Decomposing social interactions: a statistical method for estimating social impact and social responsiveness

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

[*]Corresponding author: roriwijnhorst@gmail.com

[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

Social interactions mediate the phenotypic expression of fitness-relevant traits. The expression of such labile social traits includes three distinct components: an individual's mean trait value (direct effect), its social responsiveness, and its social impact (indirect effects). Traditional methods, such as variance-partitioning or trait-based models, usually only partition individual variation into direct and indirect effects. However, individual variation in social responsiveness and its covariation with direct effects and social impact will affect responses to selection. To date, no studies have explored the performance of models that allow the decomposition of responsiveness from impact. Here, we describe a model for studying variation in phenotypic expression caused by social interactions, and we use simulations to explore its performance under various experimental designs. Our analyses show that with adequate total sample sizes ($\geq$ 3200), variance components are estimated accurately across all study designs. In contrast, covariance estimation would benefit most from including more unique individuals, followed by more unique social partners per individual, whereas repeated interactions with the same partners added the least improvement to the covariance estimation. We also found that failing to model individual variation in responsiveness, and neglecting measurement error, increases bias and imprecision in trait-based approaches. Hence, disregarding individual variation in responsiveness would ignore a key component of social behaviour, and hamper our ability to acquire unbiased estimates of indirect genetic or social effects.

Introduction

Social interactions alter selection pressures and phenotypic expression, shaping the trajectory of evolutionary change in ways that are often difficult to predict (Moore et al., 1997; Wolf et al., 1998). Phenotypes displayed by individuals or genotypes rarely emerge in isolation; rather they arise from the interaction of internal regulators and external conditions (Via & Lande, 1985; West-Eberhard, 1989). One important external factor, the social environment, consists of conspecifics that affect phenotypic expression through social interactions. The effects of the social environment can be far-reaching in traits that are solely expressed in a social context, such as cooperation, social hierarchies or parental-offspring interactions (Bailey et al., 2018; Bleakley & Brodie, 2009; Kirkpatrick & Lande, 1989; Smiseth et al., 2008; Wilson et al., 2011). Explaining (co-)variation in social traits is challenging because individuals often adjust their phenotype plastically in response to their partners' traits (Bailey & Desjonquères, 2022; Moore et al., 1997). These socially mediated effects, when heritable, are termed indirect genetic effects (IGEs) (Griffing, 1967; Moore et al., 1997). The optimal phenotype might therefore depend on other phenotypes displayed in the social environment (Maynard-Smith & Price, 1973; McNamara & Weissing, 2010), where selection could also act on an individual's competence to adjust their phenotype to a changing social environment (Martin & Jaeggi, 2022; Taborsky & Oliveira, 2012).

An often overlooked aspect of such indirect genetic effects (IGEs) is that individuals both respond to (responsiveness) and affect (impact) the phenotype of other individuals, and individuals may differ in both of these traits. Following recent proposals, social phenotypes can be decomposed into three components of individual phenotypic variation: (i) mean trait value; (ii) social responsiveness, which refers to the phenotypic response of the focal to the traits of their interacting social partners; and lastly (iii) social impact, which refers to the response an individual elicits in their social partners (Araya-Ajoy et al., 2020; de Groot et al., 2023). Previous studies in quantitative genetics have estimated population-level IGEs (reviewed by Bailey & Desjonquères, 2022), disregarding that individuals may differ in their level of social responsiveness. Common statistical models for studying social effects include the 'variance-partitioning' (Bijma, 2014; Griffing, 1967) and 'trait-based' approach (Kirkpatrick & Lande, 1989; McGlothlin et al., 2010; Moore et al., 1997; Wolf et al., 1999). The variance-partitioning approach is a type of mixed-effects model that partitions observed phenotypic variation into direct individual effects and indirect individual effects, which can be further decomposed into additive genetic components using 'animal models' (Henderson, 1984; Kruuk, 2004; Meyer, 1992; Wilson et al., 2010). The trait-based approach, mathematically equivalent to the variance partitioning approach under certain assumptions (McGlothlin & Brodie, 2009), applies a reaction norm approach to quantify social responsiveness as a slope. Both frameworks estimate the interaction coefficient ψ, which represents the population-level response and describes the magnitude and direction of phenotypic change in response to the phenotype expressed by interaction partners (see Bailey & Desjonquères, 2022; Bijma, 2014). Thus, these models typically ignore the possibility that individuals may differ in responsiveness. However, empirical evidence increasingly shows that individuals can differ in the degree to which they respond to social signals (Bailey & Zuk, 2012; Guayasamin et al., 2017; Jablonszky et al., 2022; Morand-Ferron et al., 2011; Strickland & Frère, 2019). Hence, researchers have suggested that ψ is not fixed; instead it may vary among individuals and can consequently evolve (Akçay & Van Cleve, 2012; Araya-Ajoy et al., 2020; Dingemanse & Araya-Ajoy, 2015; Kazancıoğlu et al., 2012; Wolf et al., 2008). Already there is experimental evidence that ψ can evolve under different selection regimes (Chenoweth et al., 2010). Furthermore, the covariance of social responsiveness with the mean social trait could speed up or slow down evolution through a process called 'social drive' (Bailey et al., 2021; Martin et al., 2023).

Very little is known about the extent to which variation in ψ influences social interactions for two key reasons. First, individuals may differ in their responsiveness, yet standard quantitative genetics models typically assume a fixed population-level effect. This masks important individual variation and limits evolutionary inference. Second, the traits of social partners (to which focal individuals respond) are themselves phenotypically variable, introducing both among- and within-individual variation into an environmental covariate (Araya-Ajoy et al., 2020; Dingemanse & Araya-Ajoy, 2015). This variability, as is also the case for measurement error, can attenuate estimates of ψ and underestimate true social effects. To address these problems, we use a model that incorporates both random slopes (to capture individual variation in ψ) (de Groot et al., 2023; Martin & Jaeggi, 2022) and an 'errors-in-variables' approach that corrects for noisy partner trait estimates (Dingemanse et al., 2021; Ponzi et al., 2018), which should allow an accurate estimation of social responsiveness and its evolutionary consequences.

The next challenge is to determine which study design is optimal to estimate the three components of individuality in social interactions. A common and effective laboratory approach for estimating individual differences in IGEs involves assessing individuals in pairwise assays in which individuals repeatedly interact with different social partners (e.g. Han et al., 2018; Lane et al., 2020; Santostefano et al., 2016; Wilson et al., 2009). Similar datasets have been collected through observational studies on dyadic interactions in wild populations (e.g. Brommer & Rattiste, 2008; McLean et al., 2023; Moiron et al., 2020; Tuliozi et al., 2023; Wilson et al., 2011). Several data simulation studies have explored the accuracy and precision of statistical models in estimating individual variation in labile traits (Araya-Ajoy et al., 2015; Dingemanse & Dochtermann, 2013; Martin et al., 2011; van de Pol, 2012). From these studies, we have learnt that there is a rapid increase in statistical power when more individuals are sampled, or more repeated measures per individual are taken. Furthermore, simulation studies show that resource allocation (more individuals with fewer observations per individual versus fewer individuals with more observations per individual) can matter when the total sample size is the limiting factor (Martin et al., 2011; van de Pol, 2012). We do not know how resource allocation affects the estimation of individual variation in mean social trait values, social impact or social responsiveness.

Some studies have focused on optimal study designs to estimate IGEs, comparing different group sizes or breeding designs (Bijma, 2010), but few have explored how well IGE models recover individual variation in labile traits expressed during repeated social interactions. Designing studies of social traits poses challenges not typically encountered for non-social traits, because social phenotypes must be sampled in ways that capture both the effect of the individual on its own phenotypic expression and its effects on others. Specifically, two design features are required to reliably estimate variance in social responsiveness and impact, and their covariances with mean trait value. First, individuals must be repeatedly observed interacting multiple times both as focal individuals and as social partners. This reciprocity in roles is essential to estimate covariances between how individuals behave and how they influence others (Dingemanse & Araya-Ajoy, 2015). Second, individuals must encounter sufficient variation in partner trait values, which is a prerequisite for estimating responsiveness (the slope describing the individual's phenotypic response to partner phenotypes). Without these specific design properties, many components of the multivariate structure of social traits are non-estimable. Consequently, when studying social traits under limited time or resources, researchers face critical trade-offs in allocating sampling effort across individuals, the number of partners per individual, and the number of unique pairwise interactions. One goal of this study is to explore which of these sampling axes should be prioritised to maximise precision and accuracy in estimating social impact and responsiveness.

In this study, we evaluate the accuracy and precision with which variance and covariance in mean trait value, social impact, and social responsiveness can be estimated. The impact–responsiveness framework provides a general conceptual and statistical approach and can be used to study any system in which organisms adjust their phenotype in response to others and influence the phenotypes of others, including behavioural, physiological, developmental, and ecological interactions (see de Groot et al., 2023, for a discussion of its general applicability). Here, we focus mainly on labile social traits in the context of interacting phenotypes and indirect genetic effects (Bailey et al., 2018; Bijma, 2010; Moore et al., 1997), and explore the statistical implementation of the framework. We first explore the minimally required sample size needed to obtain unbiased estimates. Next, we assess the difference in accuracy and precision of study designs that vary in the number of individuals, the number of repeated measures per individual, and the number of unique social partners. We further determine the consequences of not accounting for individual variation in responsiveness, measurement error/plasticity in partner traits, when estimating social effects. By combining these perspectives, our study aims to encourage empirical estimates of key components underlying phenotypic expression during social interactions, provide practical guidance on the optimal experimental designs and statistical analyses, and increase awareness of problems when failing to account for key sources of variation.

Methods

Data simulation

We simulated realistic social interaction data where individuals differed in their mean trait value (intercept), social impact (effect on others), and social responsiveness (response to partner phenotype). Individuals interacted pairwise, responding to a single fixed partner trait (e.g. body size) that was measured with error. Each dataset had a balanced design: all individuals interacted with the same number of partners and acted equally often as focal and partner. To answer our questions, we simulated and analysed the following datasets:

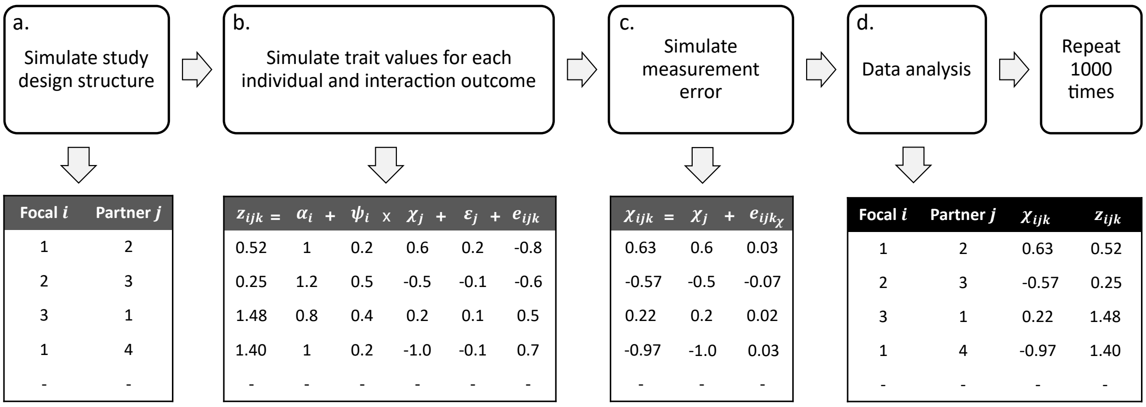
1. To assess model performance as a function of the total sample size, we simulated 1000 datasets

per sample size, varying the number of individuals while keeping the number of interactions per individual constant.

2. To evaluate the effects of sampling design on bias and precision, we simulated 1000 datasets with balanced designs that partitioned a fixed total number of observations among varying numbers of individuals, social partners per individual, and repeated interactions.

3. To compare the performance of alternative statistical models, we simulated 1000 datasets using four study designs and to each we fit the full model and several reduced models lacking particular components.

The main parameters of interest are two fixed effects (population intercept β_0 and population slope $\bar{\psi}$), and six (co)variance components: the among-individual variance in mean trait value (V_α); social responsiveness (V_ψ); and social impact (V_ϕ); and their three covariances (Table S1). We assessed the accuracy and precision of estimated parameters. We measured accuracy as relative bias, calculated as: $\frac{1}{n}\Sigma\frac{\theta - \hat{\theta}_i}{\theta} \cdot 100\%$, where θ is the true simulated value, $\hat{\theta}_i$ is the model estimate (posterior median) of the ith simulation, and n is the number of simulations. We measured precision as relative dispersion (MADm: Mean Absolute Deviation of the mean), calculated as: $\frac{1}{n}\Sigma\frac{|\bar{\theta} - \hat{\theta}_i|}{\bar{\theta}} \cdot 100\%$, where $\bar{\theta}$ is the grand mean of the 1000 posterior median model estimates. Instead of formal power analyses (less common in Bayesian settings), we report performance using bias and dispersion and provide an open-access simulation tool "socialSim" for researchers to explore expected performance under their study design of choice (Wijnhorst, 2025). The data simulation and model fitting workflow is summarised in Figure 1.



Focal i	**Partner j**
1	2
2	3
3	1
1	4
-	-

z_{ijk} =	α_i +	ψ_i x	χ_j +	ε_j +	e_{ijk}
0.52	1	0.2	0.6	0.2	-0.8
0.25	1.2	0.5	-0.5	-0.1	-0.6
1.48	0.8	0.4	0.2	0.1	0.5
1.40	1	0.2	-1.0	-0.1	0.7
-	-	-	-	-	-

χ_{ijk} =	χ_j +	e_{ijk_χ}
0.63	0.6	0.03
-0.57	-0.5	-0.07
0.22	0.2	0.02
-0.97	-1.0	0.03
-	-	-

Focal i	**Partner j**	χ_{ijk}	z_{ijk}
1	2	0.63	0.52
2	3	-0.57	0.25
3	1	0.22	1.48
1	4	-0.97	1.40
-	-	-	-

Symbol	Description
z_{ijk}	The phenotype of individual *i* when it interaction with *j* at instance *k*
α_i	Mean behaviour: individual deviation of the population mean
ψ_i	Social responsiveness: individual deviation of the population slope
χ_j	True value of the phenotype to which the focal responds
ε_j	Residual impact: partner effects caused by unknown/unmeasured traits
e_{ijk}	Residual effect/nonsocial environmental effect
e_{ijk_χ}	Measurement error or nonsocial phenotypic plasticity around the true value of χ_j

Figure 1: ***Workflow of data simulations and analyses*** *a.) A data structure is created that contains a certain number of individuals that interact with a fixed number of social partners in balanced designs. b.) We simulated the outcome of the social interaction (z_{ijk}) based on their simulated trait values of each individual and their partner drawn from a multivariate normal distribution plus residual error. c.) Before we enter our observed data into the model, we assume that the fixed opponent trait is measured with error before every social interaction. Therefore, we added or subtracted simulated measurement error to obtain the observed opponent trait χ_{ijk}. d.) The data is analysed by the model and this process is iterated a 1000 times.*

Simulated effect sizes

Effect sizes were selected to represent biologically realistic magnitudes derived from empirical estimates. The population response $\bar{\psi}$ was set to 0.3, similar to the mean (0.27) of the significant non-reciprocal, positive estimates $\bar{\psi}$ (obtained from Bailey & Desjonquères, 2022). The variance components were adjusted so that the total phenotypic variances sum to 1, with the variance explained by the focal individual at 0.3 (approximating the mean repeatability of animal behaviours of 0.37; Bell et al., 2009) and the variance explained by the social partner (variance of social impact) at 0.1 and residual variance $V_e = 0.6$. This variance explained by the social partner is representative of partner effects, however for indirect genetic effects, the effect size is expected to be somewhat smaller (6% for behavioural traits and 3% for all traits; Santostefano et al., 2024). The variance of social responsiveness was set to 0.1, giving an elevation-to-slope ratio of $1 : 0.5$, similar to empirical estimates ($1 : 0.65$) (Brommer, 2013). The measurement error for the impact trait was set to 0.1 (10% of the variance in the social partner trait χ_j). Full parameter settings are given in Supplementary Table S7.

Analyses

Total sample size

To assess model performance as a function of total sample size, we simulated 1000 datasets of social interaction data. In each dataset, the number of interactions per individual was kept constant at 8. The smallest dataset included 50 individuals (400 observations), and the number of individuals was doubled at each step to a maximum of 800 individuals (6400 observations). Our aim was to identify the minimum sample size at which the full model for estimating impact and responsiveness (I&R model) yields unbiased estimates, which was used for subsequent analyses.

Sampling design

To determine how to best allocate limited sampling resources, we assessed how different study designs influence the accuracy and precision of parameter estimates. In many ecological and evolutionary studies, researchers face logistical constraints that limit the total number of observations that can be collected. Therefore, we examined how design choices affect parameter bias and precision. Each observation represented a pairwise interaction, and individuals acted as both focal and partner. We varied sampling designs along three key axes:

- Number of individuals ('individuals'), which determines the total population of interacting individuals.
- Number of unique social partners per focal individual ('partners'), which reflects the variety of partners encountered by each individual.
- Number of repeated interactions per dyad ('repeats'), which controls the extent to which

specific focal-partner pairs are observed multiple times. A value of '1x' indicates that each focal interacts only once with each unique social partner (i.e. no repetitions of unique dyads).

We kept total observations constant while trading off two axes at a time, allowing us to assess how these factors influence estimation accuracy and precision.

Model comparison

To evaluate how different statistical modelling choices influence the estimation of social parameters, we compare several models used in the study of indirect genetic effects (IGEs) and social trait evolution. Specifically, to investigate the consequences of not accounting for certain sources of variation, we compare the full impact and responsiveness model (I&R) to incomplete models (i.e. models that lack certain parameters, Table 1). As such, I&R is compared to two reduced models, the variance-partitioning model (V-P) and the trait-based approach (Trait). Secondly, we modify the trait-based approach to include random slopes or an errors-in-variables correction to identify which missing component causes increased bias or imprecision. For this, we compare the full I&R model (equivalent to Trait+RS+EIV) against two reduced models: one including an errors-in-variable correction (Trait+EIV) and one including random slopes (Trait+RS) for individual social responsiveness. Below, we describe the models in increasing order of model completeness.

Variance-partitioning model (V-P):

The variance partitioning approach can be described as:

$$z_{ijk} = \beta_0 + \alpha_i + \phi_j + e_{ijk} \ , \tag{1}$$

where z_{ijk} denotes the phenotype of individual i after interacting with social partner j at instance k. The fixed intercept β_0 represents the population mean phenotype. The random effect α_i represents the deviation of individual i's mean trait value from the population mean. The variance of these deviations (V_α) quantifies among-individual variance, which may arise from direct genetic effects (DGEs) and permanent environmental influences. The random partner effect ϕ_j represents the deviation associated with social partner j, that is, the extent to which partner j influences the phenotype of others. The variance of these deviations (V_ϕ) quantifies variation in social impact, which may arise from indirect genetic effects (IGEs) as well as non-genetic partner effects. The residual term e_{ijk} represents unexplained deviations at the observation level. The residual variance (V_e) captures within-individual variation that is not attributable to focal identity or repeatable partner effects. When individuals interact both as focal and partner (as is common in social interaction datasets), the model can also be used to estimate the covariance ($\mathrm{Cov}_{\alpha\phi}$) between an individual's mean trait value and its social impact (Wilson et al., 2009):

$$\begin{bmatrix} \alpha \\ \phi \end{bmatrix} \sim \mathrm{MVN}(0, \Omega), \quad \Omega = \begin{bmatrix} V_\alpha & \mathrm{Cov}_{\alpha\phi} \\ \mathrm{Cov}_{\alpha\phi} & V_\phi \end{bmatrix}, \quad e \sim \mathrm{N}(0, V_e). \tag{2}$$

The covariance between focals' mean trait value and social impact ($\mathrm{Cov}_{\alpha\phi}$) is critical for predicting evolutionary change in social traits. When decomposed into genetic and environmental components, it corresponds to the DGE–IGE covariance, which can accelerate or constrain evolutionary responses depending on its sign and magnitude (Bijma et al., 2007; Wilson et al., 2009; Wolf et al., 1998). The V-P model is commonly used in IGE studies to estimate the variance attributable to social partners without explicitly modelling the partner traits through which those effects are mediated. While the model provides an estimate of the total variance of social impact (V_ϕ), it does not identify trait-based pathways or quantify individual differences in responsiveness.

Trait-based (Trait) model:

The trait-based model can be described as:

$$z_{ijk} = \beta_0 + \alpha_i + \bar{\psi}\chi_{ijk} + \epsilon_j + e_{ijk}\ . \tag{3}$$

The trait-based model adopts a reaction norm framework (Dingemanse & Araya-Ajoy, 2015; Kirkpatrick & Lande, 1989; McGlothlin et al., 2010; Moore et al., 1997; Wolf et al., 1999), modelling the focal phenotype as a function of a measured trait of the social partner, χ_{ijk} (e.g. body size). The response is estimated through $\bar{\psi}$, the interaction coefficient or slope that represents the mean response of the population to the trait values of the social partners. To account for unexplained social effects not captured by the measured trait, we include a partner identity effect ϵ_j. Although this addition is not conventional in trait-based models, it ensures mathematical equivalence to the variance-partitioning model. The total variance in social impact can then be expressed as:

$$V_\phi = \bar{\psi}^2 V_\chi + V_\epsilon + 2\bar{\psi}\mathrm{Cov}_{\chi\epsilon} \ . \tag{4}$$

Subsequently, we can again estimate the covariance matrix, the same as we derive from the V-P approach.

Trait-based model with random slopes (Trait+RS):

The trait-based model can be extended to estimate social responsiveness:

$$z_{ijk} = \beta_0 + \alpha_i + (\bar{\psi} + \psi_i)\chi_{ijk} + \epsilon_j + e_{ijk} \ . \tag{5}$$

This extension of the trait-based model includes random slopes ψ_i, which represent individual-specific deviations from the population slope $\bar{\psi}$. This allows individuals to differ in their responsiveness to partner traits. The model therefore estimates a 3×3 covariance matrix that includes mean trait value (V_α), social responsiveness (V_ψ), and residual partner effects (V_ϵ), and their covariances.

Trait-based model with measurement error correction (Trait+EIV):

To account for measurement error or labile variation in the partner trait, we supplement the trait-based model with an error correction:

$$z_{ijk} = \beta_0 + \alpha_i + \bar{\psi}\chi_j + \epsilon_j + e_{ijk} \ , \tag{6}$$

where the latent trait value χ_j is estimated by:

$$\chi_{ijk} = \beta_{0_\chi} + \chi_j + e_{ijk_\chi} \ . \tag{7}$$

This model is an extension of the Trait model, where χ_{ijk} denotes the trait value of the social partner j as observed when interacting with focal i at instance k. We partition this observed value into two components: a partner-specific effect χ_j and a residual term e_{ijk_χ}. In the context of measurement error, e_{ijk_χ} represents random error around the true partner trait value χ_j. More generally, this same structure can also be interpreted as a decomposition of the partner trait into genetic and environmental components: χ_j can be viewed as the heritable additive genetic contribution to the partner trait, whereas e_{ijk_χ} represents non-heritable influences (environmental or transient effects). Thus, the measurement-error model provides a framework that can be applied both to correct for error in trait measurements or to allow inference about the genetic basis of social impact through partner traits, thereby linking directly to IGE theory (McGlothlin & Brodie, 2009). This model also estimates a 3×3 covariance matrix that includes mean trait value (V_α), partner impact trait (V_χ), and residual partner effects (V_ϵ), and their covariances.

Impact and responsiveness model (I&R):

The complete model to estimate individual mean trait values, social responsiveness, and social impact is described as:

$$z_{ijk} = \beta_0 + \alpha_i + (\bar{\psi} + \psi_i)\chi_j + \epsilon_j + e_{ijk} \, , \tag{8}$$

$$\chi_{ijk} = \beta_{0_\chi} + \chi_j + e_{ijk_\chi} \, . \tag{9}$$

This model extends the trait-based approach by combining two components: random slopes and the error-correction framework. This complete model yields a 4×4 covariance matrix that estimates the variances and covariances of mean trait value (V_α), the partner impact trait (V_χ), social responsiveness (V_ψ), and residual partner effects (V_ϵ). From these estimates, we can derive the joint covariance structure of mean trait value (α), social impact (ϕ), and social responsiveness (ψ) (see Supplementary Equations S1, S2, S3, S4 for details):

$$\begin{bmatrix} \boldsymbol{\alpha_i} \\ \boldsymbol{\psi_j} \\ \boldsymbol{\phi_i} \end{bmatrix} \sim \mathrm{MVN}(0, \Omega) : \Omega = \begin{bmatrix} V_\alpha & \mathrm{Cov}_{\alpha\psi} & \mathrm{Cov}_{\alpha\phi} \\ \mathrm{Cov}_{\alpha\psi} & V_\psi & \mathrm{Cov}_{\psi\phi} \\ \mathrm{Cov}_{\alpha\phi} & \mathrm{Cov}_{\psi\phi} & V\phi \end{bmatrix} \qquad [e] \sim \mathrm{MVN}(0, V_e) \, . \tag{10}$$

Table 1: ***Overview of the variance components estimated in each model.*** *Total social impact variance is estimated either directly (V_ϕ) or via a combination of the variance in impact trait and residual impact ($\bar{\psi}^2 V_\chi + V_\epsilon$). The two trait models, Trait and Trait+RS do not correct for measurement error, and thus estimate impact (V_ϕ) using the variance of the partner trait that includes measurement error (χ_{ijk}. Each model estimates a covariance matrix containing all individual-level random effects and their covariances corresponding to that model specification.*

Component Model	Mean behaviour	Responsiveness	Impact trait	Total / Residual impact
V-P	V_α			V_ϕ
Trait	V_α		χ_{ijk}	V_ϵ
Trait+RS	V_α	V_ψ	χ_{ijk}	V_ϵ
Trait+EIV	V_α		V_χ	V_ϵ
I&R	V_α	V_ψ	V_χ	V_ϵ

Model specification

All simulations were implemented in R (version 4.5.1, R Core Team, 2025) and analysed in a Bayesian framework using Stan probabilistic programming language (Carpenter et al., 2017) via the ‘rstan’ package (version 2.32.2) (Stan Development Team, 2025). Each model used weakly informative priors: normal distributions (mean = 0, SD = 1) for fixed effects, and truncated normal distributions (mean = 0, SD = 1; lower bound = 0) for variance parameters. Correlation structures among random effects were estimated via Cholesky decomposition with an LKJ(1) prior. All models were run with one chain with 1000 warm-up, and 5000 iterations. The models were run in parallel on multiple processing units (up to 56) using the ’future’ and ’future.apply’ packages (Bengtsson, 2021).

Results

Total sample size

The I&R model recovered fixed effects and variances with high accuracy (Figure 2). Estimates of the population mean (β_0) and interaction coefficient ($\bar{\psi}$) showed negligible bias, for sample sizes $\geq$ 800. The dispersion decreased steadily with the sample size for all parameters. “All variance components showed minimal bias (< 5%), except a slight overestimation of social impact V_ϕ (8.2% at $n = 400$). Covariances were more difficult to estimate, with strong underestimation at small sample sizes (-17.6% to -30.2% at 400). Among covariances, $\text{Cov}_{\psi\phi}$ was most difficult to estimate, followed by $\mathrm{Cov}_{\alpha\psi}$, whereas $\mathrm{Cov}_{\alpha\phi}$ was least biased. The bias decreased consistently with larger samples, with less than 4% bias at the largest sample size of 6400 observations with 800 individuals. These results show that estimates of fixed effects and variances stabilise at moderate sample sizes (≥ 800), whereas reliable estimation of covariance components requires substantially larger datasets. Based on these patterns, we continued the subsequent analyses with a total sample size of 3200 observations.

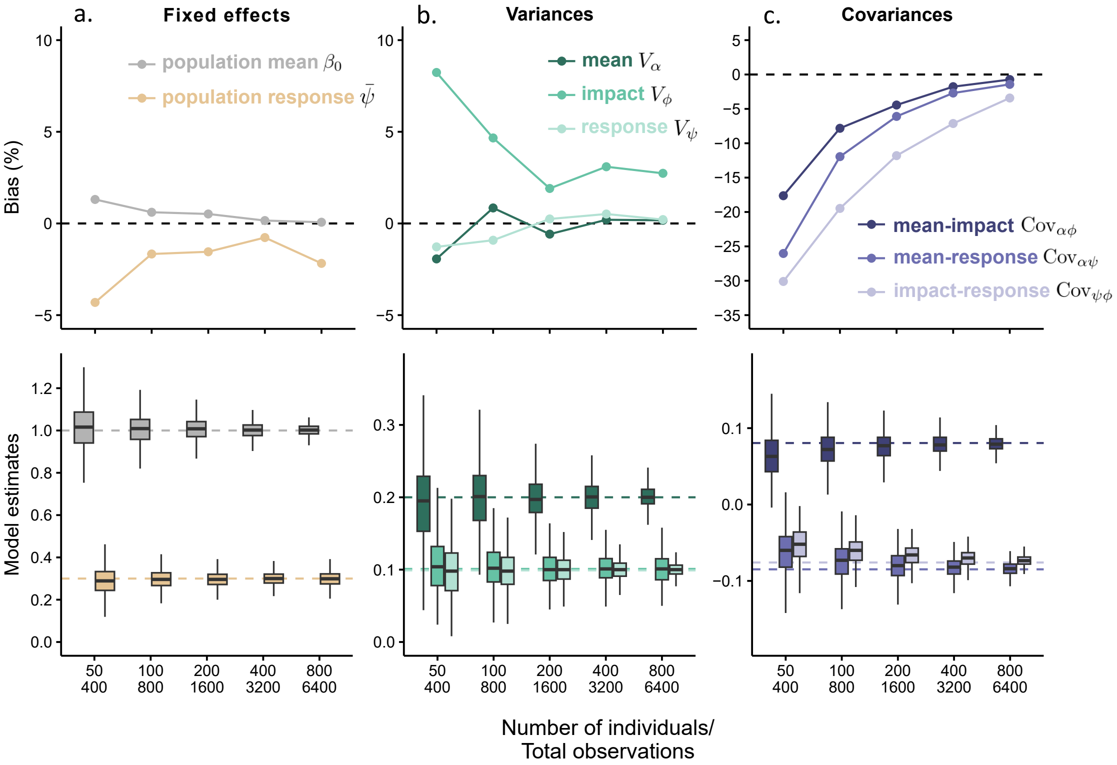


Figure 2: ***Bias and dispersion of model parameters estimated by the impact and responsiveness (I&R) model across various total sample sizes****. The top panels show the relative bias (%) of the posterior medians of a 1000 analysed simulated datasets per sample size. The bottom panel shows boxplots of all 1000 posterior medians. The left-side panels (a) show the fixed effects: population mean β_0 and interaction coefficient $\bar{\psi}$. The middle panels (b) show variances of: mean trait value V_α, social impact V_ϕ and social responsiveness V_ψ. The right-side panels (c) show the covariances: mean trait value-social impact $Cov_{\alpha\phi}$, mean trait value-social responsiveness $Cov_{\alpha\psi}$ and social impact-social responsiveness $Cov_{\psi\phi}$. The total sample sizes are increased by increasing the number of individuals. Each individual interacts with eight different social partners. The dotted lines represent the simulated 'true' estimate.*

Sampling design

We examined different sampling designs while keeping the total sample size constant at 3200 observations. Throughout the results, we only highlight changes greater than 5% across study designs; smaller differences were considered negligible. Across all designs, estimates of variance components showed minimal differences in both accuracy (bias) and precision (dispersion) (Figure 3, Table S2). Bias in variance components was generally below 5%, with variance in social impact consistently showing a slight overestimation. Variance dispersion was generally stable, except for an in increase in the dispersion for variance in means when fewer individuals and more repeats of pairs were included (Figure 3b). These results indicate that at a total sample size of 3200 all study designs appear adequate to obtain reliable variance estimates. In contrast, estimates of individual-level covariances were more sensitive to study design choices. Covariances were generally underestimated, typically by less than 10%, although the magnitude of bias varied between design choices. Moreover, covariances did not all respond similarly to trade-offs between sampling axes.

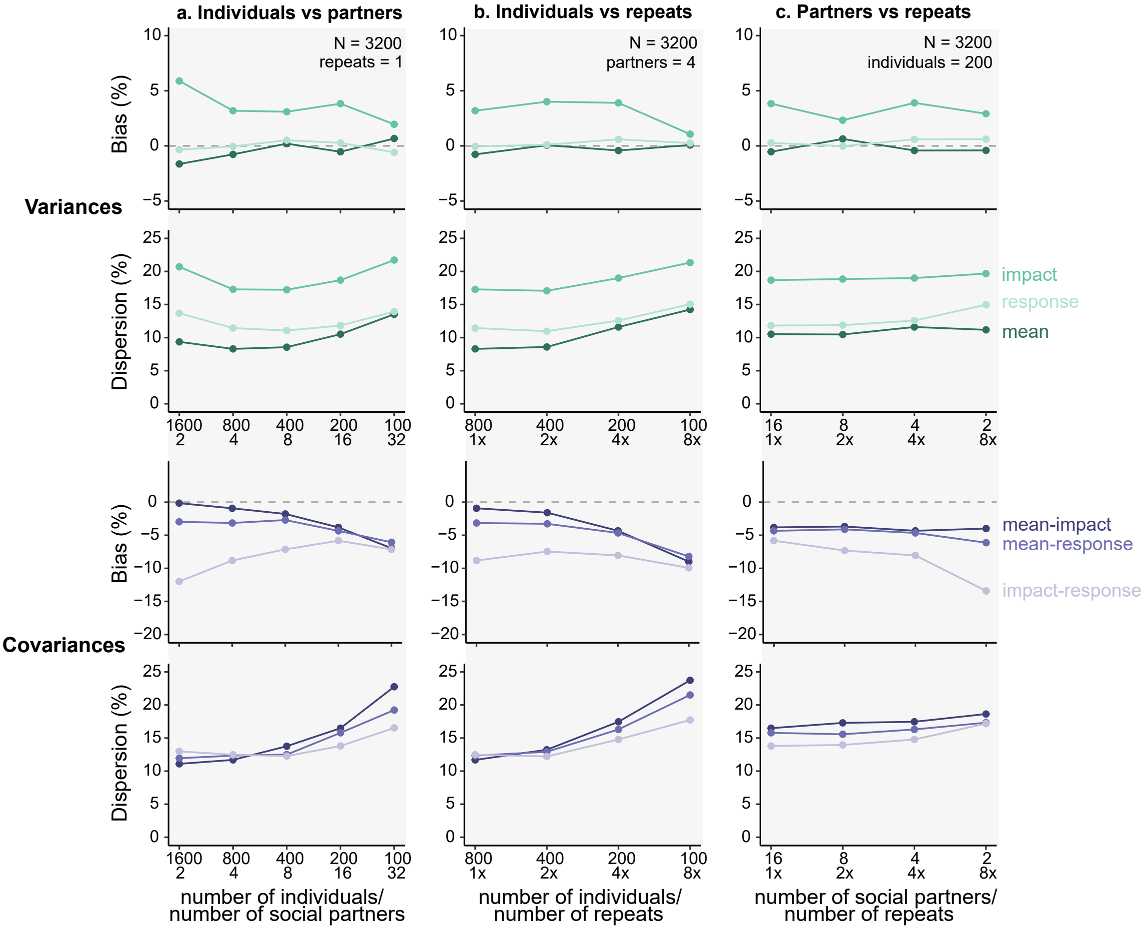


Figure 3: ***Effect of sampling design choices on bias and dispersion of variance and covariance parameters****. Relative bias and relative dispersion of the model estimates of variances and covariances over 1000 simulated datasets per study design. Each study design contains a total sample size of 3200 observations. The left-side panels (a) show the trade-off between the number of individuals and the number of social partners (each individuals interacts once with each social partner). The middle panels (b) show the trade-off between the number of individuals and repeated dyadic interactions (repeats) for a total sample size of 3200 (each individual interacts with four different social partners). The right-side panels (c) show the trade-off between interacting with more different social partners against interacting repeatedly with the same social partners for a total sample size of 3200 (each study design has 200 unique individuals). Top panels represent bias and dispersion of the variance parameters and the bottom panels the covariance parameters.*

More individuals or more unique partners per individual

To investigate this, we compared designs varying the number of individuals (100–1600) and partners per individual (2–32), keeping 3200 total observations. Increasing the number of individuals at the expense of repeated interactions reduced accuracy and precision for some covariances (Figure 3a). Specifically, bias increased for the mean–impact and mean–responsiveness covariances as fewer individuals were included. In contrast, the impact–responsiveness covariance showed the opposite trend, with bias decreasing as the number of individuals decreased and the number of partners per individual increased, reaching a minimum at 200 individuals with 16 partners each. Overall, the design with 400 individuals and 8 partners yielded the lowest average bias across the three covariances. Dispersion also increased as the number of individuals decreased, rising from 11.1% to 22.8% for the mean–impact covariance and from 12.0% to 19.2% for the mean–responsiveness covariance (Table S2). The dispersion of the impact–responsiveness covariance was largely unaffected.

More individuals or more repeated interactions with the same partners

Next, we compared designs varying the number of individuals (100–800) and repeated interactions with the same partners (1x–8x), keeping 3200 total observations (Figure 3b). Here, bias increased for the mean–impact covariance (from −0.9% to −9.0%) and the mean–responsiveness covariance (from −3.1% to −8.2%) when repeated interactions were prioritised over including more individuals (Table S2). Bias in the impact–responsiveness covariance remained stable across designs. However, dispersion increased for all three covariance estimates as fewer individuals were included. This indicates that including more individuals outweighs the benefit of having more repeated pair interactions for covariance estimation.

More social partners or more repeated interactions with the same partners

Lastly, we examined the trade-off between the number of social partners (2-16) and the number of repeated interactions with the same partner (1x-8x), keeping 3200 total observations. Bias and dispersion remained largely unchanged across study designs, except for the impact–responsiveness covariance, which showed a marked increase in bias (from −5.8% to −13.4%) as the number of unique partners decreased and repeated dyads increased (Figure 3c, Table S2). Overall, there is little evidence that prioritising more partners versus more repeated dyads affects the estimation accuracy and precision.

Taken together, these results show that estimation of variance in mean trait values, social impact, and social responsiveness are not strongly affected by study design choices for a total sample size of 3200. Covariances, on the contrary, are more sensitive to how social interaction observations are partitioned. Designs that balance moderate numbers of individuals with moderate numbers of partners are expected to perform well overall, and including more individuals outweighs the benefit of prioritising more repeated dyads. We further tested the trade-off in sampling design axes for a total sample size of 800 observations, which showed similar patterns (see Figure S1, Table S3)

Model comparison

We compared the complete I&R model to two reduced models, the variance-partitioning approach and the trait-based model. Specifically, these reduced models do not account for variation in social responsiveness, nor for measurement error in the partner trait. Our results show that models that did not account for measurement error (Trait and Trait+RS) in the partners trait showed large biases in all three model parameters (Figure 4 and 5).

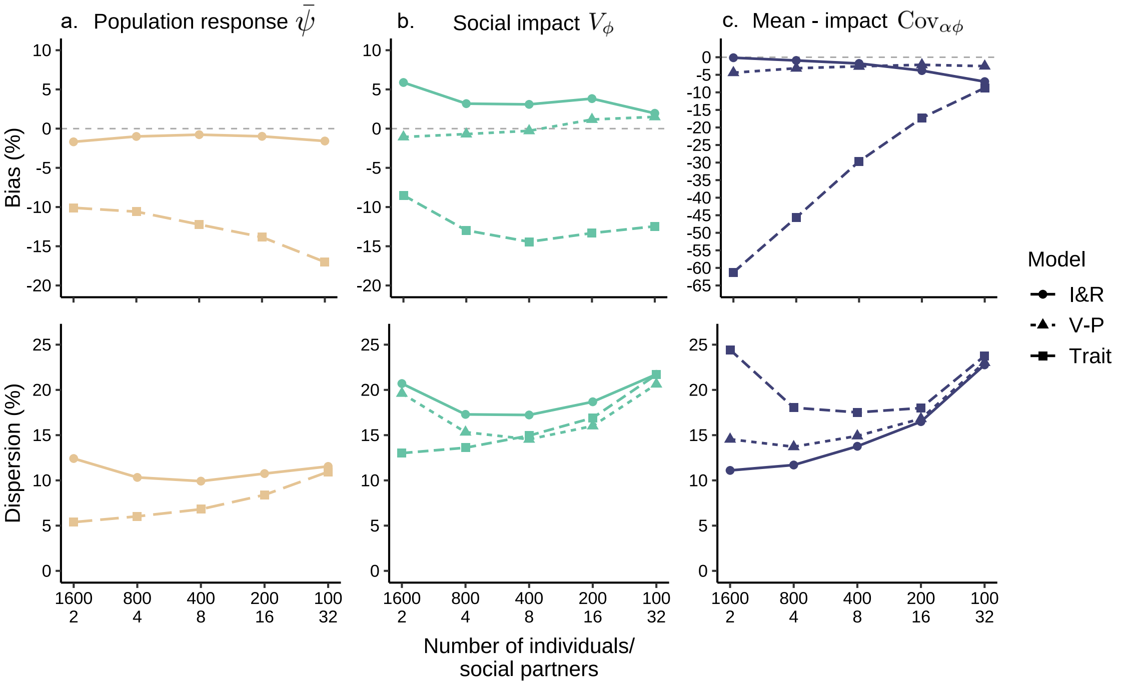


Figure 4: ***Comparison of the impact and responsiveness model (I&R) with a variance-partitioning model (V-P) and a trait-based model (Trait).*** *The figure shows the accuracy (bias) and precision (dispersion) of (a) the population response $\bar{\psi}$, (b) the variance in social impact V_ϕ, (c) the covariance between mean trait value and social impact $\mathrm{Cov}_{\alpha\phi}$. Top panels show the relative bias of three models that analysed the same 1000 simulated datasets for four different partitions of number of individuals and number of social partners per individual. Bottom panel shows the relative dispersion (MADm), expressed as a percentage.*

For the population-level response $\bar{\psi}$, the I&R model showed minimal bias across study designs (all < 2%), with dispersion between 9.9% and 12.4% (Table S4). The Trait model underestimated $\bar{\psi}$ by 10.1 - 17.0%, particularly when fewer individuals were included (Figure 4). Dispersion in the estimation of $\bar{\psi}$ was slightly higher in the I&R model compared to the Trait model across study designs. For the variance in social impact V_{ϕ}, the Trait model showed consistent underestimation between −8.5% and −14.4% compared to slight overestimation by the I&R model (2.0% to 5.9%) (Figure 4, Table S5). The variance partitioning model (V-P), however, showed no marked bias in the estimation of the variance in social impact. The dispersion of V_{ϕ} was slightly higher in the I&R model. All models underestimated the mean-impact covariance ($Cov_{\alpha\phi}$). In the I&R model, bias ranged from −0.1% to −7.0%, with dispersion ranging from 11.1% to 22.8%. In the variance-partitioning model, bias ranged from −2.2% to −4.4%, with dispersion ranging from 13.7% to 23.0%. The trait-based model showed severe underestimation, particularly when more individuals interacted with fewer social partners (−61.3%), with bias decreasing to −8.7% when fewer individuals interacted with more social partners (Figure 4, Table S6). Overall, the I&R model and V-P model performed better than the Trait model, which produced extremely biased estimates of the mean-impact covariance under most sampling conditions (which is equivalent to a DGE-IGE covariance).

To test whether the model error in the trait-based approach arose from missing the variance component that estimates responsiveness or measurement error, we compared the full I&R model to two models that each lacked one of these components. Our results show that there are very little differences in estimation bias and precision between the full impact and responsiveness (I&R) model and a model that does not account for individual variation in responsiveness, however, not modelling measurement error can cause substantial biases (Figure 5).

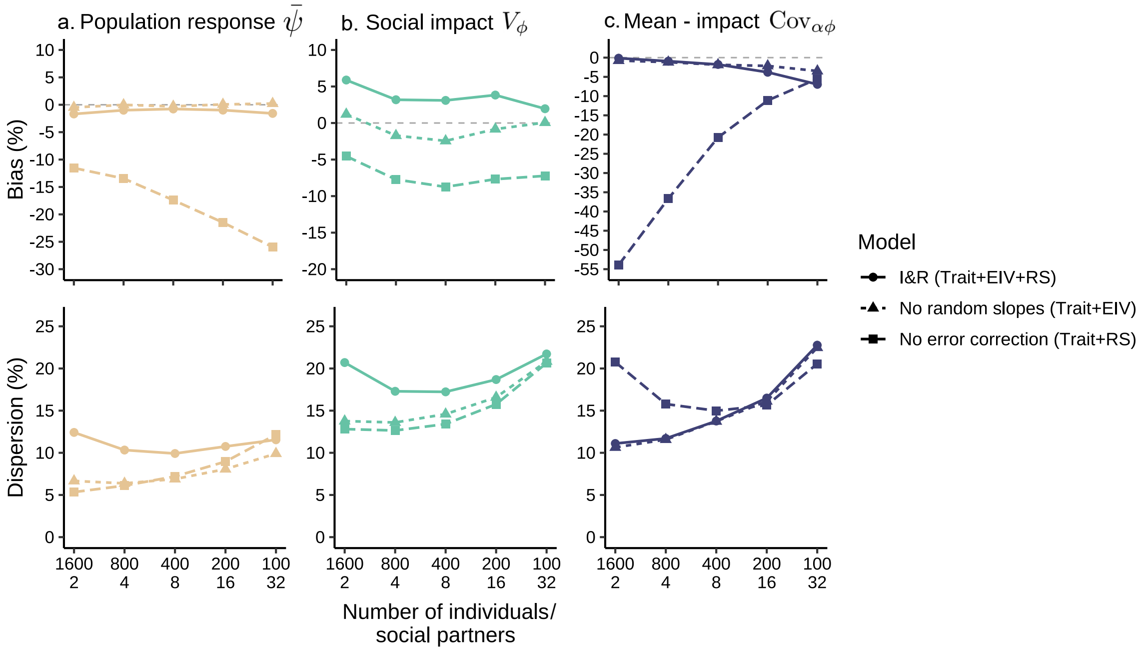


Figure 5: ***Comparison of the impact and responsiveness (I&R) model with a trait-based model without random slopes (Trait+EIV) and a trait-based model without errors-in-variables correction (Trait+RS).*** *See caption Figure 4 for details.*

For the population-level response $\bar{\psi}$, both the I&R and Trait+EIV models showed minimal bias ($<$ 2%) across all sample sizes (Table S4). In contrast, the Trait+RS model consistently underestimated $\bar{\psi}$, with bias worsening as the number of individuals decreased (up to 25.96% at 100 individual), even more bias than the Trait model (Table S4). Dispersion was similar between models (6–12%). For the variance in social impact V_ϕ, Trait+EIV model (1.2% to 2.5%) was slightly closer to unbiased estimation than the I&R. The Trait+RS model showed the largest bias (-4.5% to -8.7%). Both Trait+RS and Trait+EIV provide less biased estimates for variance in social impact than the basic Trait model. For the mean-impact covariance Cov$\alpha\phi$, the Trait+EIV model and I&R model do not show substantial estimation bias. The Trait+RS model showed large bias, particularly when more individuals interacted with fewer social partners. This bias was slightly less than the basic Trait model.

Overall, the Trait+EIV model performed comparably to the full I&R model, with even slightly more accurate and precise estimation of V_ϕ. Adding error correction thus substantially improves the Trait model. In contrast, the trait-based model with random slopes and without error correction

(Trait+RS) performed significantly worse than the I&R model. However, compared to the Trait model, the Trait+RS model performed worse in estimating the population-level response $\bar{\psi}$, slightly better in estimating V_ϕ and slightly better in estimating Cov$\alpha\phi$. We also performed the model comparison with a lower total sample size of 800 observations, which showed similar patterns (Figures S2, S3).

Discussion

Understanding how social traits evolve requires not only studying individual variation in the mean phenotype, but also accurately estimating how individuals influence, and are influenced by, their social partners. Building upon recent conceptual advances (Araya-Ajoy et al., 2020; de Groot et al., 2023), our study provides a systematic exploration of the statistical and study design challenges associated with estimating mean trait values, social responsiveness, and social impact. Our simulations demonstrate that it is possible to estimate mean trait value, social responsiveness and social impact with reasonable accuracy and precision, but that model performance can be strongly affected by both sampling design and model structure. In particular, we show that failing to account for measurement error in the traits of social partners can lead to biased estimates of key parameters, including the population-level responsiveness $\bar{\psi}$ and the covariance between mean phenotype and social impact. By systematically assessing these issues, our results provide practical guidance for studying the evolution of labile social traits in dynamic systems.

Our results highlight that accurate and precise estimation of social parameters depends strongly on total sample size. Across increasing sample sizes, the I&R model produced unbiased estimates of fixed effects and variance components, with dispersion decreasing markedly as larger datasets were used. This aligns with earlier simulation work showing that model performance improves rapidly with increasing numbers of individuals and observations per individual (Dingemanse & Dochtermann, 2013; Martin et al., 2011). Notably, covariance estimates between components of individuality (Cov$_{\alpha\phi}$, Cov$_{\alpha\psi}$, and Cov$_{\psi\phi}$) were consistently underestimated at small sample sizes, which largely disappeared when sample sizes were larger ($\geq$ 3200). This indicates that detecting and quantifying covariance structure among individuals in social traits requires substantially more data than estimating univariate effects (i.e. variances) alone. Large sample sizes are necessary to obtain unbiased estimates, especially when researchers expect even smaller effect sizes, for instance, when estimating indirect genetic effects (Santostefano et al., 2024). These sample sizes are fortunately not uncommon in studies on IGEs. Based on data extracted from Santostefano et al. (2024), we found that across 47 meta-analysed IGE studies, an average of 5023 individuals were included per study, with a mean of 1.48 observations per individual. Nonetheless, six studies were conducted with fewer than 400 individuals and fewer than 1600 total observations. Similar studies, where sample sizes are limited, might benefit the most from optimising sampling methods (see Figure S1). We found that at least 3200 total observations are needed, which aligns with the widely acknowledged difficulty in obtaining precise IGE estimates without large and well-structured datasets (Bijma, 2010; Charmantier et al., 2014). Our inspection of the full parameter set indicated that the impact–responsiveness model struggled most when estimating the residual impact parameter, which was typically overestimated by about 50% (Table S7). This bias likely arises because variation in residual impact is estimated through the random identity of the partner, which is also used to model the partner's impact trait affecting the focal individual. The model therefore has difficulty disentangling these two partner effects. We furthermore suspect that the complexity of the multivariate structure likely increases the data requirements for precise estimation. In our simulations, four correlated traits (mean phenotype, responsiveness, impact trait, and residual impact) were modelled, and the I&R model had to disentangle their variances and covariances using only the focal and partner identities and (labile) phenotypes within single assays.

Estimating social responsiveness and impact, as well as their covariances with mean phenotype, requires specific features in study design that are rarely addressed in detail. In particular, the need for individuals/genotypes to act both as focal subjects and social partners, variation in partner traits, and repeated interactions between individuals or genotypes imposes constraints that are unique to studies of social behaviour. By examining these sampling design requirements, we found that the specific design choices for datasets with a total of 3200 observations did not have extreme effects on the accuracy and precision of the estimation of variance components (changes in bias and dispersion never exceeded 5%). This means that researchers are able to compensate by investing in increasing the number of social partners per individual or repeating the same dyadic interactions if they do not have access to a large population size. Similarly, in some cases, if observations of social behaviour are often with the same social partner, researchers are forced to obtain an adequate population size of unique individuals to estimate all variance components. For example, this applies to longitudinal studies on indirect genetic effects in breeding attempts of long-lived animals that form strong pair bonds (Moiron et al., 2020; Teplitsky et al., 2010). However, for covariances, our results show that analyses using small population sizes could suffer from lower accuracy and precision in estimating individual-level covariances. To accurately estimate covariances, having more individuals is preferred over having more social partners or more dyadic repeated interactions at an equal total sample size in almost all cases. Furthermore, we also show that repeating pairwise interactions with the same individuals at the cost of using more individuals or more unique social partners is not advisable. In our analyses, increasing the repeats of the same dyads always resulted in a decrease in number of individuals or a decrease in number of unique dyads (social partners). We show that if the total sample size is kept equal having more repeats of the same dyad is either detrimental or does not improve the model estimation. However, repeating dyads with the purpose of increasing the total sample size should improve estimation accuracy and precision, but likely not as much as increasing the number of individuals or the number of unique dyads. Thus, based on our specific simulations, the order of priority for increasing sample sizes should be:

1. Increase the number of individuals;
2. Increase the number of social partners per individual (unique dyads);
3. Increase the number of repeated dyads.

However, it is important to recognise that our conclusions are relevant for the specific parameter values, effect sizes and the impact-responsiveness model used in our simulations. Consequently, we strongly encourage researchers to simulate their own datasets and analyse them to identify the sampling design most suitable for their expected effect sizes and study system. To facilitate this process, we developed socialSim, an easy-to-use R package that provides a simple workflow for designing and evaluating social interaction studies (Wijnhorst, 2025). The package includes three core functions: `simulate_data()`, which generates social interaction datasets under user-specified parameters. The function `run_model()`, where the user can choose one of the hierarchical Bayesian models tested in this article in Stan; and `summarise_results()`, which extracts the relative bias and relative precision of the estimated parameters. Importantly, socialSim can be used without any prior experience in Stan programming or Bayesian hierarchical modelling.

In order to detect the consequences of having incomplete models when we suspect complex multivariate social phenotypes, we compared a complete I&R model to several reduced models. Importantly, we show that trait-based models which lack specific components to estimate individual differences in traits may perform worse. Our comparisons show that the variance-partitioning (V-P), however, showed very little biases and low dispersion in estimating the variation in social impact and the covariance (mean trait value x social impact) under large sample sizes. This is a positive result because the variance-partitioning approach is also the most widely used method for estimating IGEs (Bailey & Desjonquères, 2022). However, we show that using a trait-based model, that does not account for variation in slopes or measurement error, can lead to an underestimation of the social effect V_ϕ and the mean-impact covariance $Cov_{\alpha\phi}$. We demonstrate the well-known effect that not accounting for measurement error leads to an attenuation of the regression coefficient ($\bar{\psi}$ in our model), which also caused an underestimation of the social effect V_ϕ. Interestingly, our model comparison indicates that adding or removing random slopes has little influence on the accuracy and precision of model estimates. Thus, including random slopes is not detrimental and may even be preferable when individual variation in responsiveness is of interest. In contrast, not accounting for random slopes when such variation is present in the data does not appear to worsen model performance. This is somewhat unexpected, as previous studies have emphasised the importance of modelling among-individual variation in slopes. For instance, omitting random slopes can bias fixed effects and inflate Type I error rates (Barr et al., 2013), or lead to overestimated between-individual variance components depending on the intercept–slope correlation (i.e. the mean behaviour–responsiveness covariance) (Schielzeth & Forstmeier, 2009). However, in our case, we do not observe such overestimation. Instead, the variance attributable to individual differences in slopes (0.1) is absorbed by the residual variance when slopes are not modelled (residual variance increases from 0.6 to 0.7). Consequently, estimates of repeatability for direct and indirect effects (calculated as the proportion of variance explained by V_α and V_ϕ, respectively) remain stable. Nevertheless, this means that within-individual variance is inflated, as variance in slopes is absorbed by the residual, masking meaningful differences in responsiveness. Therefore, given both prior evidence for the potential risks of omitting random slopes and our finding that their inclusion is at least not harmful, we recommend incorporating random slopes into IGE models to better capture individual differences in social responsiveness.

The impact-and-responsiveness framework we propose is particularly useful when the partner trait is either measured with error or varies substantially between social interactions. In the context of social effects, we are primarily interested in how repeatable individual differences in partners shape the focal individual's behaviour. These effects are not caused by non-heritable or transient expressions of a partner's phenotype during a given interaction, but by repeatable traits, such as mean levels of aggression or body size, that exert influence across multiple encounters (Bleakley & Brodie, 2009; Saltz, 2013; Wilson et al., 2009). Therefore, rather than modelling the observed phenotype expressed in a single interaction, we estimate the latent mean trait value of each partner using a double equation (errors-in-variables) model. This approach captures the repeatable among-individual variation that drives social effects and allows us to quantify its contribution to focal behaviour. Importantly, this latent partner trait (χ_j) can also be partitioned into additive genetic and permanent environmental components using an animal model. This enables the estimation of social breeding values and total genetic variance underlying IGEs using the interaction coefficient $\bar{\psi}$ (Bijma et al., 2007; McGlothlin & Brodie, 2009; Wolf et al., 1999). Thus, the model not only accounts for measurement error or stochastic expression in labile traits, but also aligns with the conceptual goal of identifying the stable genetic and/or phenotypic individual differences in partners that generate social effects.

Several theoretical papers have suggested modelling social responsiveness using random slopes in IGE frameworks (Araya-Ajoy et al., 2020; Bailey et al., 2021; Dingemanse & Araya-Ajoy, 2015; Martin & Jaeggi, 2022), which is further supported by observational and experimental evidence that individuals show repeatable differences in how they respond to the social cues (Bailey & Zuk, 2012; Chenoweth et al., 2010; Fürtbauer & Fry, 2018; Guayasamin et al., 2017; Jablonszky et al., 2022; Morand-Ferron et al., 2011; Strickland & Frère, 2019; Strickland et al., 2021). We support this perspective and show that including random slopes does not harm estimation accuracy or precision. Therefore, we recommend considering random slopes in IGE models, especially when aiming to disentangle social impact and responsiveness, two traits that can vary independently and jointly shape social phenotypes (de Groot et al., 2023). Exploring how these traits genetically covary, including with the direct effects, will be key to understanding the evolution of social behaviour (Araya-Ajoy et al., 2020; Bailey et al., 2021; Martin & Jaeggi, 2022). By assessing the utility of an impact-and-responsiveness model, we hope to provide a useful statistical tool for the study of the expression of social traits.

Data Availability Statement

The data and code used to generate the data and results of this study are available on:
https://github.com/RoriWijnhorst/Social-impact-and-responsiveness

Acknowledgements

We thank Joel Pick for lending advice on the computational challenge of running many Bayesian models. The authors also thank Alastair Wilson for insightful discussions related to the manuscript and the reviewers.

References

Akçay, E., & Van Cleve, J. (2012). Behavioral responses in structured populations pave the way to group optimality. *American Naturalist*, *179*(2), 257–269. https://doi.org/10.1086/663691

Araya-Ajoy, Y. G., Mathot, K. J., & Dingemanse, N. J. (2015). An approach to estimate short-term, long-term and reaction norm repeatability. *Methods in Ecology and Evolution*, *6*(12), 1462–1473. https://doi.org/10.1111/2041-210X.12430

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

Bailey, N. W., & Desjonquères, C. (2022). The indirect genetic effect interaction coefficient ψ : Theoretically essential and empirically neglected. *Journal of Heredity*, *113*(1), 79–90. https://doi.org/10.1093/jhered/esab056

Bailey, N. W., Desjonquères, C., Drago, A., Rayner, J. G., Sturiale, S. L., & Zhang, X. (2021). A neglected conceptual problem regarding phenotypic plasticity's role in adaptive evolution: The importance of genetic covariance and social drive. *Evolution Letters*, *5*(5), 444–457. https://doi.org/10.1002/evl3.251

Bailey, N. W., Marie-Orleach, L., & Moore, A. J. (2018). Indirect genetic effects in behavioral ecology: Does behavior play a special role in evolution? *Behavioral Ecology*, *29*(1), 1–11. https://doi.org/10.1093/beheco/arx127

Bailey, N. W., & Zuk, M. (2012). Socially flexible female choice differs among populations of the Pacific field cricket: Geographical variation in the interaction coefficient psi (ψ). *Proceedings of the Royal Society B Biological Sciences*, *279*(1742), 3589–3596. https://doi.org/10.1098/rspb.2012.0631

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: A meta-analysis. *Animal Behaviour*, *77*(4), 771–783. https://doi.org/10.1016/j.anbehav.2008.12.022

Bengtsson, H. (2021). A unifying framework for parallel and distributed processing in r using futures. *The R Journal*, *13*(2), 208–227. https://doi.org/10.32614/RJ-2021-048

Bijma, P. (2014). The quantitative genetics of indirect genetic effects: A selective review of modelling issues. *Heredity*, *112*(1), 61–69. https://doi.org/10.1038/hdy.2013.15

Bijma, P. (2010). Estimating indirect genetic effects: Precision of estimates and optimum designs. *Genetics*, *186*(3), 1013–1028. https://doi.org/10.1534/genetics.110.120493

Bijma, P., Muir, W. M., & Van Arendonk, J. A. M. (2007). Multilevel selection 1: Quantitative genetics of inheritance and response to selection. *Genetics*, *175*(1), 277–288. https://doi.org/10.1534/genetics.106.062711

Bleakley, B. H., & Brodie, E. D., III. (2009). Indirect genetic effects influence antipredator behavior in guppies: Estimates of the coefficient of interaction psi and the inheritance of reciprocity. *Evolution*, *63*(7), 1796–1806. https://doi.org/10.1111/j.1558-5646.2009.00672.x

Brommer, J. E. (2013). Variation in plasticity of personality traits implies that the ranking of personality measures changes between environmental contexts: Calculating the cross-environmental correlation. *Behavioral Ecology and Sociobiology*, *67*(10), 1709–1718. https://doi.org/10.1007/s00265-013-1603-9

Brommer, J. E., & Rattiste, K. (2008). "Hidden" reproductive conflict between mates in a wild bird population. *Evolution*, *62*(9), 2326–2333. https://doi.org/10.1111/j.1558-5646.2008.00451.x

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Charmantier, A., Garant, D., & Kruuk, L. E. B. (2014). *Quantitative Genetics in the Wild*. OUP Oxford.

Chenoweth, S. F., Rundle, H. D., & Blows, M. W. (2010). Experimental evidence for the evolution of indirect genetic effects: changes in the interaction effect coefficient, Psi (ψ), due to sexual selection. *Evolution*, *64*(6), 1849–1856. https://doi.org/10.1111/j.1558-5646.2010.00952.x

de Groot, C., Wijnhorst, R. E., Ratz, T., Murray, M., Araya-Ajoy, Y. G., Wright, J., & Dingemanse, N. J. (2023). The importance of distinguishing individual differences in 'social impact' versus 'social responsiveness' when quantifying indirect genetic effects on the evolution of social plasticity. *Neuroscience and Biobehavioral Reviews*, *144*, 104996. https://doi.org/10.1016/j.neubiorev.2022.104996

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., & Dochtermann, N. A. (2013). Quantifying individual variation in behaviour: mixed-effect modelling approaches. *Journal of Animal Ecology*, *82*(1), 39–54. https://doi.org/10.1111/1365-2656.12013

Dingemanse, N. J., Araya-Ajoy, Y. G., & Westneat, D. F. (2021). Most published selection gradients are underestimated: why this is and how to fix it. *Evolution*, *75*(4), 806–818. https://doi.org/10.1111/evo.14198

Fürtbauer, I., & Fry, A. (2018). Social conformity in solitary crabs, carcinus maenas, is driven by individual differences in behavioural plasticity. *Animal Behaviour*, *135*, 131–137. https://doi.org/10.1016/j.anbehav.2017.11.010

Griffing, B. (1967). Selection in reference to biological groups. I. Individual and group selection applied to populations in unordered groups. *Australian Journal Biological Science*, *20*, 127–142.

Guayasamin, O. L., Couzin, I. D., & Miller, N. Y. (2017). Behavioural plasticity across social contexts is regulated by the directionality of inter-individual differences [The Cognition of Fish]. *Behavioural Processes*, *141*, 196–204. https://doi.org/10.1016/j.beproc.2016.10.004

Han, C. S., Tuni, C., Ulcik, J., & Dingemanse, N. J. (2018). Increased developmental density decreases the magnitude of indirect genetic effects expressed during agonistic interactions in an insect: density-dependent indirect genetic effects. *Evolution*, *72*(11), 2435–2448. https://doi.org/10.1111/evo.13600

Henderson, C. R. (1984). *Applications of linear models in animal breeding*. University of Guelph Press.

Jablonszky, M., Canal, D., Hegyi, G., Krenhardt, K., Laczi, M., Markó, G., Nagy, G., Rosivall, B., Szász, E., Zsebők, S., & Garamszegi, L. Z. (2022). Individual differences in song plasticity in response to social stimuli and singing position. *Ecology and Evolution*, *12*(5), e8883. https://doi.org/10.1002/ece3.8883

Kazancıoğlu, E., Klug, H., & Alonzo, S. H. (2012). The evolution of social interactions changes predictions about interacting phenotypes. *Evolution*, *66*(7), 2056–2064. https://doi.org/10.1111/j.1558-5646.2012.01585.x

Kirkpatrick, M., & Lande, R. (1989). The evolution of maternal characters. *Evolution*, *43*(3), 485–503. https://doi.org/10.1111/j.1558-5646.1989.tb04247.x

Kruuk, L. E. B. (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Philos. T. Roy. Soc. B.*, *359*, 873–890. https://doi.org/10.1098/rstb.2003.1437

Lane, S. M., Wilson, A. J., & Briffa, M. (2020). Analysis of direct and indirect genetic effects in fighting sea anemones. *Behavioral Ecology*, *31*(2), 540–547. https://doi.org/10.1093/beheco/arz217

Martin, J. S., & Jaeggi, A. V. (2022). Social animal models for quantifying plasticity, assortment, and selection on interacting phenotypes. *Journal of Evolutionary Biology*, *35*(4), 520–538. https://doi.org/10.1111/jeb.13900

Martin, J. S., Jaeggi, A. V., & Koski, S. E. (2023). The social evolution of individual differences: future directions for a comparative science of personality in social behavior. *Neuroscience & Biobehavioral Reviews*, *144*, 104980. https://doi.org/10.1016/j.neubiorev.2022.104980

Martin, J. G. A., Nussey, D. H., Wilson, A. J., & Réale, D. (2011). Measuring individual differences in reaction norms in field and experimental studies: a power analysis of random regression models. *Methods in Ecology and Evolution*, *2*(4), 362–374. https://doi.org/10.1111/j.2041-210X.2010.00084.x

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

McGlothlin, J. W., & Brodie, E. D., III. (2009). How to measure indirect genetic effects: the congruence of trait-based and variance-partitioning approaches. *Evolution*, *63*(7), 1785–1795. https://doi.org/10.1111/j.1558-5646.2009.00676.x

McGlothlin, J. W., Moore, A. J., Wolf, J. B., & Brodie, E. D., III. (2010). Interacting phenotype and the evolutionary process III. Social evolution: indirect genetic effects and social selection. *Evolution*, *64*(9), 2558–2574. https://doi.org/10.1111/j.1558-5646.2010.01012.x

McLean, E. M., Moorad, J. A., Tung, J., Archie, E. A., & Alberts, S. C. (2023). Genetic variance and indirect genetic effects for affiliative social behavior in a wild primate. *Evolution*, *77*(7), 1607–1621. https://doi.org/10.1093/evolut/qpad066

McNamara, J. M., & Weissing, F. J. (2010). Evolutionary game theory. In T. Székely, A. J. Moore, & J. Komdeur (Eds.), *Social Behaviour: Genes, Ecology and Evolution*. Cambridge University Press.

Meyer, K. (1992). Bias and sampling covariances of estimates of variance components due to maternal effects. *Genetics Selection Evolution*, *24*(6), 487. https://doi.org/10.1186/1297-9686-24-6-487

Moiron, M., Araya-Ajoy, Y. G., Teplitsky, C., Bouwhuis, S., & Charmantier, A. (2020). Understanding the social dynamics of breeding phenology: indirect genetic effects and assortative mating in a long-distance migrant. *American Naturalist*, *196*(5), 566–576. https://doi.org/10.1086/711045

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

Morand-Ferron, J., Varennes, E., & Giraldeau, L.-A. (2011). Individual differences in plasticity and sampling when playing behavioural games. *Proceedings of the Royal Society B Biological Sciences*, *278*(1709), 1223–1230. https://doi.org/10.1098/rspb.2010.1769

Ponzi, E., Keller, L. F., Bonnet, T., & Muff, S. (2018). Heritability, selection, and the response to selection in the presence of phenotypic measurement error: Effects, cures, and the role of repeated measurements. *Evolution*, *72*(10), 1992–2004. https://doi.org/10.1111/evo.13573

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

Saltz, J. B. (2013). Genetic composition of social groups influences male aggressive behaviour and fitness in natural genotypes of *Drosophila melanogaster*. *Proc. R. Soc. B.*, *280*(1771), 20131926. https://doi.org/10.1098/rspb.2013.1926

Santostefano, F., Moiron, M., Sánchez-Tójar, A., & Fisher, D. N. (2024). Indirect genetic effects increase the heritable variation available to selection and are largest for behaviors: A meta-analysis. *Evolution Letters*, *9*(1), 89–104. https://doi.org/10.1093/evlett/qrae051

Santostefano, F., Wilson, A. J., Araya-Ajoy, Y. G., & Dingemanse, N. J. (2016). Interacting with the enemy: indirect effects of personality on conspecific aggression in crickets. *Behavioral Ecology*, *27*(4), 1235–1246. https://doi.org/10.1093/beheco/arw037

Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. *Behavioral Ecology*, *20*(2), 416–420. https://doi.org/10.1093/beheco/arn145

Smiseth, P. T., Wright, J., & Kölliker, M. (2008). Parent–offspring conflict and co-adaptation: behavioural ecology meets quantitative genetics. *Proceedings of the Royal Society B Biological Sciences*, *275*(1645), 1823–1830. https://doi.org/10.1098/rspb.2008.0199

Stan Development Team. (2025). RStan: The R interface to Stan [e version 2.32.7].

Strickland, K., & Frère, C. H. (2019). Individual variation in the social plasticity of water dragons. *American Naturalist*, *194*(2), 194–206. https://doi.org/10.1086/704089

Strickland, K., Mitchell, D. J., Delmé, C., & Frère, C. H. (2021). Repeatability and heritability of social reaction norms in a wild agamid lizard. *Evolution*, *75*(8), 1953–1965. https://doi.org/10.1111/evo.14298

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

Teplitsky, C., Mills, J. A., Yarrall, J. W., & Merilä, J. (2010). Indirect genetic effects in a sex-limited trait: the case of breeding time in red-billed gulls. *Journal of Evolutionary Biology*, *23*(5), 935–944. https://doi.org/10.1111/j.1420-9101.2010.01959.x

Tuliozi, B., Mantovani, R., Schoepf, I., Tsuruta, S., Mancin, E., & Sartori, E. (2023). Genetic correlations of direct and indirect genetic components of social dominance with fitness and morphology traits in cattle. *Genetics Selection Evolution*, *55*, 84. https://doi.org/10.1186/s12711-023-00845-8 van de Pol, M. (2012). Quantifying individual variation in reaction norms: how study design affects the accuracy, precision and power of random regression models. *Methods in Ecology and Evolution*, *3*(2), 268–280. https://doi.org/10.1111/j.2041-210X.2011.00160.x

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

West-Eberhard, M. J. (1989). Phenotypic plasticity and the origins of diversity. *Annual Review of Ecology, Evolution, and Systematics*, *20*(Volume 20, 1989), 249–278. https://doi.org/10.1146/annurev.es.20.110189.001341

Wijnhorst, R. E. (2025). *Socialsim: Simulate and analyse social interaction data* [R package version 0.1.6]. https://doi.org/10.32614/CRAN.package.socialSim

Wilson, A. J., Morrissey, M. B., Adams, M. J., Walling, C. A., Guinness, F. E., Pemberton, J. M., Clutton-Brock, T. H., & Kruuk, L. E. B. (2011). Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, *Cervus elaphus*. *Journal of Evolutionary Biology*, *24*(4), 772–783. https://doi.org/10.1111/j.1420-9101.2010.02212.x

Wilson, A. J., Gelin, U., Perron, M.-C., & Réale, D. (2009). Indirect genetic effects and the evolution of aggression in a vertebrate system. *Proceedings of the Royal Society B Biological Sciences*, *276*(1656), 533–541. https://doi.org/10.1098/rspb.2008.1193

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., & Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology*, *79*(1), 13–26. https://doi.org/10.1111/j.1365-2656.2009.01639.x

Wolf, J. B., Brodie, E. D., III, & Moore, A. J. (1999). Interacting phenotypes and the evolutionary process. II. Selection resulting from social interactions. *American Naturalist*, *153*(3), 254–266. https://doi.org/10.1086/303168

Wolf, J. B., Brodie III, E. D., Cheverud, J. M., Moore, A. J., & Wade, M. J. (1998). Evolutionary consequences of indirect genetic effects. *Trends in Ecology and Evolution*, *13*(2), 64–69. https://doi.org/10.1016/S0169-5347(97)01233-0

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

Supplementary material

The full 4 x 4 matrix estimated in the I&R model:

$$\begin{bmatrix} \boldsymbol{\alpha}_i \\ \boldsymbol{\psi}_i \\ \boldsymbol{\chi}_i \\ \boldsymbol{\epsilon}_i \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0},\ \boldsymbol{\Omega}\right),\ \boldsymbol{\Omega} = \begin{bmatrix} V_\alpha & \text{Cov}(\alpha,\psi) & \text{Cov}(\alpha,\chi) & \text{Cov}(\alpha,\epsilon) \\ \text{Cov}(\alpha,\psi) & V_\psi & \text{Cov}(\psi,\chi) & \text{Cov}(\psi,\epsilon) \\ \text{Cov}(\alpha,\chi) & \text{Cov}(\psi,\chi) & V_\chi & \text{Cov}(\chi,\epsilon) \\ \text{Cov}(\alpha,\epsilon) & \text{Cov}(\psi,\epsilon) & \text{Cov}(\chi,\epsilon) & V_\epsilon \end{bmatrix} e \sim \mathcal{N}(0, V_e) \tag{S1}$$

was reduced to a 3 x 3 matrix (Equation 10) estimating the (co-)variance of mean trait values, social responsiveness and social impact using the following equations:

$$V_\phi = V_\epsilon + \bar{\psi}^2 V_\chi + 2\,\bar{\psi}\,\text{Cov}(\chi,\epsilon) \tag{S2}$$

$$\text{Cov}(\alpha,\phi) = \text{Cov}(\alpha,\epsilon) + \bar{\psi}\,\text{Cov}(\alpha,\chi) \tag{S3}$$

$$\text{Cov}(\psi,\phi) = \text{Cov}(\psi,\epsilon) + \bar{\psi}\,\text{Cov}(\chi,\psi) \tag{S4}$$

Table S1: Mean percentage bias in posterior medians across 1000 simulations under different sample sizes

		Total sample size	400	800	1600	3200	6400
		Individuals	50	100	200	400	800
Parameter	Description	Sim. value			Bias (%)		
B_0	Population mean	1.00	1.31	0.61	0.52	0.16	0.07
$\bar{\psi}$	Population response	0.30	-4.31	-1.67	-1.54	-0.77	-2.18
V_α	Mean behaviour variance	0.20	-1.93	0.85	-0.58	0.20	0.17
V_ψ	Social responsiveness variance	0.10	-1.28	-0.92	0.25	0.52	0.21
V_ϕ	Social impact variance	0.10	8.24	4.66	1.91	3.09	2.74
$\text{Cov}(\alpha,\phi)$	Cov: mean × impact	0.080	-17.63	-7.82	-4.44	-1.77	-0.73
$\text{Cov}(\alpha,\psi)$	Cov: mean × responsiveness	-0.085	-26.03	-11.95	-6.10	-2.70	-1.44
$\text{Cov}(\psi,\phi)$	Cov: responsiveness × impact	-0.076	-30.10	-19.48	-11.80	-7.13	-3.42

Table S2: Bias and dispersion for key variance and covariance parameters for different sampling design (total sample size = 3200).

	Individuals	1600	800	400	200	100	800	400	200	100	200	200	200	200
	Social partners	2	4	8	16	32	4	4	4	4	16	8	4	2
	Repeats	1x	1x	1x	1x	1x	1x	2x	4x	8x	1x	2x	4x	8x
Bias (%)	Mean V_α	-1.64	-0.77	0.20	-0.54	0.67	-0.77	0.06	-0.27	0.08	-0.54	0.63	-0.27	-0.41
	Impact V_ϕ	5.88	3.19	3.09	3.83	1.96	3.19	4.01	2.66	1.06	3.83	2.33	2.66	2.92
	Response V_ψ	-0.34	-0.05	0.52	0.27	-0.59	-0.05	0.10	0.36	0.25	0.27	-0.02	0.36	0.61
	$\mathrm{Cov}(\alpha,\phi)$	-0.14	-0.92	-1.77	-3.80	-6.95	-0.92	-1.58	-4.42	-8.98	-3.80	-3.68	-4.42	-3.98
	$\mathrm{Cov}(\alpha,\psi)$	-2.96	-3.14	-2.70	-4.33	-6.06	-3.14	-3.27	-5.40	-8.20	-4.33	-4.11	-5.40	-6.12
	$\mathrm{Cov}(\psi,\phi)$	-11.98	-8.82	-7.13	-5.82	-7.14	-8.82	-7.45	-8.64	-9.93	-5.82	-7.31	-8.64	-13.41
Dispersion (%)	Mean V_α	9.36	8.30	8.55	10.52	13.53	8.30	8.58	10.62	14.24	10.52	10.47	10.62	11.17
	Impact V_ϕ	20.70	17.29	17.23	18.69	21.72	17.29	17.07	19.11	21.35	18.69	18.85	19.11	19.68
	Response V_ψ	13.67	11.44	11.07	11.82	13.93	11.44	10.97	12.80	15.05	11.82	11.88	12.80	14.97
	$\mathrm{Cov}(\alpha,\phi)$	11.10	11.70	13.77	16.49	22.76	11.70	13.23	17.79	23.74	16.49	17.29	17.79	18.63
	$\mathrm{Cov}(\alpha,\psi)$	11.95	12.35	12.50	15.79	19.23	12.35	12.92	15.72	21.51	15.79	15.57	15.72	17.33
	$\mathrm{Cov}(\psi,\phi)$	12.99	12.49	12.27	13.80	16.53	12.49	12.21	14.57	17.73	13.80	13.96	14.57	17.19

Table S3: Bias and dispersion for key variance and covariance parameters for different sampling designs (total sample size = 800).

		400	200	100	50	200	100	50	100	100	100
		2	4	8	16	4	4	4	8	4	2
		1x	1x	1x	1x	1x	2x	4x	1x	2x	4x
Bias (%)	Mean V_α	-0.01	-1.21	0.02	-2.89	-1.21	-1.39	-2.92	0.02	-1.39	-3.09
	Impact V_ϕ	14.07	5.36	2.63	1.45	5.36	5.44	1.11	2.63	5.44	8.94
	Response V_ψ	3.61	0.91	-0.42	1.04	0.91	-0.06	-2.13	-0.42	-0.06	1.18
	$\mathrm{Cov}(\alpha,\phi)$	-3.75	-6.10	-11.16	-21.00	-6.10	-10.79	-21.60	-11.16	-10.79	-9.74
	$\mathrm{Cov}(\alpha,\psi)$	-17.34	-16.12	-17.27	-23.93	-16.12	-18.90	-28.77	-17.27	-18.90	-22.14
	$\mathrm{Cov}(\psi,\phi)$	-24.53	-24.61	-24.65	-26.10	-24.61	-25.72	-32.38	-24.65	-25.72	-29.83
Dispersion (%)	Mean V_α	17.46	15.96	17.34	19.65	15.96	17.95	21.19	17.34	17.95	18.71
	Impact V_ϕ	19.65	19.81	22.05	27.95	19.81	23.43	28.64	22.05	23.43	25.99
	Response V_ψ	22.97	21.99	21.38	21.84	21.99	22.10	26.31	21.38	22.10	25.48
	$\mathrm{Cov}(\alpha,\phi)$	19.83	20.52	25.28	32.53	20.52	25.41	33.87	25.28	25.41	27.74
	$\mathrm{Cov}(\alpha,\psi)$	25.06	23.33	25.75	30.83	23.33	26.85	34.44	25.75	26.85	32.52
	$\mathrm{Cov}(\psi,\phi)$	22.75	22.22	23.89	28.53	22.22	25.72	32.29	23.89	25.72	29.81

Table S4: Fixed effect: population response $\bar{\psi}$

	Individuals	I&R	Trait	Trait+EIV	Trait+RS
Bias (%)	1600	-1.68	-10.09	-0.48	-11.54
	800	-1.00	-10.59	-0.02	-13.45
	400	-0.77	-12.23	-0.29	-17.39
	200	-0.98	-13.84	0.06	-21.48
	100	-1.58	-17.00	0.22	-25.96
Dispersion (%)	1600	12.42	5.38	6.66	5.34
	800	10.33	6.02	6.37	6.11
	400	9.92	6.81	6.89	7.20
	200	10.75	8.40	8.06	8.96
	100	11.54	10.95	9.91	12.16

Table S5: Variance: social impact V_ϕ

	Individuals	I&R	V–P	Trait	Trait+EIV	Trait+RS
Bias (%)	1600	5.88	-1.06	-8.50	1.20	-4.54
	800	3.19	-0.69	-12.97	-1.72	-7.72
	400	3.09	-0.28	-14.43	-2.45	-8.74
	200	3.83	1.16	-13.32	-0.83	-7.68
	100	1.96	1.50	-12.46	0.06	-7.23
Dispersion (%)	1600	20.70	19.61	13.01	13.76	12.81
	800	17.29	15.33	13.61	13.60	12.65
	400	17.23	14.54	14.95	14.59	13.41
	200	18.69	16.01	16.91	16.58	15.76
	100	21.72	20.63	21.68	20.85	20.64

Table S6: Covariance: mean behaviour-social impact $Cov_{\alpha\phi}$

	Individuals	I&R	V–P	Trait	Trait+EIV	Trait+RS
Bias (%)	1600	-0.14	-4.41	-61.26	-0.76	-53.91
	800	-0.92	-3.13	-45.66	-1.18	-36.66
	400	-1.77	-2.59	-29.67	-1.86	-20.76
	200	-3.80	-2.16	-17.27	-2.15	-11.14
	100	-6.95	-2.54	-8.73	-3.47	-5.51
Dispersion (%)	1600	11.10	14.55	24.43	10.64	20.75
	800	11.70	13.72	18.03	11.57	15.78
	400	13.77	14.91	17.52	13.73	14.98
	200	16.49	16.82	18.00	16.12	15.64
	100	22.76	23.02	23.76	22.47	20.55

Table S7: Mean model estimates (posterior medians) of 1000 simulated datasets under different sampling partitions (total sample size = 3200).

		Individuals	1600	800	400	200	100	400	200	100	200
		Social partners	2	4	8	16	32	4	4	4	8
		Repeats	1x	1x	1x	1x	1x	2x	4x	8x	2x
Parameter	Description	Sim. value	Model outcome								
β_0	Population mean	1.00	1.000	1.000	1.002	1.004	1.005	1.002	1.005	1.010	1.005
$\bar{\psi}$	Population response	0.30	0.295	0.297	0.298	0.297	0.295	0.299	0.296	0.293	0.295
V_α	Mean behaviour variance	0.20	0.197	0.198	0.200	0.199	0.201	0.200	0.199	0.200	0.201
V_ψ	Responsiveness variance	0.10	0.100	0.100	0.101	0.100	0.099	0.100	0.100	0.100	0.100
V_ϵ	Residual impact variance	0.01	0.017	0.015	0.014	0.014	0.014	0.015	0.014	0.014	0.014
V_χ	Impact trait variance	1.00	0.999	0.999	0.996	1.002	0.995	0.999	0.998	0.996	0.996
$r_{\alpha\epsilon}$	Corr: mean × res. impact	-0.60	-0.598	-0.589	-0.587	-0.579	-0.569	-0.585	-0.574	-0.557	-0.578
$r_{\alpha\psi}$	Corr: mean × response	0.00	0.043	0.017	0.015	-0.000	-0.000	0.012	-0.000	-0.009	0.000
$r_{\alpha\chi}$	Corr: mean × impact trait	-0.60	-0.285	-0.375	-0.451	-0.494	-0.526	-0.426	-0.451	-0.472	-0.484
$r_{\epsilon\psi}$	Corr: res. impact × response	0.60	0.599	0.595	0.590	0.584	0.569	0.590	0.583	0.567	0.585
$r_{\chi\psi}$	Corr: impact trait × response	-0.60	-0.603	-0.598	-0.590	-0.585	-0.579	-0.592	-0.584	-0.574	-0.583
$r_{\chi\epsilon}$	Corr: impact trait × res. impact	0.00	0.033	0.009	0.004	0.000	0.006	0.003	0.010	-0.001	0.007
V_e	Residual variance	0.60	0.599	0.600	0.599	0.600	0.600	0.600	0.600	0.600	0.599
V_{e_χ}	Measurement error	0.10	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100

Table S8: Mean model estimates (posterior medians) of 1000 simulated datasets under different sampling partitions (total sample size = 800).

		Individuals	400	200	100	50	100	50	100
		Social partners	2	4	8	16	4	4	2
		Repeats	1x	1x	1x	1x	2x	4x	4x
Parameter	Description	Sim. value	Model outcome						
β_0	Population mean	1.00	1.000	1.001	1.004	1.013	1.006	1.007	1.005
$\bar{\psi}$	Population response	0.30	0.298	0.297	0.296	0.293	0.297	0.291	0.296
V_α	Mean behaviour variance	0.20	0.200	0.198	0.200	0.194	0.197	0.194	0.194
V_ψ	Responsiveness variance	0.10	0.104	0.101	0.100	0.101	0.100	0.098	0.101
V_ϵ	Residual impact variance	0.01	0.024	0.016	0.015	0.014	0.015	0.015	0.017
V_χ	Impact trait variance	1.00	0.996	0.991	0.985	0.975	0.990	0.967	0.990
$r_{\alpha\epsilon}$	Corr: mean × res. impact	-0.60	-0.512	-0.525	-0.515	-0.478	-0.510	-0.460	-0.492
$r_{\alpha\psi}$	Corr: mean × response	0.00	0.021	0.030	0.010	0.001	0.021	0.004	0.019
$r_{\alpha\chi}$	Corr: mean × impact trait	-0.60	-0.113	-0.161	-0.225	-0.278	-0.197	-0.206	-0.139
$r_{\epsilon\psi}$	Corr: res. impact × response	0.60	0.584	0.571	0.550	0.511	0.548	0.513	0.561
$r_{\chi\psi}$	Corr: impact trait × response	-0.60	-0.555	-0.554	-0.539	-0.514	-0.535	-0.504	-0.522
$r_{\chi\epsilon}$	Corr: impact trait × res. impact	0.00	0.006	0.009	0.008	0.006	0.007	0.008	0.009
V_e	Residual variance	0.60	0.583	0.594	0.597	0.600	0.598	0.597	0.599
V_{e_χ}	Measurement error	0.10	0.100	0.100	0.100	0.100	0.100	0.100	0.100

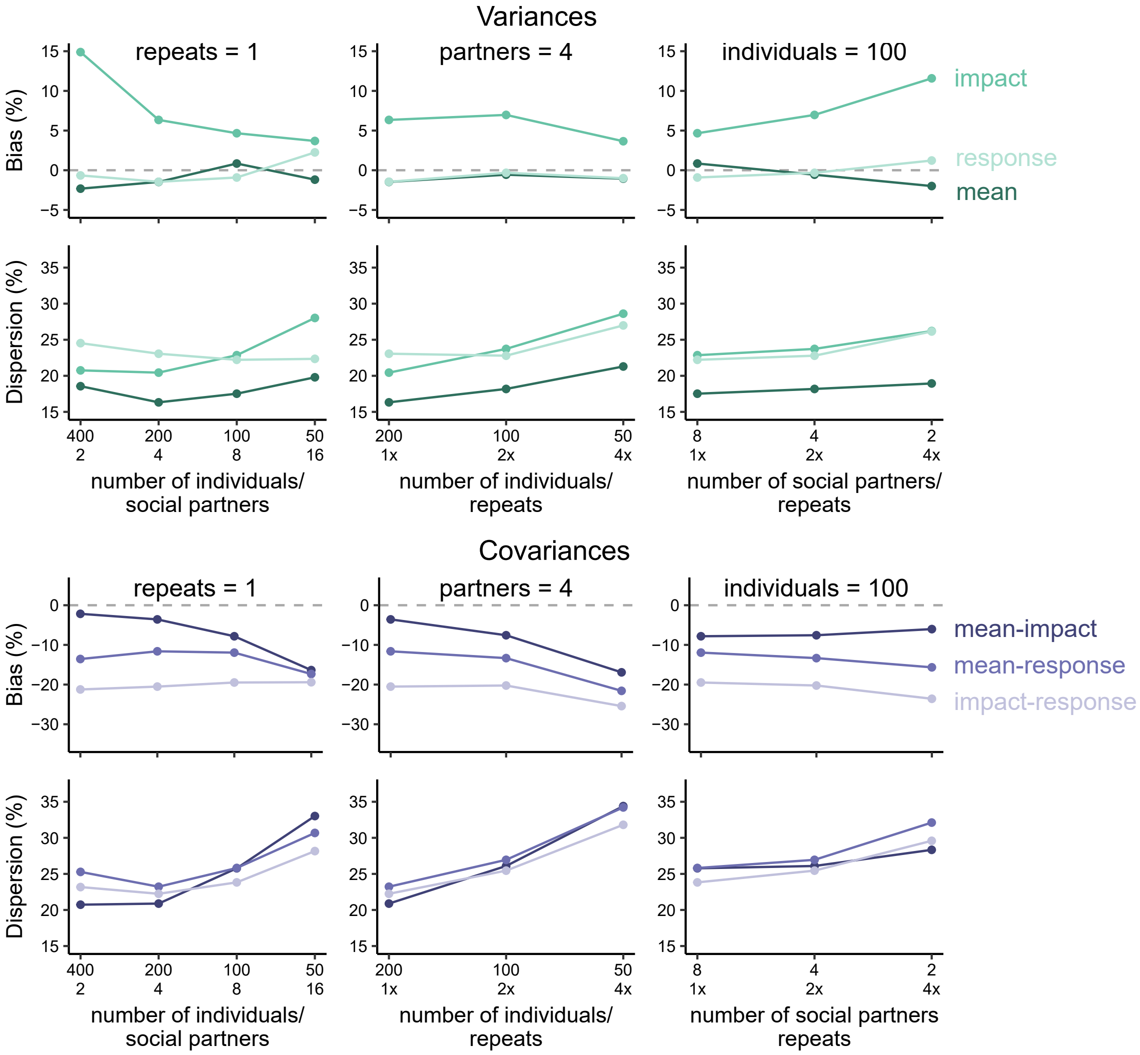


Figure S1: ***Analysis of sampling design parameters for optimal bias and precision for estimating variance and covariance parameters****. Relative bias and dispersion of the posterior median of variances and covariances components of 1000 simulated datasets per sampling design. Top and bottom left figures show the trade-off between the number of individuals and the number of social partners to obtain a total sample size of 800. Middle figures show the trade-off between the number of individuals and repeatedly interacting with the same social partners for a total sample size of 800. Right figures shows the trade-off between interacting with more different social partners against interacting repeatedly with the same social partners for a total sample size of 800. Top panels represent the bias and precision of the variance parameters and the bottom panels the covariance parameters.*

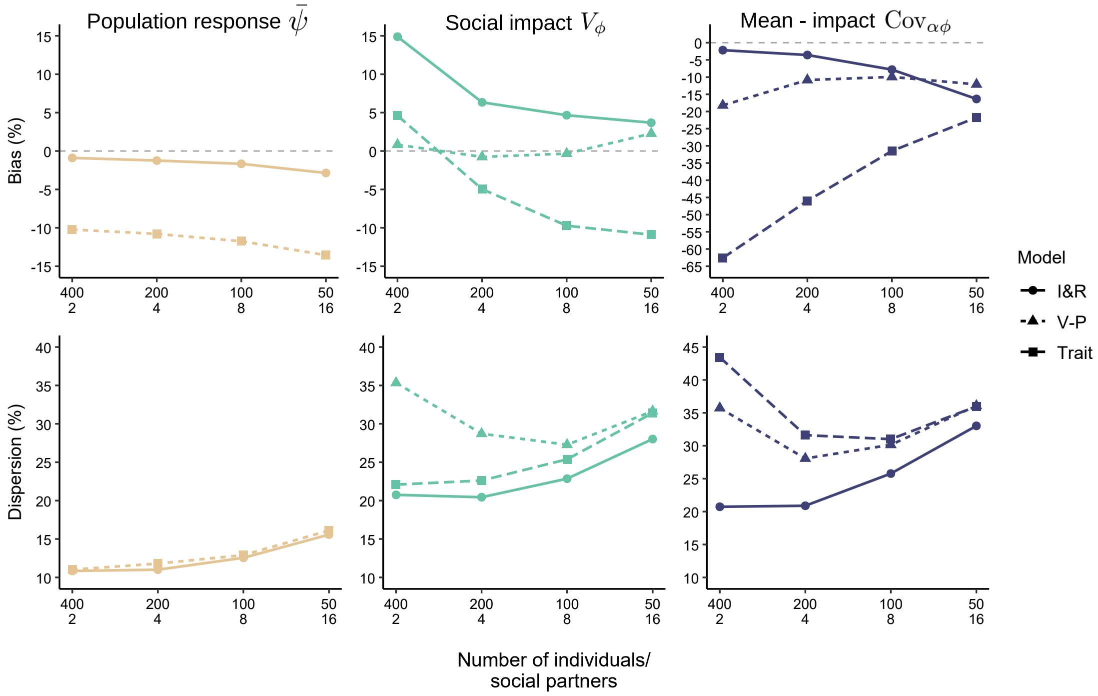


Figure S2: ***Comparison of the impact and responsiveness model (I&R) to the variance partitioning model and trait-based model for datasets with a total sample size of 800 observations.*** *Top panels show the relative bias of three models that analysed the same 1000 simulated datasets for four different partitions of number of individuals and number of social partners per individual. Bottom panel shows the relative dispersion, expressed as a percentage.*

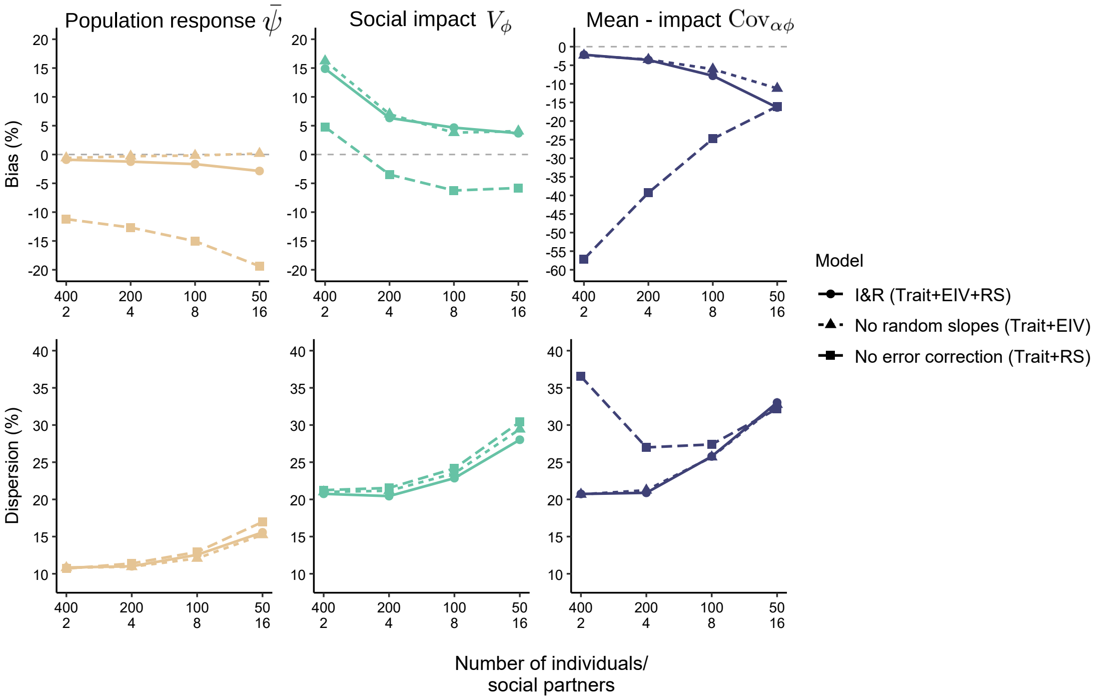


Figure S3: ***Comparison of the impact and responsiveness (I&R) model to the trait-based models without errors-in-variables correction (Trait+RS) and trait-based model without random slopes (Trait+EIV) for datasets with a total sample size of 800 observations.*** *Top panels show the relative bias of the three models after analysing the same 1000 simulated datasets for four different partitions of number of individuals and number of social partners per individual. Bottom panel shows the relative dispersion, expressed as a percentage.*