Science Advances
AAAS

Supplementary Materials for

Women's subsistence networks scaffold cultural transmission among BaYaka foragers in the Congo Basin

Haneul Jang *et al.*

Corresponding author: Haneul Jang, haneul_jang@eva.mpg.de; Daniel Redhead, daniel_redhead@eva.mpg.de

Sci. Adv. **10**, eadj2543 (2024)
DOI: 10.1126/sciadv.adj2543

This PDF file includes:

Supplementary Text
Figs. S1 to S6

Ethnographic description

H.J. and K.J. obtained informed consent from all community members (including each focal woman and their families) to allow collection of demographic data and conduct focal-follows of daily foraging trips. H.J. and K.J. conducted focal-follows of BaYaka women from one forest camp near a village along the Motaba River in the department of Likouala in the north-western region of the Republic of Congo over 236 days (March to August in 2015 and in 2016). At the beginning of the study period, H.J. and K.J. collected data on genetic relatedness, gender, and age-class of all residents during household surveys.

The mean estimated age of the five focal women was 37 years (range: $[29, 46]$). Focal women had between 2 and 6 children. During the focal-follows, H.J. and K.J. continuously recorded the foraging group composition and behaviors of the *focal* women, but not the behaviors of *all other individuals* in the group. The same focal women were followed twice for 4 consecutive weeks during the study period (between 2015 and 2016). During the study, the focal women stayed in the camp for an average of two days per month (range: 0 to 7 days per month depending on the individual). Some variation was due to illness of focal individuals or their children.

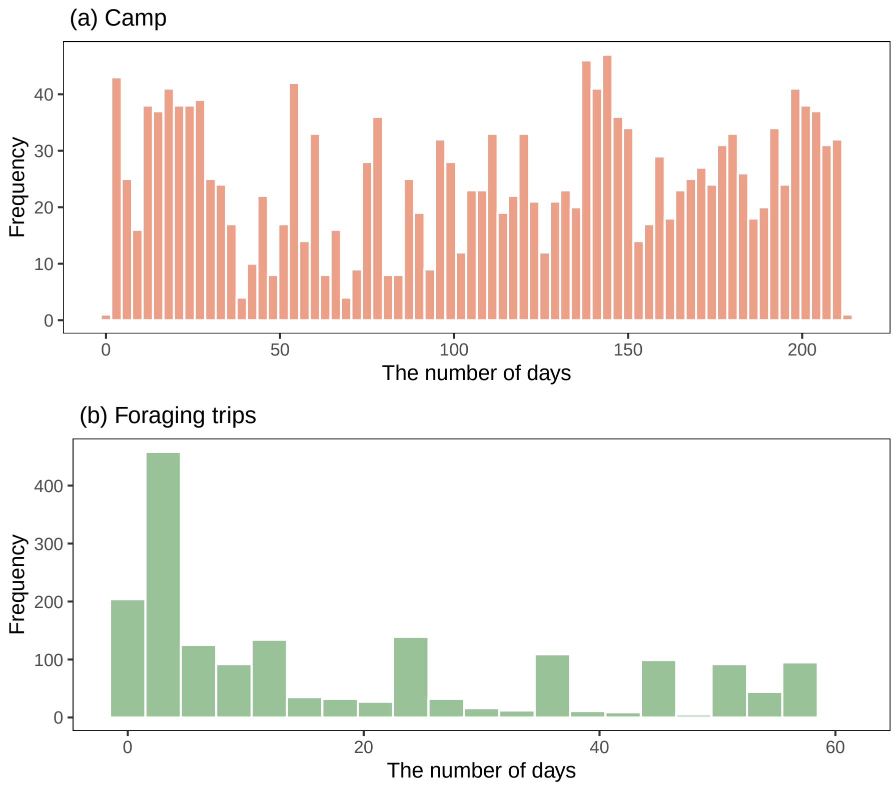
H.J. and K.J. followed the focal women's daily trips not only when they were foraging in the forest, but also when they were harvesting crops. The subsistence networks analyzed in this study used information on women's group composition during both activities. We only included the individuals who can walk in the forest independently and who can independently decide whether to join women's subsistence activities and, thus, we did not include 10 infants and toddlers (individuals 3 years of age or younger) in our analysis. In total, 60 individuals (32 females) were included in our analysis, including 10 children in early childhood (6 females), 14 children in middle childhood (5 females), 10 adolescents (6 females), and 26 adults (15 females). The average age of the 60 individuals was 31 years ($SD = 21.68$, range: $[4, 76]$ years). Across 230 days of focal-follows, the average number of days that each dyad spent within the camp was 110 days ($SD = 64.61$, range $= [1, 212]$ days; Figure S1(a)), and the average number of days that each dyad appeared together in women's subsistence activities was 18.79 days ($SD = 19.15$, range$= 0 - 59$; Figure S1(b)).

Robustness Checks and Sensitivity Analyses

When developing the two-step mask model presented in the main manuscript, we performed estimation runs on simulated data (i.e., where we know ground truth)—and results of these estimation runs suggest that our two-step model adequately recovered the ground truth values that we specified. Data and code will be maintained at: https://doi.org/10.5061/dryad.70rxwdc4t and https://github.com/haneuljangkr/bayaka-subsistence-networks

Alongside this, we performed several sensitivity analyses to examine how robust our results are to the assumptions of the two-step mask procedure that we propose as an analysis strategy. To recap, our two-step procedure adjusts for uneven sampling (i.e., focal-follows) across the study period, by only modeling outcomes for the sub-set of dyads that could have feasibly been observed conditional on the focal-follow methodology. As a robustness check, we tested a variety of other definitions of the masking tensor, Z: including no masking, an only-focal mask, a no-focal mask, and a two-step no-focal mask. In Figure S2, we illustrate how these masking layers work using a simple matrix with 21 individuals. In the true data, we have 230 such matrices of dimension 60 by 60 (one for each day of the study), but the simplified example conveys the essential information.

Figure S1: The distribution of the number of days that each dyad appeared: (a) within the camp, and (b) in women's subsistence networks across the 230-day study period.



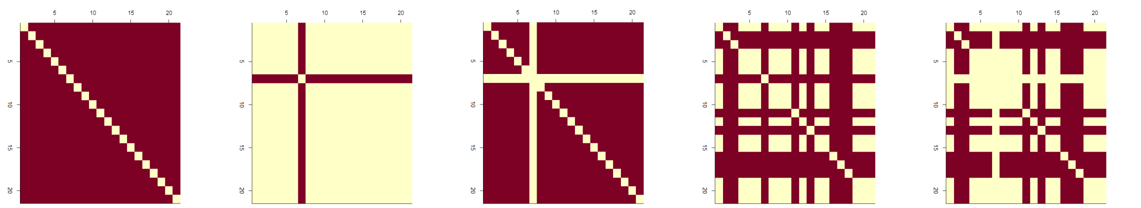
In Figure S2(a), we show the no-mask matrix. In this figure, every cell except the diagonal is colored red—indicating a value of 1 for "not-censored". This masking layer is what is generally used in social network analysis when researchers assume that the outcome matrix represents binary ties, and the network is "complete", meaning that the state of every cell in the off-diagonal of the matrix was observed and accurately measured as 0 or 1—i.e., each cell was "at risk" of being measured—and so no cells in the outcome matrix should be treated as censored.

In Figure S2(b), we show the only-focal mask matrix. In this case, individual $i = 7$ was the focal who was followed, and so only the 7^{th} column and row are red. By using this mask, we are assuming that only dyads in which the focal appears are "at risk" of being measured. This mask would be necessary if, for example, the researcher did not record ties among dyads that did not include the focal, but such ties could have occurred in the true network. However, even if the researcher did record such ties, the application of the only-focal mask is still valid, but parameter estimates from the model including the mask will only reflect patterns that hold within the sub-set of dyads that include the focal, which could be quite different from patterns that hold in other sub-sets of the data.

In Figure S2(c), we show the no-focal mask matrix. In this case, individual $i = 7$ was again the focal who was followed, and so all off-diagonal cells *except* those on the 7^{th} column and row are red. By using this mask, we are assuming that only dyads in which the focal does not appear are "at risk" of being measured accurately. This mask would be necessary if, for example, the researcher thinks that the focal-follow methodology introduced some kind of bias in which ties involving the focal are measured inaccurately, but all other ties are measured accurately. Such a case might occur

Figure S2: Examples of masking matrices, which can be used to account for various kinds of censoring as described in the text above. Each matrix features 21 fictive individuals. Individual 7 is treated as a "focal" in a focal-follow design. Individuals 2, 3, 11, 13, 16, 17, and 18 are ties of the focal. In each matrix, red cells represent observations that are "observable" under the mask, and are thus modeled. Yellow cells represent observations that are "unobservable" or "censored" under the mask, and are thus omitted from the model.

(a) No mask (b) Only-focal mask (c) No-focal mask (d) 2-step mask (e) 2-step n.f. mask



if a key informant from a study community accompanied a researcher throughout the day during a spot-check network study, and the key informant appeared to have social interactions with many more community members *by virtue of being seen with the researcher*, perhaps distorting that individual's in- or out-degree. Parameter estimates from the model including the no-focal mask will only reflect patterns that hold within the sub-set of dyads that do not include the focal, which might be more representative of the true network, if the focal-following methodology distorts the measurement of ties involving the focal.

In Figure S2(d), we show the two-step mask matrix. In this case, individual $i = 7$ was again the focal (first step) who was followed, and who encountered individuals: 2, 3, 11, 13, 16, 17, and 18 (second step). In this case, we define the masking layer so that all off-diagonal cells including either the focal or the focal's direct ties are set to 1 (i.e., colored red). This masking layer formalizes the idea that the only dyadic outcomes that can be accurately classified as ties or non-ties are outcomes which feature the focal, or somebody connected to the focal, as members. Concretely, if the researcher observes individual 2 and not individual 7 during the focal-follow, then the researcher can be sure that individuals 2 and 7 were not co-foraging. However, if the researcher observes neither individual 1 nor individual 21 during the focal-follow, then the researcher cannot exclude the possibility that individuals 1 and 21 are co-foraging, but just not in the presence of the target of the focal-follow (individual 7). As such, the application of the two-step mask ensures that "absence of evidence" is not treated as "evidence of absence" unless it is justified on the basis of the study design.

In Figure S2(e), we show the two-step no-focal mask matrix. This mask is identical to the basic two-step mask matrix from Figure S2(d), except that all cells on the 7^{th} column and row are set to zero as we saw with the no-focal mask matrix in Figure S2(c). This masking matrix would be of use in similar circumstances to the basic two-step mask matrix, but when the researcher thinks that the focal-follow methodology introduced some kind of bias in which ties involving the focal are measured inaccurately, but all other (non-censored) ties are measured accurately.

Refitting the model with different masks

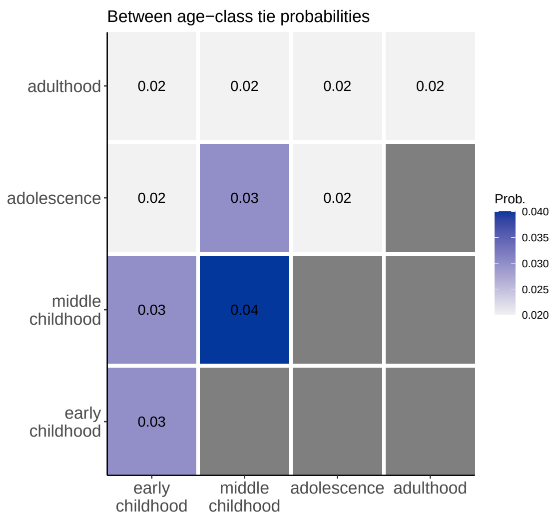
1. No masking. We estimated a base model that did not include any masking. That is, the model assumes that all dyads each day were observable and, thus, considers the (unrealistic) scenario where our focal-following design has no impact on the structure of the observed network. The results of this analysis are nevertheless qualitatively similar to results from our two-step model in the main manuscript (see Figure S3), with the qualification that the overall density of ties is estimated to be much smaller (on par with the density of recorded ties—3.6%—noted in the main text). The true density of ties, however, is much larger than 3.6%, as the point estimate of $15,008/(0.5*230*60*59) = 0.036$ assumes that no foraging ties occurred outside of the focal woman's expedition each day.

2. Only-focal mask. We also estimated a model using the only-focal mask. Here, our model is only assessing the likelihood of ties within the subset of dyads that focal women appear in. Thus, the results from this analysis could be viewed as analogous to an ego-network analysis, which examines who focal women choose to conduct their subsistence activities with. In this analysis, we find little evidence of statistically reliable preferences in relation to age-class (see Figure S4)—which may largely be attributed to the much smaller sample size of dyads that we are considering. We, however, still find that ties are primarily to kin and between females. Coresidence appears to be a stronger predictor of co-foraging in this model than in the model presented in the main text.

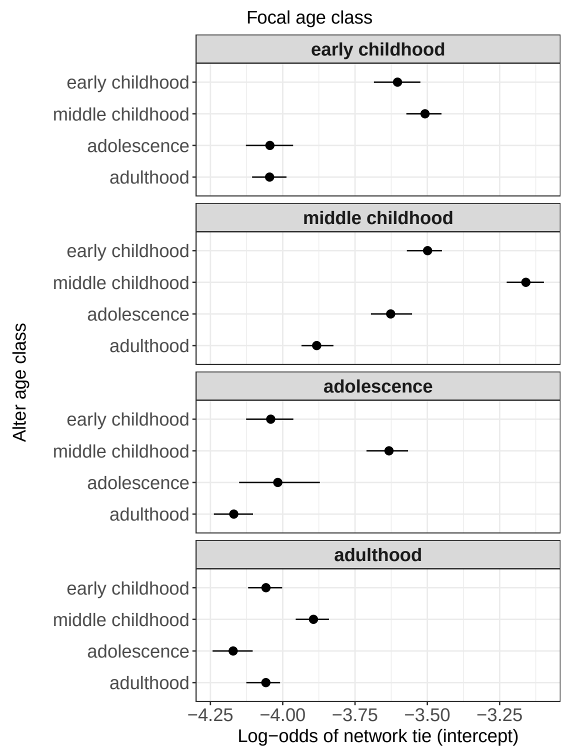
3. No-focal mask. We estimated another model that included all dyads within the sample, *other than the dyads which included focal women*, and does not include the two-step procedure that we outline in the main manuscript. As shown in Figure S5, the results of this analysis are qualitatively similar to results from our two-step model and from our no-mask model.

4. A two-step no-focal mask. For our final sensitivity analysis, we estimated a model that followed the two-step mask procedure, but also masked dyads in which the focal women appeared. As shown in Figure S6, results from this analysis are qualitatively similar to the results from our main model that uses a simple two-step mask procedure (as presented in the main manuscript).

(a) Heat map of predicted daily between age-class tie probabilities for non-coresident, unrelated, female-female dyads.



(b) Log-odds offsets (with 90% credible intervals) for between age-class ties, corresponding to the point estimates in frame (a).



(c) Dyadic parameter estimates and 90% credible intervals. The base-case of the model is for an unrelated, non-coresident, male-male dyad. As such, these estimates are indicative of the change in log-odds of a tie as a function of a unit change in the indicated predictors. Note that the female-female dyad, different sex dyad, and coresidence variables are binary indicators.

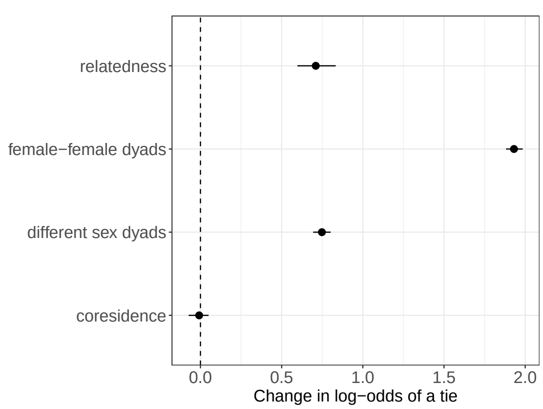
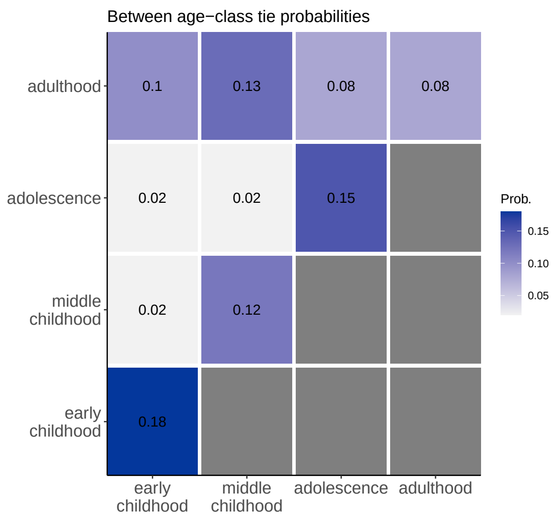
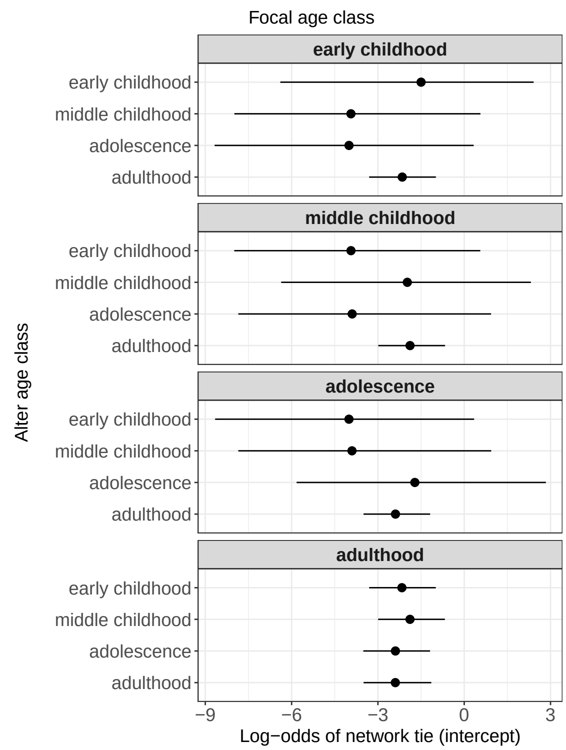


Figure S3: Posterior estimates from the **no-mask model**.

(a) Heat map of predicted daily between age-class tie probabilities for non-coresident, unrelated, female-female dyads.



(b) Log-odds offsets (with 90% credible intervals) for between age-class ties, corresponding to the point estimates in frame (a).



(c) Dyadic parameter estimates and 90% credible intervals. The base-case of the model is for an unrelated, non-coresident, male-male dyad. As such, these estimates are indicative of the change in log-odds of a tie as a function of a unit change in the indicated predictors. Note that the female-female dyad, different sex dyad, and coresidence variables are binary indicators.

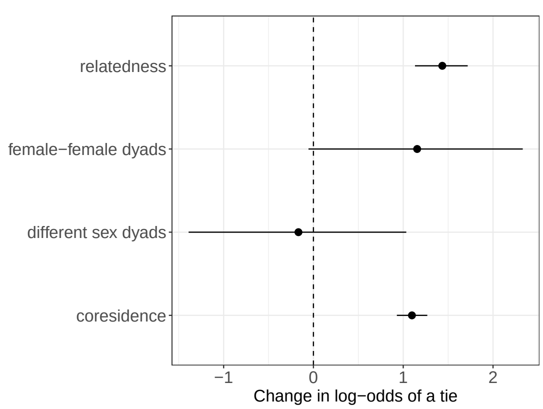
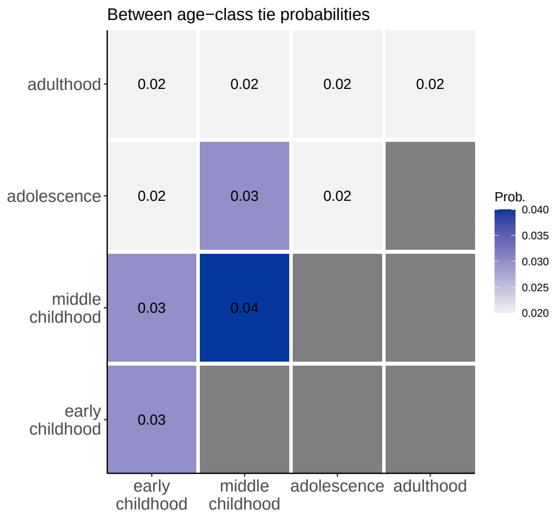
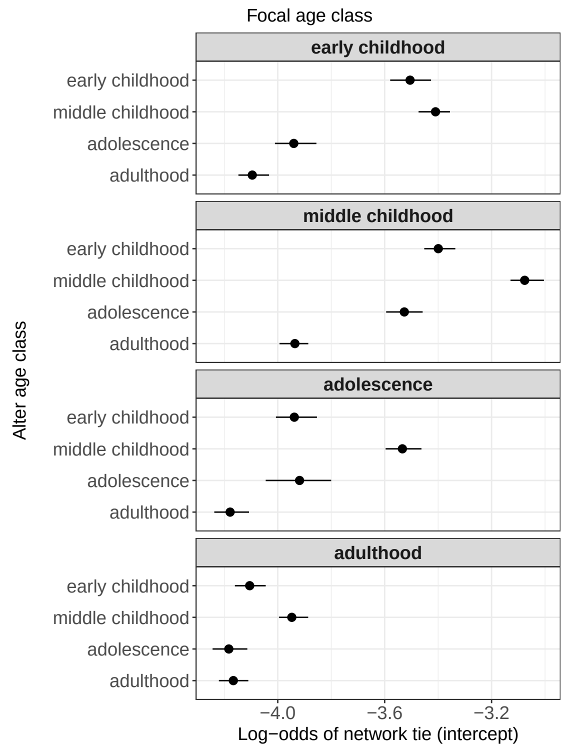


Figure S4: Posterior estimates from the **only-focal model**.

(a) Heat map of predicted daily between age-class tie probabilities for non-coresident, unrelated, female-female dyads.



(b) Log-odds offsets (with 90% credible intervals) for between age-class ties, corresponding to the point estimates in frame (a).



(c) Dyadic parameter estimates and 90% credible intervals. The base-case of the model is for an unrelated, non-coresident, male-male dyad. As such, these estimates are indicative of the change in log-odds of a tie as a function of a unit change in the indicated predictors. Note that the female-female dyad, different sex dyad, and coresidence variables are binary indicators.

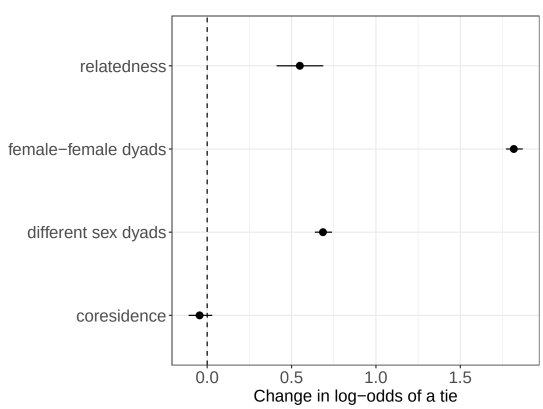
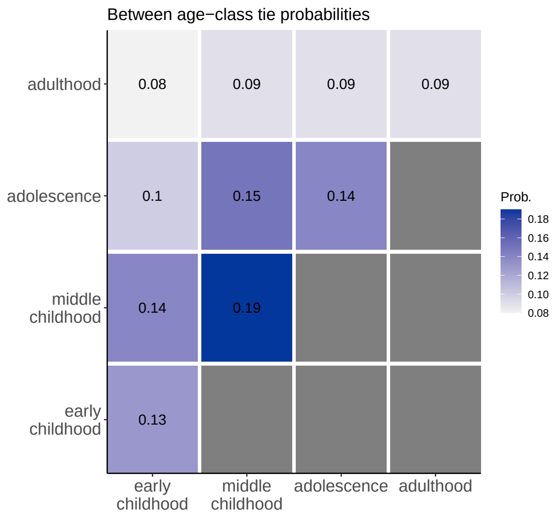
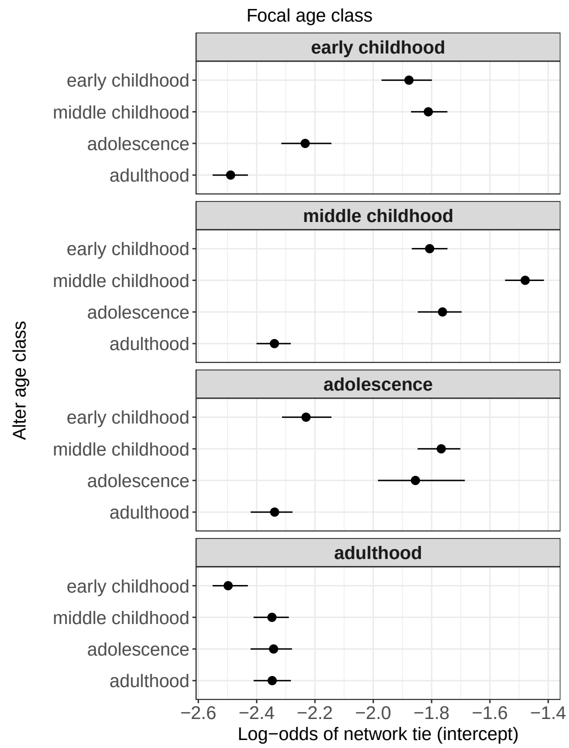


Figure S5: Posterior estimates from the **no-focal model**.

(a) Heat map of predicted daily between age-class tie probabilities for non-coresident, unrelated, female-female dyads.



(b) Log-odds offsets (with 90% credible intervals) for between age-class ties, corresponding to the point estimates in frame (a).



(c) Dyadic parameter estimates and 90% credible intervals. The base-case of the model is for an unrelated, non-coresident, male-male dyad. As such, these estimates are indicative of the change in log-odds of a tie as a function of a unit change in the indicated predictors. Note that the female-female dyad, different sex dyad, and coresidence variables are binary indicators.

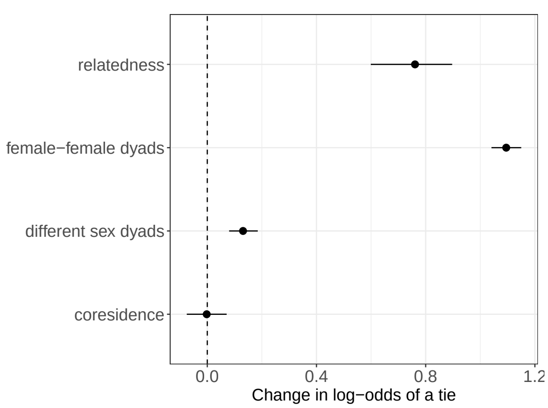


Figure S6: Posterior estimates from the **two-step no-focal model**.