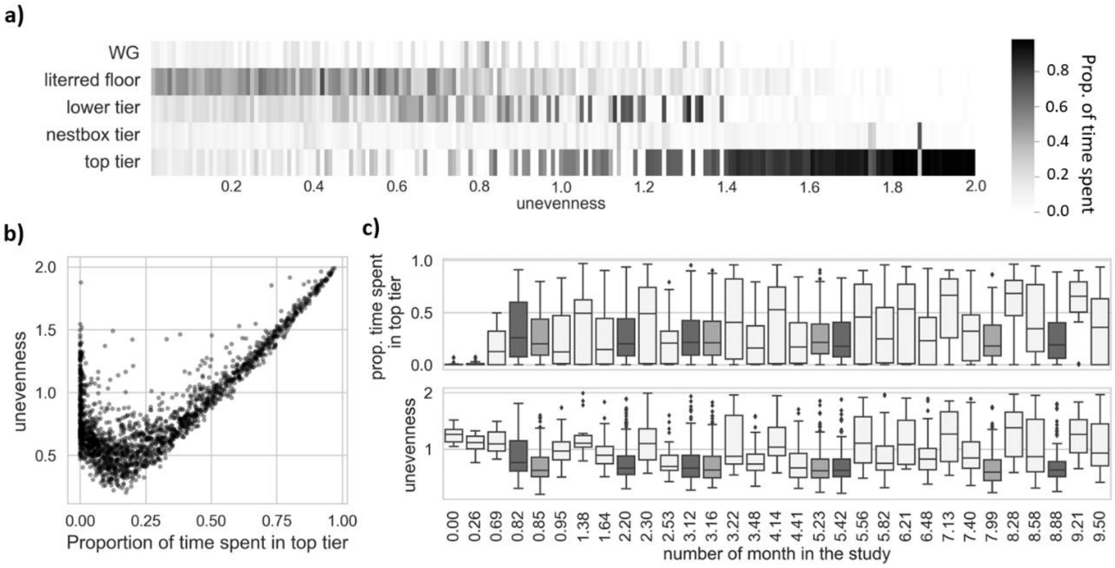


Fig 2. Illustration of the two space-use behaviours. (a) Examples of the daily proportion of time spent in each zone (darker colours for higher values) sorted by the associated daily unevenness **(b)** Scatter plot of the two behaviours with respect to

each other, where overlapping data points are represented by darker shading. **(c)** Boxplots of the two behaviours for each timestamp and dataset (Dataset1: lightest grey, Dataset2: darker grey, Dataset3: middle grey).

Statistics

We used hierarchical Bayesian continuous time dynamic models to estimate how a change in the two movement and the two space-use behaviours predicted a later change in KBF severity, and vice versa. More specifically, we fitted four latent dynamic models using one hierarchical Bayesian continuous time dynamic model [224] per spatial behaviour with the “ctsem” R package [225]. Because it is expected that KBF increases with age [21,23,226] and it is possible that spatial behaviour also changes with age, we included smooth trends in the processes (varying by hen) to limit confounding effect of age on our estimated dynamics. Therefore, for each model we estimated a trend and a dynamic fluctuation for each of the two processes, the KBF and the spatial behaviour. The trends included random initial intercepts and random slopes (varying by hen), as well as an estimated auto-effect term that varied by process. Specification of such multivariate latent process models (estimating the dynamics between two processes around their respective trends, referred in the result section as "dynKBF" for the KBF and "dyn*" for spatial behaviors, with "*" representing a spatial behavior acronym given in the results section) are described by Driver and Tomasik (2022) [227] and we give a more detailed explanation in the supplementary Text S2. Temporal effects are contained in a **drift matrix**, where the cross- (off-diagonal) and auto- (diagonal) effects can be interpreted similarly: a positive or negative effect indicates that higher values of the causal process led to rises or drops in the caused process. Therefore, the continuous-time auto-effect are expected to be negative, and more negative effects indicate a stronger force pushing the deviations back to baseline, implying that changes away from the baseline dissipate faster and the process more frequently crosses its baseline. To understand these results more intuitively, we extracted the expected auto- and cross-regression effects over time (discrete time parameters) to have an estimate of the temporal effect of the KBF severity on each spatial behaviour (and vice-versa) for specific time intervals. More specifically, while continuous-time temporal effects describe how the process is changing *at the moment*, discrete-time cross-regressions for a time interval of k represent the effect of one variable at *earlier* time $t-k$ on another variable at *later* time t. We created an empty observation row for every hen at the age of our earliest observation and set observed variables to NA when no observation was available, so that trends over time reflect trends over hens’ age. We scaled and centred the outcome variables and kept the default priors. For inference, we used a form of penalized likelihood, that is the maximum *a posteriori* estimation approach of ctsem [228]. We controlled for the two treatments and the data source identity by allowing the trend to vary based on these predictors. We used as treatment reference group hens that hatched with the standard hatchery practices (TRAN) and were not relocated in new housings during production. We used as data source reference group the Dataset2 as these hens came from a single-strain commercial flock (contrary to Dataset1) and movements were tracked with higher precision than Dataset3. To measure the MZC behaviour accurately, the tracking system should be accurately registering all transitions, even when the duration of stay in a zone is of few seconds only. For instance, the tracking system should be capable of differentiating between a hen moving directly from the top tier to the litter floor and a hen that uses the intermediate zones for only a few seconds. Therefore, we could fit the model with the MZC behaviour only using the already published dataset by Rufener et al. (2019) [10], that is Dataset1. Dataset2 and Dataset3 relied on another tracking system of lesser precision, and would require further validation to use behaviour with this level of precision. We verified normality of the observation residuals and plotted them against the time, the two observed variables, the predictors, and the predicted values of the six latent processes (called “etaprior” in ctsem). Unless otherwise specified, we reported estimates with 95% credible intervals, and deem estimates significant when the credible interval does not include zero.

Results

Movement behaviours

To improve readability, in this section we will refer to the vertical travelled distance as “VTD” and the mean-zone-crossed as “MZC”. Model estimates are given in the supplementary documents (VTD- and MZC- ctSummary_tex_fit and ctSummary_summary). We found a negative **continuous-time drift cross-effect** estimate of the KBF on the VTD (β [95% CI] = -0.19 [-0.36, -0.03]), meaning that a change in KBF lead to a change in the opposite direction of the VTD. More precisely, we found that if the standardized KBF severity increased by one, this will cause a drop in the slope of the standardized VTD by 0.19. The continuous-time drift cross-effect estimate of the KBF dynamics on the MZC dynamics was not statistically significant but tended to be positive (β [95% CI] = 0.33 [-0.03, 0.70], [90% CI] = [0.05, 0.64]), so a change in KBF tended to lead to a change in the same direction of the MZC. That is, if the severity of the standardized KBF increased by one, this will tend to cause a rise in the slope of the standardized MZC by 0.33, corresponding to a rise in the slope of the original MZC by 0.02 (= 0.33 * 0.07, where 0.07 is the SD of the original MZC). In contrast, we found that changes in either movement behaviours had no effect on subsequent level of the KBF severity (estimates of the two movement behaviour dynamics on the KBF dynamics are relatively close to zero (VTD: -0.07 [-0.18, 0.03], MZC: 0.07 [-0.12, 0.25]). In summary, the cross-effect estimates indicated that an increase in KBF severity led to a decrease in vertical travelled distance and tended to be followed by more tiers crossed within a transition (i.e., MZC), but that a change in these movement behaviours did not predict subsequent level of KBF severity.

To understand these cross-effects more intuitively we looked at the **discrete-time drift cross-effects**, showing how a change in the KBF severity predicts changes in the spatial behaviours after specific time intervals (Fig 3a-b) and vice-versa (Fig 3c-d). For instance, if the standardized KBF severity increased by one, this predicts a drop in the standardized VTD by 0.17 two months later (Fig 3a; blue line value at x = 2). In other words, we show that if the KBF severity increased by 1 standard deviation, this predicts a decrease in the VTD by 0.3 (= 0.17 * 1.79, where 1.79 is the standard deviation of VTD from supplementary Table S2) two months later. Because VTD was defined as the vertical travelled distance per hour, this decrease in VTD is equivalent to a decrease of 4.5 transitions over the lighted period of a typical day (i.e., 15 hours with natural and/or artificial light). We found that the expected effect of KBF severity on the VTD is greatest after approximately 3 months (e.g., Fig 3a, where the blue line is at its trough). It is worth mentioning that this result does not imply that any relation between the two is changing over time, but instead describes patterns of change accumulation over that time interval (Fig 3a), which can be explained by the result showing that changes in KBF and VTD are persistent in time (red lines of the Fig 3a-b for the KBF, and blue line of Fig 3c for the VTD). The discrete-time auto-effects of the MZC dynamic process indicate that unpredictable fluctuations in the MZC are expected to remain for a briefer period than those of KBF and of VTD, implying that the MZC dynamic process had the least predictive power on itself (Fig 3d blue line). In other words, the steepness of the blue line in Fig 3d suggests that the change in MZC (outside that predicted from aging) are relatively rapid and thus may not predict MZC at some future time point.

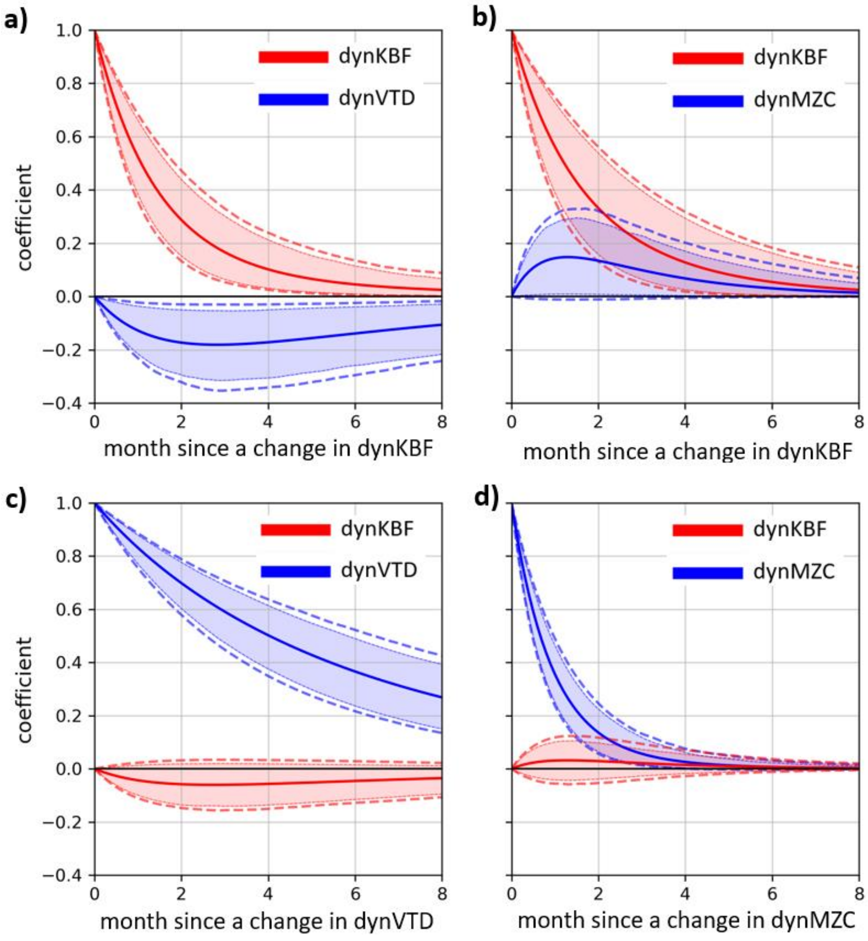


Fig 3 – Discrete-time cross-effect dynamics between the two latent processes implied by the two movement behaviour models for specific time interval (VTD: **a)** and **c)** ; MZC: for **b)** and **d)**). The mean coefficient estimates for each lag are represented by the solid line and the 95% credible intervals with the most external dashed lines, so that if the value zero is

not included in between these dashed lines the effect is significant. We also included the 90% credible intervals represented by the coloured area, to evaluate tendency (when value zero is not included in the dashed area).

Furthermore, the random disturbances influencing the KBF and the VTD dynamics processes are close to zero (0.03 [-0.06, 0.13], given in the diffusion matrices from the ctSummary_tex_fit documents), suggesting that KBF severity and VTD do not share unmeasured common causes that vary with time. A similar result applies for the MZC behaviour (-0.01 [-0.09, 0.07]).

With regard to the trend of each variable over time, we found that the VTD was diminishing for the first ~3-4 months at which point the magnitude stabilized, that the MZC followed a similar trend in the opposite direction (augmenting over time), and that the KBF increased through time until the end of the production period but at a slightly reduced speed over time (Supplementary S3 Fig). The KBF severity trend had a substantially larger range (start until end points of the trend) than the two behavioural trends, indicative of a stronger age, or time, effect (Supplementary S3 Fig).

Because we specified random intercept and slopes at the individual level, interpreting their correlation can also help to understand the nature of a behaviour - KBF association. Hens crossing on average more zones per transitions had overall more severe KBF (MZC and the KBF severity continuous intercept processes correlation: 0.47 [0.15, 0.71]), supporting the idea that KBF is positively associated with MZC. We found no such correlation between VTD and KBF (-0.05 [-0.36, 0.26]). In support of the notion that the two movement behaviours are not affecting the severity of KBF, we found that neither the initial state of the VTD trend nor the initial state of the MZC trend were significantly correlated to the KBF continuous intercept (VTD: 0.33 [-0.12, 0.66], MZC: -0.33 [-0.79, 0.36]). However, if the effect from the VTD behaviour was significant, it would mean that higher initial level of VTD associates with overall more severe KBF. The initial states of the two behaviour trend processes were not correlated to the initial state of the KBF trend process.

Space-use behaviours

To improve readability, in this section we will refer to the proportion of time spent on the top tier as "PropZ5". The model estimates are given in the supplementary documents (PropZ5- and Unevenness-ctSummary_tex_fit and ctSummary_summary). Results from the **continuous-time drift cross-effect** estimates of the KBF on the behaviours and vice-versa were all close to 0 (KBF on PropZ5: 0.06 [-0.09, 0.19], KBF on Unevenness: 0.05 [-0.10, 0.20], PropZ5 on KBF: -0.01 [-0.12, 0.11], Unevenness on KBF: -0.07 [-0.17, 0.03]). These results suggest that changes in space-use behaviours had no effect on the subsequent level of KBF severity and that changes in the KBF severity also had no effect on the subsequent space-use behaviours. These results are illustrated by the **discrete-time drift effects**, as for the movement behaviours, in Fig 4.

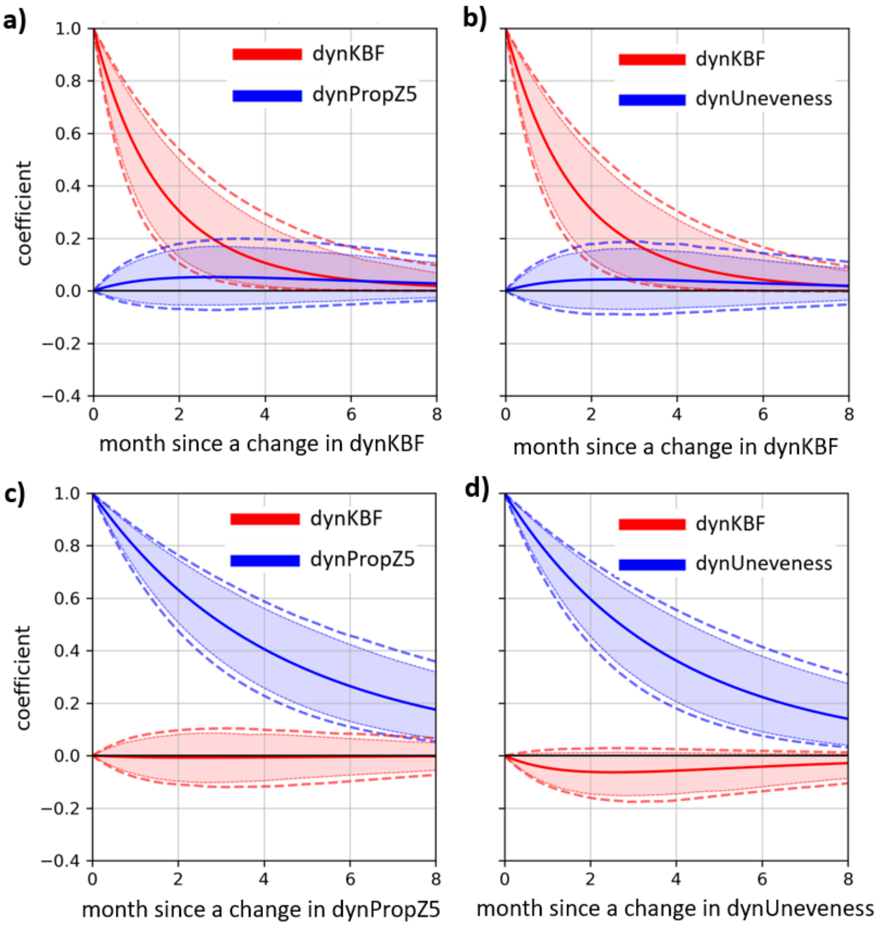


Fig 4 – Discrete-time cross-effect dynamics between the two latent processes implied by the two space-use behaviour models for specific time interval (PropZ5: **a)** and **c)** ; Uneveness: for **b)** and **d)**). The mean coefficient estimates for each lag are represented by the solid line and the 95% credible intervals with the most external dashed lines, so that if the value zero is not included in between these dashed lines the effect is significant. We also included the 90% credible intervals represented by the coloured area, to evaluate tendency (when value zero is not included in the dashed area).

Furthermore, the random disturbances influencing the PropZ5 and the KBF dynamics processes were close to zero (PropZ5: 0.00 [-0.08, 0.09]) and suggest that KBF severity and the PropZ5 behaviour did not share unmeasured common causes that vary with time. A similar result applies for the unevenness behaviour (0.03 [-0.04, 0.11])

With regard to the trend over time of each process suggested by the models, we found that the proportion of time spent on the top tier was increasing for the first ~3 months at which point the magnitude stabilized, and that the unevenness was diminishing for the first ~4 months at which point the magnitude stabilized (Supplementary Fig S3). The proportion of time spent on the top tier had a substantially larger range (start until end points of the trend) than the unevenness trend, indicative of a stronger age, or time, effect (Supplementary Fig S3).

There were two significant random effect correlations. First, the proportion of time spent on the top tier and the KBF severity continuous intercept processes were positively correlated (0.27 [0.07, 0.46]), indicating that hens that spent overall more time on the top tier had overall more severe KBF, supporting the general idea that KBF is positively associated with the time spent on the tier. Second, the unevenness behaviour and the KBF severity continuous intercept processes were positively correlated (0.31 [0.08, 0.51]), indicating that hens with a higher unevenness asymptote (suggesting an overall more uneven usage of the zones compared to what would be expected by chance) had a higher KBF slope, supporting the idea that KBF is positively associated with unevenness.

Discussion

The primary objective of this study was to gain insights into whether the severity of KBF may be affected by specific spatial behaviours and whether such fractures may, in turn, alter these behaviours. We believe this study is the first to explore both halves of this potential state-behaviour feedback loop within a single analysis framework. We found that an increase in the severity of KBF altered movement but not space-use behaviours. Specifically, an increase in KBF severity led to a decrease in vertical travelled distance and tended to be followed by a higher average number of zone crossed within a single transition. In contrast, we found no evidence that spatial behaviours affected the severity of KBF.

The decreased vertical travelled distance after an increase in the severity of KBF support our hypothesis that KBF reduce hens' activity. This result could be explained by the potential mechanical impairments in flying and walking and the perception of pain [22,24,122] resulting from such fractures. Hens experiencing pain may be less active to facilitate the healing process or to minimize associated pain. While a previous study already reported decreased vertical movements in hens with open fractures [29], Rufener et al. [10] did not find such an association. This lack of an association is likely due to a small sample size, as our study, which includes the data from Rufener et al. [10], did find an association. It is important to note that, holding all else equal, if an increase in the severity of KBF leads to a decreased vertical travelled distance, then those with more severe KBF will have a lower vertical travelled distance. However, due to the long-term consistent differences in spatial behaviours observed among hens [9,213,229], hens are not *all equal* in behavioural expression, and therefore we cannot conclude from that result that hens with higher KBF also on average exhibited lower vertical travelled distance. Instead, this result suggests that KBF induced behavioural change.

Apart from shedding light on the bidirectional relationships of the KBF severity score and animal mobility, these models also provide insights into these effects over time, which allows comparing these temporal effects across studies that used different time intervals between measurements. Here, we found that the expected effect of an increase in KBF on the vertical travelled distance to be greatest approximately three months after the change in KBF. In other words, a change in KBF was most predictive of a change in vertical travelled distance approximately 3 months later. The reason for this long-term effect is that changes in both KBF and behaviour were persistent in time. In other words, if the behaviour or the severity of KBF changes, we expect it to stay changed. A possible explanation for these enduring changes in the behaviour may be provided by the long-term individual consistency highlighted by previous studies in similar behaviours [213,229]. The enduring changes in the severity of KBF, may be attributed to the long-lasting feeling of pain, likely lasting for at least a few weeks [22,122], and healing process, which can require several months to complete [21].

We also found that an increase in the severity of KBF tended to be followed by more zones crossed within a transition, providing further evidence of altered vertical locomotion due to fractures. By using the same dataset as in Rufener et al. [10], this result reproduced the positive association previously found by that study. Additionally, the results provide insights into the causality between KBF and vertical locomotion, indicating that a change in the severity of KBF predicted a later change in the behaviour, but not vice versa, in opposition to our prediction. As tiers are not connected by ramps in our multi-tier aviary system, hens must transition between tiers by jumping or flying. Therefore, this result could suggest that hens with fractures attempt to minimize unnecessary stops between tiers, or, reduce the frequency of take-offs and/or landings. For instance, a hen at the top tier that was motivated to access the litter could make a direct transition without stopping at intermediate tiers. Prior research has showed that hens with fractures take a longer time to fly down from raised perches [27,28], which could indicate an increased reluctance toward take-offs and/or landings due to heightened pain during these movements. Alternatively, hens with fractures may be physically impaired, given that the keel is the site of muscle attachment [218] and involved in breathing [219,220], which could hinder their ability to stop in specific tiers.

The reduced vertical movements and increased mean number of zones crossed within a transition in hens with more severe KBF together indicate that hens may have difficulty in transitioning between the tiers in this multi-tier aviary system. Yet, to access all resources within our aviary, hens must move vertically, as resources are distributed throughout several tiers reaching up to 3 meters in height. Therefore, installation of structures facilitating transitions between stacked tiers, such as ramps [30,104], could reduce the prevalence of KBF and support hens to cope with fractures by ensuring improved access to all resources within the aviary. Future research could compare movement of recently fractured hens in multi-tier aviaries equipped with ramps versus those without ramps to assess whether ramps indeed facilitate the healing process of KBF.

In addition to reduced activity, we predicted that an increase in KBF would alter space-use behaviours. We expected this due to social and resource-related motivations that would lead hens to spend more time in the top tier where they would experience fewer agonistic interactions and have access to vital resources [200,217]. Results from the random effect correlations suggest that hens with overall more severe KBF used the zones more unevenly and spent more time on the top tier (during the day). The latter result supports the idea that KBF is positively associated with time spent on the top tier, as found by previous literature [10]. However, results from the temporal effects did not provide insights into whether KBF altered space-use behaviours, as we predicted, or whether space-use behaviours affected the severity of KBF. To gain a deeper understanding of the intricate relationships between KBF and space-use behaviors, future research should assess within-individual dynamics rather than relying on the effects at the population-level. That is, examining within-individual temporal effects could help determine whether several types of dynamics were at play, which could explain why we found no discernible population-level effects. For example, it is possible that some hens spent more time in the top tier due to increased severity of KBF, while others initially spent more time in the top tier for reasons other than KBF, which led to more severe KBF.

The absence of altered space-use behaviours in response to KBF may reflect different coping strategies among hens, which may stem from a variety of factors, including spatial preferences and dominance ranking. To investigate the role of individuality in response to KBF, future research relying on more observations per hen and ideally with reduced time intervals between consecutive observations could estimate within-individual dynamics between behaviours and KBF. For instance, using similar models than used in the present study, one could allow for individual variation in the temporal effect parameters [230]. Due to the limited number of observations per hen, we were unable to estimate within-individual dynamics. More data would also allow to control for additional factors that could influence the effect, such as the hybrid and flock, which we could not account for in our study.

In contrast to our result suggesting that KBF altered spatial behaviours, we found no evidence that spatial behaviours altered the severity of KBF. Specifically, we found no evidence that a change in these behaviours led to a change in the severity of KBF and no evidence that the initial level of the behaviours was related to the rate of change in the severity of KBF (from the random effects correlations). These result align with recent pathological evidence suggesting that collisions with infrastructure may not be responsible for the fractures located at the caudal tip of the bone, accounting for the majority of KBF. Instead, KBF may be attributed to the internal pressure exerted during egg-laying [126,221].

These results suggest that movement behaviours in multi-tier aviaries may not cause or exacerbate the severity of KBF. Yet, our movement behaviours likely relate to the amount (i.e., vertical travelled distance) and height (i.e., mean-zone-crossed per transition) of jumps/flights for which previous literature showed a considerable percentage of failed landings (9-21%) [231] that could result in KBF [3,30,31]. However, if KBF caused by collisions during failed landings manifest immediately after impact rather than developing gradually (e.g., micro-fractures which would gradually weaken the overall structural integrity of the bone), our methods could not have detected it. In that scenario, to determine if the observed behaviours may cause fractures, future research with smaller time intervals between consecutive observations of KBF is needed. Furthermore, we did not account for the location of the fractures, which could have hidden the potential influence of spatial behaviours in the development of fractures that are not located at the tip of the bone.

To further assess whether spatial behaviour may contribute to the maintenance and/or formation of KBF, other behaviours could be examined. For example, by studying the number of failed landings we could assess whether these landings can exacerbate KBF as generally thought, and whether KBF, in turn, can increase the risk of failing to land properly, resulting in an undesirable, positive feedback loop. The hypothesis on the second half of this feedback was stimulated by the trend we found indicating that an increase in the severity of KBF increased the likelihood of crossing multiple zones within one transition and potentially more hazardous landings. Furthermore, considering the possibility that fractures at the caudal tip of the bone may not be caused by collisions but by other processes such as egg-laying [221], future efforts should distinguish fractures from the cranial and the caudal bone surface.

With the advancements in sensor technology providing longitudinal data on welfare indicators and behaviour, we can now delve into their dynamics within commercially relevant settings. For example, by using continuous time dynamic models, instead of the common multivariate linear models, we can express the response variable of one regression equation as a predictor in another equation and thereby assess the mutual influence of the two variables within a single model. In future research, this approach could be applied to a wide range of variables extending beyond the current study's scope, including heart rate, body mass, and indicators of affective states, as well as other behaviours, including those related to social interactions, comfort, and feeding activities. By more comprehensively accounting for the existing relationships between various variables, we may better understand the long-term consequences of initial behavioural variation on animal welfare.

CHAPTER V

Based on:

Commercial hatchery practices have long-lasting effects on laying hens' spatial behaviour and health

Camille M. Montalcini[1,2], Matthew Petelle[1], and Michael J. Toscano[1]

1 Center for Proper Housing: Poultry and Rabbits (ZTHZ), Division of Animal Welfare, VPH Institute, University of Bern, Switzerland
2 Graduate School of Cellular and Biomedical Sciences, University of Bern, Switzerland

In revision in: PLOS One

Camille Montalcini participated to the experimental work needed to collect the movement and health data, conducted the analysis, and wrote the manuscript in agreement with her supervisor Dr. Michael J. Toscano and co-author Dr. Matthew B. Petelle.

Abstract

The commercial hatchery process is globally standardized and exposes billions of day-old layer chicks to stress every year. By alleviating this early stress, on-farm hatching is thought to improve animal welfare, yet little is known about its effects throughout production. This study compared the welfare and spatial behaviours during the laying period of hens hatched in an on-farm environment (OFH) to those hatched in a commercial hatchery and transferred at one day-old to a rearing barn (STAN). Specifically, we assessed how the transfer to the laying barn at 17 weeks of age, a similar stressor encountered by STAN hens early in life, affected space use and health, and determined whether effects aligned more with the 'silver-spoon' or 'environmental matching' hypothesis. We found that for the first three months post-transfer into the laying barn, OFH hens, on average, transitioned less between the aviary's tiers and spent less time on the littered floor. Because OFH hens became behaviourally more similar to STAN hens over time, these results suggest that OFH hens required a prolonged period to establish their daily behavioural patterns. Furthermore, OFH hens had more severe keel bone fractures throughout the laying period but similar feather damage and body mass to STAN hens. No differences were found in hen mortality or the number of eggs per live hen. These findings support the environmental matching hypothesis and suggest that early-life stressors may have prepared hens for later-life stressors, underscoring the importance of both early-life and adult environments in enhancing animal welfare throughout production.

Introduction

Every year, billions [37] of day-old layer chicks are exposed to stress [232] because of the globally standardized commercial hatchery process [34]. Chicks in hatcheries, are subjected to loud noise during incubation (~90dB), hatched in darkness, prevented from accessing feed, water, or litter, and subjected to sexing, vaccination, and transportation to rearing farms at one day of age [232]. The early exposure to stressors can be alleviated by hatching chicks on farm, where chicks are transported before hatching and have direct access to feed, water and litter.

Exposure to less stressful environmental conditions during development may have a positive impact on various aspects of individual fitness later on, a phenomenon called “silver-spoon” effect [233] (for a review on birds and mammals, see [5]). The silver-spoon hypothesis suggests that individuals who experience relatively better environmental conditions during development, such as characterized by reduced stressors or abundant nutritional availability, may be able to cope better when confronted with adversity later in life. Accordingly, it is not surprising that on-farm hatching has been shown to improve the well-being of broilers and laying hens later in life, in particular throughout development. Specifically, this practice has been shown to reduce total mortality [234] and footpad dermatitis [235,236] of broilers, and increase body mass [34,237,238] and reduce feather damage, comb injuries, and corticosterone reactivity during restraint [232] of laying hens.

Alternatively, by experiencing adverse environments during development, one may be prepared or adapted to handle similar adversities in the future. The “environmental matching” hypothesis suggests that environmental conditions in early life shape an individual phenotype via developmental plasticity [47], so that an individual is adapted to similar environmental conditions experienced earlier in life [239,240]. Therefore, chicks that hatched in a commercial hatchery and subsequently transported on farm could have a phenotype more adapted to aversive environments, such as transportation to a new environment, than chicks hatched on-farm. Thus, the benefit of a less aversive early life environment would depend on the environmental conditions experienced later in life [239–241]. However, studies evaluating the effect of the commercial hatchery process on adult laying hens in commercial settings are scarce, limiting our understanding of their long-term effects on hen welfare. The limited understanding is especially true for health issues that predominantly arise during adulthood or may worsen as hens age, such as feather damage [242,243] or keel bone fractures (KBF) [21,23]. Thus, a long-term approach is necessary to understand whether on-farm hatching improves hen welfare throughout production, or whether its relative benefits are eventually offset by later stressors that they are unable to manage.

In this study, we compared the severity of KBF, feather damage, body mass and spatial behaviours of laying hens hatched on-farm (OFH) to those hatched in a commercial hatchery and transferred at one day of age to the farm (STAN). Our goal was to determine if, and for how long, these two different environments experienced at one day of age could account for variations in animal welfare and behaviour during the laying period. We assessed how the transfer from the rearing to the laying barn, a similar stressor encountered by day-old STAN chicks, affected space-use and movement behaviours, as well as welfare indicators. We aimed to determine whether the observed effects aligned more with the 'silver-spoon' or 'environmental matching' hypothesis. The former hypothesis would be supported if OFH hens would display overall greater welfare compared to STAN hens, while the latter hypothesis would be supported if OFH would display overall worse welfare conditions. Better welfare could be here manifested via less severe KBF, reduced feather damage, and greater time spent in the littered floor and the winter garden. These areas provide enhanced opportunities for the expression of natural behaviours such as locomotion, exploring, foraging, scratching, and dust bathing, which are important for laying hen welfare [244,245].

Materials and methods

Ethical note

The research was conducted in accordance with the cantonal and federal regulations for the ethical treatment of experimentally used animals. All procedures were approved by the Bern Cantonal Veterinary Office (BE-45/20).

Study design

All Dekalb white chicks (*Gallus gallus domesticus*) originated from the same parent flock and began incubation off-site using standard hatchery practices. At 18 days of development, three days before hatching, 3,300 eggs were arbitrarily chosen as part of the on-farm hatch (OFH) treatment and all except 270 clear eggs transported to a commercial rearing barn at the Aviforum facilities in Switzerland. The eggs were transported in a commercial vehicle for less than 1.5 hours, which maintained a stable environmental temperature at an average of 36.4 °C. Eggs were positioned in HatchTech Setter Trays 15 cm above the littered floor, where feed, water, and litter were available. We monitored environmental conditions and temperature of 30 eggs every six hours throughout the hatching process. Specifically, we ensured that the ambient relative humidity remained above 30%, the windspeed below 0.15 m/s, ambient temperature above 32°C, and the eggshells temperature between 35-38 °C (see supplementary S1 Fig for eggshell temperature over time). Our methodology is similar to a previous study conducted on layer chicks [238].

At one day of age (DOA), OFH chicks were manually sexed by examining their wing feathers for sex-specific patterns and females were vaccinated (IB 4/91). On the same day, 1,200 chicks from the commercial hatchery were transported to the rearing barn as part of the STAN treatment. Transportation took place in a commercial vehicle over a duration of eight hours, during which a consistent environmental temperature around 28 °C was maintained. Unlike the OFH chicks, STAN chicks - in addition to being transported to new housing - hatched in darkness and were deprived direct access to food, water, or litter after hatching until arrival at the rearing barn. Similar to OFH chicks, STAN chicks were vaccinated and manually sexed by examining their wing feathers for sex-specific patterns at the hatchery facility. Although both OFH and TRAN chicks were manually sexed by the same company managing the post-hatch procedures (Prodavi SA, CH), we supervised the sorting of OFH chicks and encouraged the sexers to proceed gently, however no objective comparison was made of the handling procedures between treatment groups. STAN chicks were used to populate two rearing pens and OFH chicks were used to populate the other two pens (600 hens/pen). Males and surplus females were returned to the hatchery for humane disposal. At seven DOA all chicks were classified into a more/less explorer class. We did not use the class as an exploratory behaviour as the measurement could not be validated (Supplementary Text 1 from chapter III), though we controlled for the class in subsequent analysis. Simultaneously, 160 focal birds (40 hens/rearing pen) were selected from the 2'400 chicks. Of the 160 focal birds, 80 focal birds were classified as MEXP or LEXP (40 / class), while 80 were selected as a representative sample and used for another study that collected brain tissues throughout the laying period.

At 119 DOA, all hens were caught, put into a crate, and transported to one of eight laying pens on the same site (225 hens/pens, including 20 focal hens, four pens/treatment). Bird density was 8.1 hens per square-meter of permanent accessible area (225 hens/27.92 m^2). The laying barn contained a quasi-commercial multi-tier aviary system (Bolegg Terrace separated into 20 pens by grids illustrated in Supplementary S2 Fig; indoor length x width x height until the top tier grid floor: 7 x 2.3 x 2.69 m; previously described [30]) and an outside covered winter garden (WG ; $9.32\ m^2$) accessible by pop holes from 10:00 h to 16:00 h on most days. On the day of transfer to the laying barn the 160 focal hens were assigned a tracking device to continuously register their transitions across the indoor aviary levels and the winter garden (WG) until near the end of production (tracking period: September 2020 – July 2021). At five time points during the laying period (DOA: 127, 173, 243, 313, and 418), 16 randomly selected focal hens were killed (eight hens/treatment) to collect brain tissues as part of a separate study. Each of these time points also included welfare assessment (described in the below section), except for DOA 127, which was replaced by DOA 215 to capture more variation in animal welfare. For each hen killed, another hen from the same pen was arbitrarily selected to continuously track the same number of hens, for a total of 227 hens used in the study.

Welfare indicators

Welfare assessment included feather damage, radiographs for KBF, and body mass (digital scale in grams). During the welfare assessment, the observers were blinded to the treatment, laying pen identity, and hen class, and shown reliable in a previous effort for both feather damage and KBF severity scores [213]. The feather damage score (continuous, 0-100) was assigned using the photographs of white laying hens which we rescaled to 0-100 and took the complement to 100 so that higher scores are indicative of poorer welfare (score 1: approx. 100 – 76 depending on the extent of damage; score 2: approx. 75 – 51 ; etc.) for each body part [128]. More precisely, we assigned a score of the breast, tail, and neck, but not the back and wing feathers as these could not be reliably assessed because of the backpack containing the tracking tag (described below). We then averaged these to get an overall individual feather damage score. We assessed KBF severity (continuous, 0-100) based on the latero-lateral radiographs using the scoring methodology described by Rufener et al. [176], where the score is described as an indicator of the total amount of keel bone affected by fractures.

We excluded the first timepoint of KBF severity and feather damage, as there was little variation, with both having a median score of zero.

Spatial behaviours

We tracked individuals' transitions across five zones: the four different levels of the aviary (top tier, nestbox tier, lower tier, and littered floor) and the outside covered WG. We used a low-frequency tracking system with active tags (mass: 28.1 g) enclosed in a backpack mounted on the back of the hens (see Montalcini et al. (2022) [173] for the validation and description). Tracking data were collected from the first full day in the laying barn (DOA 119) until near the end of production (DOA 416). We excluded days with known disturbances (e.g., vaccinations or welfare assessments) and those with known tracking system malfunctions (e.g., low battery level). Subsequent analysis involved a period of 297 days of tracking, during which hens had on average 169 days tracked, with a minimum of three days tracked and a maximum of 250 days, involving a total of 227 hens and 38,303 hen-days observations.

We characterized the daily movement and space-use behaviours of each hen with six behaviours expressed while artificial light was provided. We used the (i) vertical travelled distance, defined as the total number of indoor tiers crossed, to account for the level of vertical movement. We used the proportion of the indoor time spent on the (ii) top tier, (iii) nestbox tier, and (iv) littered floor to account for indoor space-use behaviours, and (v) WG presence (yes/no) to account for the outdoor space-use behaviour. Finally, because the nestbox tier is of particular interest within commercial settings, we also used the (vi) time when a hen reached half of its nestbox tier duration, accounting only for hours where hens are expected to lay, that is between 02:00h and 08:00h, hereafter referred to as the nestbox tier timing.

Production traits

The female hatchability, i.e., the percentage of healthy female hatched, was 40% in the hatchery and 42.6% on-farm within a hatching window duration of 65 hours (for the on-farm chicks, see supplementary S1 Table for the hatching rate over time), was comparable to previously reported OFH results [238]. During the rearing period, there was a total of 11 deaths for each treatment (i.e., < 1%). We analyse production traits after the transfer to the laying barn, as hens had not laid eggs prior to that point. First, we used the number of early deaths per day in each pen during the laying phase. This dataset is right censored where the value 1 represents death, and 0 indicates being alive. Second, we used the hen daily average (average number of nest eggs per live hen) in each pen. Throughout the laying period, eggs laid inside the nestboxes were collected consistently at the same time every day and counted at the pen-level. We did not include floor eggs in our analysis as they represent approximately 0.24% of total eggs laid. In addition to the four pens per treatment group (STAN and OFH) with focal birds, we also used the data of two additional pens without focal hens but containing 205 birds from one treatment group with an additional 20 Lohmann LSL hens that hatched in the hatchery (called “special pens”).

Statistics

Welfare indicators

Statistical analyses were conducted in R version 3.6.1. To evaluate treatment effects on hen’s welfare indicators, we fitted one linear mixed-effect model from the ‘lme4’ package per welfare indicator (body mass, KBF severity, and feather damage) as a function of date (or health assessment identity), treatment, and date-treatment interaction. We controlled for class and included hen identity nested in pen identity as a random term. Pen identity was removed when fitting the KBF severity and body mass due to low variance leading to convergence issues. We scaled body mass to be within 0-1 within each welfare assessment separately prior to the model fit. Model assumptions were checked visually (normality and homoscedasticity of residuals). To assess significance of the date-treatment interaction, we compared each model with a model that did not contain the interaction variable using the function Anova from R. When the date-treatment interaction was significant ($p < 0.05$), we reported results from a post-hoc analysis with adjusted p-values (Bonferroni adjustment, package “emmeans”). When the date-treatment interaction was not significant, we removed it from the model and assessed significance of the treatment as a main effect by comparing the full and reduced models.

Spatial behaviours

To evaluate whether treatment groups differed in mean behaviours after the transfer to the laying barn, we fitted one generalized linear mixed-effect model from the package ‘glmmTMB’ [205] for each behaviour for the first month in the laying barn. To complement the findings and evaluate how long treatment groups differed in mean behaviours, we fitted models for each following month as well (10 months, six behaviours, total of 60 models), as a function of the treatment (with STAN as reference group). We chose month as the unit to analyse treatment effect over time, aiming to strike a balance between thoroughly estimating mean effects over time and avoiding potential noise linked to shorter time intervals like weeks. A previous study on the same hens, studying intra-individual variation in a composite behaviour, found that, on average, hens increased their indoor movements for 39 days after the transfer to the laying barn [213]. This previous result suggests that the transfer to the laying barn could have a long-term effect on hens’ spatial behaviours. However, we expected that any treatment differences in spatial behaviours would appear directly after the transfer to the laying barn and diminish over time. Therefore, any statistically significant treatment effect that does not follow that pattern was interpreted with care, and the results section emphasized coefficient estimates rather than p-values. We controlled for the class, time (defined as the number of days since the transfer to the laying barn), KBF severity, body mass, and number of hours with artificial lights on, by including them as fixed effects. We interpolated linearly (with monotonically increasing) both the KBF severity score and body mass for each hen separately to better control for their health between two consecutive health assessments considering that both scores exhibit an upward trend over time. The hen identity nested in the pen identity was included as a random term. To avoid convergence issues due to the very low explained variance by the pen identity, we performed a likelihood ratio test with and without the pen identity and chose the full model when the p-value was < 0.05. For the WG-related behaviour, we also controlled for the number of hours the WG was accessible and the average daily external temperature (°C), taken from the LSZB weather station (~12 km from the barn) and accessed via the Wolfram alpha API in Python. All continuous variables were scaled by two times the deviation to obtain coefficients comparable to those of binary predictors (i.e., the treatment) [246]. As the first full day in the laying barn coincided with the final day of September, we incorporated that day into the models for the first month (October).

The vertical travelled distance was modelled with a gaussian family for months 2-10 and with a zero-inflated Poisson model with the rescaled number of days in the barn as the zero-inflation parameter for month 1, as model assumption were otherwise not met. The nestbox tier usage was modelled with a gamma family (with a log-link function) and the WG presence with a binomial family (with a logit-link function). The proportion of indoor time spent on the top and nestbox tiers were both modelled with a *beta* family (with a logit-link function). The proportion of the indoor time spent on the littered floor was modelled with a gaussian family for months 2-10 and with a binomial family (with a logit-link function) to account for the excess of zeros (19.7% of observations) for month 1. Behaviours used in a beta distribution were first rescaled between 0.01 and 0.99. During the first month in the laying barn, hens had not reached the peak of production and the artificial light was turned on later in the day (days 1-9 at 9h, days 10-16 at 8h, days 17-22 at 7h, days 23-24 at 6h, days 25-30 at 5h, days 31-32 at 4h). Therefore, we did not analyse the first month in the laying barn for the two nestbox tier related behaviours because behaviours expressed during that first month would likely not be comparable with behaviours expressed in subsequent months. In addition, we did not use the first seven days when fitting the WG presence, as the opening of the WG was delayed until normal laying behavior commenced (i.e., the eighth day) as is common for commercial practices. Residuals were simulated using the 'DHARMa' package to verify model assumptions (normality and homoscedasticity of the residuals). We reported the bootstrapped coefficients [247], credible intervals, and p-values computed from the 'parameters' package using 500 iterations.

Production traits

To analyse daily mortality, we first estimated survival curves per treatments (without the special pens) and per pen (with the two special pens) using the Kaplan-Meier method with the ‘survival’ package in R [248]. Then, we used the log-rank test to test whether the treatment group (with the two special pens) differ in their survival curves. In order to account for the pen identities as a random effect and the “special pens” as a fixed effect, we also fitted a Cox proportional hazards models with both fixed and random effects, using the ‘coxme' package in R [249]. Because it is a fundamental assumption of both the log-rank test and the Cox proportion hazard test that the hazard ratio is constant over time [250], we tested the proportional hazards assumption. We found that the effect of both the pen identity, special pen effect, and the treatment covariates were not time-dependant (p<0.05). However, because there was a trend (p = 0.06) suggesting that the effect treatment had on the time until death may not be constant over time, we also split the data into two sets [251], considering the first 60 days in the laying barn (when there is a peak of mortality) and the remaining production period. The effect of treatment no longer was time dependent, and the effect of treatment remained the same. Therefore, we reported the result from the model with the full rather than split time period.

We analysed the daily average number of eggs per live hen in each pen using a sigmoid curve, $\frac{a}{1+exp\left(\frac{m-time}{s}\right)}$, where *time* represented by the number of days since being in the laying barn, *a* as the horizontal asymptote, *m* as the time point at *a*/2, and *s* as the steepness of the curve at *a*/2 [252]. Sigmoid curves were fitted using a nonlinear least square algorithm (nls function in R). We fitted a sigmoid curve for each pen separately and provided pen-level parameter estimates. More precisely, for each pen, we reported *a* as an indication of the level at which egg-production stabilises and *m* as an indication of the time point with the steepest slope, that is at the inflection point of the curve. We did so considering first the entire period and then only the onset of lay (first 60 days in the laying barn) as it is an important period in which we might observe the greatest variation between hens. Then, to statistically assess treatment effect we fitted a nonlinear mixed-effects model with the ‘nlme’ package [253] with treatment as fixed effect and pen identity as random effect on all three parameters (a, m, s). We did so when considering the entire period in the laying barn as well as when considering solely the first 60 days.

Results

Welfare indicators

The date-treatment interaction was neither a predictor of body mass ($x^2_{6,1121} = 9.49, p = 0.15$), feather damage ($x^2_{2,506} = 4.95, p = 0.08$), or KBF severity ($x^2_{3,669} = 4.34, p = 0.23$) and was therefore removed from all models. Treatment as a main effect was neither a predictor of body mass ($x^2_{1,1121} = 1.95, p = 0.16$) or feather damage ($x^2_{1,506} = 0.54, p = 0.46$), but was a predictor of KBF severity ($x^2_{1,669} = 6.35, p = 0.01$). Specifically, holding all else equal, the KBF severity score of OFH hens was, on average, 4.53 ([95% CI] = [0.99, 8.07]) points higher than that of STAN hens. Estimated marginal means (±95% CI) are presented in the supplementary information in S2 Table and model estimates and p-values are presented in S3 Table. The observed scores of each welfare indicator per treatment and date are illustrated in a violin plot (Figure 1), and their means (±SD) presented in Table 1.

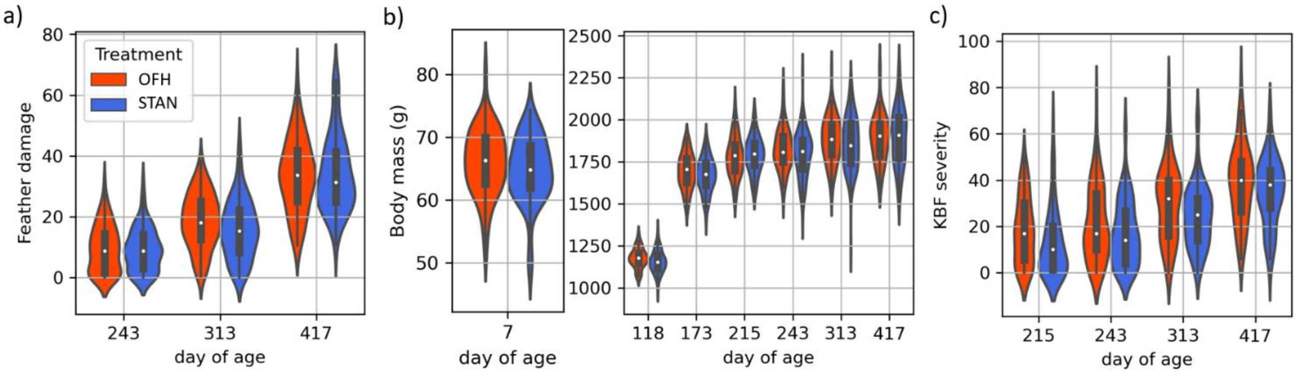


Fig 1. Violin plot of the raw welfare indicators over day of age per treatment. Red and blue colours represent the treatment groups, OFH and STAN, respectively, and feather damage is represented in (a), body mass in (b), and KBF severity in (c).

	Feather damage		Body mass (g)		KBF severity	
Day of age	OFH	STAN	OFH	STAN	OFH	STAN
7			66.04 ±5.17	64.75 ±5.12		
118			1177.75 ±57.88	1157.16 ±64.90		
173			1692.88 ±103.13	1671.03 ±99.26		
215			1780.19 ±118.01	1790.72 ±105.09	18.35 ±14.16	12.72 ±14.57
243	9.58 ±7.73	9.24 ±7.02	1822.38 ±122.70	1795.32 ±137.46	21.50 ±16.21	16.86 ±13.98
313	18.57 ±8.92	15.83 ±9.59	1877.29 ±134.15	1855.84 ±160.46	30.20 ±16.26	24.57 ±13.68
417	33.60 ±11.79	34.09 ±12.31	1888.44 ±139.77	1898.60 ±168.97	38.64 ±16.34	34.94 ±14.25

Table 1. Mean (±SD) observed values of raw welfare indicators per treatment groups.

Spatial behaviours

Model coefficients from the treatment predictor for each behaviour across months are displayed in Fig 2 (bootstrapped p-values and coefficients are detailed in supplementary S4 Table). To help further interpret the nature of the change, particularly in cases where a treatment effect was found for several months only (e.g., determining which, if any, of the treatment group had their behavioural responses *converging* toward those of the other group), the mean observed behavioural scores are presented per treatment and month in Table 2.

We can interpret exponentiated coefficients from our models with a beta family (i.e., models fitting the proportion of time spent on the top tier and the nestbox tier) as odds ratios. For example, during the first month in the laying barn, the odds of being on the top tier for OFH hens was 1.49 times that of STAN hens (95% CI [1.12, 2.04], p = 0.01; Fig 2a), indicating that on average OFH hens spent more time on the top tier. A similar effect was found for month 2 (β [95% CI] = 1.67 [1.14, 2.35], p < 0.01). Similar interpretations can be made for models with a proportion of time spent on the top tier or nestbox tier as response variables, although neither models related to the nestbox tier (figure 2c, 2f) nor to the top tier (figure 2a) had a significant effect of treatment beyond the second month.

We can interpret the exponentiated coefficients from our models with a binary response (i.e., models fitting the proportion of time spent on the littered floor (during month 1) and the WG presence responses) as odds ratios. For instance, during the first month in the laying barn, our results indicate that OFH hens were less likely to go to the littered floor compared to STAN hens. More specifically, we found an odds ratio of 0.30 (95% CI [0.12, 0.67], p < 0.01), meaning that the odds of going on the littered floor at least once during the day (during the first month) in OFH hens were 0.30 times the odds in STAN hens. A similar effect was maintained during months 2 and 3 (with a gaussian family; Fig 2b). More specifically, compared to STAN hens, OFH hens spent 7% and 5% less of their daily indoor time on the littered floor during months 2 and 3, respectively (month 2: β [95% CI] = -0.07 [-0.11, -0.02], p = 0.008; month 3: β [95% CI] = -0.05 [-0.11, -0.00], p = 0.04). Treatment was not a significant predictor of WG presence (see Fig 2d).

Furthermore, we found that OFH hens moved less vertically during the first three months (see Fig 2e). More specifically, hatching on farm was associated with a reduction of 31% in the vertical travelled distance during month 1 (exp(β) [95% CI] = 0.69 [0.59, 0.82], p < 0.001, Poisson distribution). The treatment effect persisted up to month 3, with OFH hens crossing on average 11 and 9 fewer zones per day than STAN hens during months 2 and 3, respectively (month 2: β [95% CI] = -10.62 [-16.13, -4.35], p < 0.001; month 3: β [95% CI] = -9.19 [-16.13, -2.50], p = 0.012, Gaussian distribution).

Lastly, OFH hens, on average, were slightly earlier in their nestbox tier timing during the second month than STAN hens (β [95% CI] = 0.96 [0.92, 0.99], p < 0.05; Fig 2f). The treatment showed no effect in subsequent months in the nestbox tier timing, and we found no effect of treatment on the proportion of time spent on the nestbox-tier.

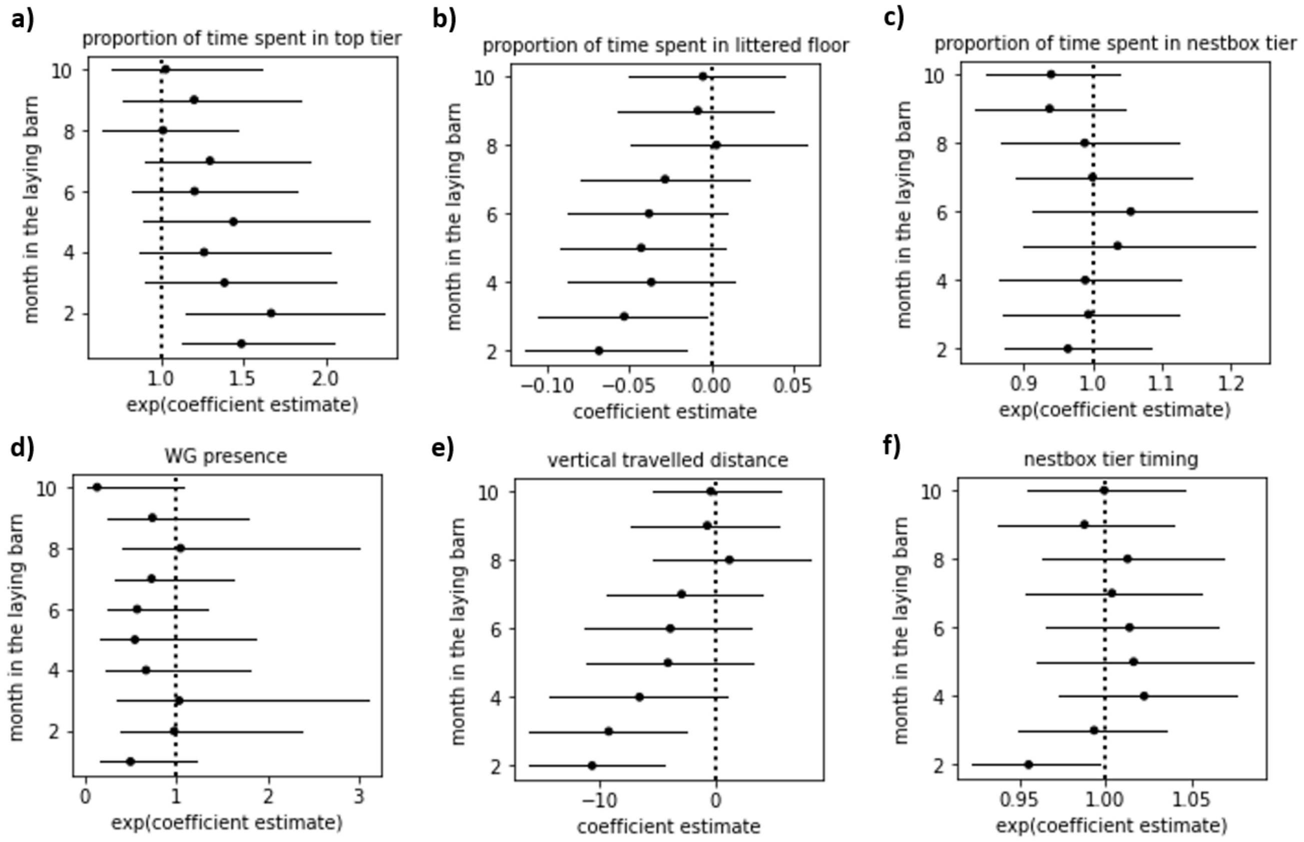


Fig 2. Coefficient estimates of treatment (with STAN as reference group) for each behaviour across months after the transfer to the laying barn. We interpreted significance whenever the confidence interval did not cross the dashed line. If the point and confidence interval lie to the right of the dashed line, it indicates that the model estimates higher behavioural response in OFH hens compared to STAN hens. Bootstrapped outputs (estimates and p-values) are detailed in supplementary S4 Table.

month in laying barn	proportion of indoor time spent			WG presence	vert. travelled distance	nestbox tier timing
	top tier	littered floor	nestbox tier			
month1	**0.53 *vs.* 0.42**	**0.26 *vs.* 0.32**		0.44 *vs.* 0.51	**28.51 *vs.* 39.02**	
month2	**0.41 *vs.* 0.30**	**0.31 *vs.* 0.39**	0.07 *vs.* 0.07	0.62 *vs.* 0.65	**47.03 *vs.* 58.67**	**5.00 *vs.* 5.31**
month3	0.34 *vs.* 0.28	**0.36 *vs.* 0.41**	0.07 *vs.* 0.07	0.66 *vs.* 0.67	**55.43 *vs.* 63.99**	4.24 *vs.* 4.29
month4	0.33 *vs.* 0.27	0.38 *vs.* 0.42	0.08 *vs.* 0.08	0.62 *vs.* 0.69	61.64 *vs.* 68.68	4.19 *vs.* 4.14
month5	0.33 *vs.* 0.28	0.37 *vs.* 0.42	0.08 *vs.* 0.08	0.61 *vs.* 0.70	63.03 *vs.* 68.94	4.16 *vs.* 4.18
month6	0.30 *vs.* 0.24	0.40 *vs.* 0.44	0.09 *vs.* 0.09	0.65 *vs.* 0.73	66.50 *vs.* 72.47	4.46 *vs.* 4.52
month7	0.31 *vs.* 0.28	0.40 *vs.* 0.41	0.09 *vs.* 0.09	0.62 *vs.* 0.69	61.38 *vs.* 65.76	4.59 *vs.* 4.62
month8	0.30 *vs.* 0.29	0.41 *vs.* 0.41	0.08 *vs.* 0.08	0.69 *vs.* 0.72	58.36 *vs.* 58.61	4.78 *vs.* 4.74
month9	0.28 *vs.* 0.26	0.43 *vs.* 0.43	0.08 *vs.* 0.09	0.71 *vs.* 0.83	57.11 *vs.* 58.38	5.05 *vs.* 5.18
month10	0.27 *vs.* 0.26	0.43 *vs.* 0.43	0.08 *vs.* 0.08	0.68 *vs.* 0.80	52.95 *vs.* 53.14	5.34 *vs.* 5.36

Table 2. **Mean observed values of raw behavioural data per treatment groups.** We highlighted in bold the months and behaviours for which we found a significant treatment effect.

Production traits

The survival probability for each treatment with 95% CI and per pen are represented in Figs 3a and 3b-c, respectively. Results from the log-rank test revealed no statistically significant difference in the survival curves between the two groups is (x^2 = 1.5, df = 1, p = 0.20). Similarly, results from the Cox proportion hazard model revealed no statistically significant difference in the hazard between the treatment groups (coefficient estimates = - 0.21, p = 0.21). The sigmoid curve fitting the daily average number of eggs per live hen for the first 60 days in the laying barn and over the full laying barn period are represented per pen in Fig 3d and 3e, respectively. Parameters are illustrated in green in Fig 3d. We provided the parameter estimates of each curve per pen in the supplementary S2 Fig. For example, from these estimates, we can observe that the levels at which the average number of daily eggs per live hen stabilised during the first 60 days in the laying barn (as measured by the *a* estimate) for OFH pens was between 0.948 and 0.964, and slightly lower for STAN pens (0.926 - 0.956). The time point (i.e., number of days since transfer to the laying barn) at the inflection point of the curve (as measured by the *m* estimate) is between 24.05 and 25.79 for OFH pens and slightly earlier for STAN pens between 23.71 and 25.00. However, we found no treatment effect on any of the sigmoidal parameters (a, m, s) when considering either the first 60 days or the entire period in the laying barn

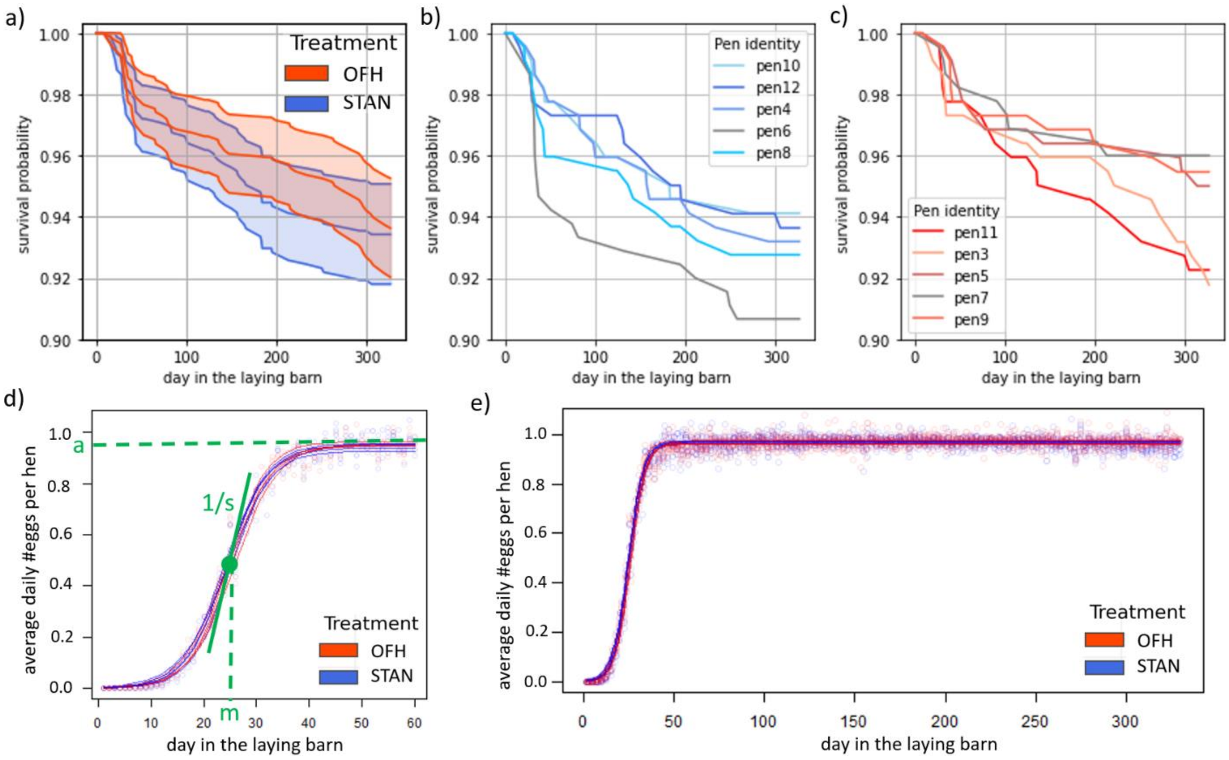


Fig 3. Survival probability and average daily number of eggs per live hens, over time and per treatment. Survival probability of OFH and TRAN hens over time in the laying barn (a), and per pen (STAN pens: (b) and OFH pens: (c), including the two special pens containing 205 animals of the one treatment group but also an additional 20 Lohmann LSL hens, highlighted in grey). Average daily number of eggs per hen (data points), with the fitted sigmoid curve for each pen, during the first 60 days in the laying barn (d) and the full period in the laying barn (e). Their associated parameters *a* (as an indication of the level at which egg production stabilise), *m* (as an indication of the time point at the inflection point of the curve), and *s* (steepness of the curve at a/2) are illustrated in green (d), and estimates given in the supplementary S2 Fig. Red and blue colours represent the treatment groups, OFH and STAN, respectively.

Discussion

In this study, we found that on average, OFH hens had more severe KBF throughout the laying period, transitioned less between the aviary tiers and spent less time on the littered floor and more on the top tier compared to STAN hens. We found no treatment effect on the daily presence in the WG, almost none on the two nestbox tier related behaviours, and none in body mass, feather damage, survival probability, and number of egg per live hens. To our knowledge, this study is the first to investigate the potential effects of the standard commercial hatchery process throughout the laying period on commercial hens' welfare and spatial behaviour, and in particular to assess the effect on KBF. These findings suggest that early-life stressors associated with the commercial hatchery environment – here, transport-related stressors and temporary post-hatch resource deprivation – may have prepared hens for their adult environment and thereby support the environmental matching hypothesis.

Welfare indicators

According to the environmental matching hypothesis, early-life stress can prepare an individual for similar adversities encountered in adulthood [240]. The adulthood of commercial laying hens is typically accompanied by aversive situations, including the transfer to the laying barn [254]. The potential stressors associated with this transfer, such as being transported to a new housing with new conspecifics, echo those encountered by STAN chicks. Therefore, the adult environment of commercial hens may better match with the environmental conditions experienced by day-old STAN chicks rather than those of day-old OFH chicks.

The result showing less severe KBF in STAN hens across the laying period compared to OFH hens provides support for the environmental matching hypothesis. Although KBF are considered as one of the greatest welfare issues in the egg production industry [20–25,123], effects of early-life stressors induced by commercial hatchery on KBF was to our knowledge not yet investigated. Various factors are thought to contribute to the prevalence of KBF, including genetic predisposition [255], nutrition [255], de-mineralized bone aggravated by the high egg laying rate [127,256,257], inactivity compromising bone health [125], and certain physical elements of the housing system that may cause pressure on the bone when hens are perching [32] or lead to trauma resulting from impact collisions [3,30,31,125]. Given that the STAN and OFH hens originated from the same parent flock, it is unlikely that genetic predisposition could explain this result. Moreover, apart from the nutritional differences after hatching, where OFH chicks had immediate access to feed and water, which is unlikely to negatively impacted their bone health, both STAN and OFH hens received the same nutrition. Because we observed minor differences in egg-production during the onset of lay, which, if anything, would suggest that OFH hens had a slower onset of lay, it is also unlikely that these differences would explain the result.

Given the limited amount of genetic and environmental differences between treatment groups outside of the treatment specific differences, we believe behavioural differences in activity and/or different usage of the housing system could explain why OFH hens had more severe KBF compared to STAN hens. For instance, our finding indicating a 31% reduction in vertical travelled distance during the first month for hens hatched on farms, may reflect a more general pattern of inactivity among OFH hens which could lead to poorer bone health and in turn higher KBF [125]. Additionally, OFH hens spent more time on the top tier, where hard metal perches are more abundant. High perch use would result in overall greater pressure on the keel bone leading to deviated keels, which in turn can weaken the keel's overall structural integrity [31]. Overall, we cannot exclude the possibility that the treatment difference we observed in the severity of KBF may be an indirect effect of the differences we observed in spatial behaviours between the treatment groups. Further research is required to replicate this finding and to understand the underlying mechanisms involved.

More recent evidence suggested that factors related to egg production, including internal pressure during the egg-laying process [126,221] and an early onset of lay [258], could favour KBF. We found no evidence of treatment effect in the average number of egg per live hens, nor did we observe treatment effect on the timing of the onset of lay. In fact, a previous study showed that on-farm hatched hens had a slower onset of egg-laying than hens hatched in commercial hatchery at 15-20 weeks of age [232]. Hence, while an early onset of laying may not explain the treatment effect on KBF observed in our study, it is plausible that other factors related to the internal pressure during the egg-laying process could be operating in the current effort.

We found that OFH hens had on average a similar body mass to STAN hens beginning with our earliest measurement at 7 DOA. We expected OFH hens to weigh more because of direct access to feed after hatching, which previous studies on day-old chicks demonstrated [34,237,238], but also because it is possible that animals may gain less weight due to early life stress [259,260]. However, beyond one day of age, the effect in previous literature is ambiguous. Studies contrasting on-farm hatching to standard commercial practices found no treatment differences at 4 or 7 DOA [34,237], a tendency for on-farm hatched chicks to weigh more up to 11 weeks of age [238], and that on-farm hatched chicks weighed less at 8, 15, 22 and 29 DOA [232]. Our study provides more evidence that treatment effects of body mass are not present during lay.

Furthermore, OFH hens had similar feather damage to STAN hens, which supports neither the silver-spoon nor the environmental matching hypothesis, but is in line with the inconclusive findings in earlier literature that reported both positive [232] and negative [237] effects. The ambiguity surrounding these results may stem from the multifactorial nature of feather damage, which includes factors such as feather pecking and abrasion resulting from different parts of the structures [118]. Therefore, the influences of the physical and social environments on feather damage could outweigh or interact with effects from early-life stressors.

Spatial behaviours

The literature on behavioural differences between hens that hatched on-farm versus in a commercial hatchery is sparse and mainly conducted in test arenas [34,232,261–263], hindering the extrapolation of the results to commercial settings, typically characterized by more complex housing systems and larger groups. In this study, we used tracking technology to monitor movements of hens within a quasi-commercial aviary system throughout the laying period.

We found that for the first three months post-transfer to the laying barn (up to 7-month-old) STAN hens spent more time on the littered floor and less on the top tier. The top tier has been shown to be used more extensively by hens with more severe keel bone fractures [10] and throughout full days following their transfer to the laying barn [213] (in a previous study on the same hens). Thus, it is possible that hens use this area over the day to offset stress or pain. Unlike the top tier, the littered floor promotes a diverse range of natural behaviours, including locomotion, dust bathing, exploring, foraging, and scratching. Therefore, these findings could indicate that OFH hens exhibited fewer natural behaviours compared to STAN hens in this early laying barn period. Furthermore, as OFH hens became behaviourally more similar to STAN hens over time (Table 2), these results further suggest that OFH hens required a prolonged period to establish their daily behavioural patterns. Overall, STAN hens may have exhibited better abilities in coping with the transfer to the laying barn, which would provide further support for the environmental matching hypothesis. Early-life stress can enhance behavioural flexibility [264], improve stress coping later in life [265], and facilitate spatial learning and memory [264,266], especially when experienced close in time and within the same context that is encountered later [267]. Therefore, increased stress induced by the commercial hatchery process near the time of transfer to the rearing barn could have induced focused attention and improved the memory of relevant information allowing STAN hens to better cope to the laying barn.

In addition to the potential effect of early-life stressors, it is possible that the STAN chicks may have exhibited faster acclimatization to their new environment due to the inherent effects of their transition from the hatchery to the farm. As suggested by Skånberg et al. [268], it is possible that environmental change during rearing plays a role in enhancing layer chicks' adaptability later in life. Specifically, STAN hens may have learned appropriate cognitive and locomotive skills during the rearing period, facilitating their adaptation to the aviary system. However, a previous study comparing the cognitive ability of layer chicks that had temporary post-hatch resource deprivation and eight-hour transport, to those that had ad libitum access to feed and water and were not transported, found no treatment effect in a cognitive test [238]. Further research is needed to investigate whether STAN hens possess better spatial-cognitive abilities in the laying period.

It is also possible that the additional early-life stressors experienced by STAN hens could have altered the functional and structural development of the Hypothalamic-Pituitary-Adrenal (HPA)-axis [269,270]. Typically, stress induced in the early postnatal period results in HPA-axis hyper-reactivity during adulthood [269–272], with enhanced depression-like behaviours and anxiety [273]. Previous studies on laying hens generally supports such alterations of the HPA-axis. It was shown that hens from a commercial hatchery had a more sensitive HPA-axis and a stronger reaction to stressors during the first weeks of life compared to on-farm hatched hens [232,262]. Here, we showed that STAN hens spent more time on the littered floor, less time on the top tier, and exhibited their typical behavioural patterns earlier after the transfer to the laying barn, in comparison to OFH hens. We believe that these behavioural pattern are not indicative of a stronger reaction to the transfer, and therefore believe that the additional early-life stressors experienced by STAN hens may not have significantly impacted their HPA-axis. Alternatively, it is also possible that the subsequent adversities encountered by all chicks in the rearing barn could have triggered the HPA axis of OFH hens to reach conditions resembling that of the STAN hens [269]. However, further research is necessary to understand how transportation and early life conditions influence the HPA-axis and its relationship to movement in laying hens.

We found almost no treatment effect on the two nestbox tier behaviours, which supports the idea of limited behavioural plasticity related to the use of the nestbox tier. Given the strong human selection for high productivity and the high motivation to use nestboxes [274], it is possible that those behaviours are more tightly correlated with physiology or with strong animal needs and thus are less plastic than others. That is, these behaviours may be less influenced by external or internal factors and more repeatable across different contexts, as suggested by previous findings [10,229].

Production traits

Until *in ovo* sexing becomes practical at commercial scale, differences between hatchery and on-farm treatments may not have practical effects relevant to animal welfare. However, with the advancement of *in ovo* sexing techniques [275], it is possible that on-farm hatching practices will become a standard method in the future. Therefore, it is important to compare both animal welfare and productivity between commercial hatchery and on-farm treatments within commercially relevant settings. We found no treatment effect on mortality, despite on-farm practices have already been shown to reduce total mortality in broilers [234]. We also found no treatment effect on the average number of egg per live hens in terms of both the level at which it stabilized (i.e., parameter *a*) and the time point at which the curve reached its inflection point (i.e., parameter *m*). To our knowledge, only two studies [232,237] have compared egg-production between such treatments, by collecting daily production data at the pen-level. The authors found that OFH hens had a slower onset of egg laying than STAN hens at 15-20 weeks of age [232] in one study and laid more and bigger eggs at 19-25 weeks of age in the second [237]. The absence of a statistical difference in our study may be attributed to the limited sample size and further research with substantially more pens or at the individual-level is needed to determine whether there are differences in egg-production between on-farm hatched hens and those hatched in commercial hatcheries.

Limitations

This study aimed to assess the potential impact of commercial hatchery practices and potential benefits of on-farm hatching for animal welfare within commercially relevant settings throughout most of the production period. Some potential early life stressors were in this study uniquely experienced by STAN animals, including food and water deprivation [276] and the transportation at one day of age to a new environment [254]. Yet, the hatchery-related procedures applied to OFH chicks might also have been aversive, including the transportation at 18 days of incubation [277] and the vaccination and sexing at one day of age. We designed the experiment so that all hatchery-related procedures should be less, or when not possible, equally, aversive for the OFH treatment compared to the TRAN treatment. However, our methodology did not allow to determine if the increased cumulative adversities encountered by STAN chicks during their development led to higher stress responses than those seen in OFH chicks. Hence, to assess the relative benefit of on-farm hatching practices for practical applications, future research should compare physiological stress and cognitive responses between treatment in the rearing and laying phase and strive to replicate on-farm hatching procedures more closely to the envisioned future practices (e.g., integrating *in ovo* sexing for the OFH treatment).

Conclusion

Hens from both treatments originated from the same parent flock yet produced different phenotypes depending on their early-life environments, suggesting the presence of developmental plasticity in our commercial hens. By providing nuance to the relative benefit of on-farm hatching for animal welfare described by previous literature, this study highlighted the critical role in designing early-life environments in accordance with the adult environment. For laying hens housed in aviary systems, our results could suggest that hatching on-farm rather than in commercial hatchery may lead to an increased mismatch between the hatching and laying environments that can be detrimental to animal welfare. We conclude that future research is needed to determine whether conventional husbandry practices, originally designed for hens hatched in commercial hatcheries, need to be adapted for hens hatched on-farm, such as implementing smoother transfers between rearing to laying barns.

General discussion

This general discussion will be divided into two sections. Based on the findings of Chapter I, I will start by briefly discussing the future use of tracking systems in monitoring individual behaviour in commercial settings and of model-based approaches to enhance accuracy while keeping original granularity of the registered data. Second, I will discuss the relevance of the findings from subsequent chapters for the welfare of farm animal, splitting the findings into the three main objectives of the thesis.

Tracking systems and data processing

To accurately track movements, it is crucial to validate tracking systems thoroughly, particularly when used within complex environments that may present challenges, such as aviaries. Although increasing sample size can minimize random errors, systematic errors caused by factors like environmental interference are more difficult to address and may lead to biases. To address the challenges arising from the high animal density and the relatively high concentration of material that can interfere with tracking signals, including metal equipment, we trained tree-based classifiers to identify false registrations, as observed via video observations (Chapter I). We used the most suitable model to the tracking data generated during the first experiment to filter potential false registrations. This model-based approach enabled us to quantify the daily number of transitions in each zone more accurately, explaining 99% of the variation compared to 91% when simply excluding registrations of duration smaller than one minute.

In addition to providing more accurate data, the model-based data cleaning approach preserved the original second-level granularity, which provided greater flexibility in subsequent analysis compared to a minute-level granularity. However, to apply the model-based approach for processing the data from the second experiment, it would have been necessary to conduct new video observations due to the system re-calibration between the two experiments. As this was not feasible, we simply excluded registrations of duration smaller than one minute, resulting in a one-minute-level granularity. This limited granularity of the data from the second experiment hindered our flexibility to extract some behaviours, such as the response to feed delivery (Chapter III), which required a level of precision down to the second due to the short duration of the feed delivery stimuli, lasting only 3 minutes. This example highlights the benefits of the model-based approach, despite requiring additional efforts including more video observations and statistical modelling efforts.

The validation of the tracking system, along with the findings from subsequent studies, underlines the efficacy of tracking systems in providing valuable insights into individual differences within commercially relevant settings. We used a low-frequency tracking system that is most comparable to the common ultra-wide band (UWB) systems (e.g., of usage: [16,108,133]), but that has lower frequencies to reduce possible interactions with the environment such as liquid and metallic materials [19]. However, a more widespread tracking system in farm animals is the radio-frequency identification (RFID) (e.g., [14,18,278]), which was also recently successfully validated in the same aviary system as used in this thesis [279]. These RFID systems are particularly attractive to researchers due to their low power consumption and operating costs [280], which has contributed to their expanding use in monitoring farm animal movements. While this thesis contributed to the collective effort of evaluating tracking systems for laying-hen farming, I believe that tracking systems based on RFID technology hold more favourable prospects than the tracking system validated in this thesis.

Implications for commercial laying hens

Spatial personality

The scientific literature on chicken behaviour is predominantly shaped by applied themes, particularly those related to improving animal welfare and productivity [6]. Yet, gaining a more fundamental understanding of chickens holds the potential to influence our perception, which could foster a positive shift in our attitudes towards these farm animals and ultimately contribute to enhanced animal welfare. Chickens (*Gallus gallus domesticus*) are the most common food-producing animals on land yet are often seen as unfamiliar [6]. They are typically perceived as cognitively simpler [7], with fewer emotions and less differentiated individual personalities [8] compared to other farm animals. However, this perception has already been shown to change when becoming more familiar with these animals [8], suggesting that unfamiliarity can negatively influence our perception. Because our perception can influence our attitudes towards animals which can in turn reinforce our initial perception [6,281], it is critical to become more familiar with chickens and their complex nature.

This thesis revealed long-term individual differences in averages of various spatial behaviours, with up to 44% and 66% of the behavioural variation attributed to individual differences, during the onset of lay and adulthood, respectively (Chapter II-III). During adulthood we found consistent individual differences in five spatial behaviours over time and across contexts, revealing their potential applicability as personality traits of commercial laying hens. Furthermore, we found that hens that travelled greater vertical distances, on average also went in the WG on more days, used the highest tier slightly less at night, and tended to use the feed-tiers more upon delivery of fresh feed. Interestingly, we found no syndrome involving the nesting-related behaviour, which suggests that its main mechanism is independent of those involved in the other behaviours and could be, for instance, the physiological rhythm. While not unexpected, these findings contribute to a growing body of research demonstrating the presence of personality in a number of avian species, including great tits (*Parus major*) [54,282,283], zebra finches (*Taeniopygia guttata*) [284–286], greylag geese (*Anser anser*) [287], and chickens (*Gallus gallus domesticus*) [58,288]. While this thesis focused on behaviours exhibited within commercial settings, we did not assess common personality traits via behavioural tests or video observations, such as aggressiveness, exploration, boldness, and sociability [53,55]. Future research should evaluate associations between similar spatial behaviours and these commonly studied traits to enhance interpretation. Although Chapter II-III did not have immediate applied purposes, they contributed to recognizing that commercial hens may have distinct personalities.

Evaluating how individuals differ in behaviours has been primarily focused on consistent individual differences in averages, restricting our understanding of important biological variation. We can obtain a more comprehensive understanding of individual differences by including residual intra-individual variation in behaviour, that is, behavioural predictability [167–169]. For example, studying individual differences in behavioural predictability following environmental changes in addition to their mean behavioural responses, may better reflect individual’s coping strategies. This is because during periods with environmental change animals may exhibit greater behavioural variation as a result from optimizing their behaviours over time and with increased experience. Since many farm animals are transferred to new environments multiple times throughout their life, behaviours reflecting coping strategies could provide valuable information in farm animal welfare studies [55,80].

In chapter II we found existence of individual differences in predictability and temporal plasticity of movements after the transfer to the laying barn. We also found that hens that were more predictable in their daily movements had more severe KBF at the end of production. Although future research is required to confirm the latter association and explore underlying mechanisms, this result underscores the potential of investigating behavioural predictability in animal welfare studies. For example, future research could investigate behavioural predictability alongside physiological responses to understand how this behavioural component relate to the proactive/reactive coping style [80]. From a neuroendocrinological perspective, more proactive animals may have less inhibitory control and in turn may be more predictable behaviourally compared to reactive animals [189]. In contrast, reactive animals would be behaviourally less predictable, more flexible and perform better under unpredictable environmental conditions [190,215]. However, due to the combination of a daily routine regulated by predictable management practices and the occurrences of unpredictable events such as diseases and vaccinations, it remains unclear which of the proactive or reactive personalities would experience higher welfare in commercial settings.

Dynamics between spatial behaviours and keel bone fractures

The state and the behaviour of animals are both dynamic and tightly coupled, influencing one another over time. Hence, there are instances where employing continuous time dynamic models may be more suitable than the standard multivariate linear models. With these dynamic models we can express the response variable of one regression equation as a predictor in another equation and thereby assess the mutual influence of the two variables within a single model. By understanding the dynamics between animal health and its behaviours, we could better comprehend when state-behaviour feedbacks may amplify welfare issues over time. In other words, such knowledge could help understanding how small initial differences between individuals could lead to large differences in state and behaviour over time.

In Chapter IV, we found that an increase in the severity of KBF altered two vertical movement behaviours, which may suggest that hens with more severe fractures may face additional difficulties transitioning between the stacked tiers of aviary systems, possibly because of additional pain [22,122] or being physically impaired [218]. Yet, to access all the resources within multi-tier aviaries, hens must move vertically, as resources are distributed throughout several tiers reaching up to 3.5 meters in height. Therefore, this result may suggest that the installation of structures facilitating transitions between the stacked tiers, such as ramps, could in addition to reduce KBF prevalence [30,104] be important for hens to cope with fractures.

In contrast to our result suggesting that KBF altered spatial behaviours, we found no evidence that spatial behaviours altered the severity of KBF. This result aligns with recent pathological evidence suggesting that collisions with infrastructure, and thus characteristic of spatial behaviours, may not be responsible for the fractures located at the caudal tip of the bone, which account for the majority of KBF. Instead, KBF may be attributed to the internal pressure exerted during egg-laying [126,221]. However, we did not account for the location of the fractures in the study, which could have hidden the potential influence of spatial behaviours in the formation and maintenance of fractures that are not located at the tip of the bone. Because previous literature has shown that KBF are associated with reduced egg laying performance [28,120,121], improvement of management practices to reduce KBF prevalence could be of interest to all pertinent stakeholders, including producers and genetic companies.

Spatial behaviours and Management practices

In addition to studying spatial behaviours in relation to animal health, investigating the effect of management practices on these behaviours may also yield practical implications for improving the welfare of commercial hens. In this thesis, we studied the long-term influence of *once-in-a-lifetime* management practices on hens' spatial behaviour (Chapter V and Chapter II) and the immediate behavioural responses to reoccurring environmental stimuli (Chapter III). In this section, I will discuss how our findings and similar efforts could be used to improve commercial practices of laying hens.

It is evident that certain aversive events, such as the transfer to a new environment (discussed in Chapter II) or stressors during development (discussed in Chapter V), can have long-lasting implications for animal welfare. Yet, potential avenues for enhancing animal welfare may also arise from studying animal responses to stimuli arising from routine management practices, such as the delivery of fresh feed, vaccination, or the daily access to a winter garden (Chapter III). We can deem a stimulus as predictable when the animals are, to some extent, able to anticipate it, whereas we can deem a stimulus as controllable, when its likelihood of occurrence depends on the behaviour of the animals [289,290]. Previous literature generally supports the idea that stress can arise from unpredictable or uncontrollable events [291,292]. Therefore, trying to increase predictability and controlling of management practices, especially aversive ones, could improve animal welfare. This could be investigated, for instance, with studies using positive reinforcement training [293], changing time intervals between their occurrences to affect the "temporal" predictability, or introducing specific sound cues before each occurrence to affect the "signalled" predictability. However, when boredom-like states are caused by a lack of novel stimulation, then it is also possible that reducing predictability of certain stimuli would help addressing such states [293]. The study of an ideal equilibrium between predictable and unpredictable routine management practices for the welfare of commercial laying hens may present an interesting avenue for research using sensor technology within commercial settings.

In this thesis, we defined a behaviour reflecting individual responses to a recurrent stimulus, the fresh feed delivery (Chapter III). This reflected an animal's tendency of being on a tier with feeders more frequently while the fresh feed is delivered than while it is not delivered. We found the behaviour to be repeatable over time and across contexts, with consistent differences between individuals explaining 33% and 24% of the variation, respectively. We also found substantially higher values of the feed delivery response than expected by chance, with on average 77±4% of individuals per day had a positive feed delivery response, indicating that most hens tended to go on a tier with feeders more frequently upon feed delivery. Although this behavioural response is based on space use and movement behaviours, the responses are expected to be predominantly driven by the presence of fresh feed or other characteristic of the stimulus, rather than specific spatial preference. It is important to acknowledge that since we cannot control for all internal or external drivers, the observed behaviour may not directly reflect animals' motivation to feed. However, it is a first step towards assessing a potential proxy of the animal's affective state in response to a recurring external stimulus based on tracking data. For example, one could explore how this feed delivery response relate to optimistic/pessimistic attitude obtained from the judgment bias test [210] and to boredom-like state. Indeed, the high level of responsiveness in hens, despite having ad libitum access to feed, could be indicative of boredom. Assessing behavioural responses to management practices within commercial settings could help understanding effects of these practices at the individual level. Ultimately, these responses could aid in promoting positive welfare state while minimizing negative ones (e.g., by increasing feeling of reward and reducing feeling of frustration or boredom).

In addition to studying responses to daily management practices, we studied the long-term influences of early-life stress induced by commercial hatchery practices on behavioural responses. We found that for the first three months post-transfer to the laying barn, which corresponds to approximately seven months of age, hens that hatched on-farm, on average, transitioned less between tiers in the aviary, compared to hens hatched in commercial hatchery under more stressful conditions. For instance, we found that on-farm hatching was associated with a decrease of 31% in the daily indoor transitions during the first month. These hens also spent less time on the littered floor, a zone with resources enabling the expression of comfort behaviours, and required more time to establish their behavioural patterns (observed later on). Altogether, the findings suggested that additional early-life stress (such as the transfer at one day of age to the rearing barn) may have prepared hens for the transfer from the rearing to the laying barn. Altogether, this thesis supports the use of tracking technologies to study daily, but also long-term, influences of human practices on hens within commercial settings.

Perspectives

In this section, I will discuss two perspectives that could stimulate new research questions and encourage the use of diverse analytical approaches on similar tracking data. First, I will discuss an alternative method that seeks to represent each hen daily raw tracking data (the hen’s series of transitions between the five zones) by a single point in a 3D space. Secondly, I will speak about how and why simulations could be considered to leverage similar datasets.

Exploring alternative behavioural traits through a 3D space

In this thesis, our selection of spatial behaviours was guided by prior assumptions regarding the significance of specific zones at particular times of the day and during specific life phases of the hens. However, we may have overlooked important behavioural traits, such as those that best uncover individuality or that are most associated with animals’ welfare condition. One reason for relying on these assumptions, was that our tracking system registered hens’ transitions between five defined zones (the three stacked tiers of the aviary, littered floor, and winter garden) instead of precise individual locations. Consequently, the data could not be treated as numerical time series, which would have enabled the computation of spatial behaviours based on distances between registered locations (instead of prior assumptions), such as activity levels, or in other set-up, home range size and tortuosity. In other words, an alternative approach could have allowed the extraction of behavioural traits based on distances without making prior assumptions on the zones.

For example, we could have represented the raw tracking data of each hen-day unique combination by a single point in a 3D space, i.e., a space with three numerical axes where distance metrics, such as Euclidean distance or cosine similarity, could then be used to measure the similarity or dissimilarity between two points’ coordinates in the space. One way to construct such a space, is to first construct a graph represented by nodes and links between nodes, where each node would correspond to the raw tracking data of a hen-day and each link between two nodes would indicate the strength of the spatial association between the two hen-days. For example, naming the two hen-day as hen *H1* - day *D1* and hen *H2* - day *D2*, we could compute the strength of their spatial association as the percentage of time hen *H1* and hen *H2* spent in the same zone, during day *D1* and day *D2*, respectively. Then, we could use an algorithm that map the nodes of the graph into a low-dimensional space of features while maximizing the likelihood of preserving the nodes neighbourhoods, such as node2vec [294]. This approach would also enable to extract centrality measures on the graph, although other methods that do not involve the construction of a graph may be computationally more efficient. By representing each hen-day tracking data in a space, we could extract other behavioural traits by examining the distances between different hen-days. Concepts from other fields, such as tortuosity or social niche [295], could inspire exploration of behavioural traits extracted from this 3D space.

To illustrate this approach, I used data from the second experiment (consisting of 1,176 unique hen-day combinations from 168 hens over a 7-day period) to represent the tracking data of each hen-day into a 3D space (Figure 1 a). In this space, I highlighted seven days of seven hens (i.e., a total of 49 hen-day) and illustrated their daily series of transition (raw tracking data) in Figure 1b using the same colour code. From this visualization, we can observe that the yellow hen-day combinations appear more isolated compared to the blue, green, and red hen-day combinations. Analysing the proximity but also the gaps between hen-days in such a 3D space may provide new behavioural traits that could, with no prior assumption on the zones, expose individual differences in their responses to environmental or internal changes.

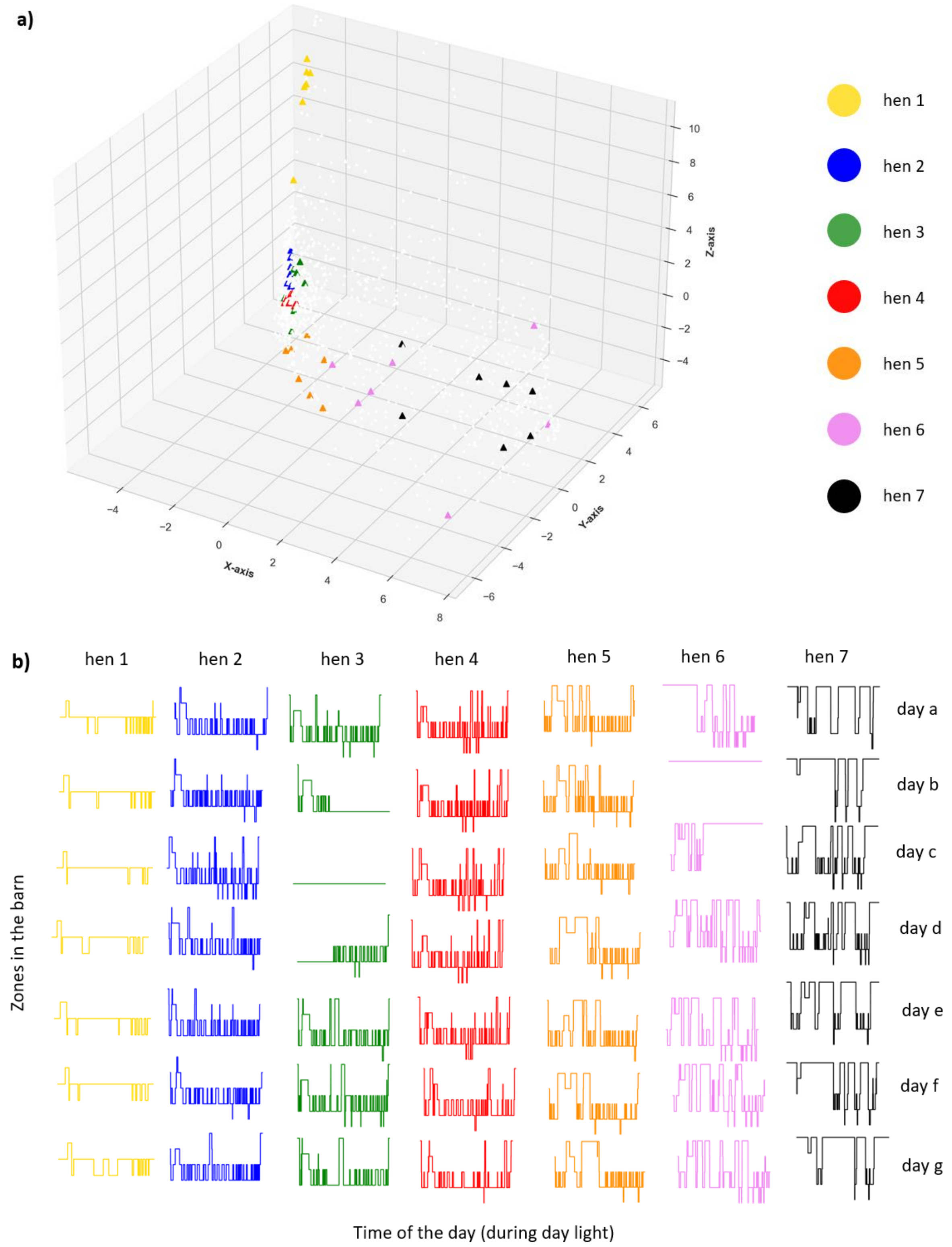


Figure 1 – Each of the hen-day is represented as a coloured triangle (the white colour represents hens used to construct the 3D space but that are not of interest in this visualization) in **a)**, and the corresponding raw tracking data for each of these triangles is illustrated in **b)**.

Simulations

In 1959, Russell and Burch [110] introduced the 3R principles (Replacement, Reduction, and Refinement), which are still used as an ethical framework for the design of animal experimentation. These principles involve three aspects to be considered when designing experiments, including the development of alternative methods allowing to replace animals (Replacement), the design of protocols to reduce the number of animals (Reduction), and the application of refinement practices (Refinement). Simulations can facilitate the implementation of these principles by generating data that may be difficult to obtain otherwise (but that are necessary to study complex systems) and by enabling researchers to explore multiple scenarios and hypotheses to improve their experimental design.

Throughout the two experiments conducted in this thesis, we have generated a large amount of data that could serve as a valuable foundation for simulating hen spatial behaviour in multi-tier aviary systems. The data could be used to generate plausible daily movement patterns and to estimate parameter ranges that can be used to compare different scenarios. I will briefly discuss two examples where simulations may be used to explore traits that may ultimately help breeding for farm animals with enhanced welfare. The first example involves the selection of new phenotypes that acknowledge hens' dual nature as both individuals and members of a flock. The second example refer to the idea of selecting for greater phenotypic variance within the next generation, in addition to targeting specific mean values.

Social facilitation occurs when an animal is more likely to engage in a particular behaviour in response to other animals engaging in the same behaviour. Social facilitation, as well as social learning, are important mechanisms shaping the behaviour of domesticated chickens, such as feeding behaviour [296], feather pecking [297], and cannibalism [298]. By using a different tracking system, one able to record animal proximity with greater precision than the system employed in this thesis, we could use metrics reflecting the tendency of an individual to follow other individuals more than being followed by them, such as the leading index introduced by Gómez et al. [299]. Consistent responses could emerge from consistent differences between individuals in their tendency to imitate the behaviour of others, but also from other mechanisms. For instance, when several individuals respond similarly to an environmental stimulus (e.g., heading towards the winter garden when it opens), consistent individual differences in their speed of expressing the behavioural response could lead to such consistent responses. Therefore, traits reflecting individuals’ tendency to imitate/spread specific behaviours should be carefully considered and validated.

Considering the role of social facilitation and learning on laying hens' behavioural expression, the inclusion of phenotypes that may reveal highly influential individuals may be an interesting avenue when studying novel selective breeding practices. For example, it is possible that selecting a smaller proportion of leaders exhibiting desirable behaviours (e.g., is among the first entering the winter garden) might prove more effective than selecting a larger proportion of the best-behaving hens (e.g., spending a lot of time in the winter garden). That is, in situations where social learning, facilitation or imitation plays a key role in shaping a negative or positive behaviour in the flock, it could be more effective to select hens not only based on their individual performance (e.g., behaviour and productivity), but also based on their effect on the rest of the flock. Because the outcome and risk of a strategy will depend on several parameters, simulations could help contrasting different strategy and bring preliminary insights into the feasibility and potential benefits of this approach. Parameters of interest may include the group size, the distribution of the behaviour (e.g., whether many hens occasionally lead the flock or if only a few hens consistently lead the flock), and the heritability of the behaviours.

Simulations may also help understanding the potential of selecting for greater phenotypic variance within the next generation, in addition to targeting specific mean values. For example, by selecting for greater variability in the timing of nestbox usage within a flock, we may reduce overcrowded nestboxes (potentially reducing smothered hens [211]) and occurrences where preferred nestboxes are occupied (potentially reducing floor eggs [212]). Previous research has already shown that some egg laying behaviours were heritable and may be used to select against floor eggs and improve egg production, including the nestbox preference, laying duration, oviposition traits, and clutch number [163,212]. Therefore, tracking technologies could be used to estimate plausible behavioural pattern that may be combined with simulations to contrast different scenarios (e.g., varying nestbox space and position of the nestboxes).

Final conclusion

This thesis provided new insights into the extent of consistent individual differences in averages behavioural expression of commercial laying hens housed in multi-tier aviaries. It also shed light on behavioural variation caused by variation in the severity of KBF and by management practices. We found long-lasting inter-individual differences in five spatial behaviours. We also found intra-individual variation underscoring the relevance of studying spatial behaviour to better understand how animals respond to environmental and internal changes. We conclude that tracking technologies hold promise as valuable tools for investigating individuality within commercially relevant settings, even in scenarios with complex designs like multi-tier aviaries. When integrated with the recognition of behaviours as signals encoding welfare-related information, tracking technologies could also allow the development of more effective approaches to improve farm animal welfare. However, further research is needed to assess relative benefits, for instance by studying similar spatial behaviours in relation to validated behavioural test and physiological responses.

References

1. 2022 Food and Agriculture Organization of the United Nations (FAO). *Statistics (Ber)*.
2. Heng Y, Peterson H, Resource XL. 2013 Consumer attitudes toward farm-animal welfare: the case of laying hens. *Journal of Agricultural and Resource Economics* **38**, 418–434.
3. Sandilands V, Moinard C, Sparks NHC. 2009 Providing laying hens with perches: fulfilling behavioural needs but causing injury? *Br Poult Sci* **50**, 395–406. (doi:10.1080/00071660903110844)
4. Weeks CA, Nicol CJ. 2006 Behavioural needs, priorities and preferences of laying hens. *Worlds Poult Sci J* **62**, 296–307. (doi:10.1079/WPS200598)
5. Hartcher KM, Jones B. 2017 The welfare of layer hens in cage and cage-free housing systems. *Worlds Poult Sci J* **73**, 767–782. (doi:10.1017/S0043933917000812)
6. Marino L. 2017 Thinking chickens: a review of cognition, emotion, and behavior in the domestic chicken. *Anim Cogn* **20**, 127. (doi:10.1007/S10071-016-1064-4)
7. Phillips CJC, McCulloch S. 2005 Student attitudes on animal sentience and use of animals in society. *J Biol Educ* **40**, 17–24. (doi:10.1080/00219266.2005.9656004)
8. Hazel SJ, O'Dwyer L, Ryan T. 2015 "Chickens Are a Lot Smarter than I Originally Thought": Changes in Student Attitudes to Chickens Following a Chicken Training Class. *Animals* **5**, 821–837. (doi:10.3390/ANI5030386)
9. Rufener C, Berezowski J, Maximiano Sousa F, Abreu Y, Asher L, Toscano MJ. 2018 Finding hens in a haystack: Consistency of movement patterns within and across individual laying hens maintained in large groups. *Sci Rep* **8**. (doi:10.1038/s41598-018-29962-x)
10. Rufener C, Abreu Y, Asher L, Berezowski JA, Maximiano Sousa F, Stratmann A, Toscano MJ. 2019 Keel bone fractures are associated with individual mobility of laying hens in an aviary system. *Appl Anim Behav Sci* **217**, 48–56. (doi:10.1016/j.applanim.2019.05.007)
11. Campbell DLM, Karcher DM, Siegford JM. 2016 Location tracking of individual laying hens housed in aviaries with different litter substrates. *Applied animal behaviour* **184**, 74–79.
12. Bari S *et al.* 2021 Relationship between Range Use and Fearfulness in Free-Range Hens from Different Rearing Enrichments. *Animals* **11**, 300. (doi:10.3390/ANI11020300)

13. Chielo LI, Pike T, Cooper J. 2016 Ranging Behaviour of Commercial Free-Range Laying Hens. *Animals* **6**, 28. (doi:10.3390/ANI6050028)

14. Larsen H, Cronin GM, Gebhardt-Henrich SG, Smith CL, Hemsworth PH, Rault J-L. 2017 Individual Ranging Behaviour Patterns in Commercial Free-Range Layers as Observed through RFID Tracking. *Animals* **7**. (doi:10.3390/ani7030021)

15. Richards GJ, Wilkins LJ, Knowles TG, Booth F, Toscano MJ, Nicol CJ, Brown SN. 2011 Continuous monitoring of pop hole usage by commercially housed free-range hens throughout the production cycle. *Veterinary Record* **169**, 338–338. (doi:10.1136/vr.d4603)

16. Stadig LM, Ampe B, Rodenburg TB, Reubens B, Maselyne J, Zhuang S, Criel J, Tuyttens FAM. 2018 An automated positioning system for monitoring chickens' location: Accuracy and registration success in a free-range area. *Appl Anim Behav Sci* **201**, 31–39. (doi:10.1016/j.applanim.2017.12.010)

17. Hartcher KM, Hickey KA, Hemsworth PH, Cronin GM, Wilkinson SJ, Singh M. 2016 Relationships between range access as monitored by radio frequency identification technology, fearfulness, and plumage damage in free-range laying hens. *Animal* **10**, 847–853. (doi:10.1017/S1751731115002463)

18. Welch M, Sibanda TZ, Vilela JDS, Kolakshyapati M, Schneider D, Ruhnke I. 2023 An Initial Study on the Use of Machine Learning and Radio Frequency Identification Data for Predicting Health Outcomes in Free-Range Laying Hens. *Animals* **13**, 1202. (doi:10.3390/ANI13071202)

19. Gharat V, Colin E, Baudoin G, Richard D. 2017 Impact of ferromagnetic obstacles on LF-RFID based indoor positioning systems. *2017 IEEE International Conference on RFID Technology and Application, RFID-TA 2017* , 284–289. (doi:10.1109/RFID-TA.2017.8098876)

20. Rodenburg T, Tuyttens F, de Reu K, Herman L, Zoons J, Sonck B. 2008 Welfare assessment of laying hens in furnished cages and non-cage systems: an on-farm comparison. *Animal Welfare* **17**, 363–373.

21. Baur S, Rufener C, Toscano MJ, Geissbühler U. 2020 Radiographic Evaluation of Keel Bone Damage in Laying Hens—Morphologic and Temporal Observations in a Longitudinal Study. *Front Vet Sci* **0**, 129. (doi:10.3389/FVETS.2020.00129)

22. Armstrong EA, Rufener C, Toscano MJ, Eastham JE, Guy JH, Sandilands V, Boswell T, Smulders T V. 2020 Keel bone fractures induce a depressive-like state in laying hens. *Sci Rep* **10**, 1–14. (doi:10.1038/s41598-020-59940-1)

23. Rufener C, Makagon MM. 2020 Keel bone fractures in laying hens: a systematic review of prevalence across age, housing systems, and strains. *J Anim Sci* **98**, S36–S51. (doi:10.1093/JAS/SKAA145)

24. Nasr MAF, Nicol CJ, Murrell JC. 2012 Do Laying Hens with Keel Bone Fractures Experience Pain? *PLoS One* **7**, e42420. (doi:10.1371/JOURNAL.PONE.0042420)

25. Harlander-Matauschek A, Rodenburg TB, Sandilands V, Tobalske BW, Toscano MJ. 2015 Causes of keel bone damage and their solutions in laying hens. *Worlds Poult Sci J* **71**, 461–472. (doi:10.1017/S0043933915002135)
26. Sih A, Mathot KJ, Moirón M, Montiglio PO, Wolf M, Dingemanse NJ. 2015 Animal personality and state–behaviour feedbacks: a review and guide for empiricists. *Trends Ecol Evol* **30**, 50–60. (doi:10.1016/J.TREE.2014.11.004)
27. Nasr MAF, Nicol CJ, Wilkins L, Murrell JC. 2015 The effects of two non-steroidal anti-inflammatory drugs on the mobility of laying hens with keel bone fractures. *Vet Anaesth Analg* **42**, 197–204. (doi:10.1111/VAA.12175)
28. Nasr M, Murrell J, Wilkins LJ, Nicol CJ. 2012 The effect of keel fractures on egg-production parameters, mobility and behaviour in individual laying hens. *Animal Welfare* , 127–135.
29. Rentsch AK, Rufener CB, Spadavecchia C, Stratmann A, Toscano MJ. 2019 Laying hen's mobility is impaired by keel bone fractures and does not improve with paracetamol treatment. *Appl Anim Behav Sci* **216**, 19–25. (doi:10.1016/j.applanim.2019.04.015)
30. Stratmann A, Fröhlich EKF, Gebhardt-Henrich SG, Harlander-Matauschek A, Würbel H, Toscano MJ. 2015 Modification of aviary design reduces incidence of falls, collisions and keel bone damage in laying hens. *Appl Anim Behav Sci* **165**, 112–123. (doi:10.1016/j.applanim.2015.01.012)
31. Stratmann A, Fröhlich EKF, Harlander-Matauschek A, Schrader L, Toscano MJ, Würbel H, Gebhardt-Henrich SG. 2015 Soft Perches in an Aviary System Reduce Incidence of Keel Bone Damage in Laying Hens. *PLoS One* **10**, e0122568. (doi:10.1371/JOURNAL.PONE.0122568)
32. Pickel T, Schrader L, Scholz B. 2011 Pressure load on keel bone and foot pads in perching laying hens in relation to perch design. *Poult Sci* **90**, 715–724. (doi:10.3382/PS.2010-01025)
33. Riber AB, Hinrichsen LK. 2016 Keel-bone damage and foot injuries in commercial laying hens in Denmark. *Animal Welfare* **25**, 179–184. (doi:10.7120/09627286.25.2.179)
34. Hedlund L, Jensen P. 2021 Incubation and hatching conditions of laying hen chicks explain a large part of the stress effects from commercial large-scale hatcheries. *Poult Sci* **100**, 1–8. (doi:10.1016/J.PSJ.2020.10.015)
35. Bell AM, Hankison SJ, Laskowski KL. 2009 The repeatability of behaviour: a meta-analysis. *Anim Behav* **77**, 771–783. (doi:10.1016/J.ANBEHAV.2008.12.022)
36. Dingemanse NJ, Kazem AJN, Réale D, Wright J. 2010 Behavioural reaction norms: animal personality meets individual plasticity. *Trends Ecol Evol* **25**, 81–89. (doi:10.1016/J.TREE.2009.07.013)
37. In press. FAO. Food and Agricultural Organization of the United Nations. See http://www.fao.org (accessed on 4 January 2023).

38. De Jager M, Weissing FJ, Herman PMJ, Nolet BA, Van De Koppel J. 2011 Lévy walks evolve through interaction between movement and environmental complexity. *Science (1979)* **332**, 1551–1553. (doi:10.1126/SCIENCE.1201187/SUPPL_FILE/IBM1201187S1.M)

39. Piersma T, Drent J. 2003 Phenotypic flexibility and the evolution of organismal design. *Trends Ecol Evol* **18**, 228–233. (doi:10.1016/S0169-5347(03)00036-3)

40. Scheiner SM. 1993 Genetics and Evolution of Phenotypic Plasticity. *Annu Rev Ecol Syst* **24**, 35–68. (doi:10.1146/ANNUREV.ES.24.110193.000343)

41. Padilla DK, Adolph SC. 1996 Plastic inducible morphologies are not always adaptive: The importance of time delays in a stochastic environment. *Evol Ecol* **10**, 105–117. (doi:10.1007/BF01239351/METRICS)

42. Nussey DH, Postma E, Gienapp P, Visser ME. 2005 Evolution: Selection on heritable phenotypic plasticity in a wild bird population. *Science (1979)* **310**, 304–306. (doi:10.1126/SCIENCE.1117004/SUPPL_FILE/NUSSEY.SOM.PDF)

43. Snell-Rood EC. 2013 An overview of the evolutionary causes and consequences of behavioural plasticity. *Anim Behav* **85**, 1004–1011. (doi:10.1016/J.ANBEHAV.2012.12.031)

44. Brönmark C, Lakowitz T, Hollander J. 2011 Predator-Induced Morphological Plasticity Across Local Populations of a Freshwater Snail. *PLoS One* **6**, e21773. (doi:10.1371/JOURNAL.PONE.0021773)

45. Wilgers DJ, Hebets EA. 2011 Complex courtship displays facilitate male reproductive success and plasticity in signaling across variable environments. *Curr Zool* **57**, 175–186. (doi:10.1093/CZOOLO/57.2.175)

46. Ouyang JQ, Sharp P, Quetting M, Hau M. 2013 Endocrine phenotype, reproductive success and survival in the great tit, Parus major. *J Evol Biol* **26**, 1988–1998. (doi:10.1111/JEB.12202)

47. Lea AJ, Tung J, Archie EA, Alberts SC. 2017 Developmental plasticity: Bridging research in evolution and human health. *Evol Med Public Health* **2017**, 162. (doi:10.1093/EMPH/EOX019)

48. West-Eberhard MJ. 2003 *Developmental Plasticity and Evolution*. See https://books.google.ch/books?hl=en&lr=&id=7DQNTPYaHlYC&oi=fnd&pg=PR15&dq=West+eberhard+2003+behaviour+plastic+traits&ots=X04UbUSaeC&sig=eQ5Ou4ndX7ryO77S5o4AXx-XVss&redir_esc=y#v=onepage&q=West%20eberhard%202003%20behaviour%20plastic%20traits&f=false.

49. Piersma T, Lindström Å. 1997 Rapid reversible changes in organ size as a component of adaptive behaviour. *Trends Ecol Evol* **12**, 134–138. (doi:10.1016/S0169-5347(97)01003-3)

50. Mathot KJ, Dingemanse NJ. 2015 Plasticity and Personality. *Integrative Organismal Biology* , 55–69. (doi:10.1002/9781118398814.CH4)

51. Stuber EF, Carlson BS, Jesmer BR. 2022 Spatial personalities: a meta-analysis of consistent individual differences in spatial behavior. *Behavioral Ecology* , 1–10. (doi:10.1093/BEHECO/ARAB147)

52. Sih A, Bell A, Johnson JC. 2004 Behavioral syndromes: an ecological and evolutionary overview. *Trends Ecol Evol* **19**, 372–378. (doi:10.1016/J.TREE.2004.04.009)

53. Réale D, Reader SM, Sol D, McDougall PT, Dingemanse NJ. 2007 Integrating animal temperament within ecology and evolution. *Biological Reviews* **82**, 291–318. (doi:10.1111/J.1469-185X.2007.00010.X)

54. Dingemanse NJ, Both C, Drent PJ, Van Oers K, Van Noordwijk AJ. 2002 Repeatability and heritability of exploratory behaviour in great tits from the wild. *Anim Behav* **64**, 929–938. (doi:10.1006/ANBE.2002.2006)

55. Finkemeier M-A, Langbein J, Puppe B. 2018 Personality Research in Mammalian Farm Animals: Concepts, Measures, and Relationship to Welfare. *Front Vet Sci* **5**, 131. (doi:10.3389/FVETS.2018.00131)

56. Dochtermann NA, Dingemanse NJ. 2013 Behavioral syndromes as evolutionary constraints. *Behavioral Ecology* **24**, 806–811. (doi:10.1093/BEHECO/ART002)

57. Cockrem JF. 2007 Stress, corticosterone responses and avian personalities. *J Ornithol* **148**, 169–178. (doi:10.1007/S10336-007-0175-8/FIGURES/6)

58. de Haas EN, Lee C, Hernandez CE, Naguib M, Rodenburg TB. 2017 Individual differences in personality in laying hens are related to learning a colour cue association. *Behavioural Processes* **134**, 37–42. (doi:10.1016/J.BEPROC.2016.11.001)

59. Boon AK, Réale D, Boutin S. 2008 Personality, habitat use, and their consequences for survival in North American red squirrels Tamiasciurus hudsonicus. *Oikos* **117**, 1321–1328. (doi:10.1111/J.0030-1299.2008.16567.X)

60. Réale D, Garant D, Humphries MM, Bergeron P, Careau V, Montiglio PO. 2010 Personality and the emergence of the pace-of-life syndrome concept at the population level. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 4051–4063. (doi:10.1098/RSTB.2010.0208)

61. Spiegel O, Leu ST, Bull CM, Sih A. 2017 What's your move? Movement as a link between personality and spatial dynamics in animal populations. *Ecol Lett* **20**, 3–18. (doi:10.1111/ELE.12708)

62. Montiglio P-O, Royauté R. 2013 Contaminants as a neglected source of behavioural variation. *Anim Behav* (doi:10.1016/j.anbehav.2013.11.018)

63. Nakayama S, Laskowski KL, Klefoth T, Arlinghaus R. 2016 Between- and within-individual variation in activity increases with water temperature in wild perch. *Behavioral Ecology* **27**, 1676–1683. (doi:10.1093/BEHECO/ARW090)

64. Tucker MA *et al.* 2018 Moving in the Anthropocene: Global reductions in terrestrial mammalian movements. *Science (1979)* **359**, 466–469. (doi:10.1126/SCIENCE.AAM9712)

65. Krause J, James R, Croft DP. 2010 Personality in the context of social networks. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 4099–4106. (doi:10.1098/RSTB.2010.0216)

66. Bell AM, Sih A. 2007 Exposure to predation generates personality in threespined sticklebacks (Gasterosteus aculeatus). *Ecol Lett* **10**, 828–834. (doi:10.1111/J.1461-0248.2007.01081.X)

67. Shettleworth SJ. 2001 Animal cognition and animal behaviour. *Anim Behav* **61**, 277–286. (doi:10.1006/ANBE.2000.1606)

68. Dingemanse NJ, Wolf M. 2010 Recent models for adaptive personality differences: a review. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 3947–3958. (doi:10.1098/RSTB.2010.0221)

69. Petelle MB, McCoy DE, Alejandro V, Martin JGA, Blumstein DT. 2013 Development of boldness and docility in yellow-bellied marmots. *Anim Behav* **86**, 1147–1154. (doi:10.1016/J.ANBEHAV.2013.09.016)

70. Bierbach D, Laskowski KL, Wolf M. 2017 Behavioural individuality in clonal fish arises despite near-identical rearing conditions. *Nat Commun* **8**, 1–7. (doi:10.1038/ncomms15361)

71. Fisher DN, Brachmann M, Burant JB. 2018 Complex dynamics and the development of behavioural individuality. *Anim Behav* **138**, e1–e6. (doi:10.1016/J.ANBEHAV.2018.02.015)

72. Stirling DG, Réale D, Roff DA. 2002 Selection, structure and the heritability of behaviour. *J Evol Biol* **15**, 277–289. (doi:10.1046/J.1420-9101.2002.00389.X)

73. Van Oers K, De Jong G, Van Noordwijk AJ, Kempenaers B, Drent PJ. 2005 Contribution of genetics to the study of animal personalities: a review of case studies. *Behaviour* **142**, 1185–1206. (doi:10.1163/156853905774539364)

74. Dochtermann NA, Schwab T, Sih A. 2015 The contribution of additive genetic variation to personality variation: heritability of personality. *Proceedings of the Royal Society B: Biological Sciences* **282**. (doi:10.1098/RSPB.2014.2201)

75. Petelle MB, Martin JGA, Blumstein DT. 2015 Heritability and genetic correlations of personality traits in a wild population of yellow-bellied marmots (Marmota flaviventris). *J Evol Biol* **28**, 1840–1848. (doi:10.1111/JEB.12700)

76. Hedlund L, Løvlie H. 2015 Personality and production: Nervous cows produce less milk. *J Dairy Sci* **98**, 5819–5828. (doi:10.3168/JDS.2014-8667)

77. Neave HW, Costa JHC, Weary DM, von Keyserlingk MAG. 2018 Personality is associated with feeding behavior and performance in dairy calves. *J Dairy Sci* **101**, 7437–7449. (doi:10.3168/JDS.2017-14248)

78. Neja W, Sawa A, Jankowska M, Bogucki M, Kręzel-Czopek S. 2015 Effect of the temperament of dairy cows on lifetime production efficiency. *Archiv Tierzucht* **58**, 193–197. (doi:10.5194/AAB-58-193-2015)

79. Richter SH, Hintze S. 2019 From the individual to the population – and back again? Emphasising the role of the individual in animal welfare science. *Appl Anim Behav Sci* **212**, 1–8. (doi:10.1016/J.APPLANIM.2018.12.012)

80. Koolhaas JM, van Reenen CG. 2016 Animal behavior and well-being symposium: Interaction between coping style/personality, stress, and welfare: Relevance for domestic farm animals. *J Anim Sci* **94**, 2284–2296. (doi:10.2527/jas.2015-0125)

81. Smith BR, Blumstein DT. 2008 Fitness consequences of personality: a meta-analysis. *Behavioral Ecology* **19**, 448–455. (doi:10.1093/BEHECO/ARM144)

82. Dingemanse NJ, Réale D. 2005 Natural selection and animal personality. *Behaviour* **142**, 1159–1184. (doi:10.1163/156853905774539445)

83. Icken W, Cavero D, Schmutz M, Thurner S, Wendl G, Preisinger R. 2008 Analysis of the free range behaviour of laying hens and the genetic and phenotypic relationships with laying performance. *Br Poult Sci* **49**, 533–541. (doi:10.1080/00071660802158357)

84. O'Malley CI, Turner SP, D'Eath RB, Steibel JP, Bates RO, Ernst CW, Siegford JM. 2019 Animal personality in the management and welfare of pigs. *Appl Anim Behav Sci* **218**, 104821. (doi:10.1016/J.APPLANIM.2019.06.002)

85. Friedrich J, Brand B, Schwerin M. 2015 Genetics of cattle temperament and its impact on livestock production and breeding – a review. *Arch Anim Breed* **58**, 13–21. (doi:10.5194/AAB-58-13-2015)

86. Fraser D. 2008 Understanding animal welfare. *Acta Vet Scand* **50**, 1–7. (doi:10.1186/1751-0147-50-S1-S1/FIGURES/1)

87. Dawkins MS. 2023 Natural Behaviour Is Not Enough: Farm Animal Welfare Needs Modern Answers to Tinbergen's Four Questions. *Animals* **13**, 988. (doi:10.3390/ANI13060988)

88. Dawkins MS. 2023 Farm animal welfare: Beyond "natural" behavior. *Science (1979)* **379**, 326–328. (doi:10.1126/SCIENCE.ADE5437/ASSET/08104670-5121-4219-A791-88A493078A03/ASSETS/IMAGES/LARGE/SCIENCE.ADE5437-F1.JPG)

89. Sirovnik J, Würbel H, Toscano MJ. 2018 Feeder space affects access to the feeder, aggression, and feed conversion in laying hens in an aviary system. *Appl Anim Behav Sci* **198**, 75–82. (doi:10.1016/J.APPLANIM.2017.09.017)

90. Huzzey JM, DeVries TJ, Valois P, Von Keyserlingk MAG. 2006 Stocking Density and Feed Barrier Design Affect the Feeding and Social Behavior of Dairy Cattle. *J Dairy Sci* **89**, 126–133. (doi:10.3168/JDS.S0022-0302(06)72075-6)

91. Estevez I, Andersen IL, Nævdal E. 2007 Group size, density and social dynamics in farm animals. *Appl Anim Behav Sci* **103**, 185–204. (doi:10.1016/J.APPLANIM.2006.05.025)

92. Proudfoot K, Habing G. 2015 Social stress as a cause of diseases in farm animals: Current knowledge and future directions. *The Veterinary Journal* **206**, 15–21. (doi:10.1016/J.TVJL.2015.05.024)

93. Tizard I. 2008 Sickness behavior, its mechanisms and significance. *Anim Health Res Rev* **9**, 87–99. (doi:10.1017/S1466252308001448)

94. Schillings J, Bennett R, Rose DC. 2021 Exploring the Potential of Precision Livestock Farming Technologies to Help Address Farm Animal Welfare. *Frontiers in Animal Science* **2**. (doi:10.3389/FANIM.2021.639678)

95. Berckmans D. 2014 Precision livestock farming technologies for welfare management in intensive livestock systems. *Rev. sci. tech. Off. int. Epiz* **33**, 189–196.

96. Herlin A, Brunberg E, Hultgren J, Högberg N, Rydberg A, Skarin A. 2021 Animal Welfare Implications of Digital Tools for Monitoring and Management of Cattle and Sheep on Pasture. *Animals* **11**, 829. (doi:10.3390/ANI11030829)

97. Weary DM, Huzzey JM, Von Keyserlingk MAG. 2009 Board-invited Review: Using behavior to predict and identify ill health in animals. *J Anim Sci* **87**, 770–777. (doi:10.2527/JAS.2008-1297)

98. Dawkins MS, Wang L, Ellwood SA, Roberts SJ, Gebhardt-Henrich SG. 2021 Optical flow, behaviour and broiler chicken welfare in the UK and Switzerland. *Appl Anim Behav Sci* **234**, 105–180. (doi:10.1016/J.APPLANIM.2020.105180)

99. Sepúlveda-Varas P, Weary DM, von Keyserlingk MAG. 2014 Lying behavior and postpartum health status in grazing dairy cows. *J Dairy Sci* **97**, 6334–6343. (doi:10.3168/JDS.2014-8357)

100. Vázquez Diosdado JA, Barker ZE, Hodges HR, Amory JR, Croft DP, Bell NJ, Codling EA. 2018 Space-use patterns highlight behavioural differences linked to lameness, parity, and days in milk in barn-housed dairy cows. *PLoS One* **13**, e0208424. (doi:10.1371/JOURNAL.PONE.0208424)

101. Matthews SG, Miller AL, Clapp J, Plötz T, Kyriazakis I. 2016 Early detection of health and welfare compromises through automated detection of behavioural changes in pigs. *The Veterinary Journal* **217**, 43–51. (doi:10.1016/J.TVJL.2016.09.005)

102. Lee HJ, Roberts SJ, Drake KA, Dawkins MS. 2011 Prediction of feather damage in laying hens using optical flows and Markov models. *J R Soc Interface* **8**, 489–499. (doi:10.1098/RSIF.2010.0268)

103. Colles FM, Cain RJ, Nickson T, Smith AL, Roberts SJ, Maiden MCJ, Lunn D, Dawkins MS. 2016 Monitoring chicken flock behavior provides early warning of infection by human pathogen Campylobacter. *Proceedings of the Royal Society B: Biological Sciences* **283**. (doi:10.1098/RSPB.2015.2323)

104. Heerkens JLT, Delezie E, Ampe B, Rodenburg TB, Tuyttens FAM. 2016 Ramps and hybrid effects on keel bone and foot pad disorders in modified aviaries for laying hens. *Poult Sci* **95**, 2479–2488. (doi:10.3382/PS/PEW157)

105. Forkman B, Boissy A, Meunier-Salaün MC, Canali E, Jones RB. 2007 A critical review of fear tests used on cattle, pigs, sheep, poultry and horses. *Physiol Behav* **92**, 340–374. (doi:10.1016/J.PHYSBEH.2007.03.016)

106. Mouchet A, Dingemanse NJ. 2021 A quantitative genetics approach to validate lab- versus field-based behavior in novel environments. *Behavioral Ecology* **32**, 903–911. (doi:10.1093/BEHECO/ARAB059)

107. Fisher DN, James A, Rodríguez-Muñoz R, Tregenza T. 2015 Behaviour in captivity predicts some aspects of natural behaviour, but not others, in a wild cricket population. *Proceedings of the Royal Society B: Biological Sciences* **282**. (doi:10.1098/RSPB.2015.0708)

108. Rodenburg TB *et al.* 2017 The use of sensor technology and genomics to breed for laying hens that show less damaging behaviour. In *8th European Conference on Precision Livestock Farming*, Nantes, France.

109. Ellen ED *et al.* 2019 Review of Sensor Technologies in Animal Breeding: Phenotyping Behaviors of Laying Hens to Select Against Feather Pecking. *Animals* **9**, 108. (doi:10.3390/ani9030108)

110. Russell W, Burch R. 1959 *The principles of humane experimental technique*. See http://117.239.25.194:7000/jspui/bitstream/123456789/1342/1/PRILIMINERY%20%20AND%20%20CONTENTS.pdf.

111. D'Eath RB, Keeling LJ. 2003 Social discrimination and aggression by laying hens in large groups: from peck orders to social tolerance. *Appl Anim Behav Sci* **84**, 197–212. (doi:10.1016/J.APPLANIM.2003.08.010)

112. McGrath N, Dunlop R, Dwyer C, Burman O, Phillips CJC. 2017 Hens vary their vocal repertoire and structure when anticipating different types of reward. *Anim Behav* **130**, 79–96. (doi:10.1016/J.ANBEHAV.2017.05.025)

113. Evans CS, Evans L. 1999 Chicken food calls are functionally referential. *Anim Behav* **58**, 307–319. (doi:10.1006/ANBE.1999.1143)

114. Collias N, Joos M. 1953 The Spectrographic Analysis of Sound Signals of the Domestic Fowl. *Behaviour* **5**, 175–188. (doi:10.1163/156853953X00104)

115. Evans CS, Evans L. 2006 Representational signalling in birds. *Biol Lett* **3**, 8–11. (doi:10.1098/RSBL.2006.0561)

116. Marx G, Leppelt J, Ellendorff F. 2001 Vocalisation in chicks (Gallus gallus dom.) during stepwise social isolation. *Appl Anim Behav Sci* **75**, 61–74. (doi:10.1016/S0168-1591(01)00180-0)

117. Steenfeldt S, Nielsen BL. 2015 Welfare of organic laying hens kept at different indoor stocking densities in a multi-tier aviary system. I: egg laying, and use of veranda and outdoor area. *Animal* **9**, 1509–1517. (doi:10.1017/S1751731115000713)

118. Heerkens JLT, Delezie E, Kempen I, Zoons J, Ampe B, Rodenburg TB, Tuyttens FAM. 2015 Specific characteristics of the aviary housing system affect plumage condition, mortality and production in laying hens. *Poult Sci* **94**, 2008–2017. (doi:10.3382/PS/PEV187)

119. Appleby MC, Hughes BO. 1991 Welfare of laying hens in cages and alternative systems: environmental, physical and behavioural aspects. *Worlds Poult Sci J* **47**, 109–128. (doi:10.1079/WPS19910013)

120. Rufener C, Baur S, Stratmann A, Toscano MJ. 2019 Keel bone fractures affect egg laying performance but not egg quality in laying hens housed in a commercial aviary system. *Poult Sci* **98**, 1589–1600. (doi:10.3382/PS/PEY544)

121. Nasr MAF, Murrell J, Nicol CJ. 2013 The effect of keel fractures on egg production, feed and water consumption in individual laying hens. *Br Poult Sci* **54**, 165–170. (doi:10.1080/00071668.2013.767437)

122. Nasr MAF, Browne WJ, Caplen G, Hothersall B, Murrell JC, Nicol CJ. 2013 Positive affective state induced by opioid analgesia in laying hens with bone fractures. *Appl Anim Behav Sci* **147**, 127–131. (doi:10.1016/J.APPLANIM.2013.04.015)

123. Riber AB, Casey-Trott TM, Herskin MS. 2018 The Influence of Keel Bone Damage on Welfare of Laying Hens. *Front Vet Sci* **0**, 6. (doi:10.3389/FVETS.2018.00006)

124. Casey-Trott TM, Widowski TM. 2016 Behavioral differences of laying hens with fractured keel bones within furnished cages. *Front Vet Sci* **3**, 42. (doi:10.3389/FVETS.2016.00042/BIBTEX)

125. Toscano MJ, Dunn IC, Christensen JP, Petow S, Kittelsen K, Ulrich R. 2020 Explanations for keel bone fractures in laying hens: are there explanations in addition to elevated egg production? *Poult Sci* **99**, 4183–4194. (doi:10.1016/J.PSJ.2020.05.035)

126. Thøfner I, Shih C-C, Villa C, Christensen JP. 2023 The egg-laying process might be a contributing factor for the development of keel bone fractures in laying hens. In *XXIInd Congress of the World Veterinary Poultry Association*, p. 228.

127. Whitehead CC, Fleming RH. 2000 Osteoporosis in Cage Layers. *Poult Sci* **79**, 1033–1041. (doi:10.1093/PS/79.7.1033)

128. Tauson R, Kjaer J, Maria GA, Cepero R, Holm K-E. 2005 The creation of a common scoring system for the integument and health of laying hens: Applied scoring of integument and health in laying hens. Final report Health from the Laywell project. See https://www.laywel.eu/web/pdf/deliverables%2031-33%20health.pdf (accessed on 22 December 2022).

129. Li L, Zhao Y, Oliveira J, Verhoijsen W, Liu K, Xin H. 2017 A UHF RFID System for Studying Individual Feeding and Nesting Behaviors of Group-Housed Laying Hens. *Trans ASABE* **60**, 1337–1347. (doi:10.13031/trans.12202)

130. Triguero-Ocaña R, Vicente J, Acevedo P. 2019 Performance of proximity loggers under controlled field conditions: An assessment from a wildlife ecological and epidemiological perspective. *Animal Biotelemetry* **7**, 1–9. (doi:10.1186/s40317-019-0186-2)

131. Dodge Y. 2008 *The concise encyclopedia of statistics*. See https://books.google.com/books?hl=en&lr=&id=k2zklGOBRDwC&oi=fnd&pg=PA1&ots=9nY1a43jnW&sig=bL6gYW8YJ9l3FubEc2cl8StQLFY.

132. Sibanda TZ *et al.* 2019 Flock use of the range is associated with the use of different components of a multi-tier aviary system in commercial free-range laying hens. *Br Poult Sci* **61**, 97–106. (doi:10.1080/00071668.2019.1686123)

133. van der Sluis M, Klerk B, Ellen ED, Haas Y, Hijink T, Rodenburg TB. 2019 Validation of an Ultra-Wideband Tracking System for Recording Individual Levels of Activity in Broilers. *Animals (Basel)* **9**. (doi:10.3390/ani9080580)

134. Liu H, Darabi H, Banerjee P, Liu J. 2007 Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* **37**, 1067–1080. (doi:10.1109/TSMCC.2007.905750)

135. Rokach L, Maimon O. 2005 Decision Trees. In *Data mining and knowledge discovery handbook* (eds O Maimon, L Rokach), pp. 165–192. New York: Springer. (doi:10.1007/0-387-25465-X_9)

136. Ebiele FMJ, Atemkeng M. 2020 Conventional Machine Learning based on Feature Engineering for Detecting Pneumonia from Chest X-rays. *dl.acm.org* , 149–155. (doi:10.1145/3410886.3410898)

137. Friedman JH. 2001 machine. *The Annals of Statistics* **29**, 1189–1232. (doi:10.1214/aos/1013203451)

138. Dorogush AV, Ershov V, Gulin A. 2018 CatBoost: gradient boosting with categorical features support.

139. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. 2017 CatBoost: unbiased boosting with categorical features. *Adv Neural Inf Process Syst* **2018-December**, 6638–6648.

140. Gorunescu F. 2011 *Data Mining: Concepts, models and techniques*. See https://books.google.com/books?hl=en&lr=&id=yJvKY-sB6zkC&oi=fnd&pg=PP2&dq=Data+Mining:+Concepts+and+Techniques+2011&ots=puJvA4pOvE&sig=PVRjdoZNMNNb254MiQ2Ze0SR_4E.

141. McNemar Q. 1947 Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **12**, 153–157. (doi:10.1007/BF02295996)

142. Good R, Fletcher HJ. 1981 Reporting explained variance. *J Res Sci Teach* **18**, 1–7. (doi:10.1002/tea.3660180102)

143. Willmott CJ, Matsuura K. 2005 Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim Res* **30**, 79–82. (doi:10.3354/cr030079)

144. Pedregosa F *et al.* 2011 Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830.

145. Bates D, Mächler M, Bolker BM, Walker SC. 2015 Fitting linear mixed-effects models using lme4. *J Stat Softw* **67**.

146. Li N, Ren Z, Li D, Zeng L. 2020 Review: Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming. *Animal* **14**, 617–625. (doi:10.1017/S1751731119002155)

147. Halachmi I, Guarino M, Bewley J, Pastell M. 2019 Smart Animal Agriculture: Application of Real-Time Sensors to Improve Animal Well-Being and Production. *Annu Rev Anim Biosci* **7**, 403–425. (doi:10.1146/ANNUREV-ANIMAL-020518-114851)

148. Rowe E, Dawkins MS, Gebhardt-Henrich SG. 2019 A Systematic Review of Precision Livestock Farming in the Poultry Sector: Is Technology Focussed on Improving Bird Welfare? *Animals 2019, Vol. 9, Page 614* **9**, 614. (doi:10.3390/ANI9090614)

149. Ahmed G, Malick RAS, Akhunzada A, Zahid S, Sagri MR, Gani A. 2021 An Approach towards IoT-Based Predictive Service for Early Detection of Diseases in Poultry Chickens. *Sustainability* **13**, 13396. (doi:10.3390/SU132313396)

150. Gómez Y, Stygar AH, Boumans IJMM, Bokkers EAM, Pedersen LJ, Niemi JK, Pastell M, Manteca X, Llonch P. 2021 A Systematic Review on Validated Precision Livestock Farming Technologies for Pig Production and Its Potential to Assess Animal Welfare. *Front Vet Sci* **8**, 660565. (doi:10.3389/FVETS.2021.660565/FULL)

151. Odintsov Vaintrub M, Levit H, Chincarini M, Fusaro I, Giammarco M, Vignola G. 2021 Review: Precision livestock farming, automats and new technologies: possible applications in extensive dairy sheep farming. *Animal* **15**, 100143. (doi:10.1016/J.ANIMAL.2020.100143)

152. Fogsgaard KK, Røntved CM, Sørensen P, Herskin MS. 2012 Sickness behavior in dairy cows during Escherichia coli mastitis. *J Dairy Sci* **95**, 630–638. (doi:10.3168/JDS.2011-4350)

153. Dochtermann NA, Schwab T, Anderson Berdal M, Dalos J, Royauté R. 2019 The Heritability of Behavior: A Meta-analysis. *Journal of Heredity* **110**, 403–410. (doi:10.1093/JHERED/ESZ023)

154. Richards GJ *et al.* 2012 Pop hole use by hens with different keel fracture status monitored throughout the laying period. *Veterinary Record* **170**, 494. (doi:10.1136/vr.100489)

155. Rodriguez-Aurrekoetxea A, Estevez I. 2016 Use of space and its impact on the welfare of laying hens in a commercial free-range system. *Poult Sci* **95**, 2503–2513. (doi:10.3382/PS/PEW238)

156. Fagan W, Lewis M, Auger-Méthé M, … TA-E, 2013 undefined. 2013 Spatial memory and animal movement. *Wiley Online Library* **16**, 1316–1329. (doi:10.1111/ele.12165)

157. Campbell DLM, Talk AC, Loh ZA, Dyall TR, Lee C. 2018 Spatial Cognition and Range Use in Free-Range Laying Hens. *Animals 2018, Vol. 8, Page 26* **8**, 26. (doi:10.3390/ANI8020026)

158. Ihwagi FW, Skidmore AK, Wang T, Bastille-Rousseau G, Toxopeus AG, Douglas-Hamilton I. 2019 Poaching lowers elephant path tortuosity: implications for conservation. *J Wildl Manage* **83**, 1022–1031. (doi:10.1002/JWMG.21688)

159. Shaw AK. 2020 Causes and consequences of individual variation in animal movement. *Mov Ecol* **8**, 1–12. (doi:10.1186/S40462-020-0197-X)

160. Matthews SG, Miller AL, Plötz T, Kyriazakis I. 2017 Automated tracking to measure behavioural changes in pigs for health and welfare monitoring. *Sci Rep* **7**, 1–12. (doi:10.1038/s41598-017-17451-6)

161. Berger-Tal O, Saltz D. 2014 Using the movement patterns of reintroduced animals to improve reintroduction success. *Curr Zool* **60**, 515–526. (doi:10.1093/CZOOLO/60.4.515)

162. Sirovnik J, Voelkl B, Keeling LJ, Würbel H, Toscano MJ. 2021 Breakdown of the ideal free distribution under conditions of severe and low competition. *Behav Ecol Sociobiol* **75**, 1–11. (doi:10.1007/S00265-020-02949-3/FIGURES/5)

163. Becot L, Bedere N, Burlot T, Coton J, Le Roy P. 2021 Nest acceptance, clutch, and oviposition traits are promising selection criteria to improve egg production in cage-free system. *PLoS One* **16**, e0251037. (doi:10.1371/JOURNAL.PONE.0251037)

164. Thompson MJ, Evans JC, Parsons S, Morand-Ferron J. 2018 Urbanization and individual differences in exploration and plasticity. *Behavioral Ecology* **29**, 1415–1425. (doi:10.1093/BEHECO/ARY103)

165. Stamps J, Groothuis TGG. 2010 The development of animal personality: Relevance, concepts and perspectives. *Biological Reviews* **85**, 301–325. (doi:10.1111/J.1469-185X.2009.00103.X/FORMAT/PDF)

166. Salinas-Melgoza A, Salinas-Melgoza V, Wright TF. 2013 Behavioral plasticity of a threatened parrot in human-modified landscapes. *Biol Conserv* **159**, 303–312. (doi:10.1016/J.BIOCON.2012.12.013)

167. Stamps JA, Briffa M, Biro PA. 2012 Unpredictable animals: individual differences in intraindividual variability (IIV). *Anim Behav* **83**, 1325–1334. (doi:10.1016/J.ANBEHAV.2012.02.017)

168. Hertel AG, Royauté R, Zedrosser A, Mueller T. 2021 Biologging reveals individual variation in behavioural predictability in the wild. *Journal of Animal Ecology* **90**, 723–737. (doi:10.1111/1365-2656.13406)

169. Biro PA, Adriaenssens B. 2013 Predictability as a personality trait: Consistent differences in intraindividual behavioral variation. *American Naturalist* **182**, 621–629. (doi:10.1086/673213)

170. Henriksen R, Höglund A, Fogelholm J, Abbey-Lee R, Johnsson M, Dingemanse NJ, Wright D. 2020 Intra-Individual Behavioural Variability: A Trait under Genetic Control. *International Journal of Molecular Sciences 2020, Vol. 21, Page 8069* **21**, 8069. (doi:10.3390/IJMS21218069)

171. Zeltner E, Hirt H. 2010 Effect of artificial structuring on the use of laying hen runs in a free-range system. *http://dx.doi.org/10.1080/00071660310001616264* **44**, 533–537. (doi:10.1080/00071660310001616264)

172. Vandekerchove D, Herdt P, Laevens H, Pasmans F. 2004 Colibacillosis in caged layer hens: characteristics of the disease and the aetiological agent. *Avian Pathology* **33**, 117–125. (doi:10.1080/03079450310001642149)

173. Montalcini CM, Voelkl B, Gómez Y, Gantner M, Toscano MJ. 2022 Evaluation of an Active LF Tracking System and Data Processing Methods for Livestock Precision Farming in the Poultry Sector. *Sensors* **22**, 659. (doi:10.3390/S22020659)

174. Revelle W. 2021 Procedures for Psychological, Psychometric, and Personality Research [R package psych version 2.1.9].

175. Kaiser HF. 1960 The Application of Electronic Computers to Factor Analysis. *Educ Psychol Meas* **20**, 141–151. (doi:10.1177/001316446002000116)

176. Rufener C, Baur S, Stratmann A, Toscano MJ. 2018 A Reliable Method to Assess Keel Bone Fractures in Laying Hens From Radiographs Using a Tagged Visual Analogue Scale. *Front Vet Sci* **5**, 124. (doi:10.3389/fvets.2018.00124)

177. Hertel AG, Niemelä PT, Dingemanse NJ, Mueller T. 2020 A guide for studying among-individual behavioral variation from movement data in the wild. *Mov Ecol* **8**. (doi:10.1186/S40462-020-00216-8)

178. Nakagawa S, Schielzeth H. 2010 Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews* **85**, 935–956. (doi:10.1111/J.1469-185X.2010.00141.X)

179. Cleasby IR, Nakagawa S, Schielzeth H. 2015 Quantifying the predictability of behaviour: statistical approaches for the study of between-individual variation in the within-individual variance. *Methods Ecol Evol* **6**, 27–37. (doi:10.1111/2041-210X.12281)

180. Bürkner P-C. 2017 brms: An R Package for Bayesian Multilevel Models Using Stan. *J Stat Softw* **80**, 1–28. (doi:10.18637/JSS.V080.I01)

181. Vehtari A, Gelman A, Gabry J. 2017 Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput* **27**, 1413–1432. (doi:10.1007/S11222-016-9696-4)

182. Gelman A, Rubin DB. 1992 Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* **7**, 457–472. (doi:10.1214/SS/1177011136)

183. Hadfield JD. 2010 MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *J Stat Softw* **33**, 1–22. (doi:10.18637/JSS.V033.I02)

184. Houslay TM, Wilson AJ. 2017 Avoiding the misuse of BLUP in behavioural ecology. *Behavioral Ecology* **28**, 948–952. (doi:10.1093/BEHECO/ARX023)

185. Hertel AG, Niemelä PT, Dingemanse NJ, Mueller T. 2019 Don't poke the bear: using tracking data to quantify behavioural syndromes in elusive wildlife. *Anim Behav* **147**, 91–104. (doi:10.1016/J.ANBEHAV.2018.11.008)

186. Occhiuto F, Vázquez-Diosdado JA, Carslake C, Kaler J. 2022 Personality and predictability in farmed calves using movement and space-use behaviours quantified by ultra-wideband sensors. *R Soc Open Sci* **9**. (doi:10.1098/RSOS.212019)

187. Moinard C, Statham P, Haskell MJ, McCorquodale C, Jones RB, Green PR. 2004 Accuracy of laying hens in jumping upwards and downwards between perches in different light environments. *Appl Anim Behav Sci* **85**, 77–92. (doi:10.1016/J.APPLANIM.2003.08.008)

188. Cordiner LS, Savory CJ. 2001 Use of perches and nestboxes by laying hens in relation to social status, based on examination of consistency of ranking orders and frequency of interaction. *Appl Anim Behav Sci* **71**, 305–317. (doi:10.1016/S0168-1591(00)00186-6)

189. Coppens CM, De Boer SF, Koolhaas JM. 2010 Coping styles and behavioural flexibility: towards underlying mechanisms. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 4021. (doi:10.1098/RSTB.2010.0217)

190. Koolhaas JM, de Boer SF, Coppens CM, Buwalda B. 2010 Neuroendocrinology of coping styles: Towards understanding the biology of individual variation. *Front Neuroendocrinol* **31**, 307–321. (doi:10.1016/J.YFRNE.2010.04.001)

191. Martin JGA, Pirotta E, Petelle MB, Blumstein DT. 2017 Genetic basis of between-individual and within-individual variance of docility. *J Evol Biol* **30**, 796–805. (doi:10.1111/JEB.13048)

192. Prentice PM, Houslay TM, Martin JGA, Wilson AJ. 2020 Genetic variance for behavioural 'predictability' of stress response. *J Evol Biol* **33**, 642–652. (doi:10.1111/JEB.13601)

193. Siegford JM, Berezowski J, Biswas SK, Daigle CL, Gebhardt-Henrich SG, Hernandez CE, Thurner S, Toscano MJ. 2016 Assessing Activity and Location of Individual Laying Hens in Large Groups Using Modern Technology. *Animals* **6**. (doi:10.3390/ani6020010)

194. Dana L.M. Campbell, Geoff N. Hinch, Jeff A. Downing, Caroline Lee. 2016 Fear and coping styles of outdoor-preferring, moderate-outdoor and indoor-preferring free-range laying hens. *Applied Animal Behaviour* **185**, 73–77.

195. Kruschwitz A, Zupan M, Buchwalder T, Huber-Eicher B. 2008 Nest preference of laying hens (Gallus gallus domesticus) and their motivation to exert themselves to gain nest access. *Appl Anim Behav Sci* **112**, 321–330. (doi:10.1016/J.APPLANIM.2007.08.005)

196. Cooper JJ, Appleby MC. 1995 Nesting behaviour of hens: Effects of experience on motivation. *Appl Anim Behav Sci* **42**, 283–295. (doi:10.1016/0168-1591(94)00543-N)

197. Ringgenberg N, Fröhlich EKF, Harlander-Matauschek A, Toscano MJ, Würbel H, Roth BA. 2015 Nest choice in laying hens: Effects of nest partitions and social status. *Appl Anim Behav Sci* **169**, 43–50. (doi:10.1016/J.APPLANIM.2015.04.013)

198. Olsson IAS, Keeling LJ. 2000 Night-time roosting in laying hens and the effect of thwarting access to perches. *Appl Anim Behav Sci* **68**, 243–256. (doi:10.1016/S0168-1591(00)00097-6)

199. Schrader L, Müller B. 2009 Night-time roosting in the domestic fowl: The height matters. *Appl Anim Behav Sci* **121**, 179–183. (doi:10.1016/J.APPLANIM.2009.09.010)

200. Odén K, Keeling LJ, Algers B. 2010 Behaviour of laying hens in two types of aviary systems on 25 commercial farms in Sweden. *Br Poult Sci* **43**, 169–181. (doi:10.1080/00071660120121364)

201. Nir I, Shefet G, Aaroni Y. 1994 Effect of Particle Size on Performance.: 1. Corn. *Poult Sci* **73**, 45–49. (doi:10.3382/PS.0730045)

202. Schiffman HR. 1968 Texture preference in the domestic chick. *J Comp Physiol Psychol* **66**, 540–541. (doi:10.1037/H0026354)

203. Portella FJ, Caston LJ, Leeson S, Enl Leeson CLJ. 1988 Apparent feed particle size preference by laying hens. *Can J Anim Sci* **68**, 915–922. (doi:10.4141/CJAS88-101)

204. Pawar S, Sajjanar B, … VL-AdvA, 2016 undefined. 2016 Assessing and mitigating the impact of heat stress in poultry. *Adv. Anim. Vet. Sci.* , 332–341. (doi:10.14737/journal.aavs/2016/4.6.332.341)

205. Brooks ME; *et al.* 2017 glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J* **9**, 378–400. (doi:10.3929/ETHZ-B-000240890)

206. Stoffel MA, Nakagawa S, Schielzeth H. 2017 rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods Ecol Evol* **8**, 1639–1644. (doi:10.1111/2041-210X.12797)

207. Niemelä PT, Dingemanse NJ. 2017 Individual versus pseudo-repeatability in behaviour: Lessons from translocation experiments in a wild insect. *Journal of Animal Ecology* **86**, 1033–1043. (doi:10.1111/1365-2656.12688)

208. Geweke JF. 1991 Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Staff Report*

209. Montalcini C, Toscano M, M Petelle. 2022 Scope of consistent inter-individual differences in mean movement behavior within a commercial aviary early on. In *Volume 2 of the Proceedings of the joint 12th International Conference on Methods and Techniques in Behavioral Research and 6th Seminar on Behavioral Methods* (ed JBA-MBGR& AS Andrew Spink),

210. Paul ES, Harding EJ, Mendl M. 2005 Measuring emotional processes in animals: the utility of a cognitive approach. *Neurosci Biobehav Rev* **29**, 469–491. (doi:10.1016/J.NEUBIOREV.2005.01.002)

211. Gray H, Davies R, Bright A, Rayner A, Asher L. 2020 Why Do Hens Pile? Hypothesizing the Causes and Consequences. *Front Vet Sci* **7**, 616836. (doi:10.3389/fvets.2020.616836)

212. Bécot L, Bédère N, Coton J, Burlot T, Le Roy P. 2023 Nest preference and laying duration traits to select against floor eggs in laying hens. *Genetics Selection Evolution* **55**, 1–10. (doi:10.1186/S12711-023-00780-8/FIGURES/3)

213. Montalcini CM, Toscano MJ, Gebhardt-Henrich SG, Petelle MB. 2023 Intra-individual variation of hen movements is associated with later keel bone fractures in a quasi-commercial aviary. *Sci Rep* **13**, 2377. (doi:10.1038/S41598-023-29587-9)

214. Pagel M, Dawkins MS. 1997 Peck orders and group size in laying hens: `futures contracts' for non-aggression. *Behavioural Processes* **40**, 13–25. (doi:10.1016/S0376-6357(96)00761-9)

215. Koolhaas JM. 2008 Coping style and immunity in animals: Making sense of individual variation. *Brain Behav Immun* **22**, 662–667. (doi:10.1016/J.BBI.2007.11.006)

216. Careau V, Thomas D, Humphries MM, Réale D. 2008 Energy metabolism and animal personality. *Oikos* **117**, 641–653. (doi:10.1111/J.0030-1299.2008.16513.X)

217. Hansen I. 2007 Behavioural expression of laying hens in aviaries and cages: Frequencies, time budgets and facility utilisation. *http://dx.doi.org/10.1080/00071669408417715* **35**, 491–508. (doi:10.1080/00071669408417715)

218. Sisson S, Grossman JD, Getty R. 1975 Sisson and Grossman's The anatomy of the domestic animals. *5th ed. Saunders* , 2095. (doi:10.3/JQUERY-UI.JS)

219. Claessens LPAM. 2009 The skeletal kinematics of lung ventilation in three basal bird taxa (emu, tinamou, and guinea fowl). *J Exp Zool A Ecol Genet Physiol* **311A**, 586–599. (doi:10.1002/JEZ.501)

220. Duncker H-R. 2000 The respiratory apparatus of birds and their locomotory and me- tabolic efficiency. *Journal für Ornithologie* **141**, 1–67. (doi:10.1007/BF01651772)

221. Thøfner I, Hougen HP, Villa C, Lynnerup N, Christensen JP. 2020 Pathological characterization of keel bone fractures in laying hens does not support external trauma as the underlying cause. *PLoS One* **15**, e0229735. (doi:10.1371/JOURNAL.PONE.0229735)

222. Casey-Trott T, Heerkens JLT, Petrik M, Regmi P, Schrader L, Toscano MJ, Widowski T. 2015 Methods for assessment of keel bone damage in poultry. *Poult Sci* **94**, 2339–2350. (doi:10.3382/PS/PEV223)

223. Driver CC, Voelkle MC. 2018 Hierarchical bayesian continuous time dynamic modeling. *Psychol Methods* **23**, 774–799. (doi:10.1037/MET0000168)

224. Driver CC, Oud JHL, Voelkle MC. 2017 Continuous Time Structural Equation Modeling with R Package ctsem. *J Stat Softw* **77**, 1–35. (doi:10.18637/JSS.V077.I05)

225. Toscano M, Booth F, Richards G, Brown S, Karcher D, Tarlton J. 2018 Modeling collisions in laying hens as a tool to identify causative factors for keel bone fractures and means to reduce their occurrence and severity. *PLoS One* **13**, e0200025. (doi:10.1371/JOURNAL.PONE.0200025)

226. Driver CC, Tomasik MJ. 2022 Formalizing Developmental Phenomena as Continuous-Time Systems: Relations Between Mathematics and Language Development.

227. Driver CC, Voelkle MC. 2021 Hierarchical continuous time modeling. *The Handbook of Personality Dynamics and Processes* , 887–908. (doi:10.1016/B978-0-12-813995-0.00034-0)

228. Montalcini CM, Petelle MB, Toscano MJ. 2023 Commercial laying hens exhibit long-term consistent individual differences and behavioural syndromes in spatial traits. *R Soc Open Sci* **10**. (doi:10.1098/RSOS.230043)

229. Driver CC, Voelkle MC. In press. Introduction to Hierarchical Continuous Time Dynamic Modelling With ctsem.

230. Campbell DLM, Goodwin SL, Makagon MM, Swanson JC, Siegford JM. 2016 Failed landings after laying hen flight in a commercial aviary over two flock cycles. *Poult Sci* **95**, 188–197. (doi:10.3382/PS/PEV270)

231. Hedlund L, Whittle R, Jensen P. 2019 Effects of commercial hatchery processing on short- and long-term stress responses in laying hens. *Sci Rep* **9**, 1–10. (doi:10.1038/s41598-019-38817-y)

232. Grafen A. 1988 On the uses of data on lifetime reproductive success. In *Reproductive success* (ed T. H. Clutton-Brock), pp. 454–471.

233. De Jong IC, van Hattum T, van Riel JW, De Baere K, Kempen I, Cardinaels S, Gunnink H. 2020 Effects of on-farm and traditional hatching on welfare, health, and performance of broiler chickens. *Poult Sci* **99**, 4662–4671. (doi:10.1016/J.PSJ.2020.06.052)

234. De Jong IC, Gunnink H, Van Hattum T, Van Riel JW, Raaijmakers MMP, Zoet ES, Van Den Brand H. 2019 Comparison of performance, health and welfare aspects between commercially housed hatchery-hatched and on-farm hatched broiler flocks. *Animal* **13**, 1269–1277. (doi:10.1017/S1751731118002872)

235. Giersberg MF, Molenaar R, de Jong IC, Souza da Silva C, van den Brand H, Kemp B, Rodenburg TB. 2021 Effects of hatching system on the welfare of broiler chickens in early and later life. *Poult Sci* **100**, 100946. (doi:10.1016/J.PSJ.2020.12.043)

236. Hedlund L, Jensen P. 2022 Effects of stress during commercial hatching on growth, egg production and feather pecking in laying hens. *PLoS One* **17**. (doi:10.1371/JOURNAL.PONE.0262307)

237. Witjes VL, Bruckmaier RM, Gebhardt-Henrich SG, Toscano MJ. 2022 Effects of on-farm hatching on short term stress indicators, weight gain, and cognitive ability in layer chicks. *Appl Anim Behav Sci* , 105692. (doi:10.1016/J.APPLANIM.2022.105692)

238. Gluckman PD, Hanson MA, Spencer HG, Bateson P. 2005 Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. *Proceedings of the Royal Society B: Biological Sciences* **272**, 671–677. (doi:10.1098/RSPB.2004.3001)

239. Monaghan P. 2007 Early growth conditions, phenotypic development and environmental change. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 1635–1645. (doi:10.1098/RSTB.2007.0011)

240. Pigeon G *et al.* 2019 Silver spoon effects are constrained under extreme adult environmental conditions. *Ecology* **100**, e02886. (doi:10.1002/ECY.2886)

241. Giersberg MF, Spindler B, Kemper N. 2017 Assessment of Plumage and Integument Condition in Dual-Purpose Breeds and Conventional Layers. *Animals 2017, Vol. 7, Page 97* **7**, 97. (doi:10.3390/ANI7120097)

242. Drake KA, Donnelly ProfCA, Dawkins MS. 2010 Influence of rearing and lay risk factors on propensity for feather damage in laying hens. *Br Poult Sci* **51**, 725–733. (doi:10.1080/00071668.2010.528751)

243. Gunnarsson S, Matthews LR, Foster TM, Temple W. 2000 The demand for straw and feathers as litter substrates by laying hens. *Appl Anim Behav Sci* **65**, 321–330. (doi:10.1016/S0168-1591(99)00065-9)

244. Hemsworth PH, Edwards LE, Hemsworth PH, Edwards LE. 2020 Natural behaviours, their drivers and their implications for laying hen welfare. *Anim Prod Sci* **61**, 915–930. (doi:10.1071/AN19630)

245. Gelman A. 2008 Scaling regression inputs by dividing by two standard deviations. *Stat Med* **27**, 2865–2873. (doi:10.1002/SIM.3107)

246. Davison A, Hinkley D. 1997 *Bootstrap methods and their application*. Cambridge University Press. See https://books.google.com/books?hl=en&lr=&id=4aCDbm_t8jUC&oi=fnd&pg=PR7&dq=Bootstrap+methods+and+their+application+(Vol.+1).+Cambridge+university+press.&ots=mZ4YjbV-_4&sig=XYnomIvGBIxRHuEDDsN88xcuNjw.

247. Therneau T. 2021 A Package for Survival Analysis in R.

248. Terry M. Therneau. 2020 coxme: Mixed Effects Cox Models.

249. Goel MK, Khanna P, Kishore J. 2010 Understanding survival analysis: Kaplan-Meier estimate. *Int J Ayurveda Res* **1**, 274. (doi:10.4103/0974-7788.76794)

250. Therneau T, Crowson C, Atkinson E. 2023 Using Time Dependent Covariates and Time Dependent Coecients in the Cox Model.

251. Caroli A, Frisoni GB, Initiative TADN. 2010 The dynamics of Alzheimer's disease biomarkers in the Alzheimer's Disease Neuroimaging Initiative cohort. *Neurobiol Aging* **31**, 1263. (doi:10.1016/J.NEUROBIOLAGING.2010.04.024)

252. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2021 nlme: Linear and Nonlinear Mixed Effects Models.

253. Mitchell MA. 2009 Chick transport and welfare. *Avian Biol Res* **2**, 99–105. (doi:10.3184/175815509X431894/ASSET/175815509X431894.FP.PNG_V03)

254. Fleming RH, Mccormack HA, Mcteir L, Whitehead CC. 2007 Relationships between genetic, environmental and nutritional factors influencing osteoporosis in laying hens. *Br Poult Sci* **47**, 742–755. (doi:10.1080/00071660601077949)

255. Randall CJ, Duff SR. 1988 Avulsion of the patellar ligament in osteopenic laying fowl. *Vet Rec* **123**, 439–441. (doi:10.1136/VR.123.17.439)

256. Fleming RH, McCormack HA, McTeir L, Whitehead CC. 2004 Incidence, pathology and prevention of keel bone deformities in the laying hen. *Br Poult Sci* **45**, 320–330. (doi:10.1080/00071660410001730815)

257. Gebhardt-Henrich SG, Fröhlich EKF. 2015 Early Onset of Laying and Bumblefoot Favor Keel Bone Fractures. *Animals* **5**, 1192–1206. (doi:10.3390/ANI5040406)

258. Yahav S, Hurwitz S. 1996 Induction of Thermotolerance in Male Broiler Chickens by Temperature Conditioning at an Early Age. *Poult Sci* **75**, 402–406. (doi:10.3382/PS.0750402)

259. Scanes CG. 2016 Biology of stress in poultry with emphasis on glucocorticoids and the heterophil to lymphocyte ratio. *Poult Sci* **95**, 2208–2215. (doi:10.3382/PS/PEW137)

260. Gabrielle L, Rebecca O, Louise H, Johanna G, Per J. 2022 Play ontogeny in young chickens is affected by domestication and early stress. *Sci Rep* **12**, 1–8. (doi:10.1038/s41598-022-17617-x)

261. Hedlund L, Palazon T, Jensen P. 2021 Stress during commercial hatchery processing induces long-time negative cognitive judgement bias in chickens. *Animals* **11**, 1083. (doi:10.3390/ANI11041083/S1)

262. Giersberg MF, Molenaar R, de Jong IC, De Baere K, Kemp B, Brand H van den, Rodenburg TB. 2023 Group level and individual activity of broiler chickens hatched in 3 different systems. *Poult Sci* , 102706. (doi:10.1016/J.PSJ.2023.102706)

263. Calandreau L, Bertin A, Boissy A, Arnould C, Constantin P, Desmedt A, Guémené D, Nowak R, Leterrier C. 2011 Effect of one week of stress on emotional reactivity and learning and memory performances in Japanese quail. *Behavioural Brain Research* **217**, 104–110. (doi:10.1016/J.BBR.2010.10.004)

264. Wagner K V., Wang XD, Liebl C, Scharf SH, Müller MB, Schmidt M V. 2011 Pituitary glucocorticoid receptor deletion reduces vulnerability to chronic stress. *Psychoneuroendocrinology* **36**, 579–587. (doi:10.1016/J.PSYNEUEN.2010.09.007)

265. Oitzl MS, Reichardt HM, Joëls M, De Kloet ER. 2001 Point mutation in the mouse glucocorticoid receptor preventing DNA binding impairs spatial memory. *Proceedings of the National Academy of Sciences* **98**, 12790–12795. (doi:10.1073/PNAS.231313998)

266. Joëls M, Pu Z, Wiegert O, Oitzl MS, Krugers HJ. 2006 Learning under stress: how does it work? *Trends Cogn Sci* **10**, 152–158. (doi:10.1016/J.TICS.2006.02.002)

267. Skånberg L, Newberry RC, Estevez I, Keeling LJ. 2023 Environmental change or choice during early rearing improves behavioural adaptability in laying hen chicks. *Sci Rep* **13**, 1–13. (doi:10.1038/s41598-023-33212-0)

268. Van Bodegom M, Homberg JR, Henckens MJAG. 2017 Modulation of the hypothalamic-pituitary-adrenal axis by early life stress exposure. *Front Cell Neurosci* **11**, 87. (doi:10.3389/FNCEL.2017.00087/BIBTEX)

269. Stoneham ET, McHail DG, Samipour-Biel S, Liehr N, Lee CM, Evans JC, Boggs K, Dumas TC. 2021 Spatial Learning Is Impaired in Male Pubertal Rats Following Neonatal Daily but Not Randomly Spaced Maternal Deprivation. *Front Cell Dev Biol* **9**. (doi:10.3389/FCELL.2021.621308)

270. De Kloet ER, Joëls M, Holsboer F. 2005 Stress and the brain: from adaptation to disease. *Nature Reviews Neuroscience 2005 6:6* **6**, 463–475. (doi:10.1038/nrn1683)

271. Ladd CO, Huot RL, Thrivikraman K V., Nemeroff CB, Plotsky PM. 2004 Long-term adaptations in glucocorticoid receptor and mineralocorticoid receptor mrna and negative feedback on the hypothalamo-pituitary-adrenal axis following neonatal maternal separation. *Biol Psychiatry* **55**, 367–375. (doi:10.1016/J.BIOPSYCH.2003.10.007)

272. Maniam J, Antoniadis C, Morris MJ. 2014 Early-life stress, HPA axis adaptation, and mechanisms contributing to later health outcomes. *Front Endocrinol (Lausanne)* **5**, 73. (doi:10.3389/FENDO.2014.00073/BIBTEX)

273. Cooper JJ, Albentosa MJ. 2003 Behavioural priorities of laying hens. *Avian and Poultry Biology Reviews* **14**, 127–149.

274. Krautwald-Junghanns ME *et al.* 2018 Current approaches to avoid the culling of day-old male chicks in the layer industry, with special reference to spectroscopic methods. *Poult Sci* **97**, 749–757. (doi:10.3382/PS/PEX389)

275. De Jong IC, Van Riel J, Bracke MBM, Brand H van den. 2017 A 'meta-analysis' of effects of post-hatch food and water deprivation on development, performance and welfare of chickens. *PLoS One* **12**, e0189350. (doi:10.1371/JOURNAL.PONE.0189350)

276. Nordquist RE, Vernooij JCM, Dull CL, Pascual A, van der Linde G, Goerlich VC. 2022 The effects of transport of 18-day old hatching eggs on physiology and behaviour of slow growing broiler chicken. *Appl Anim Behav Sci* **257**, 105789. (doi:10.1016/J.APPLANIM.2022.105789)

277. van der Sluis M, Haas Y, Klerk B, Rodenburg TB, Ellen ED. 2020 Assessing the Activity of Individual Group-Housed Broilers Throughout Life using a Passive Radio Frequency Identification System-A Validation Study. *Sensors* **20**. (doi:10.3390/s20133612)

278. Gebhardt-Henrich SG, Kashev A, Petelle MB, Toscano MJ. 2023 Validation of a Radio frequency identification system for tracking location of laying hens in a commercial aviary. *bioRxiv* , 2023.02.16.528820. (doi:10.1101/2023.02.16.528820)

279. Gharat V, Colin E, Baudoin G, Richard D. 2017 Indoor performance analysis of LF-RFID based positioning system: Comparison with UHF-RFID and UWB. *2017 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2017* **2017-January**, 1–8. (doi:10.1109/IPIN.2017.8115901)

280. Malt BC, Smith EE. 1984 Correlated properties in natural categories. *J Verbal Learning Verbal Behav* **23**, 250–269. (doi:10.1016/S0022-5371(84)90170-1)

281. Van Overveld T, Matthysen E. 2010 Personality predicts spatial responses to food manipulations in free-ranging great tits (Parus major). *Biol Lett* **6**, 187–190. (doi:10.1098/RSBL.2009.0764)

282. Groothuis TGG, Carere C. 2005 Avian personalities: characterization and epigenesis. *Neurosci Biobehav Rev* **29**, 137–150. (doi:10.1016/J.NEUBIOREV.2004.06.010)

283. David M, Auclair Y, Cézilly F. 2011 Personality predicts social dominance in female zebra finches, Taeniopygia guttata, in a feeding context. *Anim Behav* **81**, 219–224. (doi:10.1016/J.ANBEHAV.2010.10.008)

284. Krause ET, Krüger O, Schielzeth H. 2017 Long-term effects of early nutrition and environmental matching on developmental and personality traits in zebra finches. *Anim Behav* **128**, 103–115. (doi:10.1016/J.ANBEHAV.2017.04.003)

285. Schuett W, Dall SRX, Royle NJ. 2011 Pairs of zebra finches with similar ‘personalities’ make better parents. *Anim Behav* **81**, 609–618. (doi:10.1016/J.ANBEHAV.2010.12.006)

286. Kralj-Fišer S, Weiß BM, Kotrschal K. 2010 Behavioural and physiological correlates of personality in greylag geese (Anser anser). *J Ethol* **28**, 363–370. (doi:10.1007/S10164-009-0197-1/TABLES/4)

287. Favati A, Leimar O, Løvlie H. 2014 Personality Predicts Social Dominance in Male Domestic Fowl. *PLoS One* **9**, e103535. (doi:10.1371/JOURNAL.PONE.0103535)

288. Overmier JB, Patterson J, Wielkiewicz RM. 1980 Environmental Contingencies as Sources of Stress in Animals. *Coping and Health* , 1–38. (doi:10.1007/978-1-4684-1042-6_1)

289. Sambrook TD, Buchanan-Smith HM. 1997 Control and Complexity in Novel Object Enrichment. *Animal Welfare* **6**, 207–216. (doi:10.1017/S0962728600019801)

290. Wiepkema PR, Koolhaas JM. 1993 Stress and Animal Welfare. *Animal Welfare* **2**, 195–218. (doi:10.1017/S0962728600015876)

291. Lee C, Colditz IG, Campbell DLM. 2018 A framework to assess the impact of new animal management technologies on welfare: A case study of virtual fencing. *Front Vet Sci* **5**, 400604. (doi:10.3389/FVETS.2018.00187/BIBTEX)

292. Bassett L, Buchanan-Smith HM. 2007 Effects of predictability on the welfare of captive animals. *Appl Anim Behav Sci* **102**, 223–245. (doi:10.1016/J.APPLANIM.2006.05.029)

293. Grover A, Leskovec J. In press. node2vec: Scalable Feature Learning for Networks. (doi:10.1145/2939672.2939754)

294. Webber QMR, Vander Wal E. 2018 An evolutionary framework outlining the integration of individual social and spatial ecology. *Journal of Animal Ecology* **87**, 113–127. (doi:10.1111/1365-2656.12773)

295. Nicol J. 2004 Development, direction, and damage limitation: Social learning in domestic fowl. *Anim Learn Behav*

296. Zeltner E, Klein T, Huber-Eicher B. 2000 Is there social transmission of feather pecking in groups of laying hen chicks? *Anim Behav* **60**, 211–216. (doi:10.1006/ANBE.2000.1453)

297. Cloutier S, Newberry RC, Honda K, Alldredge JR. 2002 Cannibalistic behaviour spread by social learning. *Anim Behav* **63**, 1153–1162. (doi:10.1006/ANBE.2002.3017)

298. Gómez Y, Berezowski J, Jorge YA, Gebhardt-Henrich SG, Vögeli S, Stratmann A, Toscano MJ, Voelkl B. 2022 Similarity in Temporal Movement Patterns in Laying Hens Increases with Time and Social Association. *Animals* **12**, 555. (doi:10.3390/ANI12050555)

299. Jollife IT, Cadima J. 2016 Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374**. (doi:10.1098/RSTA.2015.0202)

300. Björklund M, Björklund B. 2019 Be careful with your principal components. *Evolution (N Y)* **73**, 2151–2158. (doi:10.1111/EVO.13835)

301. Hoogenboom I, Daan S, Dallinga JH, Schoenmakers M. 1984 Seasonal Change in the Daily Timing of Behaviour of the Common Vole, Microtus arvalis. *Oecologia* **61**, 18–31.

302. Spoelstra K, Wikelski M, Daan S, Loudon ASI, Hau M. 2016 Natural selection against a circadian clock gene mutation in mice. *Proc Natl Acad Sci U S A* **113**, 686–691. (doi:10.1073/PNAS.1516442113/-/DCSUPPLEMENTAL/PNAS.1516442113.SAPP.PDF)

Acknowledgements

First and foremost, Matt, I'm profoundly grateful for everything you've done as my direct supervisor. Your attentive listening, stimulating discussions, and moral support were essential for the success of this journey. You not only have been an invaluable source of support, but you also helped me grow as a young scientist and researcher, thank you!

I would like to express my sincere gratitude to my supervisor Mike for providing me with the opportunity to take my first steps in science and for introducing me to the field of animal welfare. Your passion and dedication to your work has been both a great source of motivation and inspiration.

Also, I want to thank my office colleagues, Masha, Laura, Klara, Kathy, Alex, and Daniel, for the pleasant lunches and interesting office chats, with a special thanks to Alex for being an exceptional partner in pranks!

I would also like to thank the entire ZTHZ group, especially Ari, Yamenah and Sabine for the helpful comments and discussion on the thesis, and the consistent positive energy you all brought to the lab. Sabine, thank you for always bringing your dogs to the office. Each of you has contributed to making these four years even more enjoyable.

Furthermore, I want to extend my appreciation to every person that helped collecting data, including all animals, as well as Sabrina Ruchti for taking the radiographs and Aviforum staff for collecting egg data and for the animal care. A big thanks to Abdulsatar Abdel Rahman for his technical support related to the tracking system, which required a considerable amount of creativity and patience.

Finally, I am truly thankful for my mother's constant support and to my partner for being a great source of academic inspiration.

Supplementary materials

Chapter I

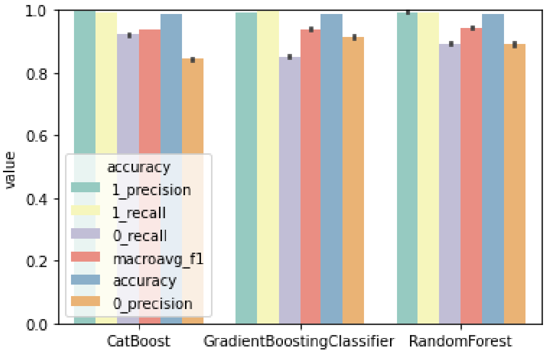


Figure S1. Precision per class (0: FR; 1: CR), recall per class (0: FR; 1: CR) and accuracy of the three classifiers over 100 random seeds.

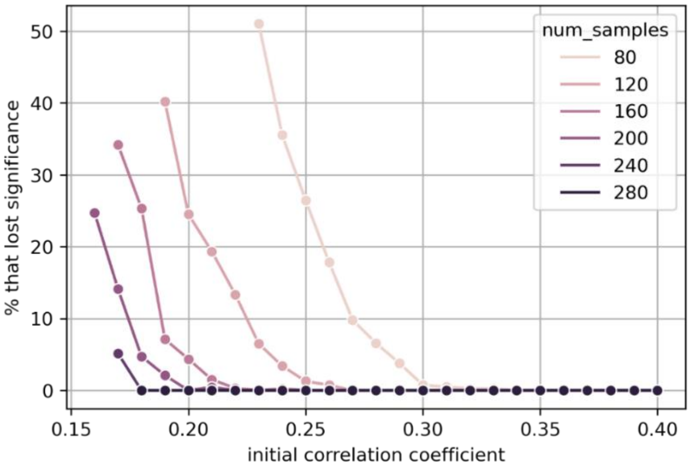


Figure S2. Percentage of simulations that lost significance (p>0.05) of its associated initial effect size (measure by pearson correlation between two simulated samples from a normal distribution: M' and H) after a change in percentage of the true variance that is recovered by M' from 0.99 to 0.91, depending on the initial effect size (varying from 0.16 to 0.4) and sample size (varying from 80 to 280).

Text S1 - Pseudo-code of the tag-algorithm

```
START every time a tag receive an LF-signal (=markers signal)  (it can be up to every half seconds in average i.e. 2sec/4markers)

List = all received signals sent from any markers of past 10 sec of the associated tracking system
previous-zone-logged = last logged transition
```

```
previous-dB-zone = dBs of last logged zone

current-dB-of-highest-zone = maximum dB entry in List
current-highest-zone = zone with maximum dB entry in List (find thanks to the modulated signal)

If    current-dB-of-highest-zone is not the unique maximum of the List:
        Don't register a new transition and repeat the procedure

current-dB-of-second-highest-zone = second highest dB entry in List
current-second-highest-zone = zone with second highest dB entry in List

If    current-highest-zone ≠ current-second-highest-zone:
        Don't register a new transition and repeat the procedure

If    previous-zone-logged ≠ current-highest-zone:
        Register a new transition to current-highest-zone and repeat the procedure

END
```

Text S2 - Tree-based classifiers

In this study we used three decision-tree based classifiers. A decision tree is a non-parametric supervised learning method that performs recursive partition of the instance space [135]. Typically, a decision tree is an acyclic directed graph and has one node with no incoming edges (root node) as well as nodes with one incoming edge. All nodes with at least one outcoming edge are called internal nodes and all nodes with no outcoming edge are called decision nodes. During training, each internal node aims to split the instance space while optimizing the classifier's performance. The Gradient boosting classifier is a 'greedy' algorithm that sequentially trains a shallow decision tree in order to correct the errors of the previously trained tree [137]. The Catboost method is a recently developed gradient boosting algorithm [138,139] that we selected in this study for its ability to process categorical features during training. More specifically, it substitutes each categorical feature with a numerical feature by using an ordered target encoding method. Target encoding commonly replace the category x_k^i of the k^{th} training example of the i^{th} categorical feature, with the estimated expected target value *y* conditioned by this category: $\mathbb{E}(y|x^i = x_k^i)$. Target encoding is known to suffer from target leakage and CatBoost tries to overcome this issue. For that purpose, it adds an artificial timeline to the training dataset by permutating the set of observation in a random order and computes each encoding value based on its own artificial history only [139]. As samples with a shorter history will have a target encoding value with higher variance, CatBoost used several random permutations during training.

Chapter II

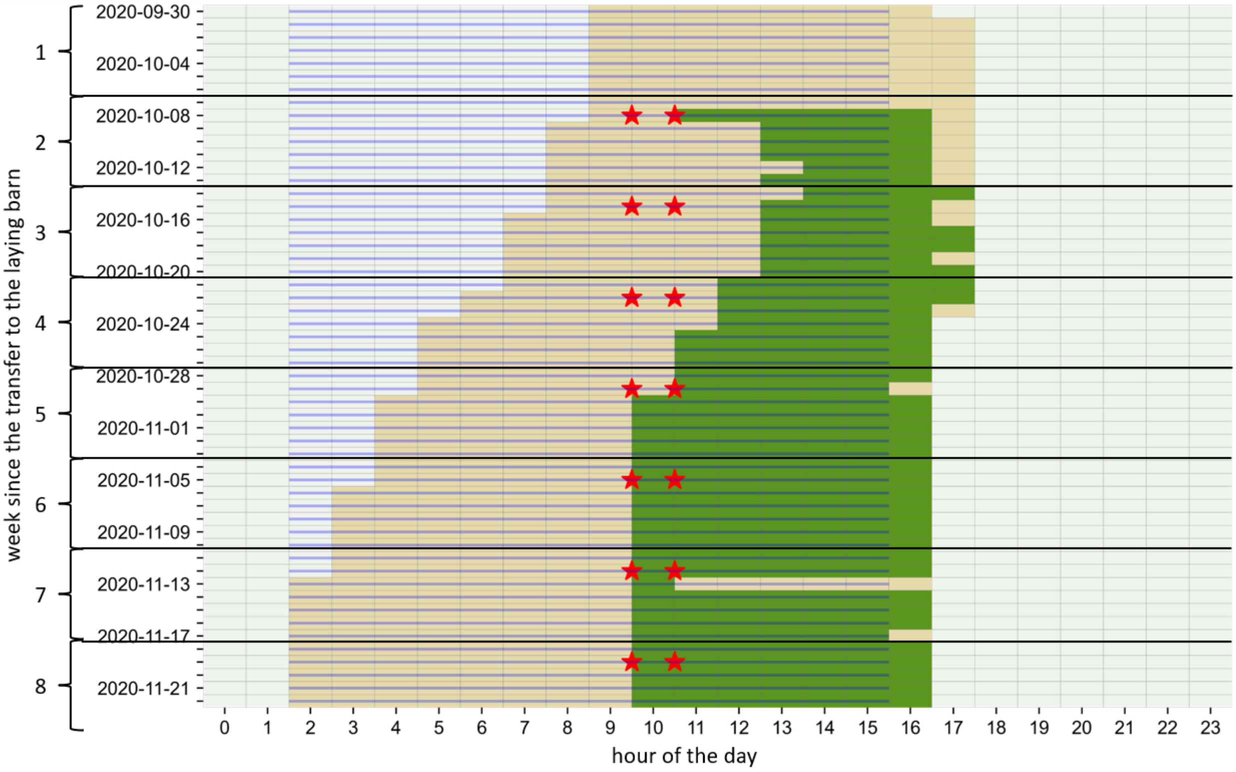


Figure S2 – Daily barn schedule where beige rectangles indicate period where the artificial light was on, red stars indicate when the manure belt was activated, and the green rectangles indicate period where the winter garden was accessible.

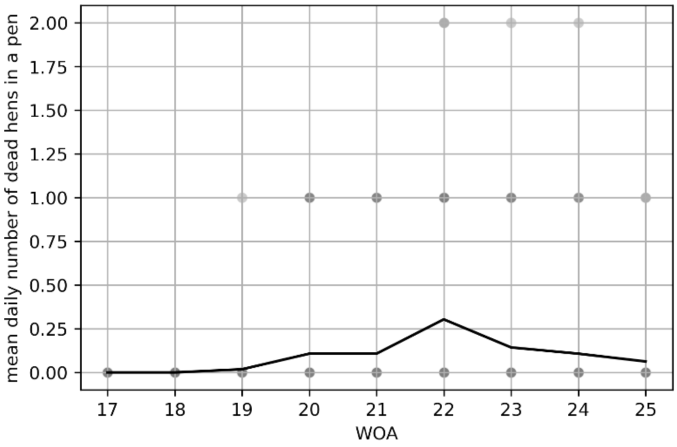


Figure S3 – The black line represents the weekly average of the daily number of dead hens across all pens and days, including a total of eight pens (each with 225 hens) and the first 54 days after the transfer to the laying barn (where WOA = week of age). Each point is a specific value for a pen on a specific day. The total number of early deaths was 46 (5 deaths/pen, except for three pens which had 6, 9 and 6 deaths respectively).

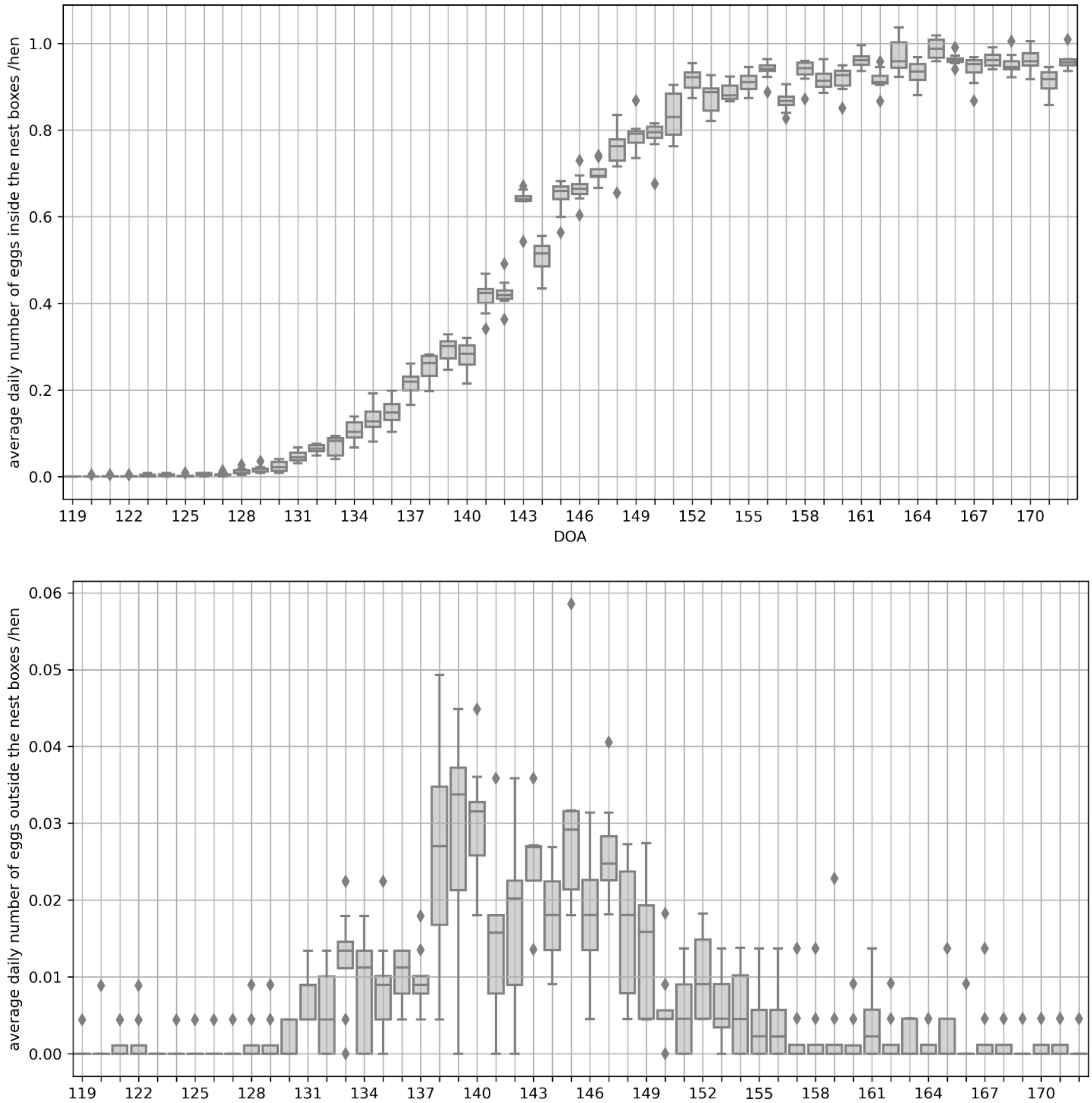


Figure S4 – Pen-level production data (for a total of eight pens), in terms of the average daily number of eggs laid (inside and outside the nest boxes) per hen over the first 54 days after the transfer to the laying barn.

Text S1 - More/less explorer chicks' class

Focal animals were part of a larger study, which classified chicks as more/less explorative (MEXP/LEXP). When chicks were one day of age, a raised platform with blue natural dye was added in each pen, so that chicks going onto the raised platform would have their feet painted in blue. We added blue dye diluted in water approximately every two hours to prevent the dye from drying out. We selected focal animals based on their feet color (blue color: MEXP, no blue color: LEXP) at 7 days of age (this corresponded to when at least 30 animals from each pen had walked onto the raised platform. In each pen, 20 focal animals were selected randomly so that half would belong to the MEXP class and half to the LEXP class (depending on the number of chicks from each class available, the number of animals per class could vary between 6-14). In addition, 20 focal animals were selected as a representative sample of the population (hereafter called OTHERS) (balanced randomization approach aiming at a same proportion of MEXP as in the overall population of that pen). In parallel to the selection process, each selected chick was attributed to a laying pen identity with a balanced randomization approach aiming at a uniform representation in each laying pen of the chicks' categories (MEXP, LEXP, OTHERS) and rearing pen identity.

Text S2 - Principal component analysis

Because not all hens and all weeks have the same amount of observation, to ensure a same weight across weeks and individuals while accounting for the variation in movements across time and individuals, we included in the PCA each first observation of a week only, for each week and each individual. The first three PCA's principal components had an eigenvalue > 1 [175] and explained 41%, 20% and resp. 14% of the total variation. We calculated loadings for each variable on a given principal component by using the correlations between each original variable and the components [300]. Among the variable loadings on the first principal component nine had absolute value > 0.4, and six had absolute value > 0.7. Principal component one was characterized by the percentage of duration in the top tier, which loaded strongly in opposite direction as the hourly realized travel distance and as the number of stay in both litter and lower perch. These loadings reflect general movement throughout the barn, with higher score associated with animals that spent more time in the floor, lower tier but also transitioned more between these zones. In contrast, PC2 had the greatest positive loadings on the number of stay in the nest box tier, which could reflect egg-laying behaviour, while PC3 was associated solely with the three WG variables (with loadings > 0.6, all other variables <0.4). This further enhances the meaning and validity of the PCA results[301]. We further evaluated the validity of the PCA by comparing the loadings of the first three principal components (detailed in Table S1 a) with the principal components generated by a PCA based on the full observations set (detailed in Table S1 b).

Equation S1 - Bivariate mixed models:

The linear model for the latent variables of the Bayesian bivariate model is defined as: $X\beta + Zu + e$, with X and Z two design matrices relating the fixed and random predictions, respectively, to the data, while β, u and e (residual) are parameters to estimate. Latter parameters are assumed to follow a multivariate normal distribution:

$$\begin{pmatrix}\beta\\ u\\ e\end{pmatrix}\sim N\left(\begin{bmatrix}\beta_0\\ 0\\ 0\end{bmatrix},\begin{bmatrix}B & 0 & 0\\ 0 & G & 0\\ 0 & 0 & R\end{bmatrix}\right) \tag{S1}$$

where the zero off-diagonal matrices imply *a priori* independence between fixed effects, random effects as well as residuals; β_0 and B are the fixed effects' prior means and respectively (co)variances; G the expected variances (V) and covariances (Cov) of the random effects:

$$G_{HenID} = \begin{pmatrix} V_{KBF\ Intercept(Int)} & Cov_{KBF\ Int,\ \ PC1\ Int} & Cov_{KBF\ Int,\ \ PC1:time} & Cov_{KBF\ Int,\ \ PC1:time^2} \\ * & V_{PC1\ Int} & Cov_{PC1\ Int,\ \ PC1:time} & Cov_{PC1\ Int,\ \ PC1:time^2} \\ * & * & V_{PC1:time} & Cov_{PC1:time,\ \ PC1:time^2} \\ * & * & * & V_{PC1:time^2} \end{pmatrix},$$

$$G_{PenID} = V_{KBF\ Intercept(Int)}$$

and R denote the expected variances of the residuals (as health scores do not have repeated measures at the individual-level we do not allow the variances to covary, and we further constrain the residual variance of health scores to be very close to zero):

$$R = \begin{pmatrix} V_{PC1\ residual} & 0 \\ 0 & V_{KBF\ residual} \end{pmatrix}.$$

Equation S2 – S3 - Bivariate random effects:

Random effects and residual output from the Bivariate (movement, KBF severity) model:

$$\begin{pmatrix}\beta\\ u\\ e\end{pmatrix}\sim N\left(\begin{bmatrix}\beta_0\\ 0\\ 0\end{bmatrix},\begin{bmatrix}B & 0 & 0\\ 0 & G & 0\\ 0 & 0 & R\end{bmatrix}\right) \tag{S2}$$

$$G_{HenID} = \begin{pmatrix} 213.518\ [150.301, 285.914] & 0.361\ [-3.818, 4.377] & 2.491\ [-0.908, 5.917] & -0.867\ [-1.734, -0.030] \\ * & 1.231\ [0.842, 1.640] & -0.659\ [-0.961, -0.381] & 0.092\ [0.031, 0.158] \\ * & * & 0.794\ [0.532, 1.091] & -0.171\ [-0.240, -0.030] \\ * & * & * & 0.047\ [0.031, 0.064] \end{pmatrix}$$

$$G_{PenID} = 2.531\ [0.038, 10.16]$$

$$R = \begin{pmatrix} 0.259\ [0.247, 0.273] & 0 \\ 0 & 0.001\ [0.001, 0.001] \end{pmatrix}$$

Random effects and residual output from the Bivariate (movement, feather damage) model:

$$\begin{pmatrix}\beta\\ u\\ e\end{pmatrix}\sim N\left(\begin{bmatrix}\beta_0\\ 0\\ 0\end{bmatrix},\begin{bmatrix}B & 0 & 0\\ 0 & G & 0\\ 0 & 0 & R\end{bmatrix}\right) \tag{S3}$$

$$G_{HenID} = \begin{pmatrix} 103.808\ [69.970, 143.350] & -0.114\ [-3.152, 2.849] & -1.423\ [-3.946, 1.154] & 0.453\ [-0.163, 1.069] \\ * & 1.231\ [0.852, 1.653] & -0.660\ [-0.965, -0.389] & 0.093\ [0.030, 0.162] \\ * & * & 0.797\ [0.540, 1.105] & -0.172\ [-0.242, -0.112] \\ * & * & * & 0.047\ [0.031, 0.065] \end{pmatrix}$$

$$G_{PenID} = 30.59\ [0.105, 86.08]$$

$$R = \begin{pmatrix} 0.259\ [0.250, 0.272] & 0 \\ 0 & 0.001\ [0.001, 0.001] \end{pmatrix}$$

	PC1	PC2	PC3	PC1	PC2	PC3
daily percentage of duration in the top tier	**-0.90**	0.08	0.29	**-0.91**	0.13	0.22
daily percentage of duration in the nestbox tier	0.16	**0.67**	-0.02	0.12	**0.69**	-0.01
daily percentage of duration in the lower tier	**0.73**	-0.17	-0.36	**0.74**	-0.23	-0.30
daily percentage of duration in the litter floor	**0.75**	-0.25	-0.36	**0.76**	-0.25	-0.34
number of stays in the top tier /h	0.10	**0.81**	0.32	0.04	**0.85**	0.28
number of stays in the nestbox tier /h	0.33	**0.88**	0.19	0.27	**0.90**	0.13
number of stays in the lower tier /h	**0.92**	0.20	-0.16	**0.91**	0.22	-0.18
number of stays in the litter floor /h	**0.93**	-0.20	-0.02	**0.92**	-0.18	0.01
sleeping height	-0.31	0.11	0.30	-0.35	0.09	0.25
vertical travelled distance /h	**0.83**	**0.48**	-0.02	**0.81**	**0.51**	-0.04
daily percentage of duration in the winter garden	**0.48**	-0.32	**0.70**	**0.40**	-0.25	**0.77**
Number of stays in the winter garden /h	**0.57**	-0.32	**0.65**	**0.52**	-0.25	**0.69**
went in the winter garden (yes/no)	**0.46**	-0.27	**0.56**	0.39	-0.21	**0.63**

Table S1 – Summary of the principal component analysis (PCA) based on: one value per week per hen for the first three columns, and the full set of observations for validation purposes, for the last three columns. Loadings with an absolute value over 0.4 are highlighted in bold.

	RI		**RS1**		**RS2**	
groups	**variance**	**std.Dev**	**variance**	**std.Dev**	**variance**	**std.Dev**
henID (intercept)	0.365	0.604	1.036	1.018	1.250	1.118
henID x time			0.155	0.394	0.794	0.891
henID x time^2					0.047	0.216
Residual	0.458	0.677	0.301	0.548	0.258	0.508
N	80		80		80	
Number of observations	3230		3230		3230	
Marginal R2	0.13		0.14		0.14	
Conditional R2	0.52		0.69		0.74	
AIC	6939		5893		5590	

Table S2 – Variance estimates with their standard deviation, R^2 and AIC values for the random intercept (RI) and the two random slope models (RS1 and RS2). RS1 is a random slope model where time is linear (hen ID x time), and RS2 includes an additional quadratic term of time (hen ID x time^2).

	RI			**RS1**			**RS2**		
	Estimate	**Std. Error**	**Pr(>|t|)**	**Estimate**	**Std. Error**	**Pr(>|t|)**	**Estimate**	**Std. Error**	**Pr(>|t|)**
(Intercept)	**-0.607**	**0.106**	**<0.001**	**-0.635**	**0.139**	**<0.001**	**-0.669**	**0.148**	**<0.001**
time	**1.064**	**0.052**	**<0.001**	**1.092**	**0.061**	**<0.001**	**1.120**	**0.107**	**<0.001**
time2	**-0.237**	**0.014**	**<0.001**	**-0.243**	**0.012**	**<0.001**	**-0.250**	**0.027**	**<0.001**
avgDay	0.056	0.069	0.420	0.062	0.070	0.383	0.049	0.070	0.481
treatment OFH	**-0.317**	**0.140**	**0.026**	**-0.313**	**0.142**	**0.031**	**-0.290**	**0.142**	**0.044**
temperature	-0.027	0.014	0.051	**-0.028**	**0.011**	**0.013**	**-0.029**	**0.010**	**0.006**
body mass	-0.056	0.070	0.424	-0.077	0.072	0.286	-0.067	0.071	0.348

Table S3 – Estimates, standard errors and p-values for random intercept (RI) and random slope models (RS1 and RS2). RS1 is a random slope model where time is linear (hen ID x time), and RS2 includes an additional quadratic term of time (hen ID x time^2). Significant effects are highlighted in bold.

		KBF severity				**Feather damage**			
response		**post.mean**	**l.95..CI**	**u.95..CI**	**pMCMC**	**post.mean**	**l.95..CI**	**u.95..CI**	**pMCMC**
PC1	**(Intercept)**	**-0.657**	**-0.948**	**-0.370**	**<0.001**	**-0.657**	**-0.940**	**-0.356**	**<0.001**
PC1	**treatment OFH**	**-0.306**	**-0.602**	**-0.026**	**0.033**	**-0.308**	**-0.601**	**-0.022**	**0.036**
PC1	**time**	**1.120**	**0.898**	**1.321**	**<0.001**	**1.120**	**0.912**	**1.332**	**<0.001**
PC1	**time ^{2}**	**-0.250**	**-0.301**	**-0.196**	**<0.001**	**-0.250**	**-0.306**	**-0.200**	**<0.001**
PC1	**temperature**	**-0.029**	**-0.050**	**-0.009**	**0.006**	**-0.029**	**-0.049**	**-0.008**	**0.006**
Health Indicator	**(Intercept)**	**54.439**	**37.079**	**72.138**	**<0.001**	**28.343**	**14.504**	**41.612**	**<0.001**
Health Indicator	**treatment OFH**	**7.214**	**0.698**	**13.633**	**0.034**	-1.591	-10.623	7.812	0.687
Health Indicator	**predictability**	**23.540**	**6.567**	**38.699**	**0.004**	-5.649	-17.305	5.822	0.343
Health Indicator	**number no-transition-days**	**5.118**	**1.237**	**8.692**	**0.007**	-0.241	-3.061	2.462	0.870
Health Indicator	**body mass**	2.172	-1.238	5.354	0.187	1.243	-1.093	3.652	0.296
Health Indicator	**class MEXP**	-2.212	-9.574	5.232	0.567	2.152	-3.387	7.383	0.444
Health Indicator	**class REXP**	-0.213	-8.431	7.694	0.965	0.705	-5.199	6.512	0.824

Table S4 – Estimates, 95% credible intervals, and pMCMC for both KBF/PC1 and feather damage/PC1 bivariate models. Significant effects are highlighted in bold.

Chapter III

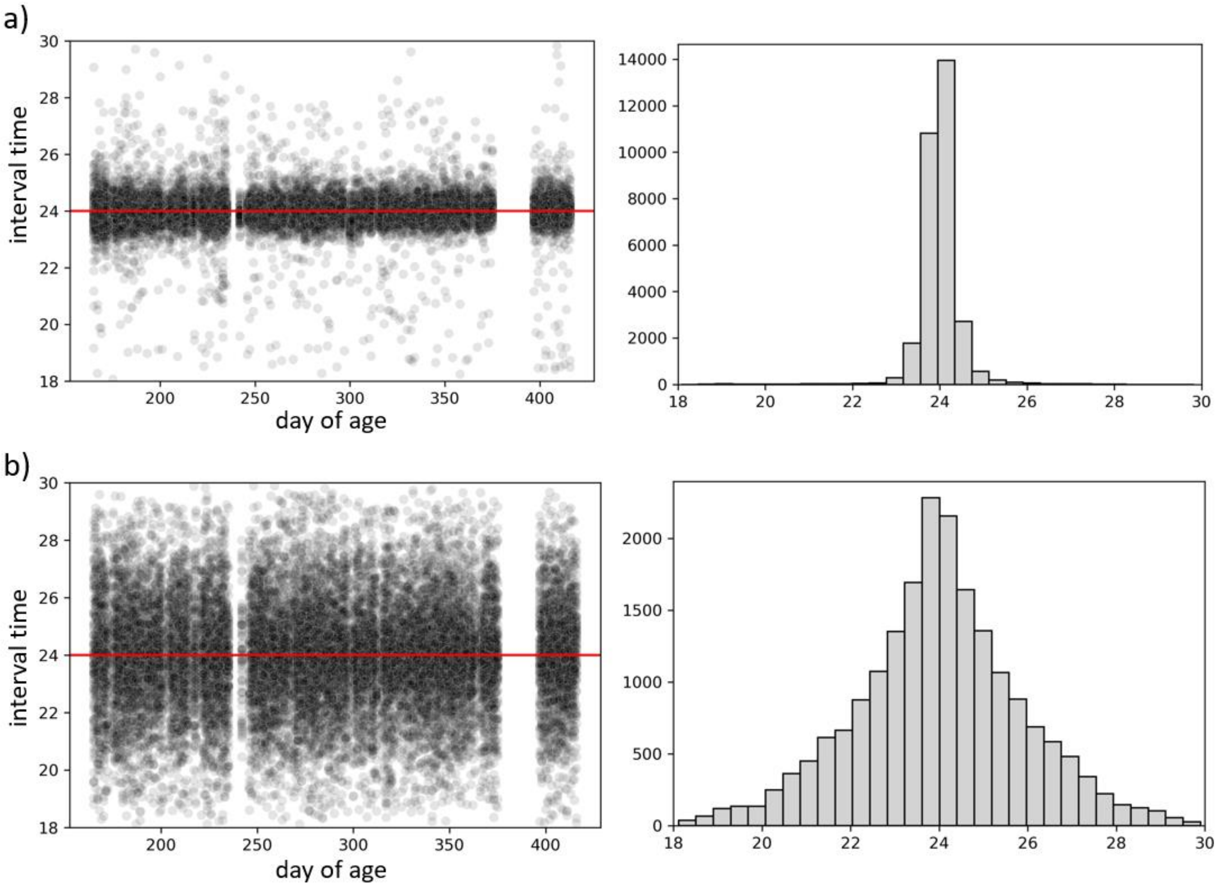


Figure S1 – Illustration of the interval time between two consecutive nestbox tier timing scores, based on the nestbox tier timing computed over the morning period 2h-8h (as used in analysis) in **(a)** and based on the nestbox tier timing computed over the following 6h hours (8h - 14h) **(b)**. For the two scenarios, we displayed the interval time over time (left side) and its density plot (right side). These visuals are extracted from the full dataset (i.e., not only including days used in subsequent analysis)

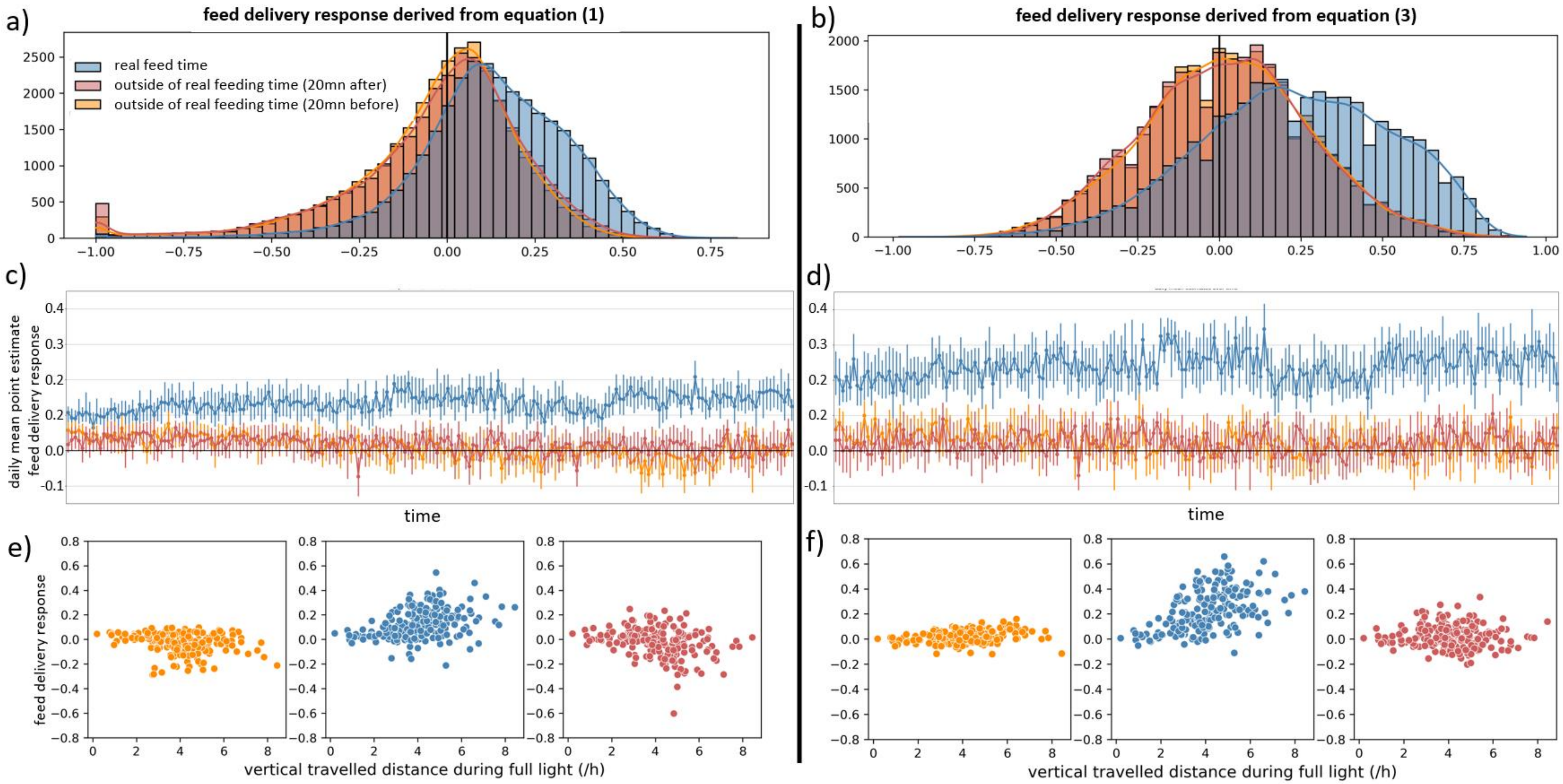


Figure S2 – Feed delivery response derived from equation (1) (left side) and from equation (3) (right side), with behavioural responses computed with the true feed delivery timing (blue) and fake feed delivery timing (orange and red). The behavioural responses are compared with a density plots (a)(b) the daily mean point estimates over time (c)(d), and the mean hens' feed delivery response in relation to their mean vertical travelled distance during full light (/h) in (e)(f).

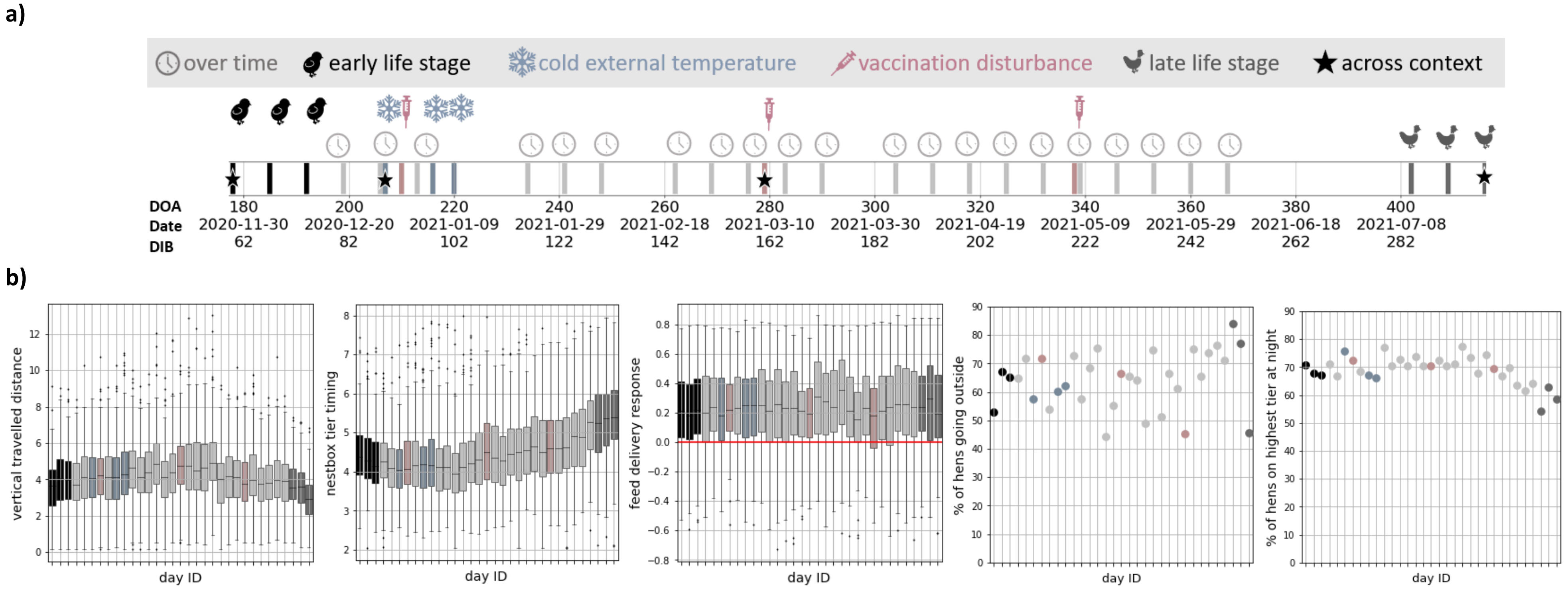


Figure S3 – Illustration of the observed behaviours over time, considering only days selected in the analysis. The timeframe of selected days is given in **a)**. The five spatial behaviours are given in the following order (in **b)** left to right): with boxplots for the vertical travelled distance per indoor hour, the nestbox timing and the feed delivery response behaviours, followed by the two binary behaviours represented by the percentage of hens going outside and the percentage of hens staying most of the night-time in the highest tier (i.e. for WG presence and sleeping tier behaviours, respectively). Colours highlight the different categories of the selected days (over time ; early and late life stages; cold external temperature and vaccination disturbances).

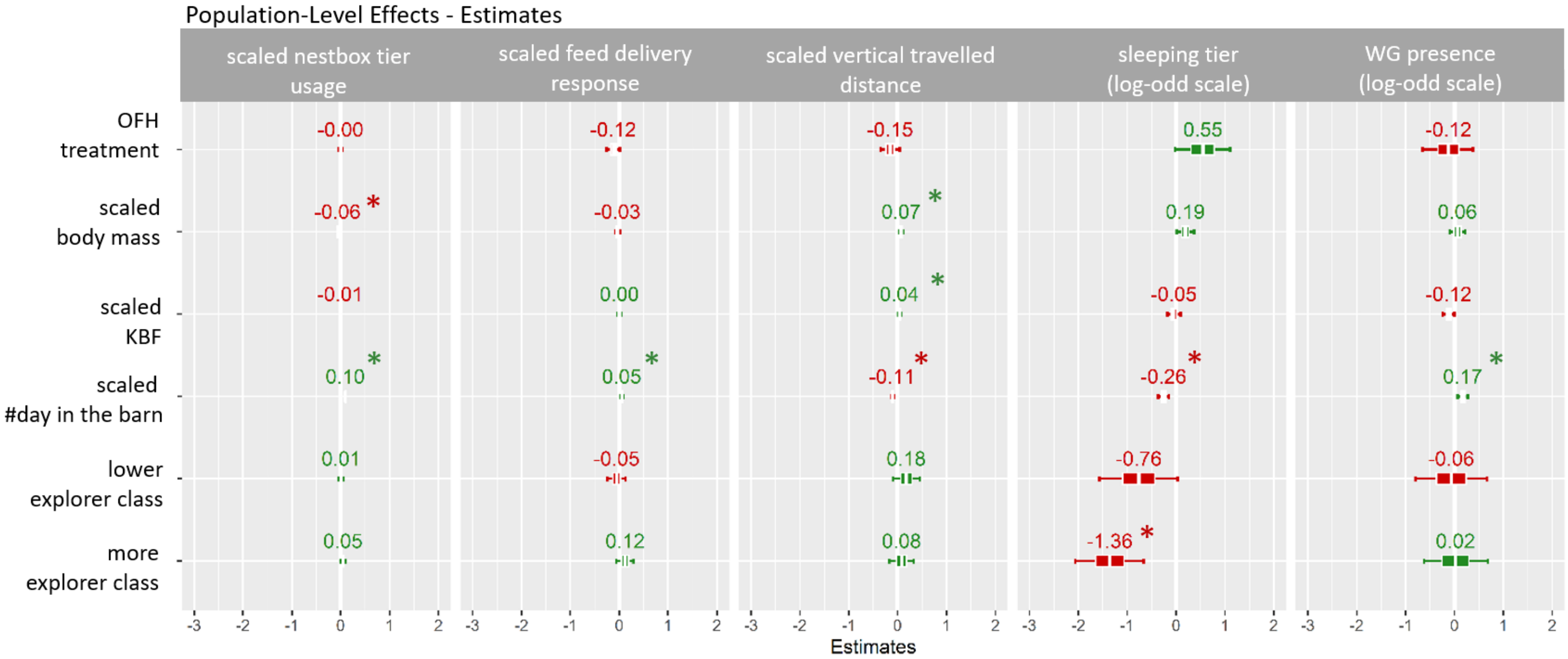


Figure S4. Population-level effects estimates (95% credible interval) of the OFH treatment, the scaled body mass, the scaled keel bone fracture severity score, the number of days in the laying barn since transfer on behaviours, and of the more and lees explorer class (in contrast to hens selected randomly), from the multivariate mixed model, which responses are provided in the columns

Text 1 – More/less explorer chicks' class not validated

We installed one raised platform with blue dye per pen when chicks were one day old, so that more explorer chicks would have blue feet. We planned to change the blue dye with a violet dye once at least 30 chicks could be classified as "more explorer", in each of the four rearing pens (600 hens / pen). By testing the dye previously on dead hens, we knew that feet colored by the blue dye and then by the violet dye, would have been distinguishable to feet colored by the violet dye only. However, there was not enough more explorer chicks within a suitable timeframe (< 5 days) to allow for the intermediate class, which we believe was needed to reliably differentiate the two classes of interest (more and less explorers). Furthermore, we added dye diluted in water approximately every two hours to prevent the dye from drying out, but we noticed on some occasions that the dye had already dried out. Therefore, class remained unvalidated, and we included it in the variable in our model only to control as we used it to help select individuals.

Text 2 - Feed delivery response

Here, we provide a detailed description of how the feed delivery response was defined. In the present manuscript, this behaviour was used to extract a daily behavioural score but note that other time units would work the same way. Let S be an external stimulus. Let B be a behavior of interest under presence of S and which can also occur under the absence of S. Following a diurnality index defined by [302] (and more recently used by [185,303]) we can define a simple descriptor of the tendency of B (e.g., number of occurrences, time spent) being greater while S occurs than while S is absent (SA):

$$(1)\ \ behavioural\ tendency\ = \frac{B_S/T_S - B_{SA}/T_{SA}}{B_S/T_S + B_{SA}/T_{SA}},$$

where B_S and B_{SA} are the behaviour B while S occurs and is absent, respectively, and T_S and T_{SA} are the corresponding exposure time to S and SA. Hence, when B is expressed exclusively in absence of S, this measure takes the value of -1, which changes to +1 when B is expressed solely while S occurs. Because the duration of each period during which S occurs can have varying length, the behaviour B should be defined at a same frequency than S (e.g. every minute), so that it is possible to have an aggregated value for behaviour B under S (B_s). Note that if the exposure time was the same under presence or absence of S for all days, then the formula would simplify to $(B_S - B_{SA})/(B_S + B_{SA})$.

However, when the exposure time is not always the same under presence or absence of S, then the measure will be comparing (subtracting and adding) two ratios that may not share the same statistical properties. Therefore, we here propose an alternative but similar measure, which compares B_S with N behavioural scores B_{SA_*}, each associated to a random exposure time T_{SA_*} that equals T_S :

$$(2)\ \ behavioural\ tendency\ = \frac{1}{N}\sum_{r=0}^{N}\left(\frac{B_S - B_{SA\_r}}{B_S + B_{SA\_r}}\right).$$

More generally, when S occurs several times over the day, for each such occurrences (called a period p, included in a set P), the associated behavioural score B_{Sp} can be compared with N behavioural scores B_{SpA_*}, each associated to a random exposure time T_{SpA_*} that equals T_{Sp} :

$$(3)\ \ behavioural\ tendency = \frac{1}{|P|}\sum_{p\in P}\left(\frac{1}{N}\sum_{r=0}^{N}\left(\frac{B_S - B_{SA\_r}}{B_S + B_{SA\_r}}\right)\right).$$

Note that with this definition, because each period when S occurs is used separately in the formula, one could also extract meta-data associated to each period and identify drivers of this behaviour. For example, using the time of the day associated to each period, we could evaluate if the behavioural tendency is in average greater in the morning than later in the day. Based on this last behavioural tendency measure, we defined the feed delivery response, explained in more detail in the main manuscript. Comparison between equation (1) and equation (3) is illustrated in the bellow Figure S1, where blue represent the feed delivery response with the real timing, and red and orange represent the feed delivery response based on a feed delivery timing with a forward and backward push of 20mn in time, respectively. Both indices differentiate with what is expected by chance, especially the one from equation (3). To control results from behavioural syndrome we further visualised the mean hens' feed delivery response in relation to their mean vertical travelled distance during full light (/h) in Figure S1 c.

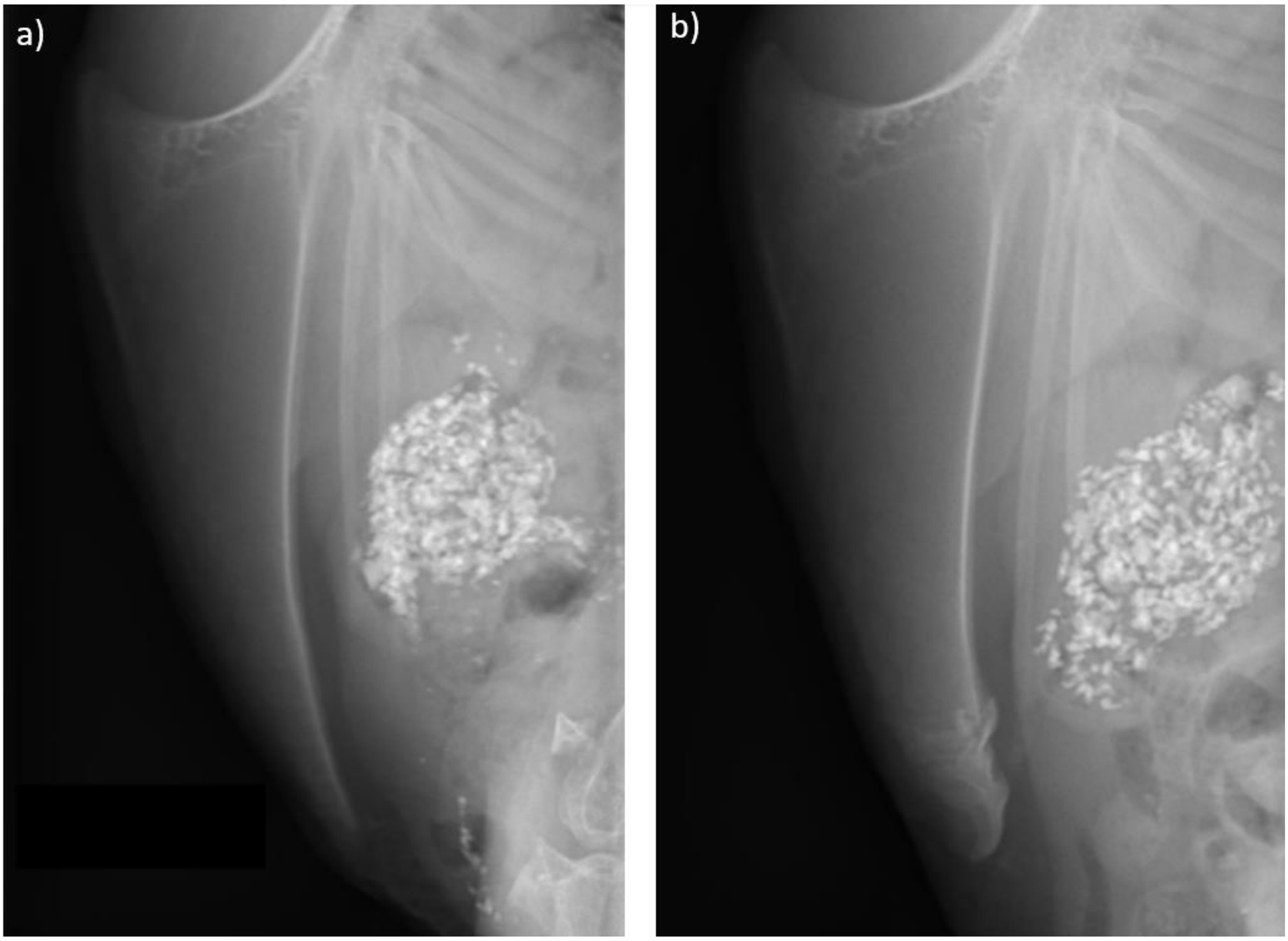
Additional details on the computation

We did not consider the first and last period of feed delivery (02:30 h and 16:15 h) to avoid overlap with behaviour related to the night. Because hens that are in the WG might not hear the sound associated to the delivery of fresh feed to a same extent as hens that are inside, only periods where a hen was not in the WG were included.

context	movement behaviour	between-individual variance	within-individual variance	trait mean
cold external temperature	sleeping tier	5.457 [2.323, 12.761]	4.730 [4.432, 5.484]	0.696
across context	sleeping tier	3.250 [1.456, 4.953]	4.676 [4.420, 5.293]	0.690
over time	sleeping tier	4.414 [3.021, 5.541]	4.784 [4.454, 5.342]	0.702
late life stage	sleeping tier	17.872 [7.716, 162.859]	4.121 [4.021, 4.515]	0.586
early life stage	sleeping tier	5.797 [2.471, 11.296]	4.642 [4.334, 5.395]	0.686
vaccination disturbance	sleeping tier	4.794 [1.825, 15.130]	4.830 [4.529, 5.708]	0.707
over time	WG presence	4.301 [3.049, 5.361]	4.377 [4.147, 4.654]	0.647
across context	WG presence	1.329 [0.479, 1.855]	4.054 [4.002, 4.185]	0.558
early life stage	WG presence	6.341 [2.779, 9.559]	4.234 [4.081, 4.642]	0.618
late life stage	WG presence	6.840 [2.639, 21.463]	4.687 [4.375, 5.271]	0.691
cold external temperature	WG presence	5.107 [2.042, 7.220]	4.170 [4.030, 4.486]	0.601
vaccination disturbance	WG presence	1.826 [0.513, 2.656]	4.208 [4.058, 4.472]	0.611
late life stage	feed delivery response	0.033 [0.027, 0.040]	0.057 [0.050, 0.066]	0.251
vaccination disturbance	feed delivery response	0.015 [0.012, 0.018]	0.050 [0.044, 0.057]	0.201
across context	feed delivery response	0.019 [0.015, 0.022]	0.058 [0.051, 0.065]	0.208

cold external temperature	feed delivery response	0.032 [0.027, 0.039]	0.044 [0.038, 0.051]	0.229
early life stage	feed delivery response	0.027 [0.023, 0.032]	0.048 [0.042, 0.055]	0.207
over time	feed delivery response	0.029 [0.025, 0.032]	0.057 [0.054, 0.060]	0.207
vaccination disturbance	nestbox tier timing	0.024 [0.018, 0.031]	0.018 [0.012, 0.024]	4.486
cold external temperature	nestbox tier timing	0.028 [0.023, 0.032]	0.011 [0.008, 0.015]	4.242
early life stage	nestbox tier timing	0.024 [0.019, 0.028]	0.011 [0.007, 0.014]	4.366
late life stage	nestbox tier timing	0.023 [0.016, 0.032]	0.016 [0.010, 0.024]	5.329
across context	nestbox tier timing	0.017 [0.013, 0.023]	0.020 [0.015, 0.025]	4.629
over time	nestbox tier timing	0.023 [0.020, 0.026]	0.021 [0.019, 0.023]	4.498
vaccination disturbance	vertical travelled distance	1.624 [1.402, 1.872]	1.472 [1.294, 1.691]	4.312
cold external temperature	vertical travelled distance	2.487 [2.242, 2.767]	0.911 [0.805, 1.044]	4.250
early life stage	vertical travelled distance	2.100 [1.908, 2.310]	0.571 [0.500, 0.653]	3.883
late life stage	vertical travelled distance	1.272 [1.155, 1.407]	0.356 [0.316, 0.408]	3.421
across context	vertical travelled distance	1.412 [1.236, 1.663]	1.535 [1.380, 1.723]	3.867

Table S1 - Each behavioural trait mean, within-, and among-individual variance within each context, over time and across context.



[redacted]

Figure S1 - Two latero-lateral radiographs, with a KBF severity score of 0 in **a)** and of 44 in **b)**.

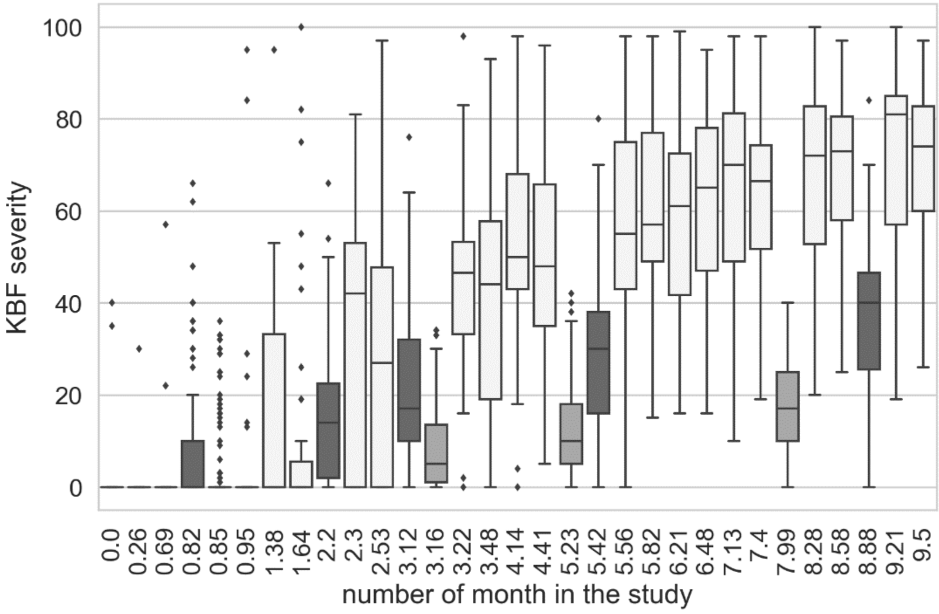


Figure S2 - Boxplots of the severity of KBF for each timestamp and dataset (Dataset1: lightest grey, Dataset2: darker grey, Dataset3: middle grey).

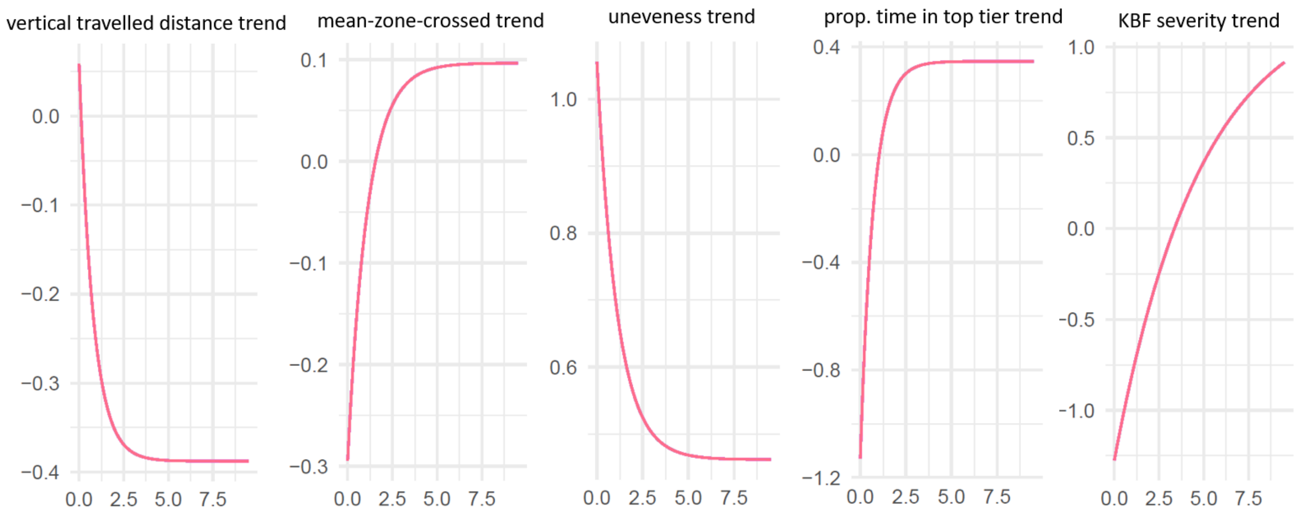


Figure S3 – Latent expectations conditional on the covariates, for the vertical travelled distance, mean-zone-crossed, uneveness, proportion of time spent in the top tier, and KBF severity trend processes. The KBF severity trend process for the four movement behaviour models were similar, and we displayed the one from the model with the mean-zone-crossed.

Text S1

The first dataset (Dataset1) used was published by Rufener et al. in 2019 [10], where they tracked 60 Lohmann Brown (LB) and 60 Lohmann Selected Leghorn (LSL) hens using an infrared tracking system. Hens in a pen were not single-strain. All dataset tracked over six days previous to assessing the keel bone. Because the two hybrids exhibited different space-use behaviours, the authors analysed the hybrids separately and suggested the differences could have been caused by a higher susceptibility to stress (due to being part of a phenotypic minority) by LSL hens leading to a more uneven usage of the zones. Therefore, we only included LB hens, with on average 10 ± 1.73 observations per hen (4-11 observations / hen, total observations: 593).

We used for the second dataset (Dataset2) data from a larger study (chapter V), where we tracked 227 Dekalb white hens with a low-frequency tracking system that we previously validated and described in [173]. Raw data were processed via the described ML-method, that uses a classifier trained to detected false registration. Half of the hens hatched-on-farm and the other half hatched in a commercial hatchery. After excluding hens that did not have enough observations for subsequent analysis (i.e. with less than three observations), we were left with 153 hens, mostly because of another study that collected epigenetic samples on focal hens. Hens from this dataset had on average 4 ± 0.76 observations (3-5 observations /hen, total observations: 658).

Finally, we used a third dataset (Dataset3) first presented in the current study, where we tracked 169 Dekalb white hens, with the same tracking system as in Dataset2 and processed the data with the SD-method [173], i.e. by filtering all transitions of short duration (less than 1 min). We selected randomly half of the focal hens from each pen and relocated them in a new identical pen three times throughout the production period, while the other half staid in their home pen during the entire laying phase. After excluding hens that had less than three observations, we were left with 163 hens (4 hens died, 2 tags were misfunctioning). Hens from this dataset had on average 4 ± 0.28 days tracked (3-4 observations /hen, total observations: 638).

Text S2

For each model we estimated a trend and a dynamic fluctuation for the two processes (the KBF and the spatial behaviour). Specification of such multivariate latent process model is described by Driver and Tomasik (2022)[227] but we give some details below:

Trend processes

We accounted for the trend of the behaviour and the KBF severity, by incorporating random initial intercepts and random slopes, as well as an estimated auto-effect term for the slopes that varied by process. We allowed varying initial intercepts across hens, by incorporating two latent "**trend" processes** (called "BHV" and "KBF") each measured with one manifest indicator (the behaviour and the KBF severity score) and estimated their initial latent covariance matrix parameters (T0VAR). We estimated their auto-regression effects of time and set their cross-regression effects to 0. To incorporate individual variation in the trend, we specified the continuous intercept parameter as random-effects by incorporating two additional latent **"continuous intercept" processes** (called "cintBHV" and "cintKBF") influencing the respective original latent processes ("BHV" and "KBF", respectively) with a cross-regression effect of 1 and by estimating the initial latent covariance matrix parameters.

Dynamics processes

To estimate the fluctuations, or dynamics, around the trends, we incorporated two additional latent **"dynamics" processes**, called "dynBHV" and "dynKBF". We estimated the auto- and cross-regression effects of these processes (with system noise) to estimate the speed and dynamic of the fluctuation, respectively. We did not allow these two processes to interact or covary with the trend processes and therefore fixed adequate elements from the temporal effects and the initial latent covariances to zero. The dynamics processes, dynBHV and dynKBF, are measured with one manifest indicator, the behaviour and the KBF severity score, though account only for the unpredictable change in the observation (i.e. what is not accounted by the trend processes). In other words, the dynamic processes solely estimate the fluctuations around the trend without interacting with the trend.

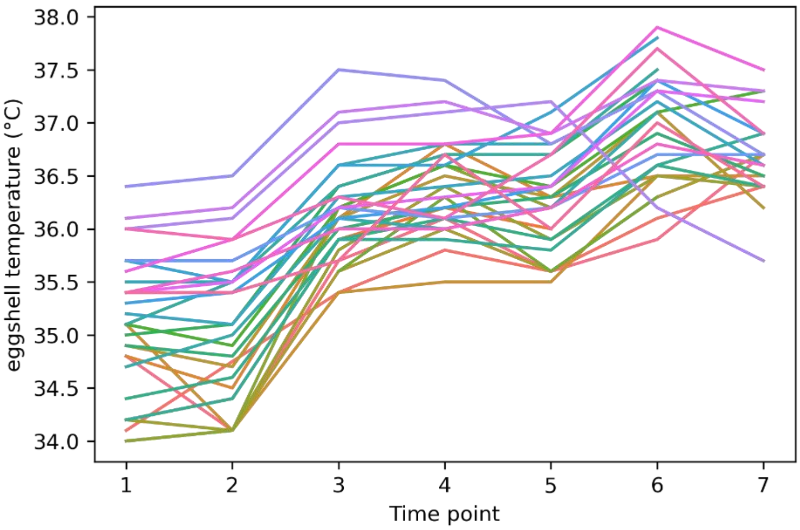
	depth (m)	width (m)	number / pen	prop. time accessible	Space (m^2)	normalized space
WG	4.050	2.300	1.000	0.400	3.726	0.110
litterred floor	7.000	2.300	1.000	1.000	16.100	0.476
lower tier	2.530	2.300	1.000	1.000	5.819	0.172
nestbox	1.130	0.510	4.000	0.933	2.152	0.064
balcony	2.300	0.320	2.000	1.000	1.472	0.044
top tier	1.970	2.300	1.000	1.000	4.531	0.134

Table S1 – Characteristics of the laying barn, including the depth (m), the width (m), the number of such element per pen, the proportion of time accessible, the space (m^2), and the normalized space. Note that the nestbox tier is composed of the nestboxes and balconies

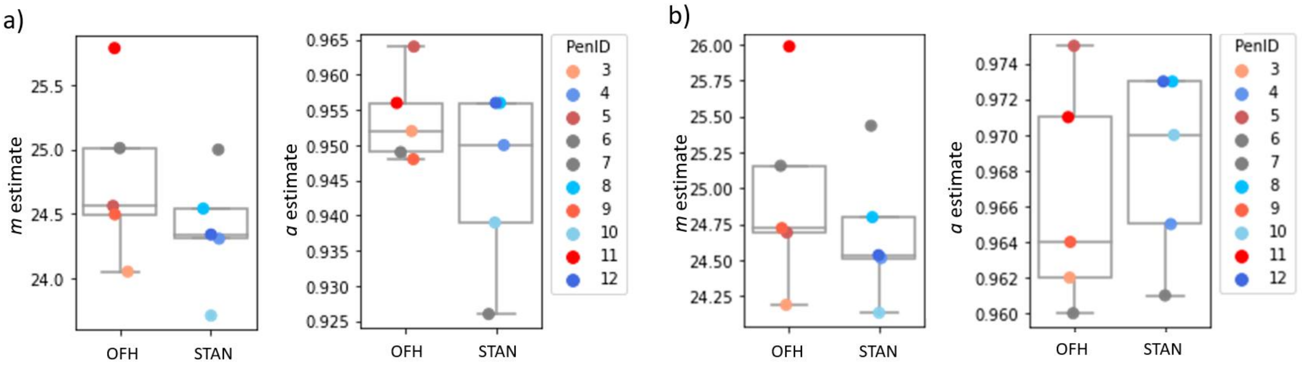
	VTD	MZC	unevenness	PropZ5	KBF severity
Dataset1	3.42 ±2.17	1.04 ±0.07	1.01 ±0.39	0.30 ±0.31	40.34 ±32.97
Dataset2	3.95 ±1.59	1.08 ±0.09	0.78 ±0.36	0.29 ±0.25	21.03 ±18.03
Dataset3	4.11 ±1.51		0.71 ±0.32	0.27 ±0.22	10.21 ±10.40
All datasets	3.84 ±1.79	1.06 ±0.08	0.83 ±0.38	0.29 ±0.26	23.44 ±25.34

Table S2 – Mean and standard deviation of the four spatial behaviour and the KBF severity score

Chapter V



S1 Fig. Eggshell temperature over time. We monitored 30 OFH eggs every six hours, for a total of 7 timestamps, until a significant proportion of the chicks hatched.



S2 Fig. Parameter estimates of the sigmoid curve per pen. Parameter *a* may be viewed as an indication of the level at which egg production stabilise. Parameter *m* as an indication of the time point at the inflection point of the curve. These estimates of the curves fitting the first 60 days and the full period in the laying barn are given in (b) and (d), respectively.

Time point	Pen : 1.1.1	Pen : 1.1.2	Pen : 1.1.3
1 june 13h	1	0	0
1 june 18h	1	0	0
1 june-2 june 00h	1	0	0
2 june 6h	2	1	1
2 june 12h	8	26	16
2 june 18h	-	-	-
2 june-3 june 00h	~15%	~40%	~25%
3 june 6h	~75%	~80%	~80%
3 june 12h	>80%	>80%	>80%
3 june 18h	>80%	>80%	>80%
3 june-4 june 00h	>80%	>80%	>80%
4 june 6h	>80%	>80%	>80%

S1 Table. Hatching rate over time for the OFH chicks.

	Feather damage		Body mass (g)		KBF severity	
DOA	OFH	STAN	OFH	STAN	OFH	STAN
7			65.54	64.70		
			[64.41, 66.67]	[63.58, 65.82]		
118			1170.91	1161.00		
			[1158.12, 1183.71]	[1148.06, 1173.94]		
173			1686.78	1673.76		
			[1670.18, 1703.39]	[1656.97, 1690.54]		
215			1792.27	1777.06	17.70	13.20
			[1772.30, 1812.25]	[1756.72, 1797.40]	[14.70, 20.80]	[10.1, 16.3]
243	9.44	8.49	1824.61	1801.59	21.30	16.70
	[6.80, 12.10]	[5.84, 11.10]	[1795.20, 1854.02]	[1771.95, 1831.23]	[18.20, 24.30]	[13.70, 19.80]
313	17.43	16.48	1884.65	1856.46	29.60	25.10
	[14.79, 20.10]	[13.82, 19.10]	[1848.58, 1920.71]	[1820.12, 1892.80]	[26.60, 32.70]	[22.00, 28.20]
417	34.12	33.17	1905.84	1884.41	39.50	35.00
	[31.46, 36.80]	[30.50, 35.80]	[1877.94, 1933.74]	[1856.34, 1912.49]	[36.40 42.60]	[31.90, 38.10]

S2 Table. **Estimated marginal means and 95% confidence intervals for welfare indicators**.

	severity		**FeatherDamage**		**weight_norm**	
Predictors	*Estimates*	*p*	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	13.29 ***	**<0.001**	8.88 ***	**<0.001**	0.51 ***	**<0.001**
	(10.15 – 16.43)		(6.53 – 11.24)		(0.47 – 0.54)	
CLASS [LEXP]	0.86	0.732	-1.97	0.217	-0.01	0.677
	(-4.08 – 5.80)		(-5.10 – 1.16)		(-0.07 – 0.04)	
CLASS [MEXP]	-1.17	0.610	0.79	0.588	0.01	0.677
	(-5.67 – 3.33)		(-2.07 – 3.64)		(-0.04 – 0.06)	
date2021-02-01	3.56 **	**0.001**			-0.05 ***	**<0.001**
	(1.39 – 5.73)				(-0.08 – -0.02)	
date2021-04-12	11.92 ***	**<0.001**	7.99 ***	**<0.001**	0.08 ***	**<0.001**
	(9.71 – 14.14)		(6.48 – 9.50)		(0.06 – 0.11)	
date2021-07-25	21.77 ***	**<0.001**	24.68 ***	**<0.001**	-0.05 ***	**<0.001**
	(19.48 – 24.07)		(23.10 – 26.25)		(-0.08 – -0.03)	

Treatment [OFH]	4.53^{*} (0.99 – 8.07)	**0.012**	0.95 (-1.95 – 3.85)	0.519	0.03 (-0.01 – 0.06)	0.166
date2020-09-29					-0.01 (-0.03 – 0.02)	0.584
date2020-11-23					0.05^{***} (0.03 – 0.08)	**<0.001**
date2021-01-04					-0.06^{***} (-0.09 – -0.03)	**<0.001**
Random Effects						
σ^2	101.10		49.72		0.01	
τ_{00}	125.66_{HenID}		$43.07_{\text{HenID:PenID}}$ 1.73_{PenID}		0.02_{HenID}	
ICC	0.55		0.47		0.58	
N	194_{HenID}		194_{HenID} 8_{PenID}		231_{HenID}	
Observations	669		506		1121	
Marginal R^2 / Conditional R^2	0.244 / 0.663		0.522 / 0.749		0.090 / 0.619	

** p<0.05   ** p<0.01   *** p<0.001*

S3 Table. Model outputs for welfare indicators.

month	proportion of time spent in top tier (exp())	proportion of time spent in litter floor	proportion of time spent in nestbox tier (exp())	WG presence (exp())	vertical travelled distance	mid-nestbox tier usage (exp())
1	**1.49 [1.12, 2.05], p=0.01**			0.50 [0.17, 1.22], p=0.13		
2	**1.67 [1.14, 2.35], p=0.00**	**-0.07 [-0.11, -0.02], p=0.01**	0.96 [0.87, 1.08], p=0.50	0.98 [0.40, 2.37], p=0.96	**-10.62 [-16.13, -4.35], p=0.00**	**0.96 [0.92, 1.00], p=0.02**
3	1.38 [0.90, 2.06], p=0.12	**-0.05 [-0.11, -0.00], p=0.04**	0.99 [0.87, 1.12], p=0.93	1.03 [0.35, 3.10], p=0.95	**-9.19 [-16.13, -2.50], p=0.01**	0.99 [0.95, 1.04], p=0.77
4	1.26 [0.86, 2.02], p=0.26	-0.04 [-0.09, 0.01], p=0.18	0.99 [0.87, 1.13], p=0.86	0.67 [0.23, 1.80], p=0.46	-6.53 [-14.38, 1.07], p=0.10	1.02 [0.97, 1.08], p=0.36
5	1.44 [0.89, 2.27], p=0.10	-0.04 [-0.09, 0.01], p=0.11	1.04 [0.90, 1.23], p=0.64	0.55 [0.17, 1.87], p=0.36	-4.08 [-11.11, 3.22], p=0.30	1.02 [0.96, 1.09], p=0.60
6	1.20 [0.82, 1.83], p=0.28	-0.04 [-0.09, 0.01], p=0.14	1.06 [0.91, 1.24], p=0.48	0.57 [0.24, 1.34], p=0.19	-3.87 [-11.33, 3.05], p=0.31	1.01 [0.97, 1.06], p=0.60
7	1.30 [0.90, 1.90], p=0.20	-0.03 [-0.08, 0.02], p=0.28	1.00 [0.89, 1.14], p=1.00	0.73 [0.33, 1.63], p=0.47	-2.90 [-9.40, 4.00], p=0.41	1.00 [0.95, 1.06], p=0.90
8	1.01 [0.63, 1.46], p=0.97	0.00 [-0.05, 0.06], p=0.90	0.99 [0.87, 1.12], p=0.88	1.05 [0.41, 3.00], p=0.88	1.22 [-5.44, 8.13], p=0.73	1.01 [0.96, 1.07], p=0.64
9	1.20 [0.77, 1.84], p=0.40	-0.01 [-0.06, 0.04], p=0.76	0.94 [0.83, 1.05], p=0.29	0.74 [0.25, 1.78], p=0.55	-0.69 [-7.22, 5.49], p=0.82	0.99 [0.94, 1.04], p=0.64
10	1.03 [0.70, 1.61], p=0.89	-0.01 [-0.05, 0.04], p=0.82	0.94 [0.85, 1.04], p=0.27	0.13 [0.01, 1.07], p=0.07	-0.38 [-5.46, 5.55], p=0.93	1.00 [0.95, 1.05], p=0.97

S4 Table. Bootstrapped estimates and p-values for the model fitting spatial behaviours. *

*For simplicity, we summarize here again the statistical models used. The proportion of the indoor time spent on the **top and nestbox tiers** were both modelled with a beta family. The proportion of the indoor time spent on the **litter floor** was modelled with a gaussian family for months 2-10 and with a binomial family to account for the excess of zeros (19.7% of observations) for month 1 (bootstrapped outputs given in main text to avoid confusion due to the different family distribution used: 0.30 [0.12, 0.67], p=0.0). The **WG presence** was modelled with a binomial family and the **nestbox tier usage** with a gamma family (log-link function). The **vertical travelled distance** was modelled with a gaussian family for months 2-10 and with a zero-inflated Poisson model with the rescaled number of day in the barn as zero-inflation parameter for month 1 (bootstrapped outputs given in main text to avoid confusion due to the different family distribution used: 0.69 [0.59, 0.82], p<0.001). The significant estimates are highlighted in bold (p<0.05).

DECLARATION OF ORIGINALITY

Last name, first name: Montalcini Camille

Matriculation number: 11-814-357

I hereby declare that this thesis represents my original work and that I have used no other sources except as noted by citations.

All data, tables, figures and text citations which have been reproduced from any other source, including the internet, have been explicitly acknowledged as such.

I am aware that in case of non-compliance, the Senate is entitled to withdraw the doctorate degree awarded to me on the basis of the present thesis, in accordance with the "Statut der Universität Bern (Universitätsstatut; UniSt)", Art. 69, of 7 June 2011.

Place, date:
Bern, 31.10.2023

Signature: