,

Neophobia across social contexts in juvenile herring gulls

Reinoud Allaert[1,2*], Sophia Knoch[2,3*], Simon Braem[1,2], Dries Debeer[4], An Martel[2,5], Wendt Müller[6], Eric Stienen[7], Luc Lens[1,2], & Frederick Verbruggen[1,2,3]

[1] Dept. Biology, Ghent University
[2] Centre for Research on Ecology, Cognition and Behaviour of Birds
[3] Dept. Experimental Psychology, Ghent University
[4] Research Support Office, Faculty of Psychology and Educational Sciences,
[5] Dept. Pathobiology, Pharmacology and Zoological Medicine, Ghent University
[6] Dept. Biology, Antwerp University
[7] Research Institute for Nature and Forest

*These authors contributed equally (shared first authorship).

Correspondence: reinoud.allaert@ugent.be

Abstract

Neophobia, the fear or avoidance of the unfamiliar, can have significant fitness consequences. It is typically assessed by exposing individuals to unfamiliar objects when they are alone, but in social species the presence of conspecifics can influence neophobia. However, previous research on the effect of group dynamics on neophobic responses has produced mixed results. Here, we explored the degree of neophobia of an individual in different social contexts in a highly social species, the herring gull. To this end, we exposed juvenile herring gulls (N = 54) to novel objects in both individual and group settings (4-5 individuals), replicating each condition twice. Individuals tested in groups were quicker to eat, and spent more time near a novel object than individuals tested alone. The results of our study suggest that the presence of group members reduces perceived individual risk, allowing individuals to behave less cautiously. Preregistered Stage 1 protocol: https://osf.io/u4b7q (date of in-principle acceptance: 17/05/2024)

Keywords: Animal Behaviour, Behavioural Inhibition, Neophobia, Social Behaviour, Herring Gull, Animal Personality

Introduction

Neophobia is the fear or reluctance to engage with new or unfamiliar objects, places or scenarios. It is often considered to be a consistent personality trait across species, affecting an individual's survival and adaptation (Both et al., 2005; Greggor et al., 2015; Kimball and Lattin, 2023; Vrublevska et al., 2015). Research into animal behaviour is increasingly focusing on neophobia because of its significance in the context of rapid environmental change. The world is rapidly urbanising, with the footprint of urban land cover expected to at least double by the end of the century (Gao and O'Neill, 2020). Many species must therefore adapt to human-induced changes in their environment, and hence, to unfamiliar scenarios (Lee and Thornton, 2021; McKinney, 2002). In such situations, neophobia can, on the one hand, serve as a survival mechanism, allowing individuals to avoid potential threats and increase their chance of survival (Greenberg and Mettke-Hofmann, 2001). On the other hand, excessive aversion to novelty can restrict exploratory behaviour, limiting an individual's ability to locate and exploit novel resources, learn from its novel environment and adapt to environmental changes (Biondi et al., 2010; Greenberg, 2003).

To assess neophobia, individuals are typically exposed to novel food, objects, or spaces (Greggor et al., 2015; Mettke-Hofmann, 2017). For example, in the 'novel object task', an individual encounters an unfamiliar object, often placed next to a food reward, in a familiar environment. The latency to approach the food (in the presence of the novel object) or to interact with the novel object itself, is then used as a measure of neophobia (Greggor et al., 2015; Miller, Lambert, et al., 2022; Vernouillet and Kelly, 2020). These measures have been used in cross-species comparisons to investigate, for example, the socio-ecological drivers of neophobia (Mettke-Hofmann et al., 2002; Miller and Lambert, 2024; Miller, Lambert, et al., 2022), or within species, to investigate both the causes and consequences of individual differences in neophobia (Greenberg and Mettke-Hofmann, 2001).

Most research on neophobia has focused on individual animals, both in laboratory and field settings. However, it is important to consider that many species are to various extents reliant on social information, so individuals can influence each other's behaviour. This is also true in the context of adapting to environmental changes and urbanisation (Lee and Thornton, 2021). For instance, when individuals encounter a new environment, they may learn from others about appropriate roosting or nesting sites, food sources, or unfamiliar predators (Harel et al., 2017; Keen et al., 2020; Loukola et al., 2012). In this context, several studies suggest that the presence of conspecifics also influences neophobia. However, the mechanisms behind this social phenomenon are still a topic of debate due to the various patterns that have been observed.

First, some studies have found that individuals in groups are generally less neophobic than when tested alone. For example, Coleman and Mellgren presented zebra finches (*Taeniopygia guttata*) with novel feeders and decorated the feeders with novel objects (Coleman and Mellgren, 1994). Individuals in a group approached and started using the new and decorated feeders more quickly than when tested alone. Other studies reported similar patterns in different species for some (but not necessarily all) measures of neophobia (Benson-Amram and Holekamp, 2012; Kareklas et al., 2018; Moretti et al., 2015; Soma and Hasegawa, 2004). Such mitigating effects of social context on neophobia may be attributed to 'risk dilution' (Krause and Ruxton, 2002) or 'social buffering' (Kikusui et al., 2006). These theories predict that neophobia, or fear responses in general, are reduced in the presence of others, as individuals in a group collectively share the potential risks associated with novel situations or threats. This shared risk perception will also lead to more uniform behaviour within the group, as individuals adapt their actions in response to the behaviour of conspecifics.

Second, some studies found the opposite pattern. For example, common ravens (*Corvus corax*) and carrion x hooded crows (hybrid; *C. corone, C. cornix*) approached novel objects faster when alone than when accompanied by a conspecific (Miller, Bugnyar, et al., 2015; Stöwe, Bugnyar, Heinrich, et al., 2006a; Stöwe, Bugnyar, Loretto, et al., 2006). Other studies have observed similar patterns in other species, including Indian mynahs, *Acridotheres tristis* (Griffin, Lermite, et al., 2013), house sparrows, *Passer domesticus* (Kelly et al., 2020), and even zebra finches (Kerman et al., 2018; St. Lawrence et al., 2021), thus failing to replicate the findings of the

aforementioned study by Coleman and Mellgren (1994). Interestingly, however, some of these studies found that once individuals reached the novel object, they spent more time interacting with it when in the presence of others (either in pairs or in groups) than when isolated (Miller, Bugnyar, et al., 2015; St. Lawrence et al., 2021; Stöwe, Bugnyar, Heinrich, et al., 2006a). It has therefore been suggested that the slower approach latencies may be due to conspecifics 'negotiating', by using behavioural cues to coordinate their actions and deciding who will approach the novel object first. Consequently, this may lead to a convergence of individual behaviours, as group members align their actions based on these cues.

Third, some studies failed to find effects of social context on average neophobic responses altogether (e.g Apfelbeck and Raess, 2008). While it is of course possible that social context does not matter for some species, it is also possible that the presence of conspecifics alters behaviour of individuals without changing the mean response. Specifically, in environments where conspecifics' behaviour serves as an indicator of appropriate responses, individuals may adjust their own behaviour to match that of others (Herbert-Read et al., 2013). This synchronisation of behaviours within the group, or 'social conformity', enhances cohesion and helps the group to adapt to their environment. Observations in a variety of species, such as zebra finches (Schuett and Dall, 2009) and gouldian finches, *Erythrura gouldiae* (King et al., 2015), show how individuals adapt their behaviour and mirror their partners' character traits. For instance, if a gouldian finch exhibited bold behaviour, the observing individual tended to become bolder as well, while if the partner displayed shyness, the observing individual mirrored this trait (King et al., 2015). Thus, this study found that the neophobic response was similar on average for individuals tested alone or in pairs, but there was less variation between individuals in the paired condition compared to the alone condition.

Current study The aim of this study is to investigate if and how the social context affects neophobia in the herring gull (*Larus argentatus*). Gulls' natural coastal habitat is rapidly disappearing, forcing them to live closer to humans in urban environments and to rely more on anthropogenic food sources (Coulson, 2015; Nager and O'Hanlon, 2016). Although reports in popular media may suggest that herring gulls are generally not neophobic due to their approach towards humans or stealing food, such anecdotes do not necessarily reflect the species' behaviour at a population level (Inzani et al., 2023). In fact, widely differing levels of neophobia as well as individual differences therein exist within populations (Inzani et al., 2023). The latter finding suggests that for some individuals, it might be easier to adapt to environmental change and urbanisation than for others. Indeed, there is considerable intraspecific variation in how herring gulls utilise urbanised areas, ranging from minimally to almost complete dependence (O'Hanlon et al., 2017; Pavlova and Wronski, 2020).

Herring gulls are a highly social species, utilising cues not only from conspecifics, but even from other species, including humans. This suggests that social learning is a key aspect of gull behaviour (Feist et al., 2023; Frings et al., 1955; Gandolfi, 2009; Goumas et al., 2020). Thus, when assessing their neophobia, it is important to do this not only in an individual context, but also in a social (group) context. Based on previous findings (as reviewed above), we predict that the distribution of neophobic responses will depend on the social context. However, the direction of the effects will depend on the social mechanisms at play. In Figure 1, we provide a template for testing the three different hypotheses of group effects, taking into account two measures, namely the average neophobic response and the variance between individuals.

Overall, we predict that there will be lower variance between individuals when they are tested in a group, compared to when they are tested alone. After all, all of the major hypotheses discussed above assume that individuals become more similar to each other by spreading risk, jointly buffering stress, negotiating with each other, or simply through social conformity. However, there are three possible scenarios regarding the average neophobic response. First, the 'risk dilution' hypothesis predicts that herring gulls will be *less* neophobic on average when in a group compared to when they are alone (scenario A in Figure 1). Second, the 'negotiation' hypothesis predicts that individuals will be *more* neophobic when in group (scenario B in Figure 1). Third, according to the 'social conformity' hypothesis, individuals will tend to mimic one another's behaviours—those who are neophobic will show a decrease in their fear of novel objects when surrounded by others who are less neophobic, and vice versa (scenario C in Figure 1). Thus, in this third scenario, there is a reduction of

variance but no change in the average neophobic response. These three predictions are contrasted with the null hypothesis that social context does not modulate variance, or group means ('Null Hypothesis', Figure 1).

To test these predictions, juvenile herring gulls were subjected to four distinct conditions: individual or group tests paired with a control or novel object. Each condition was repeated twice. The guidelines for designing neophobia tests of Greggor et al. (2015) were followed, and a within-subject design with a relatively large sample size (N = 67 individuals) was chosen to further increase the statistical power of the study. One additional reason for the inconsistent previous findings is that sample size was relatively low in many studies (see also Farrar et al., 2020). In addition, the herring gulls used in this study were raised by hand from the egg to control for sampling bias, a recurring issue when testing wild animals. After testing, they were released in the wild.

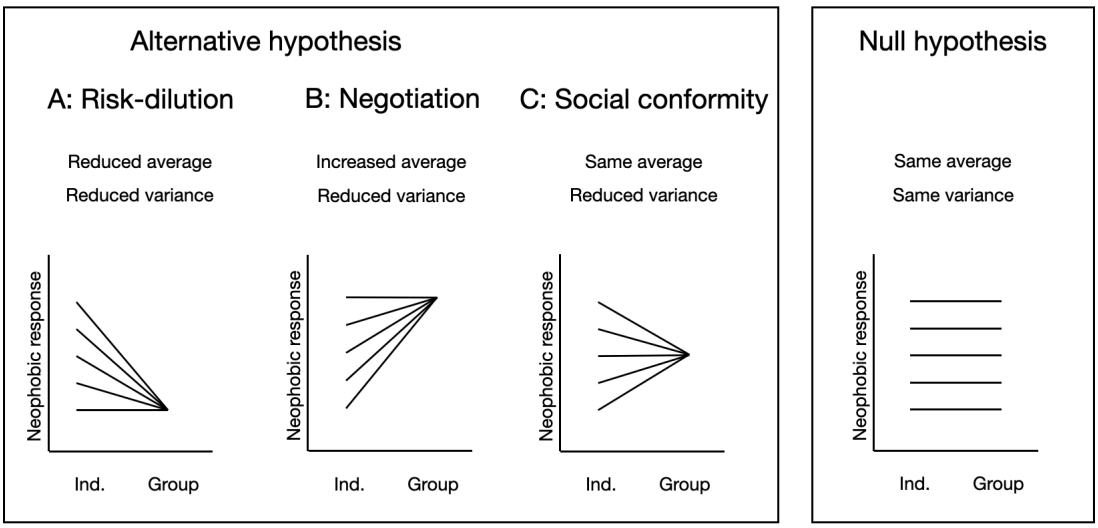

Figure 1. Overview of hypotheses

Material and methods

Sample size

We originally planned to test 80 herring gulls twice across a 2x2 design (thus eight tests per individual; see above). We performed an *a-priori* power sensitivity analysis using G*Power (Erdfelder et al., 2009), for a repeated measures MANOVA with three within-subject factors: *context* (with levels group and individual), *object* (with levels control and novel object), and *trial* (with levels 1 and 2). Our initial analysis indicated that a sample size of 80 would be sufficient to detect small main effects of *context, object, and trial* (Cohen's f effect size of 0.11 (Cohen, 2013); Power = 0.80 ; correlation among repeated measures = 0.5), as well as an interaction between context and object with small effect size (0.11; Power = 0.80 ; correlation among repeated measures = 0.5). We reared the gulls from the eggs (see the 'Subject section' below) and we anticipated that in some cases herring gull eggs would be mistaken for those of the phylogenetically and ecologically related lesser black-backed gull (LBBG) during egg collection. For logistical reasons, the chicks could only be identified to the species level after testing by visual inspection of plumage differences. To mitigate the potential reduction in sample size (due to the exclusion of LBBGs), we conducted a second *a-priori* power analysis accounting for a potential 10% dropout rate. This a-priori analysis revealed that even with a 10% reduction, our study would still have sufficient statistical power (Cohen's f effect size of 0.17) to detect significant effects.

Due to unanticipated mortality, we were only able to test 67 birds (instead of the registered 80). Of these 67 birds, 13 individuals were later identified as LBBG (a higher percentage than we had anticipated) and were

excluded from further analysis in accordance with the registered protocol as there may be differences in neophobic responses between migratory (i.e. LBBG) and non-migratory (i.e. herring gulls) species (Miller and Lambert, 2024). This further reduced our final sample size (N = 54). Although this is a significant reduction from our planned sample size (N = 72 after exclusion of LBBG), it is important to note that our sensitivity analyses were based on repeated measures MANOVAs (within subjects factors). This type of analysis does not take into account the additional flexibility offered by (G)LMMs, which are not currently covered by G*Power or most other power estimation tools. The mixed effects models used in this study (in line with the registered protocol) are more robust and better equipped to deal with unexplained variance than the fixed effects MANOVAs used in our sensitivity analysis. Thus, despite the reduction in sample size, our proposed mixed-effects models are expected to retain sufficient power to detect the effects of interest. An overview of the group composition, including the number of LBBG individuals and the sex distribution, is provided in Supplementary Table 2.

Subjects

Egg Collection and Incubation

The herring gulls used in this study are part of a larger research project and were raised and tested at the avian research facilities of Ghent University (Lab number LA1400452), located at the Wildlife Rescue Centre (WRC) in Ostend, Belgium. Eggs were collected in May and June 2024, from nests of roof-breeding parents, by the Research Institute for Nature and Forest (INBO) under the license of the Agentschap voor Natuur en Bos (ANB) and the 'gull patrol' team, authorised to remove eggs along the Belgian coasts for nuisance prevention. Collected before the pipping stage, the eggs were transported to the WRC under stable conditions for further incubation, using Brinsea Ova-Easy incubators (temperature = 37.5°C; humidity = 45%). Upon arrival, eggs were marked with a unique nest identifier and the two largest eggs, which are typically the first laid eggs of a clutch (Parsons, 1972), were incubated. They were checked twice daily for small cracks, indicating pipping. Eggs showing signs of pipping, were moved to a MS700U Hatchery (temperature = 37.2°C; humidity = 50%).

Chick Rearing

Once hatched and fully dried, the chicks received a unique combination of colour rings for identification. Feather samples were collected for sex determination via PCR, following the protocol outlined by Fridolfsson and Ellegren, 1999. This method targeted the CHD1W and CHD1Z introns using 2550F/2718R primers, with PCR conditions set to 30 cycles at an annealing temperature of 56°C. The chicks were then housed in groups of 10 in boxes with netting bottoms (size = 120 x 60 x 60cm, LWH) within heated rooms (ambient temperature= 15-25°C; humidity=40%-80%; under natural light conditions). Each box contained a heating plate (30 x 30cm). The semi-precocial chicks were hand-fed small pieces of fish and dog pellets soaked in water, supplemented with Akwavit, a complementary feed specially developed for fish eating animals (Kasper Faunafood, The Netherlands). Food was available *ad libitum*. Once the chicks were at least 5 days old and their weight exceeded 60 grams, they were moved to outside enclosures (size = 500 x 205 x 265cm, LWH), housed in stable groups of 8-10 individuals. Outside, heating plates were provided during the first few days if night-time temperatures were forecasted to drop below 5°C, or in the event of adverse weather conditions such as heavy rain or storms. Food consisted of a mixture of dog pellets soaked in water and fish, provided 4 times per day, following the default policy at the WRC. Water was provided *ad libitum*. Individuals were tested when they were approximately 30 days old, shortly before they reached fledging age. After testing, the birds were moved to a large flight cage (approximately 180m^2) for dehabituation from handling. Once they were 8-10 weeks old, birds were released in the wild, and a subset (n = 23) received a GPS-device.

Behavioural Test: Novel Object Task

Task Design: For testing purposes, each home enclosure containing 8-10 birds was pseudo-randomly divided into two stable testing groups of four to five individuals that were familiar with each other. Within these

subgroups, we ensured that nestmates were not placed in the same testing group. This arrangement allowed to maintain consistent housing conditions when not testing, while ensuring that testing sessions consistently involved the same subgroups of four to five individuals.

In the 'novel object' condition, birds were exposed to a pseudo-randomly selected novel object. Conversely, in the 'control object' condition, birds were exposed to a familiar object. By placing a familiar object behind the food plate in the control condition, we ensured that responses in the 'novel object' condition were elicited by the novelty of the object and not just the presence of the object itself (see e.g. Greggor et al., 2015, for justification). The familiar object remained in place throughout the testing and habituation period to avoid dishabituation from the familiar object. It was replaced by the novel object only during the novel object testing sessions. To preserve the integrity of the experimental design, the novel object introduced in each of the four sessions was unique, thus each bird's interaction with it marked their first encounter. The experimental timeline spanned from late June to mid-July, and lasted 8 consecutive days. We used five objects (Figure 2) of similar size (approximately the same size as a four weeks old gull), but of different colour, form and texture.

Figure 2. Novel or control objects.

Prior to the Task: In preparation of the novel object task, and following a series of cognitive tests as part of another study (three tests in total), the test setup (Figure 3) was introduced into the birds' home enclosure when the birds were not present. This setup included the pre- and post-testing pens, the start area, and one of our five pseudo-randomly selected objects, which later acted as the control object in the neophobia assessments. After having introduced the test setup, birds were allowed to accustom to the presence of the test apparatus for a period of six days. This habituation period minimised any potential stress towards a new environment, which may influence the behavioural outcome of the test trials.

In order to distinguish the birds when they were being tested in a group, each individual was given a unique marking (marker pen, Raidex) a few days before the test, which could be easily detected by a roof-mounted camera, as colour rings were not visible in the video recordings.

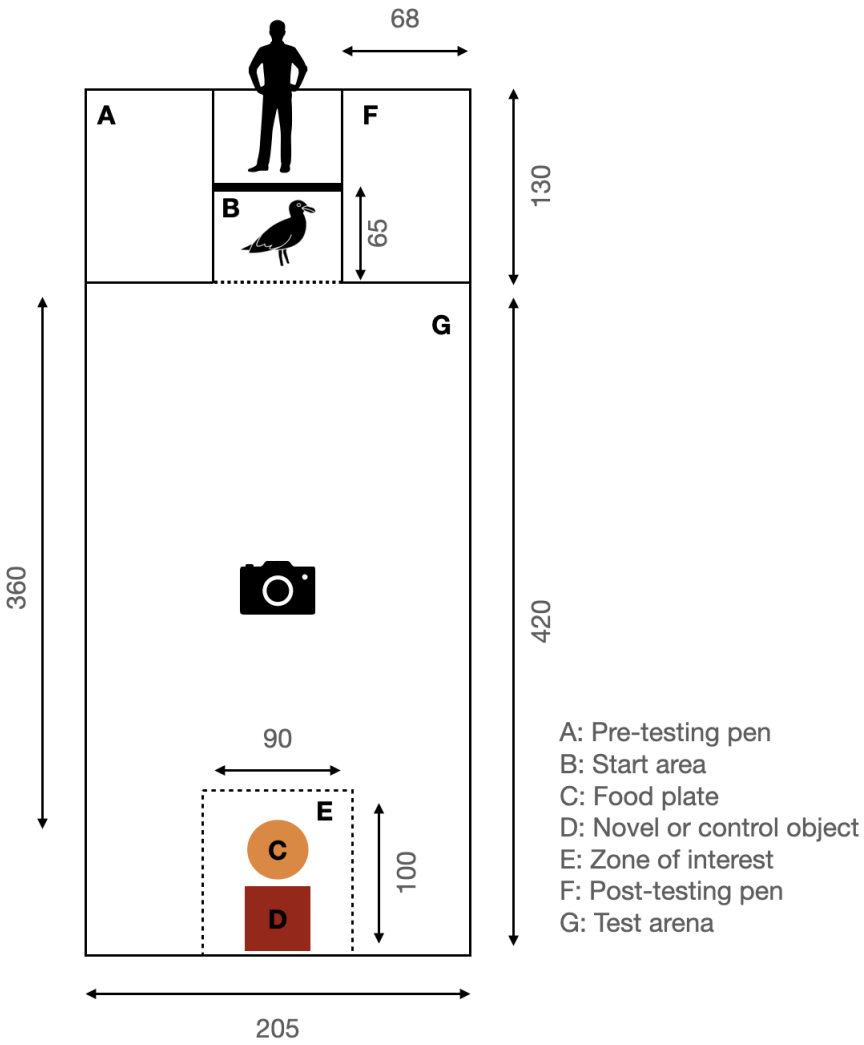

Figure 3. Test setup in home enclosure.

Testing Protocol: The testing commenced after the six-day habituation period. Order of conditions was counterbalanced to incorporate control and novel object conditions, as well as individual versus group settings, with the entire sequence being repeated twice. The animals were food deprived since their last feeding moment the evening before each test at 5:30 PM, to reduce motivational differences before testing. Testing began around 7:30 AM and was completed around 11 AM. In both group and individual settings, individuals were given a maximum of 10 minutes to leave the start area and enter the test arena. Once an individual entered, the trial duration was a fixed 10 minutes. During this period, individuals had the opportunity to feed, but the trial continued for the full 10-minute duration regardless of whether the bird first touched the food. This approach aligns with previous novel object studies (Brown and Nemes, 2008; Bruijn and Romero, 2021; Lecuelle et al., 2011).All tests were recorded with roof-mounted cameras.

Prior to testing, all the birds were moved to the pre-testing holding pen. Next, a food plate (27 cm in diameter), completely filled with fish, and an object (novel or control, depending on the condition) were placed at the back of the enclosure, with the food plate placed in front of the object to rule out directional preference. A single bird, or group of birds, depending on the social context, was placed in the start area. The tester lifted the door of the start area after 15 seconds and left, giving the bird(s) access to their home enclosure (Figure 3). The first 10-minute period started when the door began to move, the second 10-minute period started for each bird individually when it left the start area. The test session ended 10 minutes after the bird had left the

start area in individual trials, or after all birds had left the start area in group trials. Next, the tester moved the tested bird(s) to the post-testing holding pen and started a new test with a new (group of) bird(s).

Data processing and analysis

Video coding. We coded all videos using the free, open-source software BORIS (Behavioural Observation Research Interactive Software) (Friard and Gamba, 2016). Four events were coded, namely 'start of trial', 'test arena entry', 'eating', and 'zone of interest' (see Table 1 for full descriptions). Based on the coded events, we determined latencies and cumulative times. By extracting the time difference between 'start of trial' and 'test arena entry', we determined the latency to leave the start area (Figure 3). In order to determine the latency to approach the food, we extracted the time difference between 'test arena entry' and 'eating'. Time spent in the zone of interest (i.e. in proximity to the food reward and/or novel object, see Figure 3) was calculated as the cumulative time over the length of the trial. If an individual did not perform one of the target behaviours, we assigned the maximum latency, representing the full task duration (in seconds), to that behaviour. For example, the behaviour 'test arena entry' has a latency of 600 seconds if an individual did not enter the test arena. This maximum latency applies only to latency measures; for time spent in the zone of interest (ZOI), a value of 0 was recorded if a bird did not enter the ZOI. For the group tests, we followed each bird individually to code their behaviours.

Video coding was conducted collaboratively by multiple experimenters, with 20 percent of all videos being double-coded by a third experimenter to assess inter-rater-reliability (IRR) using Cohen's Kappa. This third coder was blinded to the original coding decisions and the type of the objects (control or novel), although they were not blind to the overall study aims. Our analysis resulted in a Cohen's Kappa of 0.89, which indicates strong agreement between coders (McHugh, 2012).

Table 1. Ethogram of behaviours that were coded in BORIS. The 'zone of interest' was defined as a fixed rectangle that included the object and the food bowl. To ensure comprehensive observation coverage, this area was expanded by the approximate body length of a 4-week-old gull (30 cm). This ensured that all relevant activities within and around the novel object were captured.

Action	Definition
Start of trial (Point event)	Moment the door starts moving.
Test arena entry (Point event)	When the entire bird is outside the start area.
Eating (Point event)	When the beak touches the food.
Zone of interest (State event)	The bird is considered to enter or leave the zone of interest when the front half of its body crosses the (notional) boundary of the zone.

Statistical analysis

Statistical analyses were conducted using R, version 4.4.1 (R Core Team, 2021). All package version numbers are documented and managed using the `renv` package (Ushey and Wickham, 2024). Mixed-Effects Models (LMMs) were fitted using the `lme4` package (Bates et al., 2015), and parameter estimation along with p-values were calculated using the `lmerTest` package (Kuznetsova et al., 2017), via Satterthwaite's degrees of freedom method. Model assumptions, including normality and heteroscedasticity, were assessed using the `performance` package (Lüdecke et al., 2021), and transformations (log or Box-Cox) were applied where necessary.

Initial diagnostic plots indicated non-normality of residuals and heteroscedasticity in the models. To address these violations, we followed a structured approach. Each dependent variable was first fitted using the raw data. When this did not meet assumptions, a log transformation was applied. This approach sufficiently improved model fit for the ZOI duration model, and the model was refitted accordingly. For the latency to enter and latency to eat models, however, the log transformation did not resolve assumption violations. In these

cases, a Box-Cox transformation was implemented, with optimal lambda values determined using maximum likelihood estimation. The estimated lambda values were $\lambda = -0.869$ for latency to enter and $\lambda = -0.828$ for latency to eat. The models were refitted using the Box-Cox transformed dependent variables, leading to improved model assumptions.

The primary objective of the analysis was to determine whether neophobic responses differed between individual and group trials. LMMs were fitted to different latency measures (latency to enter, latency to eat, and ZOI duration) under appropriate transformations (log or Box-Cox). Models were selected based on the best fit and diagnostics, with Type III sum of squares used to ensure appropriate partitioning of variance for the fixed effects. Key fixed effects included *object* (control vs. novel objects) and *context* (individual vs. group). We added *trial* as a fixed effect to account for repeated testing. Additionally, sex was contrast-coded, and included as a fixed effect to account for potential differences between males and females. For two individuals with missing data, one where the PCR failed and another where the sample was lost, a value of 0 was assigned. Initially, we fitted a full random effects structure (in line with the preregistered report) accounting for variability at the *NestID*, *GroupID*, and *BirdID* levels, with specific terms for individual (*indiv_dummy*) and group (*group_dummy*) conditions to capture within-subject and within-group variation. The full random effects model was:

$$\begin{aligned}\text{Latency} \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (1|\text{NestID}) \\ &+ (-1 + \text{Group\_dummy}|\text{GroupID}) \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{1}$$

However, the full random effects structure, outlined in the preregistered report, led to over-parameterisation. Consequently, non-significant interactions were dropped to simplify the model. In addition, the final models were simplified by including only random intercepts for *BirdID*, while retaining the dummy variables where the model allowed it. This approach effectively captured individual-level variability in both individual and group conditions, avoiding over-fitting. The final model structure for each latency measure is as follows:

$$\begin{aligned}\text{Box-Cox(Latency to enter)} \sim &\text{Object} + \text{Context} + \text{Trial} + \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{2}$$

$$\begin{aligned}\text{Box-Cox(Latency to eat)} \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Sex} \\ &+ (1|\text{BirdID})\end{aligned} \tag{3}$$

$$\begin{aligned}\text{Log(ZOI duration)} \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{4}$$

For models fitted on Box-Cox transformed latency data, transformation parameters were estimated using the `MASS` package (Venables and Ripley, 2002). Marginal means for the fixed effects (context and object) were computed and back-transformed to the original scale (seconds) using the Box-Cox inverse transformation, or an inverse log transformation for eating latency, with the `emmeans` package (Lenth, 2024). Random effect variances for individual (*indiv_dummy*) and group (*group_dummy*) contexts were extracted from the model outputs using the `lme4` package (Bates et al., 2015). To aid interpretation, we back-transformed the random effects by simulating random effects for 1000 individuals under both conditions (individual and group), while accounting for the covariance between *indiv_dummy* and *group_dummy* using the `mvtnorm` package (Genz et al.,

2020). These simulated random effects were then combined with the predicted fixed effects for each condition (individual vs. group, control vs. novel object) and back-transformed to the original scale. The inverse Box-Cox transformation was applied to latency to enter and ZOI duration, while the inverse log transformation was applied to latency to eat.

We also fitted a multivariate model on the combined dataset for *latency to enter* and *latency to eat*, using a contrast for behaviour type (*eat_vs_leave_contrast*) to account for potential correlations between the two outcomes. The multivariate model confirmed the findings from the univariate analyses, indicating consistent effects across both latency measures. However, for ease of interpretation, the results of the univariate models are presented here, as they allow a more straightforward interpretation of the individual effects of each predictor on the dependent variables. Although the multivariate model is not discussed in detail, a full walk-through of all intermediate models, including the preregistered version of the statistical analysis and the multivariate model results, is provided in the supplementary material. For exploratory purposes and to determine the robustness of our findings, we reran the analyses including all LBBG data. These analyses produced very similar results to those reported below (see Supplementary Table 6).

Post-hoc analyses of significant interactions were performed using estimated marginal means via the `emmeans` package (Lenth, 2024), with appropriate back-transformations applied for models with transformed dependent variables. Random effect variances were compared between individual and group trials using likelihood ratio tests to assess whether separate variance components were warranted for each condition. Binary predictors were contrast-coded as (-0.5 vs 0.5) to optimise interpretability. Multicollinearity concerns were minimal due to the balanced nature of the predictors, so variance inflation factor (VIF) assessments were not required. Finally, model assumptions were verified through diagnostic plots, and pairwise comparisons for significant findings were adjusted using Bonferroni-Holm corrections.

Results

Descriptive statistics for each dependent variable across the four experimental conditions are summarised in Table 2. The table includes the means, standard deviations, minimum and maximum values (all in seconds), and the number of non-responses (instances where birds did not perform the target behaviour) for each condition.

Table 2. Descriptive statistics for each dependent variable across the four experimental conditions.

Variable	Condition	Mean (s)	SD (s)	Min (s)	Max (s)	Non-responses
Latency to enter	Group-control	2.34	1.41	0.87	9.40	0
	Group-novel	2.58	7.52	1.08	27.92	0
	Individual-control	11.30	53.14	1.16	335.56	1
	Individual-novel	23.35	79.10	0.93	494.57	2
Latency to eat	Group-control	3.29	6.03	1.93	20.50	0
	Group-novel	18.59	43.73	1.72	239.68	2
	Individual-control	22.94	27.61	2.56	66.13	3
	Individual-novel	146.62	118.90	2.73	367.37	24
ZOI duration	Group-control	114.43	135.42	8.20	465.87	0
	Group-novel	180.59	135.55	9.83	597.23	2
	Individual-control	153.28	152.78	10.08	598.13	2
	Individual-novel	104.16	148.34	2.27	595.44	16

All results are reported using both the transformed values (Box-Cox or logarithmic) and back-transformed

values (Mean [M], SE, CI) to facilitate interpretation on the original scale. Box-Cox transformations were applied to address non-normality and heteroscedasticity for *latency to enter* and *latency to eat*, while a logarithmic transformation was used for *zone of interest (ZOI) duration*. In the main text, we report back-transformed values for ease of interpretation, and p values for the statistical tests. The complete output of the mixed-effects models can be found in Table 3.

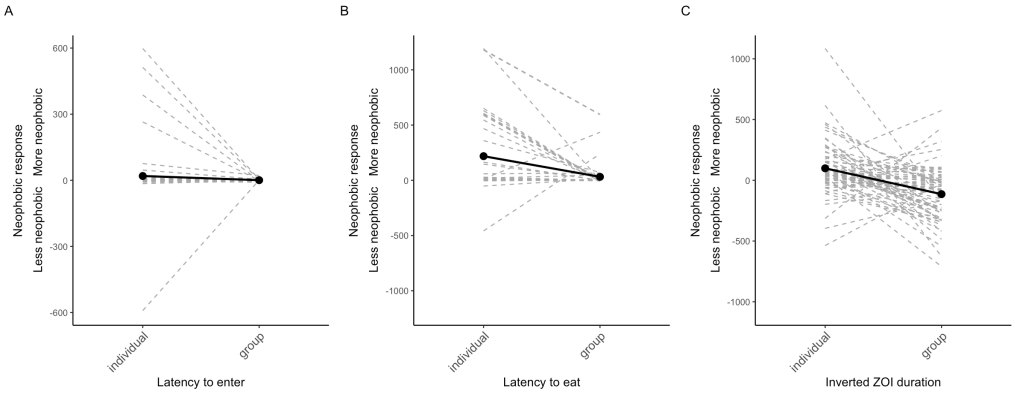

Figure 4. Difference plots of raw values illustrating changes in neophobic response from *individual* to *group* contexts for each dependent variable (i.e. latency to enter, latency to eat, ZOI duration). Black lines show the average response whereas dotted lines show the individual responses to illustrate the variance. In particular, plot A (latency to enter) shows a reduced variance but not a reduced average in neophobic responses across social contexts. Note that individual variation is partly masked, as lines are plotted on top of each other. Plot B (latency to eat) illustrates a reduced average neophobic response across social contexts, whereas we could not test for a reduction of variance. Plot C (ZOI duration) indicates a reduced average and a reduced variance in neophobic responses across social contexts.

Latency to enter

Significant effects of *context* (individual vs. group) ($p = 0.04$) and *sex* ($p < 0.01$) were found on the birds' latency to enter the test arena, while *object* (novel vs. control) did not have a significant effect ($p = 0.47$). Specifically, birds tested in the *group* context entered the arena significantly faster (back-transformed $M = 1.91$ s, $SE = 0.05$, 95% CI $= [1.81, 2.02]$) than those tested *individually* (back-transformed $M = 2.07$ s, $SE = 0.09$, 95% CI $= [1.91, 2.26]$). On average, males entered the arena more quickly than females ($M = 1.84$s, $SE = 0.068$, 95%, CI $= [1.71, 1.98]$ for males; $M = 2.15$, s, $SE = 0.089$, 95%, CI $= [1.99, 2.34]$ for females).

Variance analysis revealed greater individual variability in latency to enter the test zone when birds were tested *individually* (back-transformed $\sigma^2 = 2.72$s^2, SD = 1.65s) compared to when they were tested in a *group* (back-transformed $\sigma^2 = 1.22$s^2, SD = 1.10s). A likelihood ratio test confirmed that this reduction in variance in the group context was statistically significant ($\chi^2(2) = 11.8$, $p = 0.02$). These findings suggest that birds' behaviour was more consistent when tested in groups.

Moreover, the estimated correlation between the individual and group random effects was high (Corr = 0.82), indicating strong repeatability of behaviour across both contexts. Birds that entered quickly when tested alone also tended to enter quickly when tested in groups.

Latency to eat

Latency to eat was significantly influenced by *context* ($p < 0.001$) and object ($p < 0.001$). We also found a significant interaction between *context* and *object* ($p = 0.02$). As shown in Figure 4, the effect of the novel object on latency to eat was more pronounced when birds were tested individually (Table 3). Specifically, birds in the *group-control* condition ate the fastest (back-transformed $M = 2.96$ s, $SE = 0.12$, 95% CI $= [2.74, 3.21]$),

followed by those in the *group-novel* condition ($M = 3.52$ s, $SE = 0.17$, 95% CI $= [3.22,\ 3.88]$) and birds in the *individual-control* condition ($M = 5.20$ s, $SE = 0.35$, 95% CI $= [4.59,\ 5.98]$); birds in the *individual-novel* condition showed the longest latency to eat ($M = 9.81$ s, $SE = 1.13$, 95% CI $= [7.97,\ 12.64]$). Notably, in the *individual-novel* condition, 24 birds did not eat at all during the trial (Table 2).

In addition, we found a main effect of *sex* ($p < 0.04$), as males ($M = 4.00$s, $SE = 0.25$, 95%, CI $= [3.56, 4.56]$) were, on average, faster to eat than females ($M = 4.88$s, $SE = 0.35$, 95%, CI $= [4.27, 5.68]$).

Variance analysis indicated that the full model, which retained the random effect structure for both individual and group conditions, did not provide a better fit than the reduced model (likelihood ratio test: $\chi^2(2) = 1.04$, $p = 0.59$), indicating no significant difference in variance between the two conditions.

Zone of Interest (ZOI) Duration

Analysis of the time spent in the zone of interest (ZOI) indicated significant effects of *context* ($p = 0.003$), and *object* ($p = 0.005$), and a significant interaction between them ($p < 0.001$) (Table 3). Birds in the *group-novel* condition spent the most time in the ZOI (back-transformed $M = 111.8$s, $SE = 13.39$, 95% CI $= [88.3, 141.7]$), followed by those in the *individual-control* condition (back-transformed $M = 98.4$s, $SE = 15.44$, 95% CI $= [72.0, 134.4]$). Birds in the *group-control* condition spent slightly less time in the ZOI (back-transformed $M = 80.2$s, $SE = 9.60$, 95% CI $= [63.3, 101.6]$), while those in the *individual-novel* condition spent the least time in the ZOI (back-transformed $M = 38.8$s, $SE = 6.09$, 95% CI $= [28.4, 53.1]$). Notably, in the *individual-novel* condition, 16 birds did not enter the ZOI at all (Table 3).

We also observed a significant interaction between *object* $\times$ *sex* ($p = 0.010$), indicating that males and females responded differently to novel versus control objects. Females spent more time in the ZOI in the *control* condition (back-transformed $M = 103.1$s, $SE = 15.56$, 95% CI $= [76.4, 139.1]$), compared to the *novel* condition (back-transformed $M = 58.7$s, $SE = 8.85$, 95% CI $= [43.5, 79.2]$). In contrast, males showed a more stable response across object conditions, spending a similar amount of time in the ZOI for both control (back-transformed $M = 75.7$s, $SE = 11.84$, 95% CI $= [55.5, 103.2]$) and novel objects (back-transformed $M = 74.7$s, $SE = 11.67$, 95% CI $= [54.7, 101.8]$). This suggests that females exhibited a stronger response to object novelty than males.

Variance analysis indicated greater individual variability in ZOI duration when birds were tested *individually* (back-transformed $\sigma^2 = 15380.41\,\text{s}^2$, SD = 124.02 s) compared to when they were tested in a *group* (back-transformed $\sigma^2 = 10220.50\,\text{s}^2$, SD = 101.10 s). A likelihood ratio test confirmed that this reduction in variance in the group context was statistically significant ($\chi^2(2) = 15.815$, $p < 0.001$), suggesting that birds' behaviour was more consistent when tested in groups.

A high estimated correlation between the individual and group random effects (Corr = 0.68) was observed, suggesting that birds that spent more time in the ZOI when tested individually also tended to do so when tested in a group. This indicates consistent behaviour across both social contexts

Table 3. Results of linear mixed-effects models for all dependent variables, including variance components and t-values (with degrees of freedom) where applicable. Significant effects are highlighted in bold.

Effect	Latency to enter			Latency to eat			ZOI duration		
	Estimate (SE)	t-value (df)	p-value	Estimate (SE)	t-value (df)	p-value	Estimate (SE)	t-value (df)	p-value
Fixed Effects									
Intercept	0.697 (0.021)	t(154.22) = 33.399	< 0.001	0.837 (0.015)	t(103.08) = 55.060	< 0.001	4.098 (0.126)	t(133.95) = 32.530	< 0.001
Context (Group)	**-0.046 (0.022)**	t(52.61) = -2.080	0.042	**-0.201 (0.011)**	t(374.00) = -18.445	< 0.001	**0.427 (0.136)**	t(52.99) = 3.128	0.003
Object (Novel)	0.013 (0.018)	t(321.58) = 0.721	0.472	**0.090 (0.011)**	t(374.00) = 8.258	< 0.001	**-0.288 (0.104)**	t(320.01) = -2.776	0.006
Context $\times$ Object	–	–	–	**-0.053 (0.022)**	t(374.00) = -2.440	0.015	**1.262 (0.208)**	t(320.02) = 6.079	< 0.001
Trial	**-0.052 (0.004)**	t(344.37) = -13.030	< 0.001	-0.0012 (0.0024)	t(374.00) = -0.507	0.613	**0.068 (0.023)**	t(339.53) = 2.984	0.003
Sex	**-0.087 (0.029)**	t(52.08) = -2.970	0.004	**-0.055 (0.026)**	t(52.00) = -2.109	0.040	-0.034 (0.183)	t(52.06) = -0.187	0.852
Object $\times$ Sex	–	–	–	–	–	–	**0.550 (0.212)**	t(320.08) = 2.593	0.010
Variance components									
Variance (Individual)	**0.018 (0.133)**	–	–	**0.007 (0.085)**	–	–	**0.748 (0.865)**	–	–
Variance (Group)	**0.006 (0.075)**	–	–	–	–	–	**0.192 (0.438)**	–	–
Residual	0.036 (0.189)	–	–	0.013 (0.113)	–	–	1.163 (1.078)	–	–
Likelihood Ratio Test	$\chi^2(2) = 11.83$	–	$p = 0.003$	–	–	–	$\chi^2(2) = 15.815$	–	$p < 0.001$

Discussion

This study investigated how social context affects neophobic responses in juvenile herring gulls, specifically whether the presence of conspecifics influences both average behaviour and behavioural variability. We tested three hypotheses: (1) the risk dilution hypothesis, which posits that individuals share the perceived risk among group members, predicting reduced neophobia and reduced variance in groups; (2) the negotiation hypothesis, which suggests that individuals negotiate who will approach first, predicting increased neophobia but reduced variance; and (3) the social conformity hypothesis, which proposes behavioural synchronisation within the group, resulting in no change in average neophobic responses but reduced variance in groups.

We found that individuals tested in groups were on average quicker to enter the test arena and eat than when they were alone. They also spent more time in the zone of interest. For the latter two measures, this group effect was most pronounced when a novel object was placed behind the food (compared to a control object), suggesting reduced neophobia in a group context. In addition, we found reduced variance in group contexts for measures of entering the test arena and time in the zone of interest but not for latency to eat.

Overall, these results are consistent with the risk-dilution hypothesis, whereby perceived risk is shared among individuals, resulting in each bird perceiving a lower level of threat when in group than when alone with a novel object. Observing conspecifics in the vicinity of a novel object and feeding next to it appears to reduce individual neophobic responses, as each bird likely perceives the risk to be shared by the group. This is consistent with previous studies showing that social animals often rely on the presence of the group to make quicker decisions and engage in potentially risky situations (Keen et al., 2020; Lee and Thornton, 2021; Loukola et al., 2012). In contrast, our findings did not support either the negotiation hypothesis or the social conformity hypothesis. However, both hypotheses have been supported by other studies that have found no change or even an increase in the average neophobic response in the group context (Apfelbeck and Raess, 2008; Griffin, Lermite, et al., 2013; Kelly et al., 2020; Kerman et al., 2018; King et al., 2015; Miller, Bugnyar, et al., 2015; Schuett and Dall, 2009; St. Lawrence et al., 2021; Stöwe, Bugnyar, Heinrich, et al., 2006a,b). The inconsistency between studies could be due to a number of factors. In our study, for example, all individuals were of a similar age, size, and had very similar early-life experiences. But in natural settings, groups are not always so homogeneous. In such settings, social conformity or negotiation mechanisms may play a more important role (Ekman and Lilliendahl, 1993; Griffin, Netto, et al., 2017). For example, more experienced individuals may assume a leading role in approaching novel objects, while others may follow their lead. This hierarchical or conformist response may result in the delayed engagement of other individuals, particularly subordinates, until the perceived risk is mitigated by the actions of more experienced group members.

More generally, the effect of social context on neophobia may be dependent on the ecological niche of the species, which may explain the discrepancy between studies. Herring gulls are considered a highly social species. Species with different ecological niches or social structures may adopt alternative social mechanisms (St. Lawrence et al., 2021). Furthermore, in more solitary or less adaptable species, social learning will be less important overall, so neophobic responses may not vary as flexibly with social context at all. For example, a study conducted by Echeverria and Vassallo (2008) examined neophobic responses in house sparrows *(Passer domesticus)*. Even though birds were observed in groups, they continued to approach feeders individually, exhibiting a heightened level of neophobia. This solitary approach, even in a group setting, may reflect a social structure that prioritises individual foraging, thereby increasing perceived risk in novel situations due to the lack of social reinforcement. Thus, the effect of social context on neophobic behaviour is likely to differ between species and ecological settings.

The general pattern of our results suggests reduced neophobia in the group context. Interestingly, birds in the group context spent even more time in the zone of interest (ZOI) when the novel object was present than when the control object was present (resulting in 'negative' neophobia scores; see Figure 4). This finding contrasts with the typically observed neophobic response, where novel stimuli are typically associated with reduced time spent near the object (Greenberg, 2003). However, a similar pattern was observed in a group

of shiny cowbirds (*Molothrus bonariensis*), which spent more time near the feeder when a novel object was present compared to the control condition (Echeverría and Vassallo, 2008). Other studies have also found that once individuals reached the novel object, they spent more time interacting with it in a group context (Miller, Bugnyar, et al., 2015; St. Lawrence et al., 2021; Stöwe, Bugnyar, Heinrich, et al., 2006a). However, in these studies, there was no control condition, which makes the interpretation of the results difficult. One possible explanation for our results and those of Echeverría and Vassallo (2008) is that the presence of conspecifics reduced fear responses sufficiently to allow individuals to approach and feed despite the novel object, but not enough to reduce them altogether, resulting in some degree of vigilance. Thus, birds in groups may have balanced their reduced fear with the need for heightened vigilance (Beauchamp, 1998), leading to longer feeding times in the novel-object condition. Alternatively, being in a social context may have encouraged individuals to explore more in the novel object condition (Miller, Bugnyar, et al., 2015; Stöwe, Bugnyar, Heinrich, et al., 2006b). In the control condition, where the object was familiar, there was less need for exploration as the birds were already habituated to the object (Rankin et al., 2009). This difference in exploration could potentially also explain the increased time in the zone of interest. However, at this point, our current data analysis does not permit a definitive distinction between these possible explanations.

Although the latency to eat and the time spent in the ZOI were influenced by the presence of the novel object, no such effect of the object was observed for the latency to enter. Note that we did find a main effect of social context as juvenile herring gulls tested in groups generally exited the start area more quickly than those tested individually. The absence of an effect of object on start latency could be due to our test set up. Prior to the start of the trial, the birds were unable to see the object. Given the rapidity with which they exited the start area, it is probable that they became aware of the novel object only after leaving this start area.

Additionally, we found a strong correlation between the individual and group random effect for latency to enter and ZOI duration (note that for latency to eat, we were unable to assess repeatability as our model structure did not allow for this). This correlation suggests that the average behavioural response of an individual was consistent between social contexts, for instance, in the time spent in the ZOI. Our findings align partly with those of Stöwe, Bugnyar, Heinrich, et al. (2006a), who observed consistency in juvenile raven behaviour between individual and dyadic contexts (although this effect diminished in groups of three to six individuals). Overall, this suggests that the presence of conspecifics partially reduced, but did not completely eliminate, inter-individual variation in behaviour. From a methodological point of view, this may indicate that testing individuals alone may be useful to test intrinsic abilities; information that could then be used to further explore group dynamics (e.g. leaders vs. followers).

Finally, we found sex-specific behavioral differences in our juvenile gulls. The observed differences in neophobic responses were unexpected, as herring gulls exhibit early sexual dimorphism but do not reach sexual maturity until four years of age. Nevertheless, females spent more time near the object in the control condition compared to the novel condition, whereas males showed a more consistent response across both object conditions. This could indicate that female herring gulls are more neophobic than male herring gulls. We did not observe an interaction between sex and object condition for the other two measures, although we did find a main effect as males were generally faster than females on both novel and control trials. As our study was not specifically designed to assess sex differences, and given the constraints of our sample size, we cannot rule out the possibility that a larger dataset or a different experimental design might reveal an interaction effect for latency to eat as well. Previous research on sex differences in neophobia among avian species is limited and remains inconclusive. Although some studies suggest that males are less neophobic (Tuliozi et al., 2018), others report the opposite (Amy et al., 2017; Danel et al., 2024; Rokka et al., 2014) or find no significant differences at all (Crane and Ferrari, 2017; Schaffer et al., 2021; Schuett and Dall, 2009). These discrepancies may arise from variations in the way neophobia is assessed, making direct comparisons challenging. The functional implications of sex differences in neophobia remain unclear but could influence broader behavioral patterns in the wild, such as sex-specific foraging strategies.

A potential limitation of our study is that we worked with juvenile, hand-reared herring gulls in a cap-

tive controlled environment. This may limit the generalisability of our findings to wild populations or adult birds. For example, a meta-analysis has shown that wild-caught birds tend to have higher baseline neophobia than captive-bred individuals, probably due to the greater environmental variation encountered by wild birds (Crane and Ferrari, 2017). Although the lower baseline neophobia observed in captivity might reduce the likelihood of detecting group effects, our results demonstrate that such effects can still be identified. Furthermore, in other avian species, age and environmental familiarity have been shown to significantly influence neophobic responses (Biondi et al., 2010; Greenberg, 2003; Miller, Bugnyar, et al., 2015). However, previous research from our laboratory on captive herring gulls reared under similar conditions (Troisi et al., 2024, pre-print), and on wild-reared chicks from a neighbouring colony (Salas et al., 2024) shows that strictly controlled testing of birds at this age provides ecologically relevant data on potential behavioural expression. In addition, prior research on wild herring gulls has shown no age-related differences in latency to approach novel objects (Inzani et al., 2023), suggesting our results might extend across different age groups. Finally, all chicks in our study had very similar prior life experiences, minimising sampling bias. This approach addresses potential challenges encountered in wild populations, where tested animals may not be fully representative of the broader population due to prior habituation to specific novelties or situations, or already developed dominance hierarchies.

To conclude, our findings demonstrate that social context plays an important role in shaping neophobic behaviour in juvenile herring gulls. For social species, group living may lower the cost of learning for individuals, as they can rely on the actions of experienced conspecifics to evaluate potential threats, reducing the need for independent assessments of novel stimuli (Webster and Ward, 2011). This is under the assumption that peers provide accurate assessments of risk. Such collective risk assessment can enable more efficient exploration and engagement with the environment, mitigating the full cost of individual trial and error. Especially in rapidly changing or urbanised landscapes, where animals frequently encounter novel stimuli, the ability to draw on social cues could likely offer a distinct advantage to social species (Griffin, Netto, et al., 2017; Lee and Thornton, 2021).

Author Contributions (CREDIT)

Conceptualization: R.A., S.K., L.L., F.V.; Methodology: R.A., S.K., L.L., F.V.; Investigation: R.A., S.K., L.L., F.V.; Animal Care: R.A., S.K, S.B.; Data curation: R.A., S.K; Formal analysis: R.A., S.K; Visualization: R.A., S.K; Validation: R.A., S.K, D.D.; Writing – original draft: R.A., S.K.; Writing – review and editing: R.A., S.K., S.B., D.D., A.M., W.M., E.S., L.L., F.V.; Supervision: L.L., and F.V.; Funding acquisition: A.M., L.L., F.V.; Project administration: F.V.

Acknowledgements

We thank the Gull Patrol (Marc Verborgh, Nathan Noels) and ANB for providing the eggs, as well as the VOC staff, particularly Isabelle Allemeersch, Rijn van Maele, and Claude Velter, for their assistance with bird care and organisation. We are also grateful to Lies Baten and Angelica Alcantara-Exposito for their administrative support, and to Nathan Audenaert for his invaluable help in raising the animals.

Fundings

This research was funded by a Methusalem Project 01M00221 (Ghent University; https://www.ugent.be/en/research/funding/bof/methusalem) awarded to FV, LL, and AM, and an ERC Consolidator Grant (European Union's Horizon 2020 research and innovation programme, grant agreement no. 769595) awarded to FV. SK is funded by a FWO (Flemish Research Foundation) PhD Fellowship grant (No. 11P3G24N).

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Data, script, code, and supplementary information availability

All necessary data, scripts, and code required to replicate our study's findings will be made openly accessible at the article's OSF repository: https://osf.io/p496z/. Supplementary information, supporting our results, will also be made available at this repository. Due to the large volume of videos, totalling over 400GB, the videos will be made available upon request by contacting the senior author at frederick.verbruggen@ugent.be.

References

Amy M, D Ung, N Béguin, and G Leboucher (2017). Personality traits and behavioural profiles in the domestic canary are affected by sex and photoperiod. *Ethology* 123, 885–893. https://doi.org/10.1111/eth.12662.

Apfelbeck B and M Raess (2008). Behavioural and hormonal effects of social isolation and neophobia in a gregarious bird species, the European starling (*Sturnus vulgaris*). *Hormones and Behavior* 54, 435–441. https://doi.org/10.1016/j.yhbeh.2008.04.003.

Bates D, M Mächler, BM Bolker, and SC Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67 (1), 1–48. https://doi.org/10.18637/JSS.V067.I01.

Beauchamp G (1998). The effect of group size on mean food intake rate in birds. *Biological Reviews* 73 (4), 449–472. https://doi.org/10.1017/S0006323198005246.

Benson-Amram S and KE Holekamp (2012). Innovative problem solving by wild spotted hyenas. *Proceedings of the Royal Society B: Biological Sciences* 279, 4087–4095. https://doi.org/10.1098/rspb.2012.1450.

Biondi LM, MS Bó, and AI Vassallo (2010). Inter-individual and age differences in exploration, neophobia and problem-solving ability in a Neotropical raptor (*Milvago chimango*). *Animal cognition* 13, 701–710. https://doi.org/10.1007/s10071-010-0319-8.

Both C, NJ Dingemanse, PJ Drent, and JM Tinbergen (2005). Pairs of extreme avian personalities have highest reproductive success. *Journal of Animal Ecology* 74 (4), 667–674. https://doi.org/10.1111/j.1365-2656.2005.00962.x.

Brown GR and C Nemes (2008). The exploratory behaviour of rats in the hole-board apparatus: Is head-dipping a valid measure of neophilia? *Behavioural Processes* 78 (3), 442–448. https://doi.org/10.1016/J.BEPROC.2008.02.019.

Bruijn R de and LM Romero (2021). Chronic stress reverses enhanced neophobia following an acute stressor in European starlings. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* 335 (2), 265–274. https://doi.org/10.1002/JEZ.2431.

Cohen J (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Coleman SL and RL Mellgren (1994). Neophobia when feeding alone or in flocks in zebra finches, *Taeniopygia guttata*. *Animal Behaviour* 48 (4), 903–907. https://doi.org/https://doi.org/10.1006/anbe.1994.1315.

Coulson JC (2015). Re-Evaluation of the Role of Landfills and Culling in the Historic Changes in the Herring Gull (*Larus argentatus*) Population in Great Britain. *Waterbirds* 38, 339–354. https://doi.org/https://doi.org/10.1675/063.038.0411.

Crane AL and MCO Ferrari (2017). Patterns of predator neophobia: a meta-analytic review. *Proceedings of the Royal Society B: Biological Sciences* 284 (1861), 20170583. https://doi.org/10.1098/rspb.2017.0583.

Danel S, N Rebout, F Bonadonna, and D Biro (2024). Sex predicts response to novelty and problem-solving in a wild bird with female-biased sexual dimorphism. *Proceedings B* 291, 20242277. https://doi.org/10.1098/rspb.2024.2277.

Echeverría AI and AI Vassallo (2008). Novelty Responses in a Bird Assemblage Inhabiting an Urban Area. *Ethology* 114 (6), 616–624. https://doi.org/https://doi.org/10.1111/j.1439-0310.2008.01512.x.

Ekman JB and K Lilliendahl (1993). Using priority to food access: Fattening strategies in dominance-structured willow tit (*Parus montanus*) flocks. *Behavioral Ecology* 4, 232–238. https://doi.org/10.1093/beheco/4.3.232.

Erdfelder E, F FAul, A Buchner, and AG Lang (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods* 41 (4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149/METRICS.

Farrar B, M Boeckle, and N Clayton (2020). Replications in Comparative Cognition: What Should We Expect and How Can We Improve? *Animal Behavior and Cognition* 7 (1), 1–22. https://doi.org/10.26451/abc.07.01.02.2020.

Feist F, K Smith, and P Graham (2023). Inter-species stimulus enhancement: herring gulls (*Larus argentatus*) mimic human food choice during foraging. *Biology Letters* 19 (5), 20230035–20230035. https://doi.org/10.1098/RSBL.2023.0035.

Fox J and S Weisberg (2019). *An R Companion to Applied Regression*. Third. Thousand Oaks CA: Sage.

Fox J, S Weisberg, and B Price (2022). *carData: Companion to Applied Regression Data Sets*.

Friard O and M Gamba (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 7. https://doi.org/10.1111/2041-210X.12584.

Fridolfsson AK and H Ellegren (1999). A Simple and Universal Method for Molecular Sexing of Non-Ratite Birds. *Journal of Avian Biology* 30, 116–121. https://doi.org/https://doi.org/10.2307/3677252.

Frings H, M Frings, B Cox, and L Peissner (1955). Auditory and visual mechanisms in food-finding behavior of the Herring Gull. *The Wilson Bulletin*, 155–170. https://doi.org/10.2307/4158412.

Gandolfi G (2009). Gilberto Gandolfi (1975) Social learning in non-primate animals. *Italian Journal of Zoology* 42 (4), 311–329. https://doi.org/10.1080/11250007509431449.

Gao J and BC O'Neill (2020). Mapping global urban land for the 21st century with data-driven simulations and Shared Socioeconomic Pathways. *Nature Communications* 11, 2302. https://doi.org/10.1038/s41467-020-15788-7.

Genz A, F Bretz, T Miwa, X Mi, F Leisch, F Scheipl, and T Hothorn (2020). mvtnorm: Multivariate Normal and t Distributions.

Goumas M, NJ Boogert, and LA Kelley (2020). Urban herring gulls use human behavioural cues to locate food. *Royal Society Open Science* 7 (2), 191959. https://doi.org/10.1098/RSOS.191959.

Greenberg R (2003). The Role of Neophobia and Neophilia in the Development of Innovative Behaviour of Birds. In: *Animal Innovation*. Ed. by Reader SM and Laland KN. Oxford University Press, pp. 175–196.

Greenberg R and C Mettke-Hofmann (2001). Ecological Aspects of Neophobia and Neophilia in Birds. *Current Ornithology* 16, 119–178. https://doi.org/https://doi.org/10.1007/978-1-4615-1211-0_3.

Greggor AL, A Thornton, and NS Clayton (Dec. 2015). Neophobia is not only avoidance: improving neophobia tests by combining cognition and ecology. *Current Opinion in Behavioral Sciences* 6, 82–89. https://doi.org/10.1016/J.COBEHA.2015.10.007.

Griffin AS, K Netto, and C Peneaux (2017). Neophilia, innovation and learning in an urbanized world: a critical evaluation of mixed findings. *Current Opinion in Behavioral Sciences* 16, 15–22. https://doi.org/https://doi.org/10.1016/j.cobeha.2017.01.004.

Griffin AS, F Lermite, M Perea, and D Guez (2013). To innovate or not: contrasting effects of social groupings on safe and risky foraging in Indian mynahs. *Animal Behaviour* 86 (6), 1291–1300. https://doi.org/10.1016/j.anbehav.2013.09.035.

Harel R, O Spiegel, WM Getz, and R Nathan (2017). Social foraging and individual consistency in following behaviour: testing the information centre hypothesis in free-ranging vultures. *Proceedings of the Royal Society B: Biological Sciences* 284 (1852), 20162654. https://doi.org/10.1098/RSPB.2016.2654.

Herbert-Read JE, S Krause, LJ Morrell, TM Schaerf, J Krause, and AJW Ward (2013). The role of individuality in collective group movement. *Proceedings of the Royal Society B: Biological Sciences* 280 (1752), 20122564. https://doi.org/10.1098/rspb.2012.2564.

Inzani EL, LA Kelley, and NJ Boogert (2023). Object neophilia in wild herring gulls in urban and rural locations. *Journal of Avian Biology* 2023, e03028. https://doi.org/https://doi.org/10.1111/jav.03028.

Jaeger B (2017). *r2glmm: Computes R Squared for Mixed (Multilevel) Models*. R package version 0.1.2.

Kareklas K, RW Elwood, and RA Holland (2018). Grouping promotes risk-taking in unfamiliar settings. *Behavioural Processes* 148, 41–45. https://doi.org/10.1016/j.beproc.2018.01.003.

Keen SC, EF Cole, MJ Sheehan, and BC Sheldon (2020). Social learning of acoustic anti-predator cues occurs between wild bird species. *Proceedings of the Royal Society B: Biological Sciences* 287, 20192513. https://doi.org/10.1098/rspb.2019.2513.

Kelly, Kimball, Stansberry, and Lattin (2020). No, you go first: phenotype and social context affect house sparrow neophobia. *Biology Letters* 16, 20200286. https://doi.org/10.1098/rsbl.2020.0286.

Kerman K, KE Sieving, CS Mary, and ML Avery (2018). Social conformity affects experimental measurement of boldness in male but not female monk parakeets (*Myiopsitta monachus*). *Behaviour* 155, 1025–1050. https://doi.org/10.1163/1568539X-00003519.

Kikusui T, JT Winslow, and Y Mori (2006). Social buffering: relief from stress and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences* 361, 2215–2228. https://doi.org/10.1098/rstb.2006.1941.

Kimball MG and CR Lattin (Nov. 2023). The "Seven Deadly Sins" of Neophobia Experimental Design. *Integrative and Comparative Biology* 64, 38–54. https://doi.org/10.1093/icb/icad127.

King AJ, LJ Williams, and C Mettke-Hofmann (2015). The effects of social conformity on Gouldian finch personality. *Animal Behaviour* 99, 25–31. https://doi.org/10.1016/j.anbehav.2014.10.016.

Krause J and G Ruxton (2002). *Living in Groups*. Oxford University Press, USA.

Kuznetsova A, PB Brockhoff, and RH Christensen (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82 (13), 1–26. https://doi.org/10.18637/JSS.V082.I13.

Lecuelle S, I Bouvarel, AM Chagneau, F Laviron, P Lescoat, and C Leterrier (2011). Early visual experience of food does not appear to reduce subsequent feed neophobia in turkeys. *Poultry Science* 90 (1), 1–9. https://doi.org/10.3382/PS.2010-00882.

Lee VE and A Thornton (2021). Animal Cognition in an Urbanised World. *Frontiers in Ecology and Evolution* 9, 633947. https://doi.org/10.3389/FEVO.2021.633947.

Lenth RV (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.10.5, https://rvlenth.github.io

Loukola OJ, JT Seppänen, and JT Forsman (2012). Intraspecific social information use in the selection of nest site characteristics. *Animal Behaviour* 83 (3), 629–633. https://doi.org/10.1016/J.ANBEHAV.2011.12.004.

Lüdecke D, MS Ben-Shachar, I Patil, P Waggoner, and D Makowski (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software* 6, 3139. https://doi.org/10.21105/joss.03139.

McHugh M (2012). Interrater reliability: The kappa statistic. *Biochemia medica : časopis Hrvatskoga društva medicinskih biokemičara / HDMB* 22, 276–82. https://doi.org/10.11613/BM.2012.031.

McKinney ML (2002). Urbanization, Biodiversity, and Conservation: The impacts of urbanization on native species are poorly studied, but educating a highly urbanized human population about these impacts can greatly improve species conservation in all ecosystems. *BioScience* 52, 883–890. https://doi.org/10.1641/0006-3568(2002)052[0883:UBAC]2.0.CO;2.

Mettke-Hofmann C (2017). Neophobia. In: *Encyclopedia of Animal Cognition and Behavior*. Ed. by Vonk J and Shackelford T. Cham: Springer International Publishing, pp. 1–8. https://doi.org/10.1007/978-3-319-47829-6_908-1.

Mettke-Hofmann C, H Winkler, and B Leisler (2002). The Significance of Ecological Factors for Exploration and Neophobia in Parrots. *Ethology* 108, 249–272. https://doi.org/10.1046/j.1439-0310.2002.00773.x.

Miller R, T Bugnyar, K Pölzl, and C Schwab (2015). Differences in exploration behaviour in common ravens and carrion crows during development and across social context. *Behavioral Ecology and Sociobiology* 69, 1209–1220. https://doi.org/10.1007/s00265-015-1935-8.

Miller R and M Lambert (2024). *Evolutionary drivers of neophobia across the avian clade*. https://doi.org/10.31219/osf.io/qhy8m.

Miller R, M Lambert, A Frohnwieser, KF Brecht, T Bugnyar, I Crampton, E Garcia-Pelegrin, K Gould, AL Greggor, EI Izawa, et al. (2022). Socio-ecological correlates of neophobia in corvids. *Current Biology* 32, 74–85. https://doi.org/https://doi.org/10.1016/j.cub.2021.10.045.

Moretti L, M Hentrup, K Kotrschal, and F Range (2015). The influence of relationships on neophobia and exploration in wolves and dogs. *Animal Behaviour* 107, 159–173. https://doi.org/10.1016/j.anbehav.2015.06.008.

Nager RG and NJ O'Hanlon (2016). Changing Numbers of Three Gull Species in the British Isles. *Waterbirds* 39 (1), 15–28. https://doi.org/10.1675/063.039.SP108.

O'Hanlon NJ, RA McGill, and RG Nager (2017). Increased use of intertidal resources benefits breeding success in a generalist gull species. *Marine Ecology Progress Series* 574, 193–210. https://doi.org/10.3354/MEPS12189.

Parsons J (1972). Egg size, laying date and incubation period in the Herring Gull. *Ibis* 114, 536–541. https://doi.org/https://onlinelibrary.wiley.com/doi/10.1111/j.1474-919X.1972.tb00855.x.

Pavlova O and T Wronski (2020). City gulls and their rural neighbours: Changes in escape and agonistic behaviour along a rural-to-urban gradient. In: Nova Science Publishers, Inc. ISBN: 978-1-53618-000-8.

R Core Team (2021). R: A Language and Environment for Statistical Computing.

Rankin CH, T Abrams, RJ Barry, S Bhatnagar, DF Clayton, J Colombo, G Coppola, MA Geyer, DL Glanzman, S Marsland, et al. (2009). Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiology of learning and memory* 92, 135–138. https://doi.org/10.1016/j.nlm.2008.09.01.

Rokka K, M Pihlaja, H Siitari, and CD Soulsbury (2014). Sex-specific differences in offspring personalities across the laying order in magpies *Pica pica*. *Behavioural processes* 107, 79–87. https://doi.org/https://doi.org/10.1016/j.beproc.2014.07.019.

Salas R, W Müller, E Stienen, H Matheve, B Vanden Broecke, F Verbruggen, and L Lens (2024). Breeding density affects the movements of gull chicks, the size of their home ranges and their association with neighbours. *Royal Society Open Science* 11, 231431. https://doi.org/10.1098/rsos.231431.

Schaffer A, AL Caicoya, M Colell, R Holland, L von Fersen, A Widdig, and F Amici (2021). Neophobia in 10 ungulate species—a comparative approach. *Behavioral ecology and sociobiology* 75, 102. https://doi.org/https://doi.org/10.1007/s00265-021-03041-0.

Schuett W and S Dall (2009). Sex differences, social context and personality in Zebra Finches, *Taeniopygia guttata*. *Animal Behaviour* 77, 1041–1050. https://doi.org/10.1016/j.anbehav.2008.12.024.

Snijders TAB and RJ Bosker (2012). *Multilevel analysis : an introduction to basic and advanced multilevel modeling*. 2nd ed. Los Angeles ; London : SAGE.

Soma M and T Hasegawa (2004). The Effect of Social Facilitation and Social Dominance on Foraging Success of Budgerigars in an Unfamiliar Environment. *Behaviour* 141, 1121–1134. https://doi.org/10.1163/1568539042664560.

St. Lawrence S, I Rojas-Ferrer, and J Morand-Ferron (2021). Does the presence of a conspecific increase or decrease fear? Neophobia and habituation in zebra finches. *Ethology* 127, 1033–1041. https://doi.org/10.1111/eth.13224.

Stöwe M, T Bugnyar, B Heinrich, and K Kotrschal (2006a). Effects of Group Size on Approach to Novel Objects in Ravens (*Corvus corax*). *Ethology* 112, 1079–1088. https://doi.org/10.1111/j.1439-0310.2006.01273.x.

— (2006b). Effects of group size on approach to novel objects in ravens (Corvus corax). *Ethology* 112, 1079–1088.

Stöwe M, T Bugnyar, MC Loretto, C Schloegl, F Range, and K Kotrschal (2006). Novel object exploration in ravens (*Corvus corax*): Effects of social relationships. *Behavioural Processes* 73, 68–75. https://doi.org/10.1016/j.beproc.2006.03.015.

Troisi C, A Vernouillet, R Allaert, S Knoch, A Martel, L Lens, and F Verbruggen (2024). Beyond a Unitary Construct: Dissecting Stopping Behaviour in Two Bird Species. *bioRxiv*. in revision. https://doi.org/10.1101/2024.01.17.575695.

Tuliozi B, G Fracasso, H Hoi, and M Griggio (2018). House sparrows' (*Passer domesticus*) behaviour in a novel environment is modulated by social context and familiarity in a sex-specific manner. *Frontiers in zoology* 15, 1–14. https://doi.org/https://doi.org/10.1186/s12983-018-0267-8.

Ushey K and H Wickham (2024). *renv: Project Environments*. R package version 1.0.11, https://github.com/rstudio/renv.

Venables WN and BD Ripley (2002). *Modern Applied Statistics with S*. Fourth. ISBN 0-387-95457-0. New York: Springer.

Vernouillet A and Kelly (2020). Individual exploratory responses are not repeatable across time or context for four species of food-storing corvid. *Scientific Reports* 10, 394. https://doi.org/10.1038/s41598-019-56138-y.

Vrublevska J, T Krama, MJ Rantala, P Mierauskas, TM Freeberg, and IA Krams (2015). Personality and density affect nest defence and nest survival in the great tit. *Acta Ethologica* 18 (2), 111–120. https://doi.org/10.1007/S10211-014-0191-7/FIGURES/3.

Webster MM and AJW Ward (2011). Personality and social context. *Biological Reviews* 86, 759–773. https://doi.org/https://doi.org/10.1111/j.1469-185X.2010.00169.x.

Supplementary material

Preregistered version of the statistical analysis Statistical analyses will be conducted using R, version 4.3.X (R Core Team, 2021). Mixed-Effects Models (MMs), either linear MMs (LMMs)or generalised LMMs (GLMMs), will be fitted using the `lme4` package (Bates et al., 2015). For LLMs, parameter estimation and p-values for the estimated models will be calculated by means of the `lmerTest` package (Kuznetsova et al., 2017) via the the Satterthwaite's degrees of freedom method; for GLMMs, the `car` (Fox and Weisberg, 2019) or `carData` (Fox, Weisberg, and Price, 2022) package will be used. For the GLMM, we will use partial η-squared (η_p^2) as effect sizes, and they will be calculated by means of the `r2glmm` (Jaeger, 2017) package. Models will be fitted to the different latency measures separately, as well as combined. For the combined analysis, the approach proposed by Snijders and Bosker, 2012 will be used, which allows for the simultaneous analysis of multiple dependent variables in the case of nested data structures, thereby considering within-group and between-group variance in latency measures.

As we aim to determine whether the average neophobic response differs between individual and group trials, a (G)LMM with Type III sum of squares will be performed on the latency measures (Table 1). This analysis will include both fixed and random effects to explore the impact of different experimental conditions. The model will incorporate *Object*, *Context*, and their interaction as key fixed effects to explore how the type of object and the social setting (alone vs. in a group) interactively affect latency responses. Additionally, *Trial* will be included as a fixed effect to control for the impact of trial repeat. To specifically assess the variability in latency across individual and group trials, we will compare the estimated variance components within our mixed-effects model. Variance for individual trials will be estimated from the *Indiv_Dummy* effect at the BirdID level. For group trials, the combined estimated variances of the *Group_Dummy* effect at both the BirdID and GroupID levels will be evaluated. This comparison aims to determine whether individual differences are more pronounced in solitary compared to group settings, with an expectation that individual variances and the total variance might be higher in individual trials. Additionally, an analysis at the BirdID level between the estimated variances of the *Indiv_Dummy* and *Group_Dummy* effects will further elucidate how individual differences manifest under different trial conditions, potentially highlighting the influence of group dynamics on individual behaviour.

$$\begin{aligned}\text{Latency} \sim, &\ \text{Object} \times \text{Context} + \text{Trial} \\ &+ (1|\text{NestID}) \\ &+ (-1 + \text{Group\_Dummy}|\text{GroupID}) \\ &+ (-1 + \text{Indiv\_Dummy} + \text{Group\_Dummy}|\text{BirdID})\end{aligned}$$

In the model, *Object* refers to the stimulus presented, distinguishing between control and novel objects. *Trial* captures the two testing sessions conducted, and *Context* indicates the social environment, differentiating between individual and group settings. Random effects structures are tailored to accurately reflect the individual and group-level variability in responses. Specifically, *NestID* is included to control for similarities within nests, *Group_Dummy* identifies trials conducted in group setting, effectively marking the presence of social interactions during the test. Conversely, *Indiv_Dummy* indicates the absence of such group dynamics, highlighting trials where subjects are tested alone.

In all instances, model plots will be generated using the `performance` package (Lüdecke et al., 2021) to inspect violations of model assumptions, such as heteroscedasticity, non-normality of residuals, and the presence of outliers. Multicollinearity and autocorrelation will be evaluated, with potential model adjustments including transformation of variables or modification of the model structure (e.g., switching from LMM to GLMM). In terms of model design, binary predictors will be encoded using contrast coding (-0.5 vs. 0.5), optimizing the interpretability and efficiency of our analyses in the context of our perfectly balanced predictor

variables. Post-hoc analyses, following significant findings, will be performed with Bonferroni-Holm corrected contrasts to further explore the data. Given the balanced nature of our model predictors, concerns related to multicollinearity are minimised, negating the need for variance inflation factor (VIF) assessments traditionally used to identify redundancy among predictors.

Model overview Below is an overview of all models fitted to determine the best model for each dependent variable (DV). For each DV, the final model listed represents the one used for the analysis.

Latency to enter

$$\begin{aligned}\text{Latency to enter} \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (1|\text{NestID}) \\ &+ (-1 + \text{Group\_dummy}|\text{GroupID}) \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{1}$$

$$\begin{aligned}\text{Latency to enter} \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{2}$$

$$\begin{aligned}\log(\text{Latency to enter}) \sim &\text{Object} + \text{Context} + \text{Trial} + \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{3}$$

$$\begin{aligned}\text{Box-Cox(Latency to enter)} \sim &\text{Object} + \text{Context} + \text{Trial} + \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{4}$$

Latency to eat

$$\begin{aligned}\log(\text{Latency to eat}) \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (1|\text{NestID}) \\ &+ (-1 + \text{Group\_dummy}|\text{GroupID}) \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{1}$$

$$\begin{aligned}\log(\text{Latency to eat}) \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{2}$$

$$\begin{aligned}\text{Box-Cox(Latency to eat)} \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \tag{3}$$

$$\begin{aligned}\text{Box-Cox(Latency to eat)} \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Sex} \\ &+ (1|\text{BirdID})\end{aligned} \tag{4}$$

ZOI duration

$$\begin{aligned}\text{ZOI duration} \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \quad (1)$$

$$\begin{aligned}\log(\text{ZOI duration}) \sim &\text{Object} \times \text{Context} + \text{Trial} + \text{Object} \times \text{Sex} \\ &+ (-1 + \text{Indiv\_dummy} + \text{Group\_dummy}|\text{BirdID})\end{aligned} \quad (2)$$

Testing schedule

Table S1: Neophobia testing schedule

Note: "GC" signifies Group Control, "IC" indicates Individual Control, "GT" represents Group Test, and "IT" stands for Individual Test. The subsequent number (1 or 2) specifies whether it is the first instance or a repeat. The suffix "-X" identifies the specific object involved, numbers 1-5 corresponding to randomly assigned novel or control objects.

Day/Cage	Cage 1	Cage 2	Cage 3	Cage 4	Cage 5	Cage 6	Cage 7	Cage 8
Day 1	GC1 - 1	IC1 - 2						
Day 2	IC1 - 1	GT1 - 4						
Day 3	GT1 - 4	IT1 - 1						
Day 4	IT1 - 3	GC1 - 2	GT1 - 5	IT1 - 1				
Day 5	GC2 - 1	IC2 - 2	IT1 - 1	GC1 - 4				
Day 6	IC2 - 1	GT2 - 5	GC1 - 3	IC1 - 4				
Day 7	GT2 - 2	IT2 - 3	IC1 - 3	GT1 - 3	GC1 - 1	IC1 - 2		
Day 8	IT2 - 5	GC2 - 2	GT2 - 4	IT2 - 2	IC1 - 1	GT1 - 3		
Day 9			IT2 - 2	GC2 - 4	GT1 - 2	IT1 - 4		
Day 10			GC2 - 3	IC2 - 4	IT1 - 3	GC1 - 2	GT1 - 2	IT1 - 1
Day 11			IC2 - 3	GT2 - 5	GC2 - 1	IC2 - 2	IT1 - 4	GC1 - 4
Day 12					IC2 - 1	GT2 - 5	GC1 - 3	IC1 - 4
Day 13					GT2 - 4	IT2 - 1	IC1 - 3	GT1 - 5
Day 14					IT2 - 5	GC2 - 2	GT2 - 1	IT2 - 2
Day 15							IT2 - 5	GC2 - 4
Day 16							GC2 - 3	IC2 - 4
Day 17							IC2 - 3	GT2 - 3

Table 5. Overview of species composition and sex distribution (only herring gulls) per experimental group.

Group ID	Total	LBBG	Male HG	Female HG	Unknown HG
Cage 1A	4	0	1	3	0
Cage 1B	4	0	1	3	0
Cage 2A	4	0	3	1	0
Cage 2B	4	0	3	1	0
Cage 3A	3	1	1	2	0
Cage 3B	2	0	0	2	0
Cage 4A	4	0	4	0	0
Cage 4B	4	1	1	3	0
Cage 5A	3	2	0	3	0
Cage 5B	3	2	0	2	1
Cage 6A	3	1	2	1	0
Cage 6B	4	1	1	3	0
Cage 7A	4	1	2	1	1
Cage 7B	4	0	2	2	0
Cage 8A	4	1	4	0	0

Table 6. Results of linear mixed-effects models for all dependent variables, including variance components and t-values (with degrees of freedom) where applicable. Significant effects are highlighted in bold. This analysis includes data from the lesser black-backed gulls tested.

Effect	Latency to enter			Latency to eat			ZOI duration		
	Estimate (SE)	t-value (df)	p-value	Estimate (SE)	t-value (df)	p-value	Estimate (SE)	t-value (df)	p-value
Fixed Effects									
Intercept	0.695 (0.020)	t(178.16) = 34.917	< 0.001	0.809 (0.013)	t(129.47) = 63.236	< 0.001	4.033 (0.109)	t(168.10) = 37.051	< 0.001
Context (Group)	-0.033 (0.020)	t(65.26) = -1.613	0.112	**-0.174 (0.009)**	t(464.00) = -18.939	< 0.001	**0.360 (0.115)**	t(66.01) = 3.134	0.003
Object (Novel)	0.010 (0.017)	t(399.22) = 0.571	0.569	**0.081 (0.009)**	t(464.00) = 8.841	< 0.001	**-0.296 (0.090)**	t(398.06) = -3.282	0.001
Context × Object	–	–	–	**-0.046 (0.018)**	t(464.00) = -2.528	0.012	**1.161 (0.180)**	t(398.07) = 6.452	< 0.001
Trial	**-0.053 (0.004)**	t(428.97) = -14.432	< 0.001	-0.0022 (0.0020)	t(464.00) = -1.116	0.265	**0.071 (0.020)**	t(422.58) = 3.584	< 0.001
Sex	**-0.112 (0.028)**	t(65.07) = -4.058	< 0.001	**-0.061 (0.022)**	t(65.00) = -2.805	0.007	-0.146 (0.161)	t(65.03) = -0.905	0.369
Object × Sex	–	–	–	**-0.038 (0.019)**	t(464.00) = -2.045	0.041	**0.625 (0.183)**	t(398.07) = 3.420	< 0.001
Variance components									
Variance (Individual)	**0.022 (0.150)**	–	–	**0.006 (0.079)**	–	–	**0.633 (0.796)**	–	–
Variance (Group)	**0.004 (0.061)**	–	–	–	–	–	**0.212 (0.460)**	–	–
Residual	0.038 (0.194)	–	–	0.011 (0.107)	–	–	1.085 (1.041)	–	–
Likelihood Ratio Test	$\chi^2(2) = 19.951$	–	$p < 0.001$	–	–	–	$\chi^2(2) = 14.343$	–	$p < 0.001$

Question	**Hypothesis**	**Sampling plan**	**Analysis Plan**	**Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis**	**Interpretation given different outcomes**	**Theory that could be shown wrong by the outcomes**
Does the individual degree of neophobia differ across social contexts in a highly social species, the herring gull?	We hypothesise that the distribution of neophobic responses will change in a group context. Specifically: a.) There is a reduction of the variance in group tests. b.) The average response differs between group/individual tests, depending on the social mechanism at play	We will test 80 herring gulls twice across a 2x2 design. These four distinct conditions are: individual or group tests paired with a control or novel object. Each condition will be repeated twice. In the 'novel object' condition, birds are exposed to a pseudo-randomly selected novel object. Conversely, the 'control object' condition involves a familiar object, previously placed in their home enclosure for six days before testing. Testing trials will be randomised, see Supplementary table 1 in the main manuscript	A (G)LMM with Type III sum of squares will be performed on the different latency measures. Models will be fitted to the different latency types separately as well as combined. For the combined analysis, we will use the approach proposed by Snijders and Boskers (2012), which allows for the simultaneous analysis of multiple dependent variables in the case of nested data structures, thereby	*A-priori* power sensitivity analyses were conducted in G*Power (Erdfelder et al., 2009), using a MANOVA. This indicated that our sample size of 80 animals is sufficient to detect a small effect of *Context, Group* and *Trial*. However, we will analyse our data with (G)LMMs, which are currently not covered by G*Power or most other power-estimation tools. These models are	If social context fails to modulate variance, or group means, it could suggest that social contexts hold little significance for neophobic responses among herring gulls.	Social context may either modulate the group mean, the variance, or both. The risk dilution hypothesis suggests that being in a group will reduce both the mean and the variance of neophobia. Conversely, the negotiation hypothesis predicts an increase in mean neophobia but a decrease in within-group variance. The social conformity hypothesis predicts no change

		for a detailed testing schedule. Testing groups comprise 5 individuals by semi-randomly allocating gulls to one group. We will split nest mates across groups. Sexing is unfeasible prior to testing. While we will consider sex differences in our statistical analyses, we do not expect an effect of sex since herring gulls only reach sexual maturity at 4-years of age. Groups may also include a lesser black-backed gull. We will include all gulls for testing but will remove the lesser black-backed gulls prior to conducting the statistical analysis.	considering within-group and between-group variance in latency measures. The model will incorporate *Object*, *Context*, their interaction and *Trial* as fixed effect. A random slope for *Group* associated with each *GroupID* will be included focusing on the variability of the group effect. Moreover, the variability in the individual response due to being in a group or not will be modelled as random effects within *BirdID.*	more flexible in assigning variance as they allow for the specification of both fixed and random effects. However, by accounting for unexplained variance, our proposed mixed-effect models are more powerful than the fixed-effect MANOVAs used in our sensitivity analyses.		in mean neophobia, but a decrease in variance. The design of our study allows us to validate or refute each of these hypotheses.