Foraging group size affects repeatable individual variation and strategic social plasticity

Rori E. Wijnhorst[1*], Corné de Groot[1], Yimen G. Araya-Ajoy[2], Henrik Jensen[2], Jonathan Wright[2], Niels J. Dingemanse[1]

[*]Corresponding author:

Email: roriwijnhorst@proton.me

Affiliations:

[1]Faculty of Biology, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany

[2]Department of Biology, Norwegian University of Science and Technology (NTNU), Trondheim N-7491, Norway

Abstract

Although game theory successfully predicts how negative frequency dependence can maintain behavioural variation through stable equilibria, it does not predict whether this variation reflects repeatable differences among individuals or variation within individuals. Because social group size influences how individuals respond to social information, it may shape how individual variation is distributed around equilibria. Using a producer-scrounger game model, we therefore tested whether group size (three versus six individuals) affects within- and among-individual variation in foraging tactics in wild house sparrows (*Passer domesticus*). While average scrounging frequency did not change with group size, individuals were less socially responsive and among-individual differences in strategies increased in larger groups. Furthermore, strategies were strongly correlated across treatments, indicating individual consistency in foraging strategies. Non-linear selection analyses revealed negative frequency-dependent selection in small groups, with a locally stable equilibrium (ESS) at a scrounging proportion of 0.332, but no evidence for selection in larger groups. Our findings illustrate that group size can shape within- and among-individual variation in behaviour, thereby altering the individual-level processes through which evolutionary equilibria are maintained. This highlights the need to include continuous among-individual variation and context-dependent social responsiveness in game theoretical models to better predict evolutionary dynamics in natural populations.

Introduction

The fitness consequences of behavioural tactics often depend on the behaviour of social partners (Dingemanse & Wolf, 2013; Maynard-Smith, 1982; Taborsky & Oliveira, 2012). Negative frequency-dependent selection, whereby the fitness of a tactic decreases as it becomes more common, can maintain multiple alternative tactics within populations. Evolutionary game theory formalises these interactions by predicting how the pay-offs of alternative tactics change with their frequency, thereby identifying equilibrium tactic frequencies (evolutionarily stable strategies; ESSs) (Maynard-Smith & Price, 1973; Maynard-Smith, 1982). However, the same equilibrium frequencies can be generated through different behavioural processes (see Figure 1). For example, in the hawk-dove game, a population could be at an equilibrium state consisting of a fixed proportion of individuals that consistently adopt one fixed strategy (e.g. 50% of individuals are

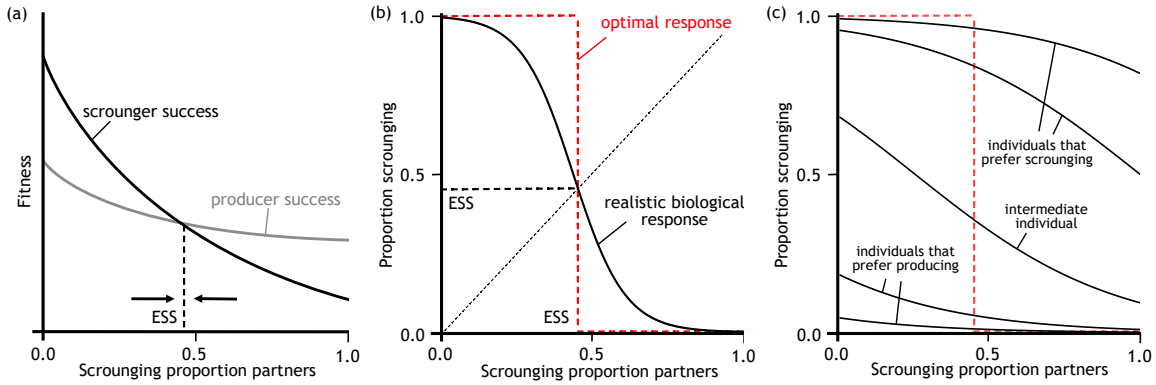

Figure 1: ***Conceptual framework of behavioural responses in the producer-scrounger game.*** *(a) Pay-offs of producing and scrounging depend on the frequency of scroungers, generating a stable equilibrium (ESS) where both tactics yield equal fitness. (b) Individuals are expected to adjust tactic use according to the relative pay-offs of the two tactics. Rather than switching instantaneously at the ESS, biological responses are expected to be gradual. As the pay-off difference between the tactics increases, the probability that an individual chooses the more rewarding tactic increases, resulting in a sigmoidal (e.g. logistic) relationship between the social environment and tactic use. The diagonal dotted line (x=y) represents tactic use frequency of the focal and the social environment are equal. The intersection of the optimal response line (in red) with this diagonal line, is where the pay-offs are equal for all, which represents the predicted equilbrium. (c) Individuals may differ consistently in their average tendency towards a tactic (individual strategy), represented by differences in intercepts of the behavioural response function. The above figures were inspired by Giraldeau and Beauchamp (1999) and McNamara and Leimar (2020).*

always hawk and 50% always dove) or individuals show flexible random variation in tactic use with fixed probabilities (e.g. all individuals display the hawk tactic 50% of the time and the dove tactic 50% of the time; Wolf & Weissing, 2010). More generally, populations may reach the same ESS because individuals differ consistently in their average tactic use (among-individual variation), because individuals plastically adjust their tactic (within-individual variation) or through a combination of both (e.g. half of individuals adopt hawk 70% and dove 30% of the time, while the other half adopt hawk 30% and dove 70% of the time; Figure 1c; Dubois et al., 2010; Wolf & Weissing, 2010; Wolf et al., 2008, 2010). Although game theory successfully predicts the population-level equilibria, it does not explain when equilibria are expected to arise through among-individual differences versus within-individual plasticity, or what determines the balance between these sources of variation.

An important, but largely unexplored, question is whether the balance between among- and within-individual variation around these equilibria depends on the social environment in which interactions occur. Within-individual variation can reflect strategic responses to changing environmental or social conditions. A key environmental factor is group size, which may alter within-individual behavioural variation by changing how individuals perceive and respond to their social environment (King & Cowlishaw, 2007; Liker &

Bókony, 2009). When individuals modify their tactic use in response to the phenotype of social partners, this behavioural plasticity is referred to as social responsiveness (Araya-Ajoy et al., 2020; Dingemanse & Araya-Ajoy, 2015; Taborsky & Oliveira, 2013; Wolf et al., 2008). Quantitative genetics offers a framework for quantifying the genetic basis of such social responsiveness, as an indirect genetic effect (Moore et al., 1997). Game theory has proposed that, as groups become larger, individuals may have fewer opportunities to learn about specific group members, thereby hindering individuals' ability to accumulate information about their social partners, such as their personality or fighting ability (McNamara & Leimar, 2020). This could lead to an increase in choosing the erroneous (less rewarding) tactic, and thereby weaken the social response. For example, it has been shown that individuals respond less strongly to alarm calls in larger groups, because false alarms occur more often (Gray & Webster, 2023). Thus, larger groups may weaken social responsiveness by reducing the reliability of social information.

Group size may also alter the expression of among-individual behavioural variation. Larger groups may favour greater individual differences when consistent behavioural specialization becomes increasingly beneficial, as observed in systems exhibiting division of labour (Amador-Vargas et al., 2015; Holbrook et al., 2013; Ulrich et al., 2018). This may be particularly relevant when individuals interact with multiple social partners whose behaviour varies over time. In larger groups, the behaviour of any single individual has a smaller influence on the group-level social environment (Hadfield & Wilson, 2007; Heidaritabar et al., 2019), while individuals may encounter a greater number of social partners and potentially more heterogeneous social information. When social information provides limited predictive value about the behaviour of individual partners, consistently expressing a particular tactic may become relatively more advantageous than adjusting behaviour to each social partner (Wolf et al., 2010). Thus, group size may alter the amount of among-individual variation and, consequently, the extent to which among-individual variation contributes to reaching population-level equilibria.

Importantly, among- and within-individual behavioural variation may not evolve independently. At the population level, these components represent alternative ways of partitioning total behavioural variation, such that greater expression of within-individual plasticity necessarily reduces the relative contribution of among-individual differences, and *vice versa* (Nakagawa & Schielzeth, 2010). The relative expressions of these components, however, are shaped by biological processes (Araya-Ajoy et al., 2023; Westneat et al.,

2015). General theory on phenotypic plasticity predicts that plasticity is favoured when environmental conditions fluctuate and provide reliable information on the selective environment (Botero et al., 2015; Martin et al., 2025b; Via & Lande, 1985). Under these conditions, individuals are able to respond to environmental cues and maintain a phenotype that matches environmental optima. In contrast, when environmental conditions are unpredictable or cues provide limited information about future selective conditions, plastic responses become less reliable and consistent phenotypes may be favoured (Dewitt et al., 1998; Tufto, 2015). Understanding how different levels of behavioural variation contribute to population-level equilibria therefore requires simultaneous estimation of among- and within-individual variation under different social conditions and relating these patterns to the underlying frequency-dependent fitness landscape.

Ultimately, selection pressures that act on the population-level social responsiveness as well as the repeatable variation among individuals will determine the organisation of behavioural variation. Whereas game theory often focusses on discrete tactics (for example producing-scrounging or hawk-dove), adaptive-dynamics theory provides a framework for analysing the evolution of continuously varying quantitative traits and for characterising the evolutionary properties of population-level equilibria (Dieckmann & Law, 1996; Geritz et al., 1998; Metz et al., 1992), and its predictions can be reconciled with quantitative-genetic models under certain assumptions about genetic architecture (Débarre et al., 2014). Although adaptive dynamics was developed to describe evolutionary change, its framework can also be applied to behavioural dynamics to characterise the costs and benefits of alternative labile phenotypes. In particular, adaptive dynamics theory identifies singular strategies and characterises their evolutionary properties. Convergence stability describes whether evolutionary trajectories are attracted towards a singular strategy, whereas evolutionary stability describes whether a singular strategy is resistant to invasion by alternative strategies (Geritz et al., 1998). Strong attraction towards a singular strategy may promote social responsiveness, since responsive individuals can adjust their phenotype to environmental fluctuation around the equilibrium and thereby remain closer to the fitness optimum (Botero et al., 2015; Taborsky & Oliveira, 2012). Thus, adaptive dynamics can be useful for predicting the location and stability of equilibrium frequencies in continuously expressed phenotypes and for characterising selection pressures around those equilibria, while simultaneous estimation of within- and among-individual variation can tell us how behavioural variation around those equilibria is expressed.

In this study, we tested the effect of group size on the partitioning of tactic variation within and among individuals by repeatedly assaying wild-caught house sparrows (*Passer domesticus*) in small (groups of three) versus larger groups (groups of six). The design allowed us to test whether the behavioural mechanisms that maintain the coexistence of producer and scrounger tactics are affected by group size. First, we used multivariate mixed-effects models to decompose behavioural variation into average social responses, individual-level tactic tendencies ("individual strategy"), and residual within-individual variation in both group-size contexts. Next, we quantified the fitness landscape of the game using foraging pay-offs to determine whether tactic pay-offs were frequency-dependent, whether a stable equilibrium (ESS) was present, and whether these relationships differed across group sizes. Finally, we tested whether individual tactic tendencies were consistent across the two social contexts. Together, these analyses allowed us to examine how group size affects the behavioural mechanisms through which population-level tactic frequencies are maintained.

Methods

Study system and population

We conducted social foraging experiments in the winter of 2023 as part of a larger two year study (2022 and 2023) with wild house sparrow (*Passer domesticus*) (de Groot et al., 2026a). This meta-population has been monitored since 2012 and consists of four flocks associated with dairy farms within the Åfjord municipality in Norway. After capture using mist-nets, each sparrow was equipped with a passive integrated transponder (PIT)-tag ring on their legs and a small barcode backpack on their backs for visual identification from above. Consequently, the sparrows were assigned to a group of six sparrows, housed together and habituated to foraging on feeder boards over two consecutive days. For this experiment, we selected 70 individuals (out of 140 total captures in 2023) to create twelve groups of six individuals; at least two groups from each flock. One of the twelve groups consisted of four 'inexperienced' sparrows and two sparrows that had just undergone the experiment as part of a different group. These two sparrows were reused, because one flock only consisted of 10 sparrows that year. These two individuals spent an additional day habituating with their second group after their first round of assays and before their second round of assays commenced. See

de Groot et al. (2026a) for further details on catching, housing, habituation and training of house sparrows.

Experimental set-up

After habituation, house sparrows were subjected to foraging assays where they interacted in smaller groups of three individuals (small group assays), where all individuals interacted using all 20 possible combinations of three from the larger group of six. These same individuals were subjected to assays where they interacted with all members of the full group of six individuals (large group). To be able to account for treatment order, groups within each flock were evenly assigned to treatment order, with half of the groups starting with the small group assays and the other half with the large group assays. The small group assays consisted of two days of behavioural testing to comply with the study design of the larger two year study of de Groot et al. (2026a) and de Groot et al. (2026b). We used RFID feeders with 36 wells equipped with four antennae per well, which could register four birds at the same time. We used 12 different random patterns to distribute 22 cups filled with sand into the wells and 14 cups filled with seeds hidden under a thin layer of sand. These 14 cups contained in total 12 g of millet seeds. In group trials, two RFID feeders were joined together (Figure S1). We used the same protocol to fill the single feeders, resulting in double the amount of food distributed over the larger feeder system for the large group assays. Specifically, there were 24 grams of seeds divided over 28 wells in the combined feeder, with 44 wells filled with only sand. Therefore, if individuals move randomly across the enlarged feeder they have the same chance of finding food or encountering other individuals as in the small group assays.

Small group assays

All small group assays started at approximately 08:00. Two groups of six would be assayed on the same day. Each round, three groups of three would be assayed simultaneously in three separate cages equipped with RFID feeder board. During the assay round, sparrows would forage for 15 minutes in a group of three. Then, after a 25-minute break, the next round of assays would start until all 20 combinations of three sparrows would have been assayed. In each round, one triad of sparrows from a group would not participate and have an extended break to maintain motivation to forage, every round alternating between the two groups. The last assay would typically start around 16:00.

Large group assays

Two groups of six were assayed on the same day, alternating every half an hour. One group started at 08:00 with the first trial of the day, and the second group started at 08:30. Six individuals foraged for 15 minutes inside the feeder cage with one-hour intervals. The average inter-assay interval per bird was slightly longer than in the small group assays (45 minutes vs 35 minutes), due to the time required for preparing subsequent assays. The last assay for the first group would start around 17:00 and for the second group around 17:30.

Behavioural scoring

The RFID systems registered each individual as they moved across the board. Using these recordings, we acquired rule-based scores of producing and scrounging (P/S). We tested whether these RFID-derived scores correlated with producing and scrounging scores obtained by human observers from a subset videos analysed using BORIS behavioural scoring software. To obtain the video observation P/S scores, seven observers were trained until they achieved inter- and intra-observer reliability exceeding 0.85 for all observed behaviours. Next, the observers scored all individuals in a subset of 68 videos for small groups (totalling 204 P/S scores) and all individuals in a subset of 12 videos for large group assays (72 P/S scores). The RFID and video scores of 2023 showed a Pearson correlation of 0.71 for producing events and 0.72 for scrounging events. The RFID-quantified number of consumed seeds showed a correlation of 0.80 with the measured number of seeds based on weighing after the assays. In all subsequent analyses we used RFID-derived scores, as we expect the RFID measures to be unbiased and standardised across all assays. For details on RFID quantification and video scoring see de Groot et al. (2026a).

Statistical analysis

We used two bivariate models to compare the differences between contexts in: (i) the tactic use of individuals in response to partners' tactic use; and (ii) the pay-offs (foraging success) of focal individuals in relation to their own and their partners' tactic use. In both models, we accounted for variables arising from the experimental design, including the sequential order of an assay within an experimental group (Seq; modelled as 1–10 and mean-centred) and whether the large group assays were conducted first (Grp; coded -0.5/0.5). Because the small group assays occurred on consecutive days, we accounted for the effect of whether the small group assay took place on the first or second testing day (Day; coded -0.5/0.5). In all our analyses, we included observations if focal individuals had at least one foraging event. Observations from focal individuals were excluded when no social partners in the assay had at least one foraging event. In total, this yielded 1153 observations from small group assays, and 462 observations from large group assays.

Behavioural response model

In the first model, we modelled the probability that a focal individual adopted the scrounging tactic as a binomial response with a logit link, based on the number of scrounging events relative to the total number of foraging events (producing and scrounging events) within an assay. The linear predictors included fixed effects describing the experimental design and the average scrounging proportion of the focal individual's social partners (population-level social responsiveness). Random intercepts were fitted for individual identity, trial and experimental group. Individual random intercepts quantified consistent differences in average tactic use (individual strategy). Random intercepts for individual identity were fitted jointly across the small and large group contexts in a bivariate framework, allowing estimation of among-individual variance in strategies within each social context and their covariance across contexts. Random intercepts for trial and experimental group accounted for the sampling structure of the experiment.

The number of scrounges n given the total foraging events N of focal individual i in assay t in group g are modelled as follows:

$$
\begin{aligned}
\text{logit}(p_{S_{itg}}) &= \mu_S + \text{ID}_{S_i} + \beta_S \bar{s}_{T,-i} + u_{S_g} + v_{S_t}, \\
\text{logit}(p_{L_{itg}}) &= \mu_L + \text{ID}_{L_i} + \beta_L \bar{s}_{G,-i} + u_{L_g} + v_{L_t}, \\
n_{S_{itg}} &\sim \text{Binomial}(N_{S_{itg}}, p_{S_{itg}}), \\
n_{L_{itg}} &\sim \text{Binomial}(N_{L_{itg}}, p_{L_{itg}}).
\end{aligned}
\tag{1}
$$

$$
\begin{pmatrix} \text{ID}_{S_i} \\ \text{ID}_{L_i} \end{pmatrix} \sim \text{MVN}\left(\mathbf{0}, \begin{pmatrix} \sigma^2_{\text{ID}_S} & \sigma_{\text{ID}_S\text{ID}_L} \\ \sigma_{\text{ID}_S\text{ID}_L} & \sigma^2_{\text{ID}_L} \end{pmatrix} \right).
$$

Here, S and L denote the small and large group assays, respectively. The parameter μ represents the population mean tactic use, ID_i is the individual-specific deviation in average tactic use (individual strategy), β is the population-level effect of partner scrounging frequency (social responsiveness), s_{-i} is the mean scrounging proportion of the social partners (all individuals within the assay except the focal individual), and u_g and v_t are random intercepts for experimental group and assay, respectively. Fixed effects accounting for the experimental design were omitted from Equation 1 for clarity.

Relative tactic pay-offs model

In the second model, we quantified the relationship between foraging success and the tactic use of both the focal individual and its social partners. Foraging success was modelled as a Gaussian response, which represented the number of seeds an individual consumed in the assay. The fixed effect predictors included linear and quadratic effects of focal and partner scrounging frequency, together with their linear interaction, thereby capturing frequency-dependent effects and allowing the fitness landscape of the producer–scrounger game to be estimated. Additional fixed effects accounting for the experimental design were included as described above. Furthermore, we accounted for the total number of foraging events (number of producing plus scrounging events) of both the focal individual and the partners. In this way, the fitness outcome represents the pay-off per used-tactic which aligns with game theoretical models (Barnard & Sibly, 1981):

$$
\begin{aligned}
Y_{S_{itg}} &= \mu_S + f(s_i, \bar{s}_{-i}) + \mathrm{ID}_{S_i} + u_{S_g} + v_{S_t} + \varepsilon_{S_{itg}}, \\
Y_{L_{itg}} &= \mu_L + f(s_i, \bar{s}_{-i}) + \mathrm{ID}_{L_i} + u_{L_g} + v_{L_t} + \varepsilon_{L_{itg}},
\end{aligned}
\tag{2}
$$

$$
f(s_i, \bar{s}_{-i}) = \beta_1 s_i + \beta_2 \bar{s}_{-i} + \beta_3 s_i^2 + \beta_4 \bar{s}_{-i}^2 + \beta_5 s_i \bar{s}_{-i}.
$$

$$
\begin{pmatrix} \mathrm{ID}_{S_i} \\ \mathrm{ID}_{L_i} \end{pmatrix} \sim \mathrm{MVN} \left(\mathbf{0}, \begin{pmatrix} \sigma^2_{\mathrm{ID}_S} & \sigma_{\mathrm{ID}_S \mathrm{ID}_L} \\ \sigma_{\mathrm{ID}_S \mathrm{ID}_L} & \sigma^2_{\mathrm{ID}_L} \end{pmatrix} \right).
$$

Here, Y denotes the number of seeds obtained per foraging event (foraging success) transformed into a relative fitness measure by dividing by the arithmetic mean of individual foraging success across all assays. This standardized measure allows comparison of average foraging success across assay types. Here, μ is the population mean foraging success, $f(s_i, \bar{s}_{-i})$ describes the fitness landscape as a function of the focal individual's scrounging proportion (s_i) and the mean scrounging proportion of all other individuals within the assay ($\bar{s}_{-i}$). We report doubled quadratic regression coefficients so that they represent stabilising/disruptive selection gradients (Endler, 1986; Stinchcombe et al., 2008). The individual-specific deviation in average foraging success is described by ID_i, and u_g and v_t are random intercepts for experimental group and assay, respectively. The residual error is represented by ε. Individual random intercepts were modelled jointly across contexts using a bivariate normal distribution, allowing estimation of among-individual variances in average foraging success within each context and their covariance across contexts. Fixed effects accounting for the experimental design, together with the total number of foraging events performed by the focal individual and its social partners, were omitted from Equation 2 for clarity.

Adaptive dynamics analyses

We estimated the optimal response function of the focal individual for a given social environment ($\bar{s}_{-i}$) by differentiating the fitted fitness landscape with respect to the focal individual's scrounging proportion and solving for the tactic that maximized expected fitness for each scrounging proportion of the social partners:

$$\frac{\partial f(s_i, \bar{s}_{-i})}{\partial s_i} = \beta_1 + 2\beta_3 s_i + \beta_5 \bar{s}_{-i} = 0. \quad (3)$$

$$\text{optimal response}(\bar{s}_{-i}) = -\frac{\beta_1 + \beta_5 \bar{s}_{-i}}{2\beta_3}. \quad (4)$$

The singular strategy point s^* was calculated as the intersection between the optimal response function and the line $s_i = \bar{s}_{-i}$ (Figure 1; Maynard-Smith, 1982; McNamara & Leimar, 2020). Next, we determined whether the singular strategy point represents an evolutionary stable equilibrium using standard adaptive dynamics criteria: convergence stability and evolutionary stability (Christiansen, 1991; Geritz et al., 1998). The convergence criterion assesses whether evolutionary change in the population is directed towards the singular strategy. The evolutionary stability criterion assesses whether alternative strategies can invade when the population is at the singular strategy point. For the fitted fitness function, the convergence stability criterion reduces to $2\beta_3 + \beta_5 < 0$, whereas evolutionary stability requires $2\beta_3 < 0$ (Geritz et al., 1998). Together, these criteria identify whether the singular strategy represents a locally attracting and invasion-resistant monomorphic equilibrium. Pairwise invasibility plots (PIPs) were generated by evaluating the relative success of focal strategies (s_i) across a range of social environments ($\bar{s}_{-i}$), with both ranging from 0 to 1 (Geritz et al., 1998). Relative success ("invasion fitness") was calculated as the difference between the predicted success of a focal individual using strategy s_i and that of an individual using the same strategy as the social environment, $I(s_i, \bar{s}_{-i}) = f(s_i, \bar{s}_{-i}) - f(\bar{s}_{-i}, \bar{s}_{-i})$. Positive values indicate that the focal strategy achieves higher success than the strategy matching the social environment. PIPs were generated using the posterior median estimates of the selection coefficients and are presented as a descriptive visualisation (Figure S2).

Preregistration

The planned analyses for this study were pre-registered on OSF (Wijnhorst et al., 2024). Similarly to the performed analyses, we proposed a bivariate model to investigate among- and within-individual variation between the two contexts. In the proposed plan, we suggested analysing individual differences in average foraging tactic and predictability (among-individual variation in within-individual variation). However, we

changed the model to logistic regression, which better fits the binomial nature of discrete tactic use and the game-theoretical framework. Since, predictability is defined as among-individual variation in residual variation, and residual variation is not estimated in binomial models, we could no longer estimate this component. Instead, we initially considered among-individual variation in social responsiveness, which also represents a component of within-individual plasticity. The estimated among-individual variation in social responsiveness was however negligible, and retaining this random-slope component in a bivariate framework would have unnecessarily increased model complexity. We therefore excluded this component from the final model. The final analyses consequently focus on among-individual differences in average tactic use and population-level social responsiveness.

The revised analyses represent a deviation from the preregistered analysis plan. The change in model choice was made before any analyses of the study data were conducted and was motivated by the distribution of the response variable and the aim of directly relating individual behavioural responses to game-theoretical predictions. The subsequent exclusion of the random-slope component was based on the absence of meaningful among-individual variation in social responsiveness and the resulting risk of overparametrisation.

Model specification and reporting

All models were implemented in the Stan Bayesian probabilistic programming language using the rstan package (v 2.32.7, Stan Development Team, n.d.) in R (v 4.5.1, R Core Team, 2025). Posterior distributions were estimated using four Markov chains, each run for 8,000 iterations with 2,000 warm-up iterations. Fixed effects were assigned weakly informative normal priors ($\mu = 0, \sigma = 1$), and standard deviation parameters were assigned exponential priors ($\eta = 2$). Correlation matrices were assigned an LKJ prior with shape parameter 2. Model convergence was assessed using trace plots with r package shinystan (version 2.6.0, Gabry & Veen, 2025), effective sample sizes, and $\hat{R}$ diagnostics. All parameters showed good convergence ($\hat{R} < 1.01$). We report posterior median estimates and 95% credible intervals.

Results

Individual tactic use in response to partners' tactics

Using two types of foraging group contexts, we tested whether group size influences average producer-scrounger tactics and variation dynamics. We found that the average amount of scrounging did not differ between contexts. By converting the log-odds estimate of the population intercept into a probability, we calculated that the population average proportion of scrounging behaviour was 0.262 (CI: 0.191, 0.351) in the small group context, and 0.311 (CI: 0.225, 0.407) in the large groups. The proportion showed no credible difference, therefore foraging group size did not systematically influence the proportion of scrounging.

We found that individuals strongly decreased their scrounging in response to the amount of scrounging used by their partners in the small groups (β_S = -6.946; CI: -7.973, -5.944, Figure 2, Table S1)). This social response was still present in the large groups, however markedly decreased in strength (β_L = -3.667; CI: -5.553, -1.719, Δ_{S-L} = -3.285; CI: -5.455, -1.115).

We found clear evidence for individual differences in strategies (intercepts) in both contexts. Importantly, these individual differences were preserved across contexts. Individual-level intercepts were very strongly correlated by 0.790 (CI: 0.418, 0.964), indicating that individuals that scrounged more in one context also tended to do so in the other (Figure 3). This provides strong evidence that individual strategy is a stable trait across social contexts.

Interestingly, group-size context had a strong influence on the distribution of phenotypic variation. Among-individual variation was much greater in the large groups, where it explained 14.4% (CI: 6.6%, 26.0%) of the variation in tactic use, compared to the small groups, where it only explained 1.7% (CI: 0.4%, 4.4%) of the variation (Δ_{S-L} = -12.6%; CI: -24.2%, -4.6%). Furthermore, social responsiveness contributed a substantially larger proportion of the total phenotypic variation in small groups (32.1%; CI: 27.5%, 36.4%) than in the large groups (8.7%; CI: 2.3%, 16.2%, Δ_{S-L} = 23.4%; CI: 14.6%, 31.1%). In contrast, residual within-individual variation was higher in the large groups. Total within-individual variation (social response variance plus residual variance) was nevertheless nearly identical across contexts (Δ_{S-L} = 0.2%; CI: -11.8%, 11.7%), explaining 61.6% and 61.4% of the total phenotypic variation in the small and large groups, respectively.

Fixed effects controlling for study design parameters were generally small, with all (except one) of the credible intervals overlapping zero. We found that individuals scrounged more when the large group assays were conducted first in the small group context (β_S = 0.476; CI: 0.048, 0.907). This showed that "treatment" order mattered for the average producing-scrounging behaviour in the small groups, but it did not for the large groups.

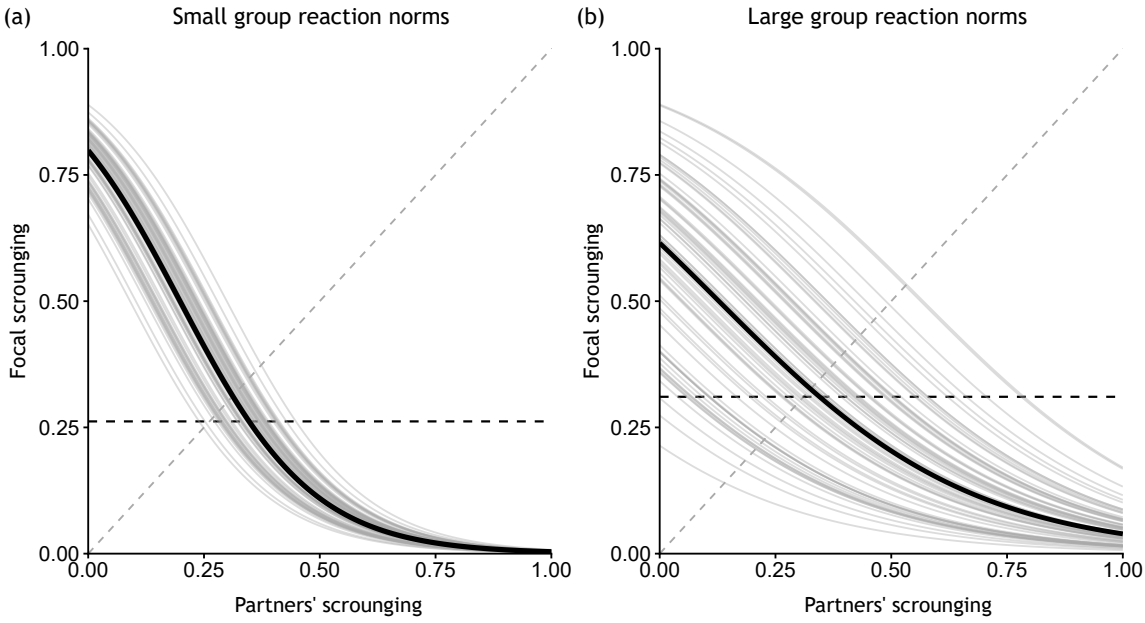

Figure 2: ***Individual reaction norms in small groups of three (a) and large groups of six (b).*** *The bold line in black represent the population average response based on the posterior median estimate of the population slope and intercept. Each grey line represents the response of one individual based on the best linear unbiased predictors (BLUPs) of their individual intercept deviation, in combination with the population-level slope and intercept. The black horizontal dotted line represents the estimated population average scrounging. The diagonal dotted line represents the x=y line, which are potential equilibria.*

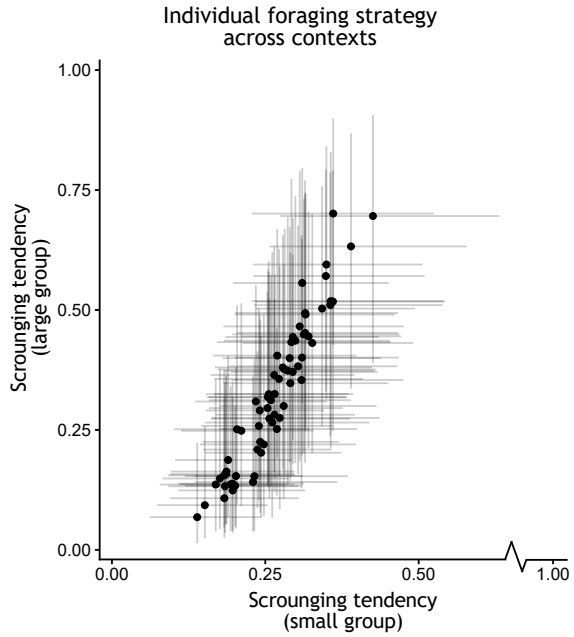

Figure 3: ***Among-individual correlation in tactic use between the two group-size contexts.*** *Each point represents the posterior median of the intercept deviation of an individual (BLUP), which was transformed to represent their average scrounging proportions. Horizontal and vertical bars show the 95% credible interval of the posterior, representing uncertainty around each estimate. The correlation was strongly positive (r = 0.790 CI: 0.418, 0.964).*

Relative tactic pay-offs in relation to partners' tactics

We compared relative tactic pay-offs between group-size contexts by assessing the pay-off landscapes generated from the two experimental treatments. We observe that the average tactic yielded a higher pay-off in small groups than in large groups ($\Delta_{\text{S-L}}$ = 0.148; CI: 0.002, 0.295, Table S2). Furthermore, foraging success naturally increased when individuals performed more foraging events per trial. Interestingly, foraging success also increased as partners performed more foraging events, and this increase was stronger in small groups than in large groups ($\Delta_{\text{S-L}}$ = 0.066; CI: 0.044, 0.089). Thus, partner foraging activity was associated with increased focal pay-offs, particularly in small groups, suggesting that focal individuals benefited from increased partner foraging activity.

We found stabilising selection for the proportion of focal scrounging (β_S = -0.748; CI: -1.332, -0.170), as well as a negative interaction between proportion of focal and partner scrounging (β_S = -0.438; CI: -0.823, -0.054) in the small groups. Thus, the fitness consequences of focal tactic use depended on partner tactic

use, with scrounging becoming less rewarding as partners scrounge more. In the large groups, we did not find strong evidence for selection on focal or partner scrounging, since all credible intervals overlapped zero. The visualised dome-shaped landscape (Figure 4b) was based on the posterior median estimates, and should therefore not be considered evidence for stabilising selection. Furthermore, none of the selection estimates showed clear differences between the two foraging-group-size contexts.

Among-individual differences accounted for 37.5% (CI: 27.0%, 49.0%) of the variance in foraging success in large groups, compared with 16.0% (CI: 10.7%, 23.4%) in small groups ($\Delta_{S\text{-}L}$ = -21.2%; CI:-32.9%, -9.9%). Conversely, a greater proportion of variance was attributable to differences among assays in small groups than in large groups ($\Delta_{S\text{-}L}$ = 0.141; CI: 0.055, 0.219).

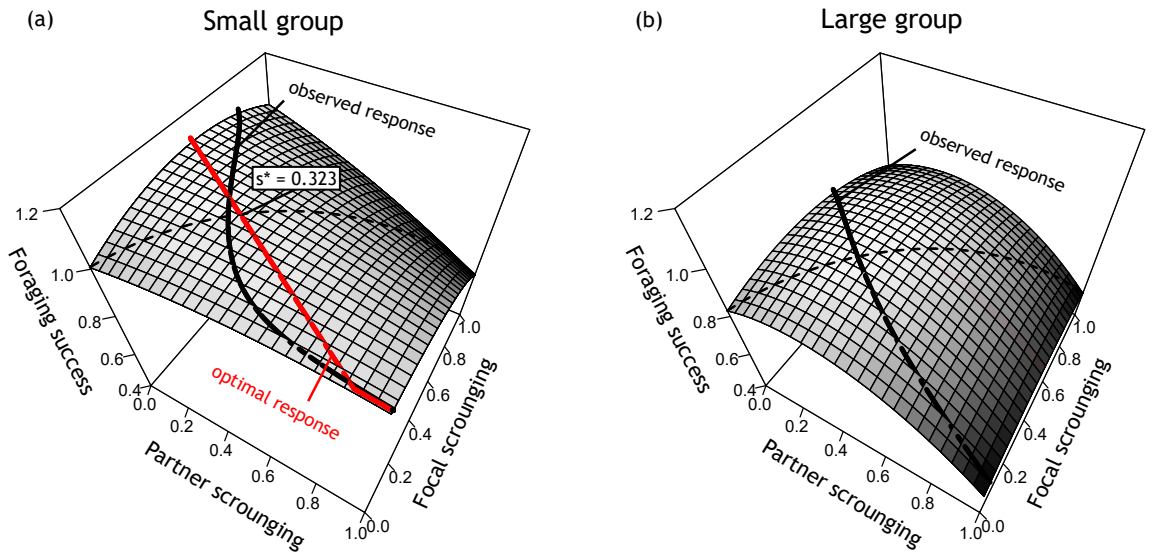

Figure 4: ***Relative tactic pay-offs for the two group-size contexts.*** *The estimated pay-off landscapes are described by the posterior median estimates of the selection coefficients. In the small group treatment, the red line describes the optimal response of focal individuals for each scrounging proportion of the partners. The dotted line represents the x=y line, where focal phenotype is the same as the partner's proportion. At the intersection of these two lines lies the estimated location of the singular point s*. All credible intervals of the selection coefficients of the large group overlapped zero, resulting in high uncertainty in the estimated optimal response and singular point; these are therefore not shown for the large group. The black line represents the population-average response from Figure 2, estimated in the first analysis and projected onto the landscape.*

Adaptive dynamics analysis

Using the selection estimates, we calculated a singular point at a scrounging proportion of 0.332 (CI: 0.152, 0.418) for small groups. This point met both the convergence stability and evolutionary stability criteria,

as both estimates and their credible intervals were lower than zero. This suggests that the singular point represents a locally attracting and invasion-resistant equilibrium.

For large groups, there was insufficient evidence for a singular strategy, since the credible intervals exceeded the biologically possible range between 0 and 1 (CI: -0.178, 1.113). Consequently, quantification of convergence stability and evolutionary stability is not reported (Table 1).

Table 1: ***Posterior estimates of the singular strategy and its evolutionary properties.***

Assay	**Singular point s^***	**Convergence stability**	**Evolutionary stability**
Small group	0.332 [0.152, 0.418]	-1.188 [-1.956, -0.425]	-0.748 [-1.332, -0.170]
Large group	0.445 [-0.178, 1.113]	-	-

Discussion

This study aimed to reveal how foraging group size affects the expression of within- versus among-individual variation in social foraging behaviour. We found that population-level producing and scrounging frequencies remained similar between small and large groups, despite substantial differences in the behavioural processes underlying tactic use. Using an adaptive-dynamics framework, we showed that the payoff landscape in small groups contained a locally stable equilibrium, whereas evidence for a corresponding equilibrium was absent in larger groups. Individuals differed consistently in their strategies and largely maintained their average strategies across group-size contexts. Individuals also showed strong negative frequency-dependent responses to partner tactic use, but these responses were weaker in larger groups. Thus, group size was associated with a shift in the relative contribution of social responsiveness versus among-individual differences to tactic variation, without substantially altering population-level tactic frequencies.

Contrary to previous studies, we did not find a change in the average tactic when the group size increased. Many producer-scrounger models predict that scrounging increases with group size (Barnard & Sibly, 1981; Coolen et al., 2007; Giraldeau & Caraco, 2000; Vickery et al., 1991). Empirical studies, however, have not consistently shown the same effect of group size. Although some have reported increases in scrounging in larger group (Aplin & Morand-Ferron, 2017; Coolen, 2002), others have reported decreases in scrounging (Haan & Kooreman, 2002; Isaac et al., 1994), as well as no effect of group size on scrounging (Liker & Bókony, 2009). Dubois and Richard-Dionne (2020) offer an explanation for the discrepancy between

the theoretical predictions and the empirical findings. They show in a simulation study that when the rate of resource discoveries increases with group size, the proportion of individuals playing scrounger should also decrease. This suggests that the effect of group size depends on how group size affects opportunities for scrounging. Importantly, an exploratory inspection of our data indicated higher average productivity in the large groups, both in terms of seeds consumed per individual and number of foraging events per individual. Furthermore, our results show that as partners were more active (higher number of producing and scrounging events), the foraging success of their group mates increased. This indicates that there was no strong competition for resources in our system and could explain why we did not find an increase in scrounging in larger groups. Thus, the predicted increase in scrounging with group size may depend on the extent to which increasing group size alters resource availability or competition among foragers.

The equilibrium point, where the pay-offs of producing and scrounging are equal for all, would in theory predict the population-average proportion of producing and scrounging (Barnard & Sibly, 1981). In the small group-context, we identified a locally stable equilibrium at a scrounging proportion of 0.332 (CI: 0.152–0.418), which overlapped with the observed average scrounging proportion of 0.262 (CI: 0.191–0.351). The fitness landscape further revealed stabilising selection on focal tactic use and negative frequency-dependence between focal and partner scrounging. Thus, in small groups, the fitness landscape revealed a stable equilibrium and individuals showed strong negative frequency-dependent responses that closely aligned with the predicted optimal response. Interestingly, the observed response was even more strongly negative frequency-dependent than the response predicted to maximise pay-offs, suggesting that additional benefits or constraints not captured by our fitness measure may contribute to the strength of social responsiveness. In the large groups, by contrast, we did not find evidence for selection on either focal or partner scrounging. Importantly, however, none of the estimated selection coefficients showed clear differences between the two group size contexts, meaning that we cannot conclude that the underlying fitness landscape differed between small and large groups. The weaker evidence for selection in large groups could therefore reflect reduced statistical power, rather than a genuinely weaker selective landscape. Nevertheless, individuals continued to show negative frequency-dependent responses in large groups. One possibility is that past experience created persistent expectations about the rewards associated with tactics in specific environments, such that individuals respond to their expected rather than their current rewards. Previous work shows that foraging strategies can persist following changes in pay-offs, as demonstrated in nutmeg

mannikins (*Lonchura punctulata*) (Morand-Ferron & Giraldeau, 2010) and that learning in adults and early life experience influence foraging strategies in house sparrows (Belmaker et al., 2012; Katsnelson et al., 2008, 2010). So, it is possible that our individuals based their tactic choice on the negative-frequency dependent pay-offs learned throughout their lives and in our experimental set-up. This could explain why the response is still present in large groups, but just decreased in strength.

Moreover, our study suggests that in smaller groups negative frequency-dependent social responsiveness plays a larger role in reaching this equilibrium point as it made a larger contribution to total phenotypic variation, whereas in larger groups consistent individual variation in tactics was more pronounced. This could be because in larger groups individuals can display their preferred tactic more consistently and the need for adaptive switching is reduced. This could be caused by a type of dilution effect, where the total impact of individuals on specific group mates becomes smaller as group size increases (Bijma, 2010; Hadfield & Wilson, 2007). Since the average scrounging proportion is determined by more group members, fluctuations around the equilibrium frequency are expected to become smaller in larger groups. Indeed, our analyses show that the overall amount of variation in the social environment slightly decreased in large groups. When variation in environmental covariates decreases, their contribution to the total expressed phenotypic variation decreases, even if the response function remains stable (Schielzeth & Nakagawa, 2022). As such, we can expect that social responsiveness contributes less to total phenotypic variation when individuals respond to group averages of larger groups, independent of changes in how animals respond to ESS fluctuations.

Furthermore, larger groups also have the potential to reduce the effectiveness of social responsiveness. For example, it could be that changes in the average tactic frequency are less observable to sparrows in larger groups, and it becomes more difficult to optimally respond to such apparently small changes. Increased group size may therefore increase uncertainty in an individual's assessment of the social environment (McNamara & Leimar, 2020). If individuals base their assessment of the population tactic frequency on a limited number of spatially nearby partners, then this assessment will become less precise in larger group sizes. Moreover, if group members randomly interact, repeated interactions with specific partners become less frequent in larger groups, reducing opportunities to gain information about specific partners and their behavioural tendencies (McNamara & Leimar, 2020). Both these processes can cause a decrease

in the strength of social responsiveness, due to an increased chance of choosing the wrong tactic given the uncertainty created by a more complex multi-individual social environment. This explanation aligns with our results where we notice an increase in residual variance (i.e. apparently unexplained tactic choices) along with a decrease in the social responsiveness in larger groups.

Furthermore, individuals differed more strongly in their strategies in larger groups. Specifically, we found that consistent among-individual variation in producing versus scrounging was substantially larger in large groups than in small groups, while the population average remained similar between contexts. Together with the strongly positive cross-context correlation, this indicates that individuals largely maintained their relative behavioural tendencies while becoming more divergent in larger groups (Figure 3). Specifically, individuals with a strong tendency to scrounge in small groups tended to scrounge even more in the large groups, whereas individuals that scrounged least in small groups tended to scrounge even less in large groups. This raises the possibility that larger groups can promote greater tactic specialisation among individuals. Research in support of this idea has found larger social groups are associated with greater individuality across several types of acoustic signalling in avian, chiropteran and rodent species (Mathevon et al., 2003; Medvin et al., 1993; Pollard & Blumstein, 2011; Wilkinson, 2003). However, it is worth noting that many other studies in animal behaviour and psychology have shown an opposite effect, namely a decrease in individuality and an increase in behavioural conformity in larger groups (Bond, 2005; Day et al., 2001; Herbert-Read et al., 2013; Insko et al., 1985). It would therefore be interesting to investigate the mechanisms that determine when increasing group size leads to behavioural diversification versus conformity.

Adaptive dynamics provides a useful framework for studying phenotypic variation, since it specifies conditions under which diversification is expected, particularly through evolutionary branching (Geritz et al., 1998). For traits under strong stabilising selection, we might expect variation to decrease and converge around a fitness optimum (Araya-Ajoy et al., 2023). Our results showed evidence of stabilising selection in small groups that supported the evolutionary stability of the equilibrium. Simultaneously, we found potentially weaker stabilising selection in large groups (Figure 4). Relaxation of stabilising selection could therefore have allowed the repeatable strategies of individuals to diverge further compared to the individual differences observed in smaller groups. At the phenotypic level, processes described by adaptive dynamics, such as evolutionary branching or reduced stabilising selection could manifest as increasing variation in

tactic use; however, this increase in variation could occur both within and among individuals. Our study addresses these processes at the level of behavioural dynamics, and shows that group size can influence the expression of individual-level phenotypic variation. These same processes may influence subsequent evolutionary dynamics when behavioural differences are heritable. As such, adaptive-dynamics theory could be applied in combination with multilevel modelling approaches to investigate how selection specifically acts on different sources of behavioural variation, including genetic differences in breeding values, among-individual differences in mean strategies, and within-individual differences in reaction-norm slopes (Araya-Ajoy et al., 2023; Martin et al., 2025a).

In our experimental design, individuals in the small group treatment interacted with multiple combinations of group members, whereas individuals in the larger six-member group treatment repeatedly interacted with the same partners. Therefore, the individuals in larger groups experienced a more stable social environment. As a consequence, repeated interactions with the same social partners may have had an influence on the estimation of among-individual variation. So, we cannot exclude the possibility that the greater among-individual variation observed in larger group size treatment was partly attributable to the increased temporal stability of the social environment rather than to group size alone. By contrast, we expect our inference regarding social responsiveness to be less sensitive to this limitation, because it was estimated directly from the expressed phenotypes of social partners rather than from their identities. Our design nevertheless also provides several important advantages. First, all individuals were repeatedly assayed in both social contexts, allowing individual cross-context comparisons. Second, persistent differences among the groups can be and were accounted for statistically. Finally, analyses of the complete dataset found that partner identity explained only a small proportion of behavioural variation (1.3% and 2.9% per partner for producing and scrounging behaviour, respectively; de Groot et al., 2026a), suggesting any effects of repeatedly interacting with the same partners are likely to be small, especially since these indirect effects are often diluted in larger groups (Hadfield & Wilson, 2007; Heidaritabar et al., 2019).

This study revisited classical game theory in the context of contemporary research in behavioural ecology. By integrating game-theoretical predictions with modern approaches to partitioning within- versus among-individual behavioural variation, we provide an example of how theoretical predictions can be more thoroughly examined and tested using experimental datasets. This approach allowed us to estimate

population-level equilibrium frequencies and investigate the behavioural processes through which these equilibria are maintained in natural populations. Using adaptive-dynamics theory, we characterised the pay-off landscape and equilibrium properties of a social game under different social conditions. We found that group size can have a strong influence on how behavioural variation is structured without affecting mean tactic use: social responsiveness was weaker in larger groups, while variation in repeatable individual strategies increased. Furthermore, differences in selection between social contexts may have contributed to changes in the distribution of behavioural variation. Finally, our analyses provide an empirical application of adaptive dynamics theory and illustrate how group size can influence the way negative-frequency dependent selection maintains variation in social behaviour in natural populations.

Acknowledgements

We would like to thank the students involved in data collection and behavioural scoring in 2023: Augustin Birot, Ozan Demir, Sophie Larsson, Rok Movh, Claudius Voith, and Laura von Winterfeld. We also thank Merit Pokriefke for commenting on early drafts of the manuscript and Alastair Wilson for discussions about the direction of the study. All animal procedures were conducted in accordance with relevant ethical guidelines for the capture and use of birds in research and received approval from the Norwegian Food Safety Authority (FOTS ID 29007) and the Norwegian Bird Ringing Centre. This research was supported by funding provided by the German Research Foundation (DFG; DI 1694 5–1) to NJD and the Research Council of Norway (RCN; 325826) to YGA and (RCN; 302619) to HJ. Additional support was provided through the RCN Centres of Excellence funding scheme (RCN; 223257).

References

Amador-Vargas, S., Gronenberg, W., Wcislo, W. T., & Mueller, U. (2015). Specialization and group size: Brain and behavioural correlates of colony size in ants lacking morphological castes. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1801), 20142502. https://doi.org/10.1098/rspb.2014.2502

Aplin, L. M., & Morand-Ferron, J. (2017). Stable producer–scrounger dynamics in wild birds: Sociability and learning speed covary with scrounging behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1852), 20162872. https://doi.org/10.1098/rspb.2016.2872

Araya-Ajoy, Y. G., Dingemanse, N. J., Westneat, D. F., & Wright, J. (2023). The evolutionary ecology of variation in labile traits: Selection on its among- and within-individual components. *Evolution*, *77*(10), 2246–2256. https://doi.org/10.1093/evolut/qpad136

Araya-Ajoy, Y. G., Westneat, D. F., & Wright, J. (2020). Pathways to social evolution and their evolutionary feedbacks. *Evolution*, *74*(9), 1894–1907. https://doi.org/10.1111/evo.14054

Barnard, C. J., & Sibly, R. M. (1981). Producers and scroungers: A general model and its application to captive flocks of house sparrows. *Animal Behaviour*, *29*(2), 543–550. https://doi.org/10.1016/S0003-3472(81)80117-0

Belmaker, A., Motro, U., Feldman, M. W., & Lotem, A. (2012). Learning to choose among social foraging strategies in adult house sparrows (Passer domesticus). *Ethology*, *118*(11), 1111–1121. https://doi.org/10.1111/eth.12013

Bijma, P. (2010). Multilevel Selection 4: Modeling the Relationship of Indirect Genetic Effects and Group Size. *Genetics*, *186*(3), 1029–1031. https://doi.org/10.1534/genetics.110.120485

Bond, R. (2005). Group size and conformity. *Group Process & Intergroup Relations*, *8*(4), 331–354. https://doi.org/10.1177/1368430205056464

Botero, C. A., Weissing, F. J., Wright, J., & Rubenstein, D. R. (2015). Evolutionary tipping points in the capacity to adapt to environmental change. *Proceedings of the National Academy of Sciences*, *112*(1), 184–189. https://doi.org/10.1073/pnas.1408589111

Christiansen, F. B. (1991). On conditions for evolutionary stability for a continuously varying character. *The American Naturalist*, *138*(1), 37–50. https://doi.org/10.1086/285203

Coolen, I. (2002). Increasing foraging group size increases scrounger use and reduces searching efficiency in nutmeg mannikins (Lonchura punctulata). *Behavioral Ecology and Sociobiology*, *52*(3), 232–238. https://doi.org/10.1007/s00265-002-0500-4

Coolen, I., Giraldeau, L.-A., & Vickery, W. (2007). Scrounging behavior regulates population dynamics. *Oikos*, *116*(3), 533–539. https://doi.org/10.1111/.2006.0030-1299.15213.x

Day, R. L., MacDonald, T., Brown, C., Laland, K. N., & Reader, S. M. (2001). Interactions between shoal size and conformity in guppy social foraging. *Animal Behaviour*, *62*(5), 917–925. https://doi.org/10.1006/anbe.2001.1820

de Groot, C., Wijnhorst, R. E., Nafstad, Å. M., Wilson, A. J., Araya-Ajoy, Y. G., Jensen, H., Wright, J., & Dingemanse, N. J. (2026a). Among-trait covariance and cross-year repeatability for direct and indirect individual effects in producer–scrounger behaviour in wild house sparrows. *Journal of Evolutionary Biology*, *39*(4), 485–498. https://doi.org/10.1093/jeb/voag003

de Groot, C., Wijnhorst, R. E., Araya-Ajoy, Y. G., Wilson, A. J., Jensen, H., Wright, J., & Dingemanse, N. J. (2026b). Multivariate social strategies buffer the fitness consequences of phenotype–environment mismatches [Preprint]. *EcoEvoRxiv*. https://doi.org/10.32942/X2N09X

Débarre, F., Nuismer, S. L., & Doebeli, M. (2014). Multidimensional (co)evolutionary stability. *The American Naturalist*, *184*(2), 158–171. https://doi.org/10.1086/677137

Dewitt, T., Sih, A., & Wilson, D. S. (1998). Dewitt tj, sih a, wilson ds. costs and limits of phenotypic plasticity. trends evol ecol 13: 77-81. *Trends in Ecology Evolution*, *13*, 77–81. https://doi.org/10.1016/S0169-5347(97)01274-3

Dieckmann, U., & Law, R. (1996). The dynamical theory of coevolution: A derivation from stochastic ecological processes. *Journal of Mathematical Biology*, *34*(5), 579–612. https://doi.org/10.1007/BF02409751

Dingemanse, N. J., & Araya-Ajoy, Y. G. (2015). Interacting personalities: behavioural ecology meets quantitative genetics. *Trends in Ecology and Evolution*, *30*(2), 88–97. https://doi.org/10.1016/j.tree.2014.12.002

Dingemanse, N. J., & Wolf, M. (2013). Between-individual differences in behavioural plasticity within populations: Causes and consequences [Special Issue: Behavioural Plasticity and Evolution]. *Animal Behaviour*, *85*(5), 1031–1039. https://doi.org/10.1016/j.anbehav.2012.12.032

Dubois, F., Morand-Ferron, J., & Giraldeau, L.-A. (2010). Learning in a game context: Strategy choice by some keeps learning from evolving in others. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1700), 3609–3616. https://doi.org/10.1098/rspb.2010.0857

Dubois, F., & Richard-Dionne, É. (2020). Consequences of multiple simultaneous opportunities to exploit others' efforts on free riding. *Ecology and Evolution*, *10*(10), 4343–4351. https://doi.org/10.1002/ece3.6201

Endler, J. A. (1986). *Natural selection in the wild*. Princeton University Press.

Gabry, J., & Veen, D. (2025). *Shinystan: Interactive visual and numerical diagnostics and posterior analysis for bayesian models* [R package version 2.7.0].

Geritz, S., Kisdi, E., Mesze´NA, G., & Metz, J. (1998). Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, *12*(1), 35–57. https://doi.org/10.1023/A:1006554906681

Giraldeau, L.-A., & Beauchamp, G. (1999). Food exploitation: Searching for the optimal joining policy. *Trends in Ecology & Evolution*, *14*(3), 102–106. https://doi.org/10.1016/S0169-5347(98)01542-0

Giraldeau, L.-A., & Caraco, T. (2000, June). *Social foraging theory*. Princeton University Press.

Gray, L., & Webster, M. M. (2023). False alarms and information transmission in grouping animals. *Biological Reviews*, *98*(3), 833–848. https://doi.org/https://doi.org/10.1111/brv.12932

Haan, M., & Kooreman, P. (2002). Free riding and the provision of candy bars. *Journal of Public Economics*, *83*(2), 277–291. https://doi.org/https://doi.org/10.1016/S0047-2727(00)00157-2

Hadfield, J. D., & Wilson, A. J. (2007). Multilevel Selection 3: Modeling the Effects of Interacting Individuals as a Function of Group Size. *Genetics*, *177*(1), 667–668. https://doi.org/10.1534/genetics.107.075622

Heidaritabar, M., Bijma, P., Janss, L., Bortoluzzi, C., Nielsen, H. M., Madsen, P., Ask, B., & Christensen, O. F. (2019). Models with indirect genetic effects depending on group sizes: A simulation study assessing the precision of the estimates of the dilution parameter. *Genetics Selection Evolution*, *51*(1), 24. https://doi.org/10.1186/s12711-019-0466-6

Herbert-Read, J. E., Krause, S., Morrell, L. J., Schaerf, T. M., Krause, J., & Ward, A. J. W. (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

Holbrook, C. T., Kukuk, P. F., & Fewell, J. H. (2013). Increased group size promotes task specialization in a normally solitary halictine bee. *Behaviour*, *150*(12), 1449–1466. https://doi.org/10.1163/1568539X-00003104

Insko, C. A., Smith, R. H., Alicke, M. D., Wade, J., & Taylor, S. (1985). Conformity and group size: The concern with being right and the concern with being liked. *Personality and Social Psychology Bulletin*, *11*(1), 41–50. https://doi.org/10.1177/0146167285111004

Isaac, R. M., Walker, J. M., & Williams, A. W. (1994). Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups. *Journal of Public Economics*, *54*(1), 1–36. https://doi.org/https://doi.org/10.1016/0047-2727(94)90068-X

Katsnelson, E., Motro, U., Feldman, M. W., & Lotem, A. (2008). Early experience affects producer–scrounger foraging tendencies in the house sparrow. *Animal Behaviour*, *75*(4), 1465–1472. https://doi.org/10.1016/j.anbehav.2007.09.020

Katsnelson, E., Motro, U., Feldman, M. W., & Lotem, A. (2010). Individual-learning ability predicts social-foraging strategy in house sparrows. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1705), 582–589. https://doi.org/10.1098/rspb.2010.1151

King, A. J., & Cowlishaw, G. (2007). When to use social information: The advantage of large group size in individual decision making. *Biology Letters*, *3*(2), 137–139. https://doi.org/10.1098/rsbl.2007.0017

Liker, A., & Bókony, V. (2009). Larger groups are more successful in innovative problem solving in house sparrows. *Proceedings of the National Academy of Sciences*, *106*(19), 7893–7898. https://doi.org/10.1073/pnas.0900042106

Martin, J. S., Araya-Ajoy, Y. G., Dingemanse, N. J., Wilson, A. J., & Westneat, D. F. (2025a). Estimating (non)linear selection on reaction norms: A general framework for labile traits. *Ecology and Evolution*, *15*(10), e72298. https://doi.org/https://doi.org/10.1002/ece3.72298

Martin, J. S., Westneat, D. F., Wilson, A. J., Dingemanse, N. J., & Araya-Ajoy, Y. G. (2025b). Frequency dependence favours social plasticity and facilitates socio-eco-evolutionary feedback in fluctuating environments. *Functional Ecology*, *39*(11), 3061–3082. https://doi.org/https://doi.org/10.1111/1365-2435.70132

Mathevon, N., Charrier, I., & Jouventin, P. (2003). Potential for individual recognition in acoustic signals: A comparative study of two gulls with different nesting patterns. *Comptes Rendus. Biologies*, *326*(3), 329–337. https://doi.org/10.1016/S1631-0691(03)00072-6

Maynard-Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature*, *246*(5427), 15–18. https://doi.org/10.1038/246015a0

Maynard-Smith, J. (1982). *Evolution and the theory of games*. Cambridge University Press.

McNamara, J. M., & Leimar, O. (2020, September). *Game theory in biology: Concepts and frontiers*. Oxford University Press.

Medvin, M. B., Stoddard, P. K., & Beecher, M. D. (1993). Signals for parent-offspring recognition: A comparative analysis of the begging calls of cliff swallows and barn swallows. *Animal Behaviour*, *45*(5), 841–850. https://doi.org/https://doi.org/10.1006/anbe.1993.1105

Metz, J., Nisbet, R., & Geritz, S. (1992). How should we define 'fitness' for general ecological scenarios? *Trends in Ecology Evolution*, *7*(6), 198–202. https://doi.org/https://doi.org/10.1016/0169-5347(92)90073-K

Moore, A. J., III Brodie, E. D., & Wolf, J. B. (1997). Interacting phenotypes and the evolutionary process I: Direct and indirect genetic effects on social interactions. *Evolution*, *51*(5), 1352–1362. https://doi.org/10.1111/j.1558-5646.1997.tb01458.x

Morand-Ferron, J., & Giraldeau, L.-A. (2010). Learning behaviorally stable solutions to producer–scrounger games. *Behavioral Ecology*, *21*(2), 343–348. https://doi.org/10.1093/beheco/arp195

Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews*, *85*(4), 935–956. https://doi.org/10.1111/j.1469-185X.2010.00141.x

Pollard, K. A., & Blumstein, D. T. (2011). Social group size predicts the evolution of individuality. *Current Biology*, *21*(5), 413–417. https://doi.org/10.1016/j.cub.2011.01.051

R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.

Schielzeth, H., & Nakagawa, S. (2022). Conditional repeatability and the variance explained by reaction norm variation in random slope models. *Methods in Ecology and Evolution*, *13*(6), 1214–1223. https://doi.org/10.1111/2041-210X.13856

Stan Development Team. (n.d.). RStan: The R interface to Stan [R package version 2.32.7].

Stinchcombe, J. R., Agrawal, A. F., Hohenlohe, P. A., Arnold, S. J., & Blows, M. W. (2008). Estimating nonlinear selection gradients using quadratic regression coefficients: Double or nothing? *Evolution*, *62*(9), 2435–2440. https://doi.org/https://doi.org/10.1111/j.1558-5646.2008.00449.x

Taborsky, B., & Oliveira, R. F. (2012). Social competence: an evolutionary approach. *Trends in Ecology and Evolution*, *27*(12), 679–688. https://doi.org/10.1016/j.tree.2012.09.003

Taborsky, B., & Oliveira, R. F. (2013). Social competence vs responsiveness: Similar but not same. A reply to Wolf and McNamara. *Trends in Ecology & Evolution*, *28*(5), 254–255. https://doi.org/10.1016/j.tree.2013.02.005
doi: 10.1016/j.tree.2013.02.005.

Tufto, J. (2015). Genetic evolution, plasticity, and bet-hedging as adaptive responses to temporally autocorrelated fluctuating selection: A quantitative genetic model. *Evolution*, *69*(8), 2034–2049. https://doi.org/10.1111/evo.12716

Ulrich, Y., Saragosti, J., Tokita, C. K., Tarnita, C. E., & Kronauer, D. J. C. (2018). Fitness benefits and emergent division of labour at the onset of group living. *Nature*, *560*(7720), 635–638. https://doi.org/10.1038/s41586-018-0422-6

Via, S., & Lande, R. (1985). Genotype-environment interaction and the evolution of phenotpic plasticity. *Evolution*, *39*(3), 505–522. https://doi.org/10.1111/j.1558-5646.1985.tb00391.x

Vickery, W. L., Giraldeau, L.-A., Templeton, J. J., Kramer, D. L., & Chapman, C. A. (1991). Producers, scroungers, and group foraging. *The American Naturalist*, *137*(6), 847–863. https://doi.org/10.1086/285197

Westneat, D. F., Wright, J., & Dingemanse, N. J. (2015). The biology hidden inside residual within-individual phenotypic variation. *Biological Reviews*, *90*(3), 729–743. https://doi.org/10.1111/brv.12131

Wijnhorst, R. E., de Groot, C., Ajoy, Y. A., Jensen, H., Wright, J., & Dingemanse, N. (2024, November). Individual variation in social foraging behavior across social contexts. https://doi.org/10.17605/OSF.IO/EDA9N

Wilkinson, G. S. (2003). 12. social and vocal complexity in bats. In F. B. M. de Waal & P. L. Tyack (Eds.), *Intelligence, culture, and individualized societies* (pp. 322–341). Harvard University Press. https://doi.org/doi:10.4159/harvard.9780674419131.c24

Wolf, M., Van Doorn, G. S., & Weissing, F. J. (2010). On the coevolution of social responsiveness and behavioural consistency. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1704), 440–448. https://doi.org/10.1098/rspb.2010.1051

Wolf, M., van Doorn, G. S., & Weissing, F. J. (2008). Evolutionary emergence of responsive and unresponsive personalities. *Proceedings of the National Academy of Sciences*, *105*(41), 15825–15830. https://doi.org/10.1073/pnas.0805473105

Wolf, M., & Weissing, F. J. (2010). An explanatory framework for adaptive personality differences. *Philosophical Transactions of the Royal Society B*, *365*, 3959–3968. https://doi.org/10.1098/rstb.2010.0215

Supplement

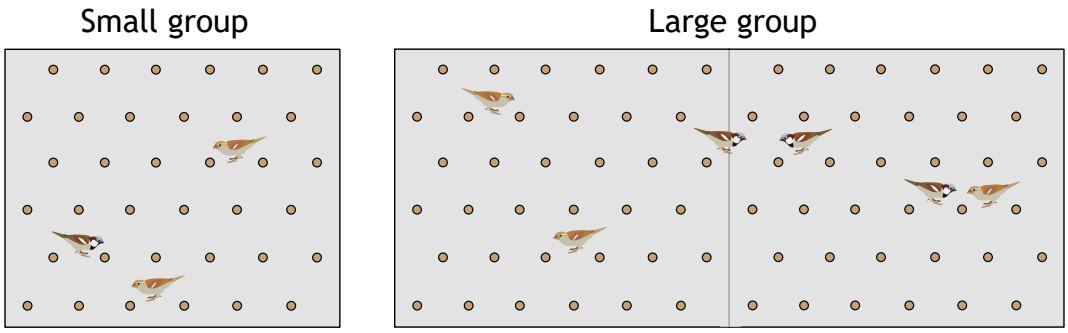

Figure S1: ***All individuals were subjected to foraging assays in groups of three ("small group") and six ("large group").*** *Large-group assays were created by combining two RFID feeders, resulting in twice the available area, number of feeding wells, and food quantity compared with the small-group assays*

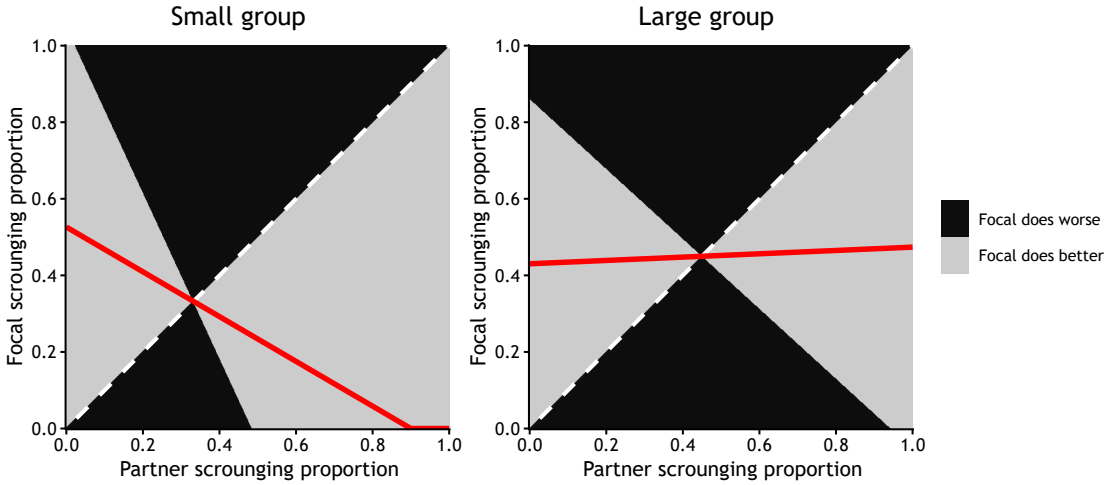

Figure S2: ***Pairwise invasibility plot of small and large social foraging groups.*** *The x-axis shows the scrounging proportion of the social environment, and the y-axis shows the scrounging proportion of the focal individual. Light grey regions indicate combinations where the focal strategy achieves higher predicted success than a strategy matching the social environment, whereas black regions indicate lower predicted success. The white dashed line represents equal focal and social-environment strategies, and the red line shows the optimal response strategy for each social environment.*

Table S1: Model for the bivariate logistic regression of the focal individual's scrounging response to partner scrounging proportion in the two assay types. Repeatabilities are reported under variance proportions, calculated as the variance explained by a random effect divided by the total variance. Within-individual variance is calculated as the variance explained by the fixed effect of the social response plus the variance of the error distribution ($\pi^2/3$), divided by the total variance. The cross-context correlation quantifies the consistency of individual-specific intercepts. We report posterior medians and 95% credible intervals and the difference in posterior estimates between small and large groups.

Parameter	Small group	Large group	Difference (S–L)
Fixed effects			
Intercept	-1.036 [-1.443, -0.614]	-0.797 [-1.234, -0.377]	-0.238 [-0.791, 0.341]
Controls			
Test day (1st or 2nd) (Day)	0.086 [-0.350, 0.530]	–	–
Assay order (Seq)	0.042 [-0.035, 0.121]	0.134 [0.024, 0.261]	-0.092 [-0.239, 0.044]
Large group assay first (Grp)	0.873 [0.089, 1.603]	-0.488 [-1.240, 0.302]	1.359 [0.274, 2.362]
Day*Seq	0.088 [-0.066, 0.244]	–	–
Day*Grp	-0.595 [-1.437, 0.236]	–	–
Seq*Grp	-0.037 [-0.190, 0.121]	-0.059 [-0.296, 0.170]	0.023 [-0.252, 0.309]
Day*Seq*Grp	0.013 [-0.298, 0.321]	–	–
Average response to partner tactic			
Partner scrounging	-6.946 [-7.973, -5.944]	-3.667 [-5.553, -1.719]	-3.285 [-5.455, -1.115]
Random effects			
V_{ID}	0.187 [0.047, 0.462]	0.899 [0.445, 1.652]	-0.700 [-1.448, -0.201]
V_{AssayID}	3.776 [2.627, 5.264]	1.381 [0.474, 2.878]	2.377 [0.516, 4.159]
V_{GroupID}	0.266 [0.008, 1.074]	0.035 [0.000, 0.557]	0.193 [-0.317, 1.004]
Variance proportions			
$VC_{\text{Among-individual}}$	0.017 [0.004, 0.044]	0.144 [0.066, 0.260]	-0.126 [-0.242, -0.046]
$VC_{\text{Social response}}$	0.321 [0.275, 0.364]	0.087 [0.023, 0.162]	0.234 [0.146, 0.311]
VC_{Residual}	0.293 [0.239, 0.360]	0.524 [0.396, 0.671]	-0.229 [-0.385, -0.087]
$VC_{\text{Within-individual}}$	0.616 [0.557, 0.670]	0.614 [0.511, 0.718]	0.002 [-0.118, 0.117]
VC_{AssayID}	0.337 [0.278, 0.396]	0.221 [0.092, 0.356]	0.116 [-0.031, 0.257]
VC_{GroupID}	0.024 [0.001, 0.089]	0.006 [0.000, 0.083]	0.014 [-0.059, 0.080]
Correlation			
$\text{cor}(\text{ID}_{small}, \text{ID}_{large})$	0.790 [0.418, 0.964]		–

Table S2: Model output for the bivariate Gaussian model of foraging success as a function the focal individual's scrounging and partner's scrounging proportion in two assay types. The cross-context correlation quantifies the consistency of individual-specific intercepts. We report posterior medians and 95% credible intervals and the difference in posterior estimates between small and large groups.

Parameter	Small group	Large group	Difference (S–L)
Fixed effects			
Intercept	1.089 [0.981, 1.194]	0.941 [0.821, 1.057]	0.148 [0.002, 0.295]
Experimental controls			
Test day (1st or 2nd) (Day)	-0.023 [-0.088, 0.041]	–	–
Assay order (Seq)	0.014 [0.003, 0.025]	0.022 [0.006, 0.038]	-0.008 [-0.027, 0.012]
Large group assay first (Grp)	0.139 [-0.035, 0.321]	-0.034 [-0.178, 0.117]	0.174 [-0.047, 0.387]
Day*Seq	-0.037 [-0.059, -0.015]	–	–
Day*Grp	-0.050 [-0.178, 0.078]	–	–
Seq*Grp	-0.017 [-0.039, 0.006]	0.004 [-0.028, 0.036]	-0.021 [-0.060, 0.018]
Day*Seq*Grp	0.020 [-0.024, 0.063]	–	–
Focal foraging events	0.123 [0.093, 0.152]	0.089 [0.054, 0.122]	0.034 [-0.011, 0.079]
Partner foraging events	0.088 [0.069, 0.108]	0.022 [0.010, 0.034]	0.066 [0.044, 0.089]
Effects of tactic use			
Focal scrounging	-0.032 [-0.136, 0.073]	0.046 [-0.079, 0.174]	-0.077 [-0.242, 0.086]
Partner scrounging	-0.079 [-0.210, 0.052]	-0.069 [-0.286, 0.149]	-0.010 [-0.262, 0.244]
Focal scrounging^2	-0.748 [-1.332, -0.170]	-0.682 [-1.434, 0.084]	-0.065 [-1.029, 0.874]
Partner scrounging^2	-0.262 [-1.030, 0.494]	-1.281 [-2.617, 0.097]	1.024 [-0.557, 2.565]
Focal × Partner	-0.438 [-0.823, -0.054]	0.029 [-0.433, 0.489]	-0.466 [-1.070, 0.133]
Random effects			
V_{ID}	0.042 [0.027, 0.066]	0.083 [0.054, 0.128]	-0.040 [-0.083, -0.008]
V_{AssayID}	0.049 [0.033, 0.066]	0.009 [0.000, 0.025]	0.039 [0.018, 0.059]
V_{GroupID}	0.011 [0.001, 0.049]	0.004 [0.000, 0.037]	0.006 [-0.026, 0.043]
V_{Residual}	0.158 [0.143, 0.177]	0.122 [0.103, 0.143]	0.037 [0.010, 0.063]
V_{Total}	0.264 [0.236, 0.309]	0.223 [0.186, 0.278]	0.041 [-0.017, 0.095]
Variance proportions			
VC_{ID}	0.160 [0.107, 0.234]	0.375 [0.270, 0.490]	-0.212 [-0.329, -0.099]
VC_{AssayID}	0.183 [0.125, 0.244]	0.041 [0.000, 0.109]	0.141 [0.055, 0.219]
VC_{GroupID}	0.044 [0.003, 0.165]	0.019 [0.000, 0.148]	0.020 [-0.104, 0.139]
VC_{Residual}	0.601 [0.504, 0.684]	0.547 [0.426, 0.664]	0.053 [-0.090, 0.195]
Correlation			
$\text{cor}(\text{ID}_{small}, \text{ID}_{large})$	0.618 [0.364, 0.798]		–