NTNU

Norwegian University of
Science and Technology

Covarying Behaviours and Innovation in a House Sparrow Metapopulation

Implications for Social Foraging

John Hammerås

MSc in Biology

Submission date: May 2016

Supervisor: Jonathan Wright, IBI

Co-supervisor: Henrik Jensen, IBI
Henrik Pärn, IBI
Bernt Rønning, IBI

Norwegian University of Science and Technology
Department of Biology

John Hammerås

Covarying Behaviours and Innovation in a House Sparrow Metapopulation:

Implications for Social Foraging

Master's thesis in Biology

Trondheim, June 2016

Supervisor: Professor Jonathan Wright

Norwegian University of Science and Technology
Faculty of Natural Sciences and Technology
Department of Biology

NTNU
Norwegian University of
Science and Technology

Acknowledgements

This master thesis was carried out at the Department of Biology at the Norwegian University of Science and Technology in Trondheim, Norway, from August 2013 to May 2016. The field work was carried out in cooperation with the House Sparrow Project at the Centre of Biodiversity Dynamics at NTNU, and I am grateful for the help that my co-supervisors – associate professors Henrik Jensen and Thor-Harald Ringsby and researchers Henrik Pärn and Bernt Rønning – along with the rest of the good people at CBD has provided in regards to various aspects of the work with this thesis, and also for the good atmosphere in the field and at lunch meetings during this time. I am also grateful to fellow master student Sindre L. Sommerli for our cooperation on the field work.

I am very fortunate to have had Professor Jonathan Wright as my main supervisor, not only because of the valuable help he has provided in regards to the thesis, but also for introducing me to the exiting field of behavioural ecology, and help me learn what sort of scientist I want to be. I have also been fortunate in getting to know fellow master students Mirja C. Olsen and Mette H. Finnøen and to have had a great cooperative relationship with them in the work on our respective theses. In addition, I am grateful for the help that postdoc Yimen G. Araya-Ajoy and phd-student (at North Dakota State University) Monica Anderson Berdal provided in the work with the statistical analyses.

This degree took me one year longer to complete than what was originally planned. I realised partway through that I did not cope at all well with the work, and that as a result I was not working effectively or happily. These were strange things to realise when I was working with things that I am interested in. Anyway, I sought help and found it in Jonathan Wright, the health services of the Student Welfare Organisation in Trondheim, friends and fellow students. I took an extension and worked out my issues with the result that I am now more aware of how I function best and how I can better cope with difficult situations and circumstances. I am grateful for the understanding and patience that the people around me have shown when I have not been fine, and allowed me to take the time to discover valuable insights into myself. The learning value of this thesis has been about more than just the science.

Abstract

Individuals in populations of many species have been found to differ consistently in suites of correlating behavioural measures. These phenomena appear to be heritable and subject to selection, thus making it as important to study the structure of behavioural correlations as it is to study each behaviour in isolation. In this part of a two-part study, flocks of house sparrows were observed at a communal feeder in captivity. We examined the individual variation and co-variance between the observed variables by the use of univariate models and hypothesis-testing with structural equation modelling (SEM). The model that fitted the data the best included correlations between variables thought to represent activity at the feeder. We also found repeatable individual differences in the specific use of the feeder, that is, individuals were found to visit the feeder for either many and short or few and long visits, suggesting alternative individual strategies for feeder use. One possible explanation for these findings in the use of the feeder is that they are expressions of true personality (i.e., consistent across contexts), while another is that they are a reflection of consistent social interactions within the flocks (e.g. dominance hierarchy). The second part of this two-part study examines the behaviour of the same individuals in a solitary context, and the combination of the two studies might distinguish between these possible explanations. Our results underline the importance of studying behaviours at the flock-level, and this points towards an interesting interplay between behavioural and population ecology in the future study of these phenomena.

Sammendrag

I mange arter har man funnet at individer innad i populasjoner har konsekvente forskjeller i hele strukturer av korrelerende atferder. Dette ser ut til å være arvelig og selektert for. Derfor er det like viktig å studere strukturen av disse korrelerende atferdene som det er å studere hver enkelt atferd isolert. I denne studien, som er en del av et prosjekt i to deler, observerte vi flokker i fangenskap mens de besøkte en matstasjon. Vi undersøkte individuell variasjon og kovarians mellom de observerte variablene med en kombinasjon av univariat-modeller og hypotese-testing med strukturell likningsmodellering (SEM). Modellen som passet best til dataene hadde en korrelasjonsstruktur som inkluderte variabler som representerte aktivitetsmål i matstasjonen. Vi fant også repeterbare individuelle forskjeller i bruken av matstasjonen, nærmere bestemt at individer enten hadde mange og korte eller få og lange besøk. En mulig forklaring på disse funnene er at de er utrykk for ekte personlighet (altså konsekvent atferd over flere kontekster), mens en annen forklaring er at de er et utrykk for konsekvente sosiale interaksjoner innad i flokkene (f.eks. på grunn av dominanshierarkier). Den andre delen av dette todelte prosjektet ser på atferden til de samme individene mens de er alene, og kombinasjonen av disse to studiene kan hjelpe til med å skille mellom de to nevnte mulige forklaringene. Våre resultater understreker viktigheten av å studere atferder på populasjons-nivå, og dette peker mot spennende et samspill mellom atferds- og populasjonsøkologi i videre studier av disse fenomenene.

Table of Contents

Introduction

Animal personality and behavioural syndromes

In almost all studies to date, individuals have been found to differ consistenly from each other over time and across environmental contexts in their average behavioural expression (Gosling 2001; Sih et al. 2004; Réale et al. 2007). Such differences, referred to as *animal personality*, implies that each individual does not express the entire range of phenotypes that exist in the population for a labile behavioural trait (meaning that there are limits to flexibility), thus creating apparently unexplained individual variation in behaviours within and between populations and species (Sih et al. 2004; Réale et al. 2010; Dingemanse et al. 2010a).

Behaviours have also been found to correlate (Sih et al. 2004; Réale et al. 2007), both between and within individuals (Dingemanse et al. 2010b). The former is often termed *behavioural syndromes* and the latter *correlational plasticity* (Dingemanse et al. 2010b). An effect of behavioural syndromes is similar to that of animal personality: an individual is not expected to exhibit optimal levels of each and every behaviour due to the constraints imposed by the correlations between behaviours. This is of research interest because it appears contrary to the usual assumption of adaptive optimality that has been popular and successful in behavioural ecology since the 1960s (Sih et al. 2004; Owens 2006). The existence of behavioural syndromes thus offers a powerful explanation for expression of behavioural traits that for traditional behavioural ecology theory may appear suboptimal or maladaptive (Bell 2007).

Behavioural syndromes can be documented by subjecting the same set of individuals repeatedly to two or more independent behavioural assays (Dingemanse et al. 2010b). It is important to repeat assays in order to find out how much of the observed differences are due to actual inherent differences between individuals. Behaviours are said to be *repeatable* when they have low within-individual variance compared to high between-individual variance across time (Bell et al. 2009). In addition, when studying consistent individual differences in behaviours in a social context one must distinguish the inherent effects of the individual from the effects that other individuals has on it. Therefore, it is important to study the same individuals in isolation as well, and then test for correlations between the two assays. This

should be done in captivity where experimental conditions and the state of each individual can be controlled and kept constant (Nelson et al. 2008).

An adaptative explanation for behavioural syndromes is that they should only evolve in populations where natural selection has favoured them (Dingemanse et al. 2007). An example of differences between populations in patterns of correlation is the studies done on the three-spined sticklebacks (*Gasterosteus aculeatus*) by Dingemanse et al. (2007, 2010b). Populations from ponds with fish predators showed tight correlations between aggression, exploration of novel food sources and altered environments, whereas populations from ponds without predators had only weak to entirely lacking correlations between these behaviours (Dingemanse et al. 2010b). Knowledge about the environments that different populations live in is thus central for the understanding of how and why behaviours correlate. Furthermore, behaviours that have evolved to correlate should perhaps be studied as one functional unit rather than just separate independently varying behaviours (Stadler et al. 2001). Hence, behavoural syndromes might exist not just because of parallel evolution on the level of each behaviour, but because it is the actual correlations between the behaviours that are favoured by selection in some situation and not others. A competing non-adaptive explanation would therefore be that such correlations exist because of constraints on the architecture of mechanisms that produce the different behaviours (Stamps 1991, but see Bell 2005). However, this would predict no variation in syndromes across populations, which is what makes the stickleback example so compeling from an adaptationist's perspective.

Innovation

Innovation is generally understood to concern novel or modified behaviours in either familiar or novel environments (Reader and Laland 2003). Anecdotal evidence for innovative behaviour abound, but common for the most well-known examples is that they involve exploration of novel food or a novel way of processing food (Reader and Laland 2003, Lefebvre and Bolhuis 2003). For instance, great tits (*Cyanistes caeruleus*) in the Midlands, UK, figured out how to open the aluminium top of milk bottles to get to the fat rich cream that had settled right underneath, and this behaviour subsequently spread to much of the UK (Fisher and Hinde 1949). This example also illustrates an additional important point with regards to innovation: it is not enough for just one individual to have figured out a novel behaviour or how to exploit a novel food source; the behaviour must also spread through the

population by means of social learning for it to be proper innovation. This means that not only must there be at least one individual in the population that is able to innovate, but the rest of the population must also have the capacity for social learning.

Species that have an opportunistic lifestyle (e.g., feed in a generalist manner or readily explore novel environments), or live in unpredictably varying environments, are likely to exhibit innovative behaviour (Reader and Laland 2003). Species and perhaps particular individuals are thus expected to be more innovative when they are more likely to be subject to novel environments, as shown by higher levels of innovation in invasive bird species (Sol et al. 2005). In addition, differences in dispersal between individuals have been found to be associated with other behavioural traits that show personality (e.g., exploration in captivity in great tits, *Parus major*, Dingemanse et al. 2003). Therefore, we might expect individuals that have more need for innovation to also exhibit greater levels of exploration and to more readily approach novel objects and foods in individual personality assays, perhaps as part of a wider innovation behavioural syndrome (Reader and Laland 2003).

Approaches to studying behavioural correlations

Structural Equation Modelling has facilitated the adoption of an approach where one is only considering a restricted subset of *a priori* models of correlations between measures. Its success is limited by the models considered, but comparisons between models (using Akaike Information Criterion) and between subsets of a data set (e.g., populations) has been made easy. For instance, one could here make a model inspired by which measures that might be aspects of the same behaviour (either in the same or different contexts) with stronger correlations between these measures and less with others, and then test if this model is a better fit to the data than models where all measures are correlating equally strong or not at all. An example of this is a study on great tits (*Parus major*) which demonstrated that measures in different contexts thought to reflect aggression correlated more strongly than they did with other behavioural measures (Araya-Ajoy and Dingemanse 2014). Such knowledge has predictive power, since one can now measure an aspect of aggressive behaviour in great tits in one context and know with some certainty that this carries over into other contexts where it might not always be possible to measure it. Other groups of behaviours could benefit from this type of analysis, for instance shyness-boldness (Bell 2007) and innovation/exploration.

Only considering a subset of *a priori* SEM models also makes it easier to inform the study of behavioural correlations with prior knowledge of ecological factors affecting the organisms. Since behavioural ecologists are especially attuned to relevant ecological factors affecting phenotypic expression, using this approach might be the way in which behavioural ecologists can contribute the most to the study of these phenomena (Bell 2007). The previously mentioned studies on three-spined sticklebacks by Dingemanse et al. (2010b), is an example of the use of this approach.

This approach involving SEM is different from the theory-free approach that uses PCA. PCA has the benefit that it might discover correlations between behaviours that are unexpected given current knowledge about functional or mechanistic relations (Bell 2007). However, possible drawbacks include the increased risk of making type I errors (detecting a pattern that is not really there), and the difficulty in comparing studies (Budaev 2010). One weakness of SEM is that it is highly dependent on the strength of the correlations between measures, and also that it requires a large sample size (Dochtermann and Jenkins 2010). The smaller the sample size, the lower is the power to detect patterns in the data, thus increasing the risk of making a type II error (failing to detect a pattern when there really is one).

Objectives and predictions

This thesis is one part of a two-part project where the objective is to see whether animal personality and behavioural syndromes, assayed in a solitary and a social context, exist in a house sparrow metapopulation. The behaviours assayed in the social context is the focus of this thesis, while the solitary context assays were the focus of the master's thesis of Sindre L. Sommerli. Predictions for the results of combining these two theses will be provided in the Discussion. For the social assays, we predict repeatable individual differences in behaviours assayed in a social context in captivity (at a communal feeder), including: order of arrival at an empty feeder, aggression towards conspecifics, and length and number of visits to the feeder. We further predict that these behaviours covary such that they form a behavioural syndrome (see below for the specific hypothesised models).

Methods

Study species

We used the house sparrow (*Passer domesticus*) as our study species. Adults of this small, passerine bird typically feed on seeds, while chicks are fed with insects, but the species is often described as being remarkably opportunistic and innovate in its choice of food, and can feed on many other items, such as the pellets of food commonly fed to cows on dairy farms (Anderson 2006). House sparrows are also very invasive, which is exemplified by its worldwide distribution and close association with man-made environments (Anderson 2006). It is also a very social species, with individuals breeding in loose colonies and almost always feeding and moving around in flocks (Anderson 2006).

Study area

Our study area included the three islands of Leka (65.088 N 11.675 E), Vega (65.655 N 11.963 E) and Vikna (64.913 N 11.001 E) off the coast of Northern Norway (coordinates given as mean coordinates within each main locality, see Jensen et al. 2013). On these islands, the house sparrows live in close association with human settlements. For the most part, this means in and around the dairy farms, although some sparrows also feed in gardens and roost in nest boxes on houses not connected with farms (e.g. on Vega). It is assumed that the sparrows captured on these farms in the winter constituted relatively discrete populations, with only limited movement between the farms, between the islands, and to and from the mainland. This seems justified given that most dispersal events in these populations are natal dispersal (Tufto et al. 2005). When the term "flock" is used in the following text, this is meant as all individuals used in the trials here that were captured on the same farm during the winter field work of February and March 2014 (see Table 1 for an overview).

This metapopulation has been used as a research system by the Centre for Biodiversity Dynamics (CBD) and its predecessors at the NTNU continuously since the early 2000s (much like a similar one on the Helgeland archipelago further north since the mid-1990s). Our project was thus a part of a larger research programme on the population ecology, genetics and evolution of these sparrow populations. All birds included in our assays underwent the same processing in the same order after capture: ID marking, measurement of morphological

features, blood sampling, BMR assay, and finally behavioural assays. The three procedures were carried out by certified individuals at the CBD, while the BMR assay was carried out by Dr. Bernt Rønning. The behavioural assays (both the solitary context and the social context assays) included here were carried out in collaboration with master student Sindre L. Sommerli. The social context assays is the focus of this thesis, while the solitary context assays was the focus of the 2015 master's thesis of Sommerli.

Table 1. Overview of the populations assayed in the experiment. Asterisks denote the flocks that were run through the assays twice in order to test for repeatability of the observed variables. The numbers in parentheses in the "Trial flock size" column denotes the number of individuals not observed in the feeder during the trial. The column "Estimated flock size" shows the estimated size of the flocks in the wild.

Island	Flock	Date of trial	Trial flock size	Estimated flock size
Leka	1*	11.feb	14	15
Vega	2	21.feb	8 (1)	8
	3	22.feb	14	14
	4	23.feb	12 (1)	15
	5*	24.feb	8	8
	6	25.feb	7 (1)	7
Vikna	7	15.mar	16 (2)	
	8	16.mar	16 (3)	18
	9	17.mar	14 (4)	16
	10*	18.mar	11	12
	11	21.mar	12 (1)	15

Catching, keeping and handling

In February and March of 2014 all house sparrows on the islands were captured using mist-nets (either inside or directly outside farm buildings) and held in captivity on a central barn on each island for marking, measurements and assays. The central barn on each island where the birds were kept in captivity was a disused one that was no longer used for holding livestock. The main room was sealed with plastic sheets, cardboard and wooden boards on walls, ceiling and floor wherever needed to prevent sparrows escaping. Windows were covered with

transparent plastic sheets so that light came in, but prevented the birds hurting themselves if they flew at them. Lamps were also lit at night to give the birds the opportunity to feed. The temperature in the barn was maintained with electric heaters at what was deemed to be normal for cow sheds in the area at that time of the year (~15°C).

The barn was divided with non-transparent plastic sheets into compartments for holding birds at different stages in the sequence of measurements and assays that we put them through. Each compartment had an electric heater and at least one lamp in it, and also branches fixed to the walls for the birds to perch. The separate room of the barn that had been used as milking house while the farm was still active, was used by us as the area where the initial individual markings, morphological measurements and blood samplings were carried out. This was in order to keep noise and movement by humans at a minimum inside the barn proper.

The initial stage of handling after the birds had been captured was marking them with individual IDs - combinations of metal and coloured plastic rings were unique for each individual, with a seven-digit number engraved on the metal ring. The next stage was measurement of morphological features. Tarsus length, beak length and beak height were measured with a digital Mitutoyo calliper (to the nearest 0.01 mm). In addition, mask length, and badge width and height (both actual and "potential", i.e., the black colouration that was under the winter feathers), were measured on males with the same calliper. An overall qualitative categorisation on the badge and beak colour was also made. Wings were measured with a ruler (to the nearest mm), and body mass was measured with a Pesola spring balance (to the nearest 0.1 g). Blood samples (25 μL) were extracted from under the wing where the brachial vein crosses the wing bone, and stored in 1 mL 96 % ethanol for later DNA tests (e.g. determining phylogenetic relationships).

A subset of captured individuals were put through BMR assays on the evening or night after capture. The BMR assays were carried out in a house separate from the barn in order to provide as stable conditions as possible by limiting noise and movement from other birds and humans. The BMR setup could accommodate up to 8 individuals during the evening and an additional 8 individuals during a second run in the night, totalling 16 overall per evening/night. We timed the capture so that as large a proportion of birds from the same flock

were put through the BMR assay during the same evening/night. We then used this same subset of the flock in our behavioural assays. Thus, the behavioural assays (see below) included the whole of a flock of 16 birds or smaller that had all been captured together on the same day. We coordinated the behavioural assays with the BMR assay in this way both because we wanted the BMR data for our behaviourally assayed birds, but also to put our birds through the same processing so that we would make sure as much as possible that they were in the same state. The BMR data were not used in this thesis.

After the BMR assays, the birds were put in a compartment in the barn proper separate from all other birds for resting and *ad libitum* feeding for a day. The feeder in the resting compartment was similar to the one later used in the social context assays so that they could become familiar with it before the assays. At about 22:30 on the evening of the resting day, the food was removed and the birds were food-deprived until the start of the behavioural assays at 08:00 the morning after. This was done both to ensure that the birds were in as similar a state as possible and because we wanted them to be sufficiently hungry to want to visit the feeder during the behavioural assay the morning after.

Experimental setup

The setup for the behavioural assays consisted of two parts: one where the behaviour of each individual bird was assayed in a solitary context (i.e., alone in a cage), and one where all birds from the same captured flock interacted at a communal feeder. The latter setup was the focus of this thesis, but a description of the former setup is also included here. Both trials were filmed with Sony high-resolution colour CCD cameras (model NC1381W). The videos were stored on 16 Mb ScanDisk memory cards by a H-264 Portable Mini Video Surveillance Recorder from LUPUS TEC (Recording resolution and frame rate: 704x576@25 FPS, 352x80@25 FOS (PAL)). The room where these experiments were carried out in measured 2.8 x 3.0 m on Leka, 3.5 x 3.5 m on Vega, and 3.0 x 3.0 on Vikna. One flock on each island (totalling 33 individuals, see Table 1) was subjected to the experimental setup twice in order to test for repeatability. The repeat assays were carried out two days after the original assays, with a day of *ad libitum* feeding and subsequent food-deprivation in between (just like they had been subjected to before the original behavioural assays after the BMR assays).

Solitary context assay

The setup of the solitary context assay consisted of a small cage with perches arranged in one of two configurations (see Figure 1), a box on the side of the cage, and a shelf on the cage wall opposite from the box onto which the novel object and novel food item would be introduced. A total of eight such cages could be used at the same time. The cages were separated with black plastic sheets so that the birds had no visual communication with each other.

The bird was put into the box on the side of the cage and left there to calm down for five minutes before the first trial started, upon which point a paper sheet was removed from between the box and the cage so that the bird was free to emerge and explore the cage as it pleased. After 20 minutes, a novel object was introduced (either a toy crab or a toy penguin, see Figure 2) onto the shelf (see Figure 1) for the bird to approach and explore. After an additional 20 minutes, novel food (dyed green or purple sunflower seeds) was introduced onto the same shelf for the bird to explore and eat. The experiment was ended after a final 20 minutes, thus the assay lasted for an hour and involved three trials. This assay was usually initiated at about 08:00 in the morning, then run again for an additional eight birds (totalling 16 birds), before being finished at around 12:00. Then followed a period of preparing the room for social context assays, in which the cages were either moved out of the room (Leka) or put in a corner and covered with non-transparent plastic sheets (Vega, Vikna).

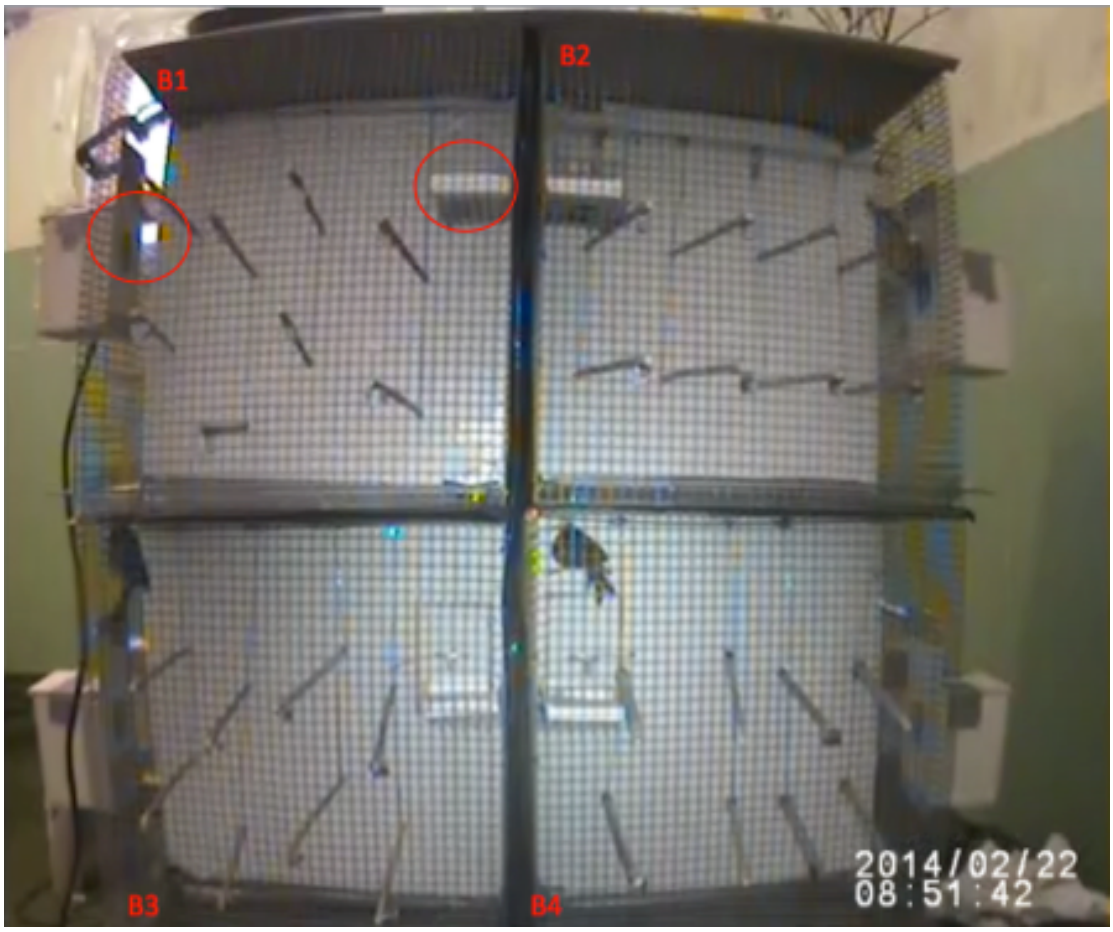


Figure 1. The experimental setup used in the solitary context assays. This is a screenshot of the camera pointing at four of the eight cages (the other four are on the back of the cages visible here, with another camera having the same view as this one). The circle on the upper left shows the entrance to the cage from the resting box, and the circle above middle in the picture shows the shelf onto which novel objects and novel food items were introduced through a small door in the cage wall. Note that cages B1 and B3 have one configuration of perches, while B2 and B4 has another. Size of each cage: 50 cm x 50 cm x 25 cm.

Figure 2. The novel objects used in the solitary context assays: a toy penguin (left) and a toy crab (right).

Social context assay

The setup of the social context assay consisted of a room with branches taped to the walls and a feeder in the corner down to which the birds could visit and feed in as they pleased. Sand was used as non-edible matrix and the majority of seeds were put in a clumped distribution and covered with sand to encourage searching and producing-scrounging behaviour. A few seeds were evenly distributed on top. Seeds were weighed before they were put in the feeder before each trial, and the remaining seeds were sifted out from the sand and weighed after each trial had ended.

The feeder was covered with plexiglass on walls and roof, exept for on one side that served as the entrance. A perch was fixed along the lower edge of the entrance with a camera pointing along it in order to facilitate identification of the birds visiting by their ID rings. Another camera filmed from above through a hole in the ceiling, and a third camera was put about 30 cm away from the entrance pointing directly at it (see Figure 3.)

Figure 3. The feeder used in the social context assays. This is a screenshot from a video obtained with the camera that is pointing directly at the entrance of the feeder. The two other cameras used in the video analysis are visible here; one to the left of the entrance, pointing along the entrance perch, and one just visible at the top, pointing down on the feeder through a hole in the ceiling. Note that all sides except the entrance were covered with plexiglass. Size of the feeder: 60 cm x 50 cm. The floor of the feeder was raised up almost to the entrance perch.

Video analysis of the feeder assays

The video analysis of the feeder assays consisted of noting the entry and leave time of all visits to the feeder, the identity of the individual who made the visit, any aggressive encounters that the individual was involved in (without determining the direction or intensity of the aggression), and calculating the visit duration as leave time minus entry time. The video files of the first flock that was analysed (flock 2, see Table 1) was reanalysed at the end of the video analysis in order to test intra-observer reliability. The correlations between these different measures ranged from r = 0.97 to 1.00.

Statistical analysis

All the statistical analyses were carried out in R version 3.1.2 (R Core Team 2014), and all analyses except for the repeatability tests were carried out on the data from the original assays with the data from the repeat assays left out. Six behavioural variables were derived from observations in the social context assays (see Table 2 for list). In addition, body mass measured on the morning of the day of the behavioural assays was used.

The distribution of number of aggressive encounters during visits was skewed towards lower values, and therefore, this was made into a binomial measure and the estimated probability that an individual's visit would involve an aggressive encounter was calculated from it. This estimate was obtained from a generalised linear-mixed effects model with visits involving aggressive encounters and total number of visits as response, and bird ID as random intercept, with residual distribution specified as binomial. The variable "change in visit duration" was obtained by extracting each individual's slope from a model with log visit duration as response, visit entry time as predictor and random slope, and bird ID as random intercept. In SEMs where flock or sex had to be controlled for, versions of these two variables were made from models that included flock or sex, respectively, as a fixed effect.

All variables were log-transformed to ensure normal-distribution, except for first visit entry time, which was disrupted out of its normal-distribution by log-transformation. Prior to SEM, all variables were also scaled (that is, the variance standardised to 1).

Table 2. Description of variables derived from observations in the social context assays. All variables were log-transformed to ensure normal-distribution, except for first visit entry time, which was disrupted out of its normal-distribution by log-transformation.

Variable	Description
First visit entry time	The time since the trial started that the individual made its first visit to the feeder.
Mean visit duration	Mean duration of all visits that the individual made to the feeder during the trial.
Change in visit duration	Change in visit duration during the trial (see main text for details).
Total visit duration	The sum of durations of all visits that the individual made to the feeder during the trial.
Number of visits	Total number of visits that the individual made to the feeder during the trial.
Aggression	The probability that an individual's visit will involve an aggressive encounter with another individual (see main text for details).
Body mass	Body mass measured on the morning of the day of the behavioural assays.

ANCOVA

Univariate ANCOVAs were carried out separately on the six behavioural variables to explore any effects of sex, flock and mass prior to the SEMs. Linear mixed-effect models were used to account for hierarchical organisation in the data set, fitted with the lmerTest package in R in order for the degrees of freedom to be estimated with Sathertwaite approximation and the sum of squares to be calculated with the type III approach (that is, testing for the main effect while controlling for all other effects).

Within- and between-individual effects were explored in these univariate models using the method described by van de Pol and Wright (2009). A within-subject centred version of the original observation variable (that is, made by subtracting the individual's mean value from each observation value) was used as a variable representing the within-individual effect. A variable including each individual's mean value was used to represent the between-individual effect. All models in this procedure used individual ID nested in flock ID as random intercept. The first model involved just the variable involving the original observation values as a covariate, while the second model had this variable decomposed into the between- and within-individual effects to see which effect they had on the response, and whether any of them were significant. The third model involved the between-individual effect and the original variable to see whether the between- and within-individual effects were significantly different from each other. The reasoning behind this last model is that since it tests for each effect while controlling for all other effects in the model, looking at the effect of the original variable is just testing again for any within-individual effects, while controlling for any between-individual effects. The between-subject effect in this model will then represent the difference between the between- and within-individual effects. The within- vs. between-flock effects were explored using the same approach, though here with only flock ID in as random intercept as there was only one observation per individual in each variable.

Models fitted for tests exploring effects on only the flock-level (i.e., there being only one observation per flock in the response variable) were fitted with a linear non-mixed model, using the car package in R, again to ensure that the sum of squares were calculated using the type III approach.

Sex had no detectable effect in any of the models (with p-values well above 0.25), and its removal was justified in every case by drops in AIC-values by more than two units from a full to a reduced model.

Repeatability

For the sub-set of flocks and individuals that were tested in the set-up twice (three flocks with a total of 33 individuals, see Table 1), repeatability was calculated by dividing the between-individual variance by the total variance:

$$r = \sigma^2_{between} / (\sigma^2_{between} + \sigma^2_{within})$$

The above equation assumes Gaussian distribution and that repeated measures were taken under the same conditions (Dingemanse and Dochtermann 2013). The former assumption was not met for the aggression variable, which was binomial, so a logit link was added to the denominator of the above equation to accommodate that (Nakagawa and Schielzeth 2010). Order was controlled for by including it in the mixed models as a fixed effect.

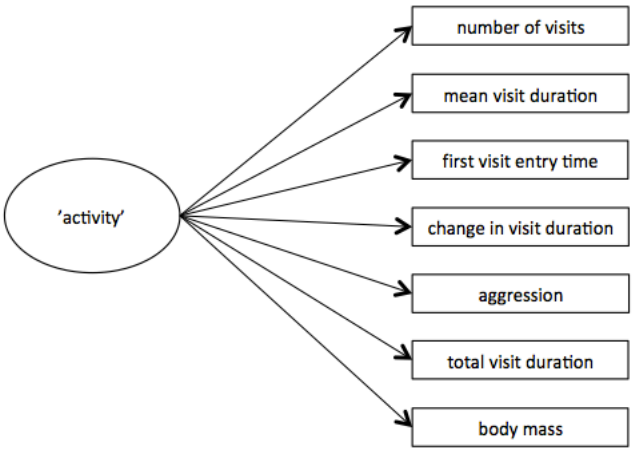
Structural Equation Modelling (SEM)

Prior to the SEMs, bivariate correlations between the seven variables were calculated using Pearson's product-moment correlation coefficients. In order to study the structure of covariances between the observed variables, structural equation modelling (fitted with the lavaan package and visualised with the semPlot package in R) was used to examine six different *a priori* hypothesised covariance structures (see Figure 4).

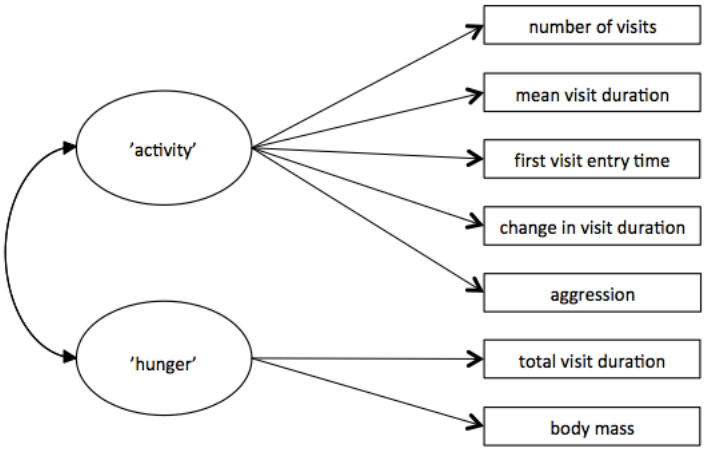
Model H0 proposes a scenario where the variables vary independently and do not produce a latent variable, while H1 is the opposite, i.e., all variables covaries such that they produce one latent variable together. H2 includes two latent variables, one for observed variables that are thought to be expressions of activity, and one for the ones that are thought to be more expressions of hunger. The two latent variables in this model covary. An alternative model where the two latent variables do not covary is not included here as that model would have involved the 'hunger' latent variable being produced by only two observed variables while not covarying with anything else, thus making it underidentified (i.e., it would have had more parameters to estimate than there was non-reduntant information in the data, Beaujean 2014). H3 proposes the same scenario as H2, though with aggression producing the 'hunger' latent variable and not 'activity'. The two latent variables in this model covary. An alternative model, H3b, was also considered where the two latent variables do not covary. H4 proposes that the latent variable is being produced by observed variables thought to reflect activity (with aggression being one of them), while the rest vary independently. H5 is the same model as H4, although with aggression varying independently as well. H6 has observed variables thought to reflect hunger (with aggression being one of them) producing one latent variable, while the rest of the observed variables are varying indepentently. An alternative model where aggression was not one of the variables producing the 'hunger' latent variable was not considered, as the model would have been underidentified (Beaujean 2014).

H0 - no correlations

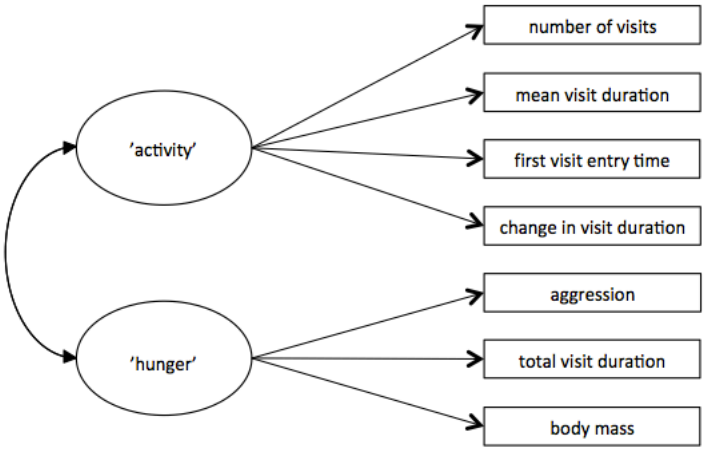
H1 – all variables correlate



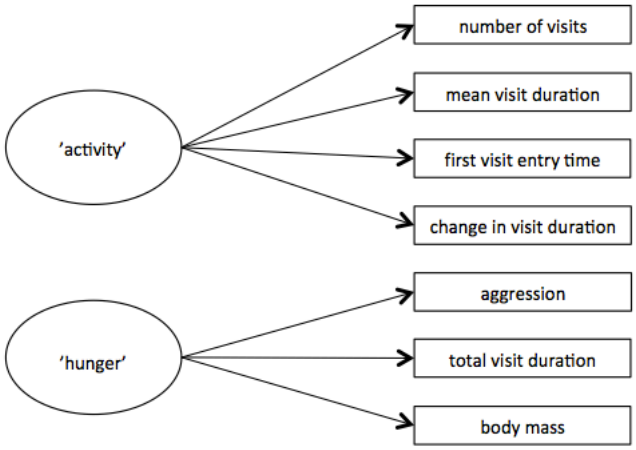
H2 – activity and hunger (aggression in activity)



H3 – activity and hunger (aggression in hunger)



H3b – activity and hunger (aggression in hunger)



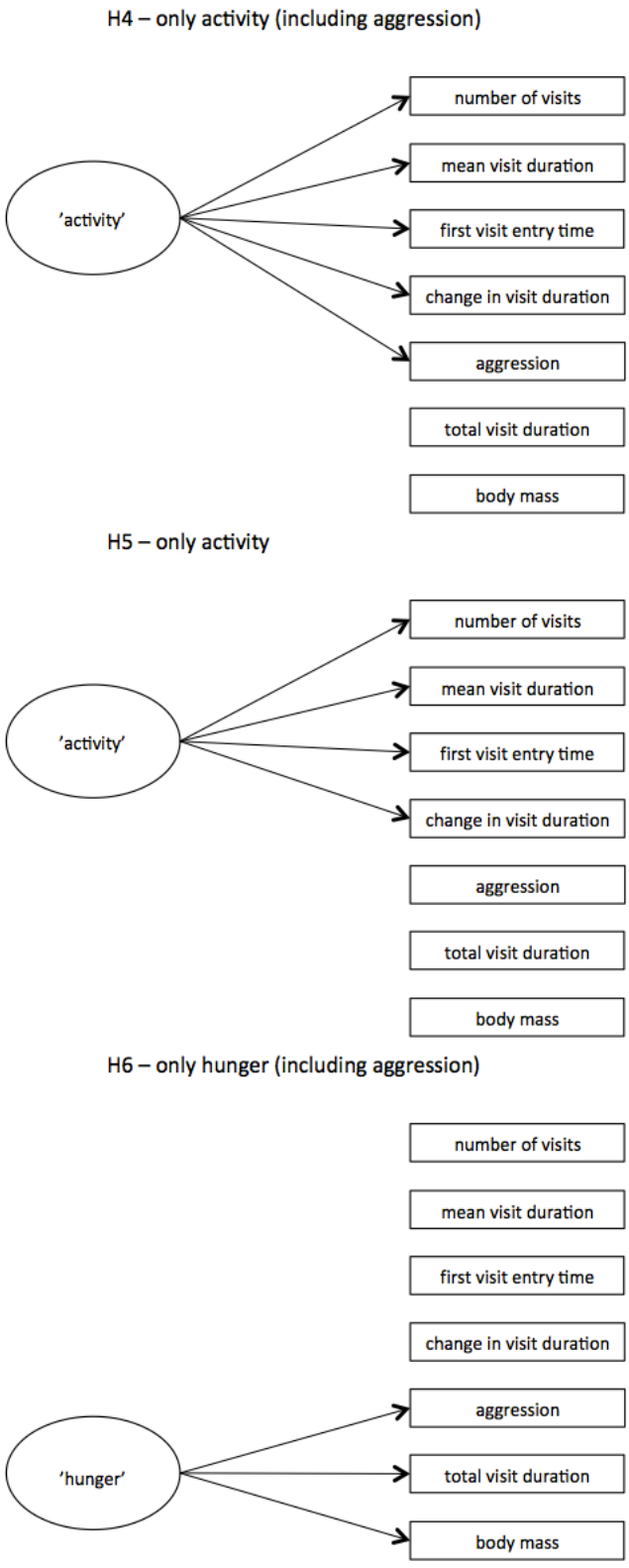


Figure 4. Diagrams of models (hypotheses) explaining the covariance structure between the observed variables. Latent variables are in circles and observed variables are in boxes. Single-headed arrows indicate a causal relationships, whereas double-headed arrows indicate covariances. See main text for further description of the hypotheses.

Two additional *post-hoc* models inspired by the correlation matrices were also examined (Figure 5). Model H7 proposes that hunger (represented by body mass) is driving the visit duration variables, that aggression increases with visit duration (which in turn covaries with body mass) and is thus also part of this syndrome, and that first visit entry time is only related to the rest of the variables through number of visits. H8 is a variation on the same theme, though with the difference that change in visit duration is related to the latent variable only through body mass. Total visit duration was also removed from all models to see if its strong relationship to some of the other variables (e.g., number of visits and mean visit duration, see below) was causing problems for the model fitting, though when this turned out not to be the case, it was left in. Models were ranked according to the Akaike information criterion (AIC).

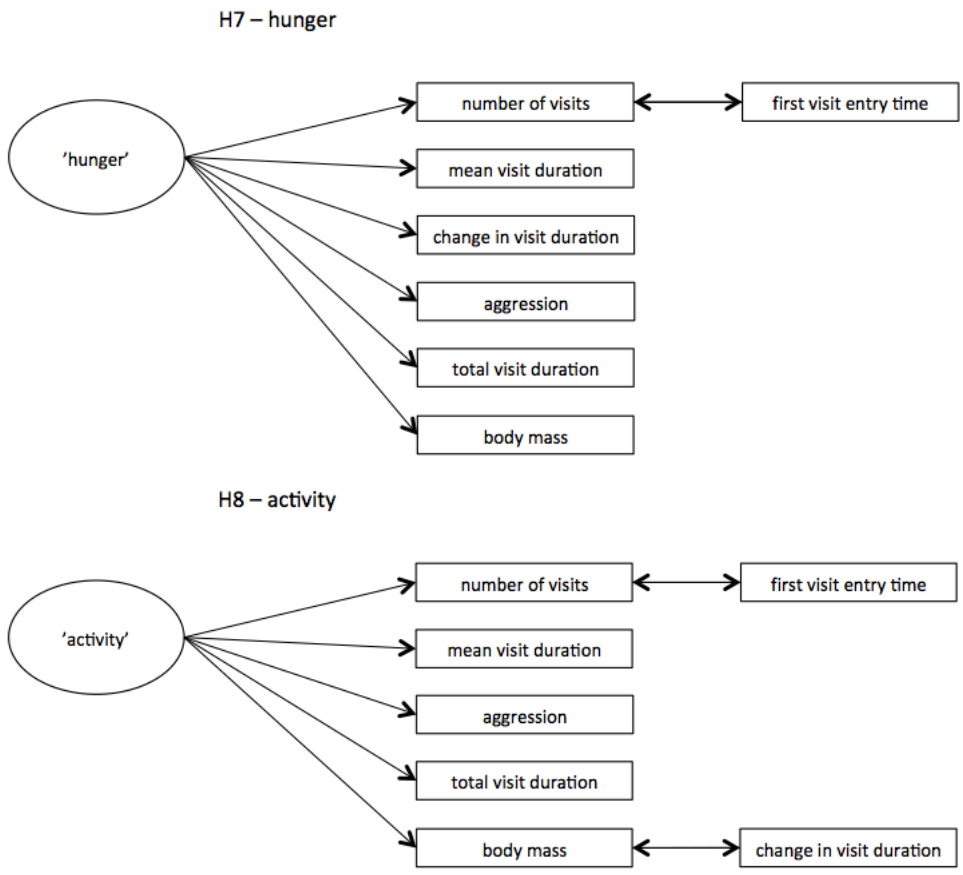


Figure 5. Diagrams of two additional models (hypotheses) explaining the covariance structure between the observed variables, inspired by the correlation matrices. Latent variables are in circles and observed variables are in boxes. Single-headed arrows indicate a

causal relationships, whereas double-headed arrows indicate covariances. See main text for further description of the hypotheses.

All models were fitted on the raw variables and also on variables controlling for either flock or sex, in order to see whether factor loadings and amount of unexplained variance attached to each observed variable differed greatly within the same models, whether it was different models that fitted the data properly for the three different versions of the variables and whether it was different models that were the best fit to the data. If one or several of these scenarios turned out to be the case it would suggest that there were differences between flocks or sexes that were driving the structure of correlations between the observed variables, rather than just the hypothesised individual differences.

Results

Body mass, nutritional state and total visit duration

Pre-trial body mass had a negative but marginally non-significant effect on mass change during the day of the individual and flock trials ($F_{1,110.54} = 3.588$; $p = 0.061$, slope = -0.089 ± 0.047). Interestingly, the between-flock effect here was significant (Figure 6; $F_{1,10.50} = 9.006$; $p = 0.013$, slope = -0.600 ± 0.200) while the within-flock effect was not ($F_{1,105.10} = 2.119$; $p = 0.148$, slope = -0.070 ± 0.048). These two effects were also significantly different from each other ($F_{1,11.73} = 6.646$; $p = 0.025$). This suggests that, as expected, the lightest birds consumed the most during the feeder trials, and hence we assume foraged more intensively because they were the most hungry. However, this effect was almost completely due to differences in hunger and foraging behaviour between flocks rather than within them, although the significance and the effect size here is perhaps relatively small (Figure 6).

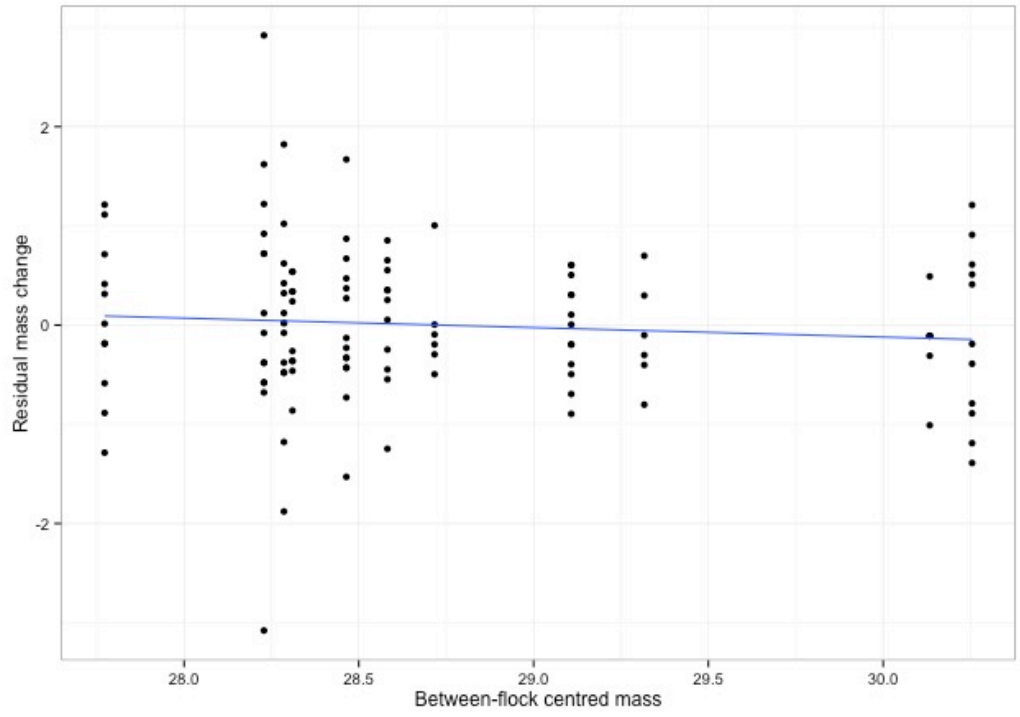


Figure 6. The effect of between-flock centred body mass on residual individual mass change (i.e. controlling for random between-flock effects). Equation of the fitted line: y = 2.741 - 0.095x.

Flocks that spent more total time in the feeder consumed more seeds ($F_{1,7} = 8.642$; $p = 0.022$, slope = 0.0004 ± 0.0001). In combination with the flock-wide relationship between

pre-trial body mass and mass change (above), it seems safe to assume that visit duration here is a good measure of individual seed consumption (i.e. birds were not spending time at the feeder carrying out any other behaviours). We therefore expected the lightest (i.e., the hungriest) birds to spend the most total time in the feeder, and this was indeed the case (Figure 7; $F_{1,114.93} = 13.924$; $p < 0.001$, slope = -0.222 ± 0.060). Both between- and within-flock effects were significant and negative ($F_{1,10.965} = 13.531$; $p = 0.004$, slope = -0.644 ± 0.175, and $F_{1,106.161} = 8.907$; $p = 0.004$, slope = -0.185 ± 0.062, respectively). The smaller effect here within flocks was significantly different from the stronger between-flock effect ($F_{1,13.799} = 6.118$; $p = 0.027$), again suggesting a larger role for between-flock differences in this regard.

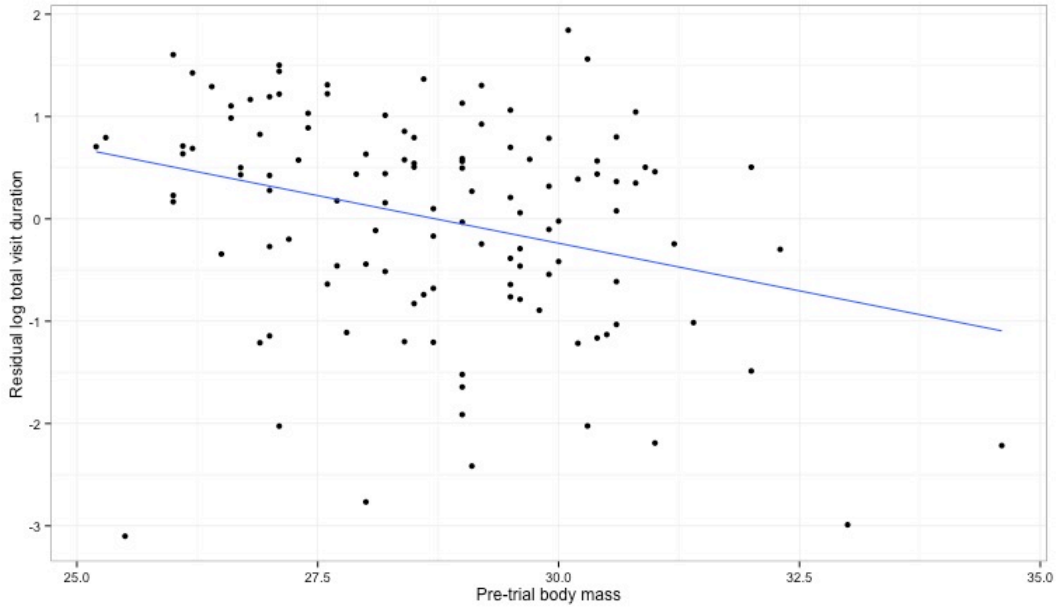


Figure 7. The negative effect of pre-trial body mass on residual log total visit duration (i.e., controlling for random between-flock effects). Equation of the fitted line: y = 5.348 - 0.186x.

Number of visits and visit duration

The birds with the most visits to the feeder also had the shortest mean visit durations (Figure 8; $F_{1,117.99} = 4.088$; $p = 0.045$; slope = -0.289 ± 0.143). The within-flock effect here was marginally non-significant and negative ($F_{1,107.675} = 3.278$; $p = 0.073$, slope = -0.269 ± 0.149), while the between-flock effect was non-significant and negative ($F_{1,7.939} = 0.943$; $p = 0.360$, slope = -0.519 ± 0.534), and the two were not significantly different from each other ($F_{1,9.226} = 0.202$; $p = 0.664$). Therefore, individual birds within flocks either used the feeder for fewer longer visits or for many shorter visits, but this effect was not particularly strong and might

just reflect chance differences in visit durations to the feeder per individual. However, the possibility exists for alternative individually repeatable strategies of many short visits versus fewer longer visits to the feeder.

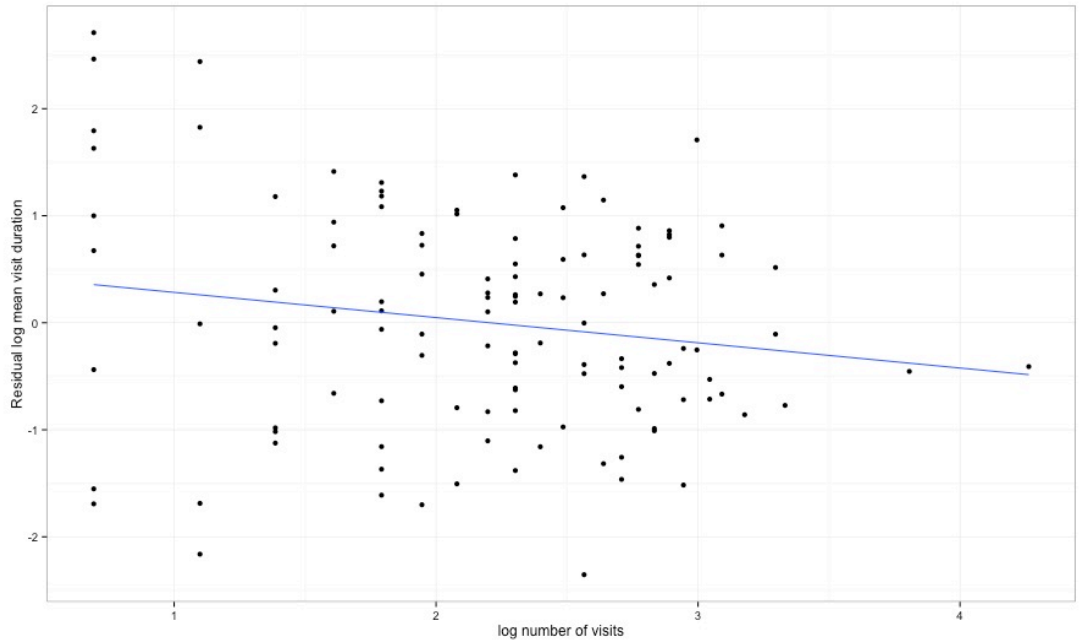


Figure 8. The negative effect of log number of visits to the feeder during the trial on residual log mean visit duration (i.e., controlling for random between-flock effects). Equation of the fitted line: y = -0.235x + 0.519.

As might be expected, number of visits to the feeder per individual had a significant and positive effect on total visit duration ($F_{1,117.99} = 42.161$; $p < 0.001$, slope = 0.923 ± 0.142). The within-flock effect was significant and positive ($F_{1,107.75} = 40.678$; $p < 0.001$, slope = 0.945 ± 0.148), while the between-flock effect was not significant, though also positive ($F_{1,8.00} = 1.605$; $p = 0.241$, slope = 0.673 ± 0.531), and the two were not significantly different from each other ($F_{1,9.298} = 0.244$; $p = 0.633$). Also as expected, mean visit duration also had a significant and positive effect on total visit duration ($F_{1,117.78} = 149.55$; $p < 0.001$, slope = 0.850 ± 0.070). Both between- and within- flock effects were significant and positive ($F_{1,10.435} = 11.522$; $p = 0.006$, slope = 0.775 ± 0.228, and $F_{1,108.808} = 137.099$; $p < 0.001$, slope = 0.859 ± 0.073), and they were not significantly different from each other ($F_{1,12.658} = 0.122$; $p = 0.733$). So, although number of visits and mean visit duration were negatively related to each other (see above), they both obviously contributed to the total foraging time per individual at the feeder.

First visit entry time

Pre-trial body mass did not have a significant effect on individual first visit entry time ($F_{1,107.34} = 2.775$; $p = 0.099$). This was not due to between- and within flock effects masking each other, as they were both also non-significant ($F_{1,9.145} = 1.051$; $p = 0.332$, slope = 689.20 ± 672.18, and $F_{1,105.055} = 2.445$; $p = 0.121$, slope = 111.53 ± 71.33), and they were also not significantly different from each other ($F_{1,9.352} = 0.730$; $p = 0.414$). Therefore, although pre-trial body mass had some effect on total feeder use once the birds had started feeding (see above), it did not seem to have an effect on the timing of bird and flock first visits to the feeder – i.e. initial boldness to use the feeder.

Number of visits during the trial had a significant and negative effect on individual first visit entry time (Figure 9; $F_{1,111.02} = 36.351$; $p < 0.001$, slope = -963.53 ± 159.81). Both between- and within-flock effects were also significant and negative ($F_{1,8.809} = 8.703$; $p = 0.017$, slope = -3330.86 ± 1129.07, and $F_{1,108.015} = 33.207$; $p < 0.001$, slope = -927.76 ± 161.00, respectively), and they were marginally non-significantly different from each other ($F_{1,9.171} = 4.440$; $p = 0.064$). This suggest that active individuals (and to some extent flocks) that came and went from the feeder more often were the ones that were more likely to visit the feeder early in the first place.

Mean visit duration had a marginally non-significant effect on first visit entry time ($F_{1,111.67} = 3.37$; $p = 0.069$, slope = -210.23 ± 114.52). The within-flock effect was marginally non-significant and negative ($F_{1,108.060} = 3.099$; $p = 0.081$, slope = -203.47 ± 115.58), while the between-flock effect was non-significant and also negative ($F_{1,9.169} = 0.423$; $p = 0.532$, slope = -576.56 ± 886.92). The two were not significantly different from each other ($F_{1,9.483} = 0.174$; $p = 0.686$). Therefore, the effect of visit number on first visit did not carry over into an indirect effect on visit duration (i.e. due to the effect shown in Figure 8).

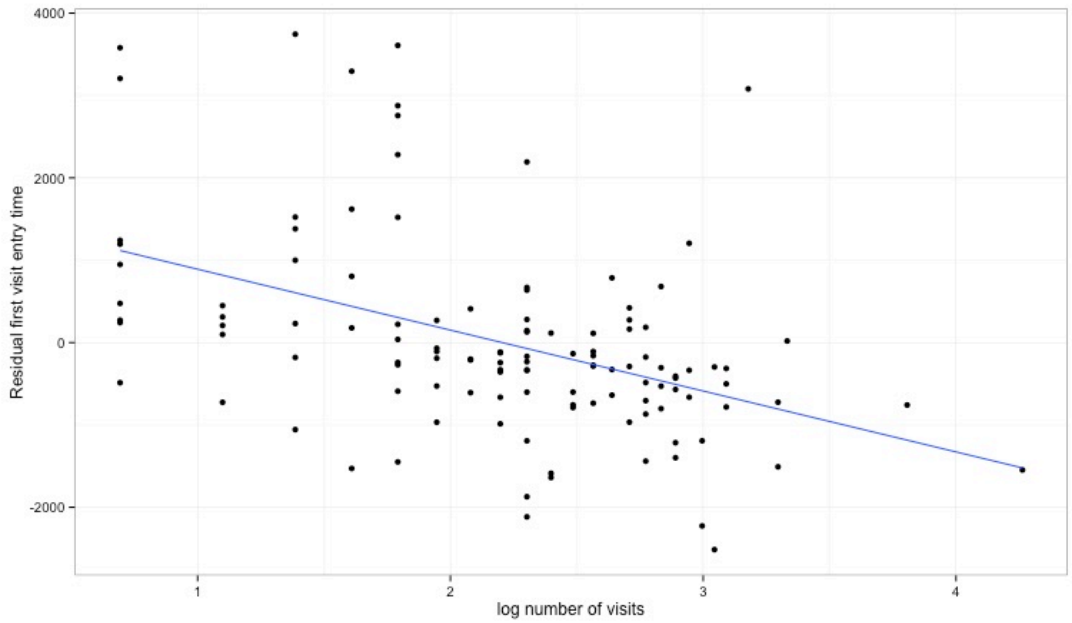


Figure 9. The negative effect of number of visits to the feeder during the trial on residual first visit entry time (i.e., controlling for random between-flock effects). Equation of the fitted line: y = 1630.3 - 739.5x.

Last visit entry time

Pre-trial body mass did not have a significant effect on last visit entry time ($F_{1,114} = 0.345$; $p = 0.558$, slope = 0.005 ± 0.008). This was not caused by the between- and within-flock effects of mass having opposite effects on last visit entry time as neither was significant, and both were positive ($F_{1,113} = 0.571$; $p = 0.452$, slope = 0.015 ± 0.020, and $F_{1,113} = 0.089$; $p = 0.766$, slope = 0.003 ± 0.009, respectively), and they were not significantly different from each other ($F_{1,113} = 0.316$; $p = 0.575$).

Number of visits per individual had a significant and positive effect on last visit entry time ($F_{1,115.31} = 46.083$; $p < 0.001$, slope = 0.225 ± 0.033). The within-flock effect was positive and significant ($F_{1,107.245} = 48.408$; $p < 0.001$, slope = 0.235 ± 0.034), while the between-flock effect was negative, though non-significant ($F_{1,7.808} = 0.012$; $p = 0.917$, slope = -0.016 ± 0.148), and the two effects were not significantly different from each other ($F_{1,8.659} = 2.759$; $p = 0.132$).

Mean visit duration did not have a significant effect on last visit entry time ($F_{1,116} = 0.010$; $p = 0.919$, slope = -0.001 ± 0.013). The between-flock effect was non-significant and negative ($F_{1,9.60} = 0.128$; $p = 0.728$, slope = 0.001 ± 0.015), while the within-flock effect was non-significant and positive ($F_{1,106.54} = 0.008$; $p = 0.931$, slope = 0.001 ± 0.015), and they were not significantly different from each other ($F_{1,16.686} = 0.126$; $p = 0.727$).

Therefore, the birds that had the most visits were more likely to have later last visit entry time, which perhaps suggests that last visit does not provide a useful measure of individual foraging behaviour within a flock, because it mostly reflects the artefact of stopping these flock trials at an arbitrary point in time whilst the flocks were still foraging.

Change in visit duration

Visits increased in duration during the trial ($F_{1,1228.4} = 7.228$; $p = 0.007$, slope = 0.197 ± 0.073). This was caused by a strongly positive within-subject effect (Figure 10, $F_{1,1139.18} = 9.805$; $p = 0.002$, slope = 0.236 ± 0.075), while the between-subject effect was negative and non-significant ($F_{1,\ 93.75} = 2.385$; $p = 0.126$, slope = -0.473 ± 0.306). These two effects were significantly different from each other ($F_{1,105.43} = 5.054$; $p = 0.027$, slope = -0.709 ± 0.316). Therefore, although there is a lot of scatter in the data compared to the effect size, in general birds spent longer at the feeder per visit as the trial went on within each flock (Figure 10).

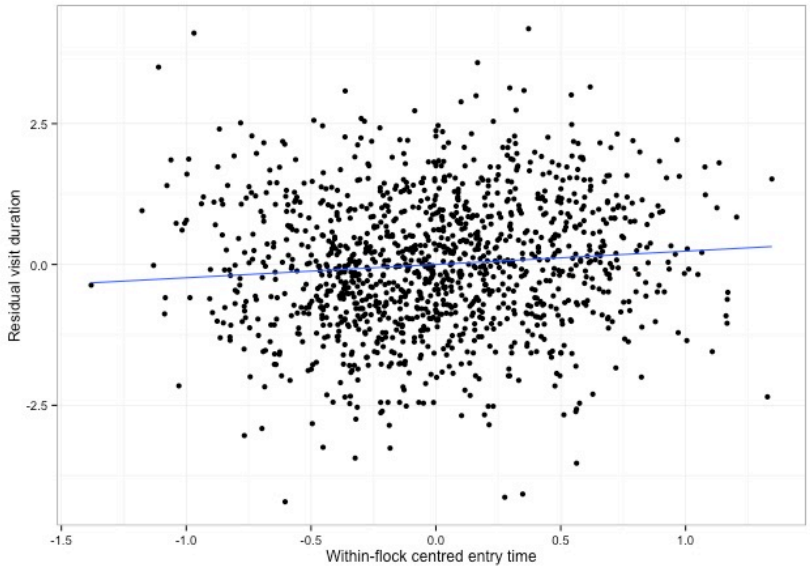


Figure 10. Change in visit duration during the trial. Visit duration is here represented by the residuals from a model with visit duration as response and bird ID nested in flock ID as

random intercept. Entry time is within-subject centred to control out flock effects. Equation of the fitted line: y = 0.236x.

Repeatability

The repeatability of all of these variables was calculated for the sub-set of individuals and flocks that were subjected to a repeat trial of the set-up. Order was controlled for by having it in the models as a fixed effect.

Table 3. Repeatability (R) of the different behavioural variables assayed in the social context assays.

Behaviour	**Repeatability (with 95% CI)**
Number of visits	0.627 (0.224 – 0.765)
Mean visit duration	0.641 (0.325 – 0.806)
Total visit duration	0.764 (0.550 – 0.872)
Change in visit duration	<0.001 (0.000 - 0.089)
First visit entry time	0.001 (0.000 - 0.124)
Aggression	0.240 (0.166 – 0.318)
Body mass	0.954 (0.894 – 0.976)

The repeatability value of the first visit entry time variable was very low, suggesting that it was not the same individuals that came down to the feeder in the repeat trial as it was in the original trial. Pre-trial body mass, on the other hand, was extremely strong, meaning that the lightest individuals before the repeat trial were the same individuals that were the lightest before the original trial. The repeatability of the variables number of visits, mean visit duration and total visit duration were also strong (see Table 3). Aggression had a lower repeatability value.

Correlation matrix

The correlation matrix (Table 4) showed generally significant and large correlations between the different variables, except for the correlations between number of visits and change in visit duration, number of visits and body mass, aggression and body mass, first visit entry time and aggression, and first visit entry time and mean visit duration.

Table 4. Pair-wise correlations between the seven behaviours (raw variables). Significant correlations are marked with grey background.

	First visit entry time	**Total duration**	**Mean duration**	**Change in duration**	**Body mass**	**Aggression**
Number of visits	r=-0.54 p<0.001	r=0.50 p<0.001	r=-0.21 p=0.024	r=0.08 p=0.378	r=-0.07 p=0.452	r=-0.26 p=0.004
First visit entry time	1	r=-0.51 p<0.001	r=-0.16 p=0.075	r=0.30 p<0.001	r=0.19 p=0.039	r=0.09 p=0.348
Total duration		1	r=0.74 p<0.001	r=-0.42 p<0.001	r=-0.37 p<0.001	r=0.37 p<0.001
Mean duration			1	r=-0.55 p<0.001	r=-0.35 p<0.001	r=0.61 p<0.001
Change in duration				1	r=0.29 p<0.001	r=-0.31 p<0.001
Body mass					1	r=-0.18 p=0.058
Aggression						1

Structural Equation Modelling (SEM)

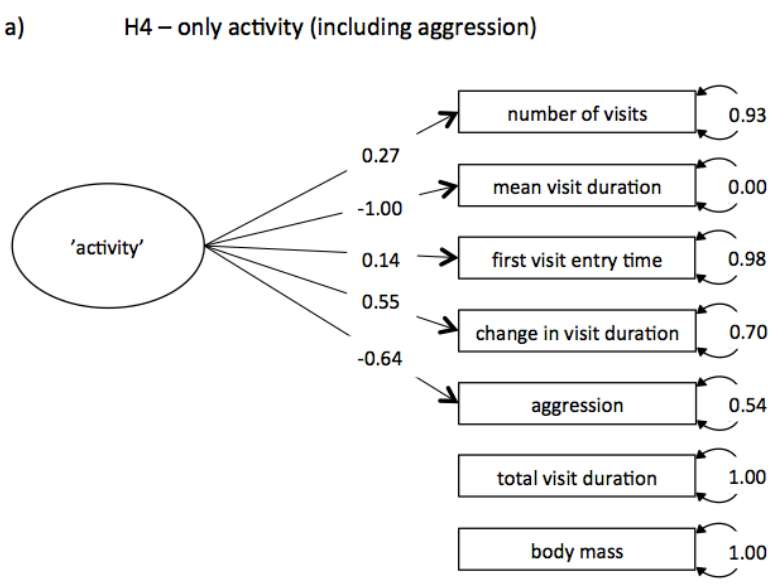
Out of the eight distinct models hypothesised to reflect the covariance structure in the data, only three converged. This suggests that despite all the significant correlations in Table 4, there were insufficient covariance in these structures, given the sample sizes and number of covariances, and it can therefore be concluded that the data is unlikely to conform to these non-converged models (see Fig.4). The models that did converge appropriately were: H0 – no correlations, H4 – activity (including aggression), and H6 – hunger (including aggression).

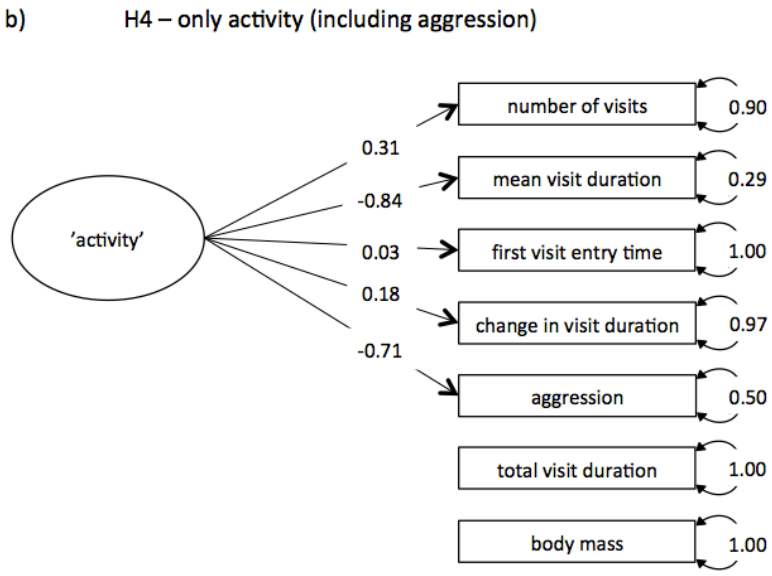
For both the raw and the within-flock centred variables, H4 – activity (including aggression) had by far the lowest AIC value (see Figure 11), thus being the best fit to the data (see Table 5). H6 – hunger (including aggression) had the second best fit, while H0 – no

correlations had the worst. For the variables where sex was controlled for, H2 – activity and hunger (aggression in activity) was the best model, while the three other models ranked in the same order as for the raw and flock-controlled variables.

Table 5. Comparison of the SEM models that converged, using AIC values.

Model	Raw variables			Controlling for flock			Controlling for sex		
	D.f.	AIC	ΔAIC	D.f.	AIC	ΔAIC	D.f.	AIC	ΔAIC
H4	12	2205.56	0.00	12	2236.75	0.00	12	2203.78	0.00
H6	10	2276.59	71.03	10	2269.70	32.95	10	2275.66	71.88
H0	7	2309.98	104.41	7	2290.75	54.00	7	2311.35	107.58





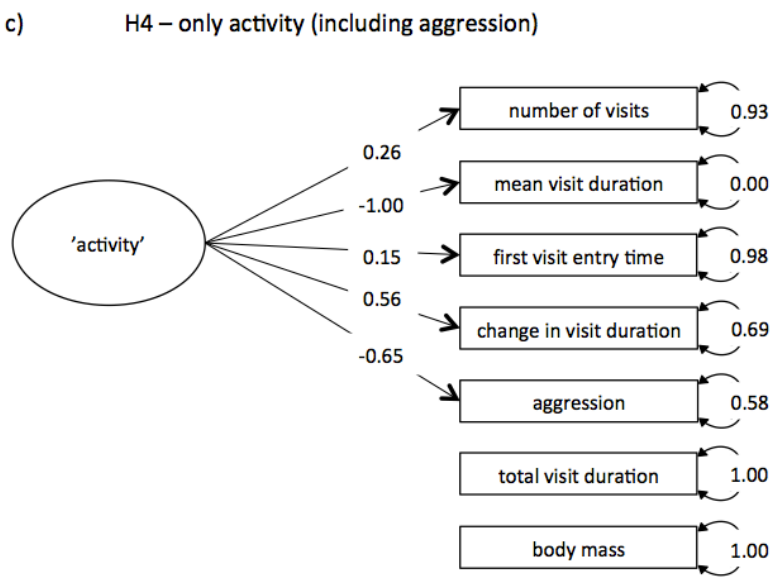


Figure 11. Causal diagrams of the model that fit the data the best: model H4 – only activity (including aggression), fitted with the raw variables (a), the variables where flock was controlled for (b), and the variables where sex was controlled for (c). The single-headed arrows that connect the latent variable (circle) to the observed variables (boxes) indicate causal relationships, and is shown with factor loadings. The double-headed arrows with values next to each of the observed variables show the error variance left unexplained by the latent variable.

In model H4 (see Figure 11), the factor loadings show that the causal relationship between the latent variable and the observed variable number of visits was positive, while the relationship to mean visit duration was negative. This is in accordance with the significant and negative effect of number of visits on mean visit duration found in the univariate models, meaning that individuals had either many and short or few and long visits to the feeder. Note that these two tactics had the same outcome in terms of feeding, as both number of visits and mean visit duration each had significant and positive effects on total visit duration, and hungrier (i.e., lighter) individuals had longer total visit durations (see above). The models where total visit duration and body mass were included in order to see if activity was driven by them did not converge (see for instance models H7 and H8, see figure 5).

The factor loadings and amount of unexplained variance in each observed variable in the best-fitting model were not much different between models fitted with raw and within-flock

centred variables, respectively (see Figure 11). It was also the same three models that fit both the raw and the within-flock centred variables properly, and they ranked from best to worst in the same order (see Table 5). This suggests that there were not any large differences between flocks in the structure of correlations between behaviours. The models fitted on the variables where sex was controlled for were also not much different from the models fitted on the raw variables, and also here was it the same three models that fit the data properly, and in the same order from best to worst. This suggests that there were not any large differences between the sexes in the structure of correlations between behaviours.

Discussion

In this study, we examined the structure of correlations between behavioural measures of birds that were feeding in a social context. By comparing different models, we found that the model which fit the data the best was the one where variables thought to be measures of activity correlated with each other (model H4 – only activity, including aggression, see figure 11).

We found with the univariate models that number of visits had a significant and negative effect on mean visit duration. Therefore, the opposite signs of the factor loadings in the best model are not surprising. Both were also highly repeatable, and all this suggests that these two variables represent alternative individual strategies of feeder use (fewer and longer visits versus many and shorter visits). One possible explanation for this finding is that it represents true personality (i.e., consistent individual differences across contexts). However, Wilson et al. (1994) cautioned that even highly reversible variables can appear just like animal personality in environments that reinforce individual differences. A study by Nelson et al. (2008) found for instance that behaviours that were highly consistent in social contexts were not consistent when assayed in solitary contexts, and they suggested that the crucial environmental factor that caused the consistency in the social context was dominance. Therefore, our findings of alternative individual strategies might not reflect true personality, but rather consistent social interactions within the flocks. It will therefore be very interesting to combine these results with the ones from the solitary context assays on the same individuals (the subject of the 2016 master's thesis of Sindre L. Sommerli), since behaviours assayed in the latter context were ones that we expected there to be consistent individual differences in. These individual differences in solitary behaviour might therefore shed light on whether there actually were individual behavioural differences in the social context assays. The behaviours assayed in the solitary context assays included exploration of novel environment, of a novel object, and of novel food, all of which are measures that either have been termed part of, or are thought to correlate strongly with, "innovative" behaviour (Reader and Laland 2003). There were also measures of general baseline activity (movement around the individual cage), since it could be that some individuals appear to explore or innovate more than others just by virtue of being more active. The social context assays data presented here suggests that there are correlations between behaviours at the feeder and repeatability in

something like activity (i.e. number versus duration of visits to the feeder), and so we might expect general activity in the solitary context to correlate strongly with the measures thought to represent activity in the social context (i.e. the latent variable 'activity' in H4, the best model). It might also be that the individual assay exploration/innovative behavioural measures correlate with this latent variable, but it will be important to factor out any indirect covariance in this case since more active individuals that move around more may appear to explore the novel environment faster or approach more closely objects in it that are novel just as a result of that activity. On the other hand, it might be that we will find consistent differences between individuals in the exploration/innovation measures that cannot be described by the general activity level (or hunger level) of the individuals. This would mean that there actually were some behavioural measures that showed personality in something more that simple activity, but that this was not captured by the current analyses of the social context assays.

There did not appear to be much difference between models fitted to the raw variables and the variables where flock was controlled for. This suggests that the flocks were not different from each other in their behavioural correlations. Therefore, since our findings suggest that the number of visits and mean visit durations do reflect different alternative strategies in feeder use, then flocks have either similar consistent social interactions (see above), or there is little difference between flocks in the true personality of the individuals in them. Possible explanations for the latter can be that the assumption of the flocks constituting discrete entities was not fulfilled, or that their respective wild environments on each farm are not different enough to cause divergent evolution of behavioural correlations in these contexts.

Included in the best model were the variables first visit entry time and aggression, which, of all the behaviours assayed in this experiment, were originally perhaps the two strongest candidates for representing true “personality”. Although first visit entry time might have represented a measure of boldness, especially given the lack of a significant effect of body mass on it, it had a very low repeatability estimate and thus does not fulfil the criterion of a personality trait that it should be consistent across time (Dingemanse and Dochtermann 2013). First visit entry time had a positive and significant effect on total visit duration, which suggests that the ones that came down to the feeder earliest, were also the ones to stay the longest. Similarly, although aggression had a strong correlation with all of the measures of

visit duration in the correlation matrix, and had a high-valued factor loading in the best model, it also had a low repeatability estimate. That aggression correlates so strongly with the measures of visit duration might be because the probability of being involved in aggressive encounters during a visit should presumably increase with visit duration. We used the same type of feeder in the compartment that the birds were held in on the day before the assays as we used in the assays themselves in order to avoid the effect of unfamiliarity with the feeder on the behavioural measures. Therefore, the non-repeatability of first visit entry time, aggression and change in visit duration is not likely to be an effect of a reaction to a novel environment.

As found in the univariate models, the lighter individuals are likely to be the hungrier ones, given the positive effect of total visit duration on seeds eaten on the flock level, negative effect of body mass on mass change, and negative effect of body mass on total visit duration. Body mass was not a part of the best model, and the models where it was included to explain activity (see models H7 and H8 in Figure 5) did not converge, which might suggest that hunger does not explain the activity at the feeder. However, one weakness with SEM is that it is highly dependent on the strength of correlations between measures. We consider it largely unlikely that this was the problem for most of the variables since the correlation matrix showed generally strong correlations between many of them. However, body mass had a weak correlation with both aggression and number of visits, and this can be a reason for the activity models including body mass not converging. It has also been pointed out that SEM requires large sample sizes. The smaller the sample size, the lower is the power to detect patterns in the data, thus increasing the risk of making a type II error (failing to detect a pattern when there really is one). Although there is an ongoing debate on exactly how large the sample size should be, our sample size of 119 individuals is regarded as adequate (see Dochtermann and Jenkins 2010).

A further next step in this work with social behaviour in house sparrows could be to further connect it with ecology and population dynamics, for which this study system is suitable, given the well-developed field logistics and a wealth of data on these individuals. For instance, we could explore whether some behavioural measures are associated with dispersal, as a study on great tits found with the association between dispersal and exploration in captivity (Dingemanse et al. 2003). If it is the case that dispersers in this house sparrow

study system have a structure of correlations in their behaviours that are different from non-dispersers, it would be interesting to see how this affects the dynamics of the flock that receives dispersers. Likewise, it would be interesting to see which structures of behavioural correlations characterises inbred flocks, since a study on this study system found signs of inbreeding depression but not inbreeding avoidance on one island population (Billing et al. 2012).

Conclusion

Our main conclusions from this study are that repeatable individual differences are present in measures of activity at the feeder, and that these measures correlate with each other. More specifically, individuals differed in their specific use of the feeder. Although it is a possibility that these findings are reflections of true ”personality” (i.e., consistent across contexts), this must be confirmed by comparing the results from this study on behaviours in a social context with the results from another study on behaviours in a solitary context on the same individuals. Another possible explanation for these differences is that they are due to consistency of social relationships within these flocks, which could be studied further by exploring its association with for instance dominance hierarchies. This study emphasises the importance of studying behaviours at the flock-level for such social species as the house sparrow, and points towards an interesting interplay between behavioural and population ecology in the future study of these phenomena.

References

Anderson, T. R. 2006. *The Biology of the Ubiquitous House Sparrow: From Genes to Populations.* Oxford University Press, Inc.

Beaujean, A. A. 2014. *Latent Variable Modelling Using R.* Routledge.

Bell, A. M. 2005. *Behavioural differences between individuals and two populations of stickleback (*Gasterosteus aculeatus). Journal of Evolutionary Biology **18**: 464-473.

Bell, A. M. 2007. *Future directions in behavioural syndromes research*. Proceedings of the Royal Society B: Biological Sciences **274**:1611, 755-761.

Bell, A. M., Hankison, S. J., and Laskowski, K. L. 2009. *The repeatability of behaviour: a meta analysis*. Animal Behaviour **77**: 771-783.

Billing, A. M., Lee, A. M., Skjelseth, S., Borg, Å. A., Hale, M. C., Slate, J., Pärn, H., Ringsby, T.-H., Sæther, B.-E., and Jensen, H. 2012. *Evidence of inbreeding depression but not inbreeding avoidance in a natural sparrow population*. Molecular Ecology **21**: 1487-1499.

Budaev, S. V. 2010. *Using principal components and factor analysis in animal behaviour research: caveats and guidelines*. Ethology **116**: 5, 472 - 480.

Dingemanse, N. J., Both, C., van Noordwijk, A. J., Rutten, A. L., and Dent, P. J. 2003. *Natal dispersal and personalities in great tits (*Parus major). Proceedings of the Royal Society B: Biological Sciences **270**: 741-747.

Dingemanse, N. J., Wright, J., Kazem, A. J. N., Thomas, D. K., Hickling, R., and Dawnay, N. 2007. *Behavioural syndromes differ predictably between 12 populations of three-spined stickleback.* Journal of Animal Ecology **76**: 1128-1138.

Dingemanse, N. J., Kazem, A. J. N., Réale, D., and Wright, J. 2010a. *Behavioural reaction norms: animal personality meets individual plasticity*. Trends in Ecology and Evolution **25**:2, 81-90.

Dingemanse, N. J., Dochtermann, N., and Wright, J. 2010b. *A method for exploring the structure of behavioural syndromes to allow formal comparison within and between data sets*. Animal Behaviour **79**: 439-450.

Dingemanse, N. J., and Dochtermann, N. 2013. *Quantifying individual variation in behaviour: mixed-effect modelling approaches.* Journal of Animal Ecology **82**: 39-54.

Docthermann, N. A., and Jenkins, S. H. 2010. *Multivariate methods and small sample sizes*. Ethology **117**: 95-101

Fisher, J., and Hinde, R. A. 1949. *The opening of milk bottles by birds.* British Birds **42**: 347-357

Gosling, S. D. 2001. *From mice to men: what can we learn about personality from animal research?* Psychological Bulletin **127**:1, 45-86

Jensen, H., Moe, R., Hagen, I. J., Holand, A. M., Kekkonen, J., Tufto, J., and Sæther B.-E. 2013. *Genetic variation and structure of House sparrow populations: is there an island effect?* Molecular Ecology **22**: 1792-1805

Lefebvre, L. and Bolhuis, J. J. 2003. *Positive and negative correlates of feeding innovations in birds: evidence for limited modularity.* In: *Animal Innovation* (ed. Reader, S. M. and Laland, K. N.). Oxford University press, Inc.

Nakagawa, S. And Schielzeth, H. 2010. *Repeatability for gaussian and non-gaussian data: a practical guide for biologists.* Biological Reviews **85**: 935-956.

Nelson, X. J., Wilson, D. R., and Evans, C. S. 2008. *Behavioural syndromes in stable social groups: an artifact of external constraints?* Ethology **114**: 1154-1165.

Owens, I. P. F. 2006. *Where is behavioural ecology going?* Trends in Ecology and Evolution **21**:7, 356-361

van de Pol, M. and Wright, J. 2009. *A simple method for distinguishing within- vs between-subject effects using mixed models.* Animal Behaviour **77**: 753-758.

R Core Team. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/.

Reader, S. M., and Laland, K. N. 2003. *Animal innovation: an introduction*. In: *Animal Innovation* (ed. Reader, S. M. and Laland, K. N.). Oxford University press, Inc.

Réale, D., Dingemanse, N. J., Kazem, A. J. N., and Wright, J. 2010. *Evolutionary and ecological approaches to the study of personality*. Philosophical Transactions of the Royal Society B: Biological Sciences **365**: 3937-3946.

Réale, D., Reader, S. M., Sol, D., McDougall, P. T., and Dingemanse, N. J. 2007. *Integrating animal temperament within ecology and evolution.* Biological Reviews **82**: 291-318

Sih, A., Bell, A., and Johnson, J. C. 2004. *Behavioural syndromes: an ecological and evolutionary overview.* Trends in Ecology and Evolution **19**:7, 372 – 378.

Sol, D., Duncan, R. P., Blackburn, T. M., Cassey, P., and Lefebvre, L. 2005. *Big brains, enhanced cognition, and response of birds to novel environments.* Proceedings of the National Academy of Sciences of the United States of America **102**: 15, 5460-5465

Stadler, B. M. R, Stadler, P. F, Wagner, G. P, and Fontana, W. 2001. *The topology of the possible: formal spaces underlying patterns of evolutionary change*. Journal of Theoretical Biology **213**: 241-274

Stamps, J. A. *Why evolutionary issues are reviving interest in proximate behavioural mechanisms.* American Zoologist **31**: 338-348.

Tufto, J., Ringsby, T.-H., Dhondt, A. A., Adriaensen, F. and Matthysen, E. 2005. *A parametric model for estimation of dispersal patterns applied to five passerine spatially structured populations.* The American Naturalist **165:** 1, E13-E26.

Wilson, D. S., Clark, A. B., Coleman, K. and Dearstyne, T. 1994. *Shyness and boldness in hmans and other animals.* Trends in Ecology and Evolution **9**: 442-446