Supporting Information

Zerubavel et al. 10.1073/pnas.1802176115

Participants

Participants were 16 healthy young adults (13 females; 2 males; 1 transgender) with mean age of 21.2 (SD = 1.2). They were recruited from a summer program in which 20 undergraduate students and recent graduates (from 10 US colleges) agreed to spend 9 wk together to organize workers and collect oral histories. There were initially 20 volunteers enrolled in this summer program (19 completed it), of which 18 were willing and able to participate in the T1 fMRI scan session. We were unable to collect T1 liking ratings from two of these individuals and consequently could not compute social relations modeling (SRM) relationship effects for any T1 liking relations involving either of them (26). Thus, the present study analyzes data collected from the remaining 16 group members—those who were scanned at T1 and provided liking ratings both at T1 and at T2—and the 120 dyadic relationships (i.e., 240 interpersonal sentiments) between them. Participants were, for the most part, unacquainted before this summer program; in the vast majority (82%) of dyads, both individuals reported not knowing each other at T1 (see *Robustness Checks* below for analyses conducted to verify that such variables did not confound our results).

All participants provided informed consent, were English-speaking, and had normal or corrected-to-normal vision. They were screened for a history of serious neuropsychiatric disorders, head injury, and other conditions that prevented scanning (e.g., a pacemaker) before taking part in the fMRI scanning session.

Preprocessing/General Linear Model Parameters

Functional data were preprocessed with SPM8 software (Wellcome Department of Cognitive Neurology, University College London), including slice time correction, motion correction, realignment, coregistration of functional and anatomical data, normalization to a standard template (Montreal Neurological Institute) using segmentation parameters, 3-mm^3 isometric voxels, and spatial smoothing with a 6-mm Gaussian kernel. fMRI data were subjected to a first level of regression, separately for each subject, using an ordinary–least-squares general linear model (GLM) implemented with Neuroelf, version 1.0, software (neuroelf.net). The GLM included one regressor per face stimulus (i.e., representing the 10 repetitions of each), created by convolving the canonical hemodynamic response function with a series of boxcars representing the 1,000-ms intervals during which a particular face was presented. In addition, the GLM included six motion parameters as estimated during realignment as well as a discrete cosine transform-based basis set covering low frequency up to 1/80 Hz to account for signal variability introduced by head motion and temporal drifts. The output of these first-level regressions was a series of parameter estimate (beta) maps for subsequent analysis.

Isolating Relational Components of Liking and Neural Valuation Using SRM Analysis

Our round-robin design—which structured both the liking assessments and fMRI face-viewing paradigm—allowed us to partition sources of variance underlying group members' liking ratings and neural valuations, specifically, to disambiguate relationship-specific effects from generalized individual-level effects. Consider the variability in group members' liking ratings: some of this variance is attributable to individual differences, that is, group members varying in how much they (*i*) generally like others, and (*ii*) are generally liked by the rest of the group; however, the predominant source of liking variance is relationship specific or relational, that is, attributable to group members having unique attractions toward one another (1). SRM isolates this relational component of interpersonal attraction from person-level confounds, that is, distilling how much Anita uniquely likes Buddy by taking into account Anita's general tendency to like others as well as Buddy's general tendency to be liked by others.

Capturing the relationship-specific component of liking is particularly important for understanding affective reciprocity as a truly dyadic phenomenon of relationships: in the context of raw liking measures, reciprocity could be driven by generalized person-level effects (i.e., correlation between Anita's popularity—generally being liked by group members—and her overall tendency to like other group members); by contrast, in the context of relational liking measures, the SRM construct of dyadic reciprocity reflects correlation between Anita's particular attraction to Buddy and Buddy's particular attraction to Anita (1, 24, 26).

We used the R package TripleR (26) to compute SRM relationship effects (1) for neural reward value, T1 liking, and T2 liking. The relationship effects for the neural variable were computed using our aggregate measure of reward system ROI activity, that is, the simple average of the two ROIs' activations. Model results for ROI-specific analyses (i.e., using relationship effects computed using only the vmPFC ROI or only the VS ROI, respectively) are reported in Table S1.

This allowed us to model T2 liking outcomes as a function of T1 neural reward responses and initial liking using these variables' uniquely relational components. We conducted this analysis using the actor–partner interdependence model (APIM), an established method for dyadic analysis of SRM relationship effects (24). Note that we also conducted a version of the APIM analysis using the raw measures of liking and neural valuation rather than their relational components; as described in *Results*, these analyses replicated our neural findings that T1 reward activity predicted both liking and being liked at T2.

Implementation of APIM Analysis Using SEM

Our analyses aimed to test whether group members' T2 liking outcomes were predicted by their T1 neural valuations, even controlling for their baseline attractions at T1. The fact that these T2 liking observations evidence reciprocation—statistical association between dyad members' outcomes—violates standard regression models' assumption of independence. This violation would lead us to underestimate SEs and perform overly lenient significance testing. Therefore, we adopted the APIM analytic framework to assess—rather than ignore—such dyadic interdependence of T2 liking outcomes and quantify how much of this reciprocation is explained by the model.

We implemented these analyses using the Stata 14 (StataCorp, 1985–2015) *sem* function for structural equation models and the *mlmv* estimation method for maximum likelihood with missing values. This approach allowed for missing data and offered several other advantages. First, by estimating the entire model at once, we could simultaneously estimate path coefficients for actor and partner effects (while controlling for each other and any other covariates) as well as correlations between predictor variables (24, 25). In addition, these structural equation analyses allowed us to impose specific restrictions on model parameters such that dyad members would share common (i.e., exhibit equal) actor effects, partner effects, predictor means, predictor variances, outcome intercepts, and residual variances (Fig. S1). These equality constraints were necessary since dyad members in this study were fundamentally interchangeable, meaning that they could both equivalently fulfill either the actor or partner roles (in contrast to distinguishable dyads such as husband–wife or child–parent role pairings). Finally, this analytic approach enabled us to incorporate clustered SEs, that is, to specify the SE calculation using dyad as a cluster variable. This method allows error terms to be correlated within dyad clusters and accordingly adjusts for this dyadic nonindependence of T2 liking outcomes in computing SEs and significance tests (25).

More formally, the APIM with interchangeable dyad members can be depicted schematically as in Fig. S1 or expressed as a multivariate regression:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} i \\ i \end{bmatrix} + \begin{bmatrix} a & p \\ p & a \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix},$$

where a is actor effect, p is partner effect, i is outcome intercept, x_1 is person 1's predictor, x_2 is person 2's predictor, y_1 is person 1's outcome, y_2 is person 2's outcome, ε_1 is person 1's outcome disturbance, and ε_2 is person 2's outcome disturbance.

For further details on the APIM and its estimation using SEM, see refs. 24 and 25, respectively.

Missing Data

For the T2 liking outcome measure, our data included 235 observations or 98% of the 240 possible directed relations between 16 participants. The five missing values were due to one participant not fully completing the T2 sociometric assessment administered via an online survey. Turning to our T1 predictor variables, we acquired neural measures of reward value for all 240 directed relations and initial liking ratings for 197 (82%) of them. These missing values were due to a programming error in the T1 sociometric assessment instrument that resulted in 43 randomly drawn omissions from the full roster of 240 possible liking relations.

Our analytic approach was designed to handle these missing data, both in the computation of relationship-specific effects and the APIM implementation using SEM. TripleR temporarily imputes missing values (by averaging row and column means) to calculate SRM effects; subsequently, imputed relationship effects are turned back to missing values. Simulation studies have demonstrated that, for groups with 10 or more members, relatively little deviation from true values can be expected even with 20% or more missing values (26). In terms of the structural equation analyses, we used the *sem* function in Stata 14 and specified the *mlmv* estimation method for maximum likelihood with missing values.

Robustness Checks

We conducted a series of robustness checks to verify that our neural findings were not driven by other mechanisms, in particular, structural antecedents of liking and affiliation ties identified by sociologists. Considering the mechanism of homophily (i.e., similarity breeds interpersonal attraction), our robustness checks included measures of similarity for each of the following variables: gender, race, ethnicity, household income, college affiliation, having already graduated from (versus still being enrolled in) college, dispositional narcissism, each of the Big Five personality dimensions (extraversion, agreeableness, neuroticism, conscientiousness, and openness to experience), and belonging to the same "team" of participants (i.e., the summer program organized its volunteers into five teams of four individuals each). We also tested an additional variable that designated whether (at least one of) the dyad members reported knowing each other before the summer program (22 dyads = 18%). Each iteration of the APIM analysis included one of these potential confounds as a covariate along with the same four predictor variables described before (i.e., actor's and partner's T1 liking ratings and neural responses). As none of these sociological variables demonstrated statistically significant results in our models (all values of $P > 0.2$), they were not included in other analyses.

The variables incorporated in these robustness checks concerned potential confounds at the dyadic level. There was no need for the robustness checks to include individual-level covariates because our analysis solely modeled relationship-specific effects (i.e., measures from which person-level effects of actors and of partners had already been removed).

Analyses by Individual ROIs

As explained in *Methods*, we used a functional localizer task in a separate participant sample to independently define neural reward system ROIs in vmPFC and VS. We averaged together both ROIs' activations during the face-viewing task for a composite neural measure of reward value. The rationale for aggregating vmPFC and VS ROIs is that these brain regions are anatomically interconnected (13) and functionally coactivated in processes underlying valuation and reward (9–15); furthermore, these ROIs were identified together from the same localizer task.

Table S1 reports model results of three separate APIM analyses that incorporated (*i*) the aggregate neural measure of reward value averaged across ROIs, (*ii*) only the vmPFC ROI, or (*iii*) only the VS ROI. As with the aggregate reward system measure, the vmPFC ROI demonstrated both the actor effect ($\beta = 0.151$; $P < 0.05$) and the partner effect ($\beta = 0.155$; $P < 0.05$). Although the VS ROI similarly exhibited the partner effect ($\beta = 0.134$; $P < 0.05$), we did not find evidence for the corresponding actor effect there ($\beta = 0.041$; $P > 0.4$). As we had no specific predictions about these prognostic effects manifesting more strongly or exclusively in either one of the reward-related ROIs, we hesitate to offer post hoc interpretation of any such differences. The observed pattern of results suggests that the neural actor effect may be primarily driven by vmPFC, but further research will be needed to clarify the relative predictive strengths of vmPFC and VS activity as a neural basis for forecasting different kinds of interpersonal sentiments and social relations.

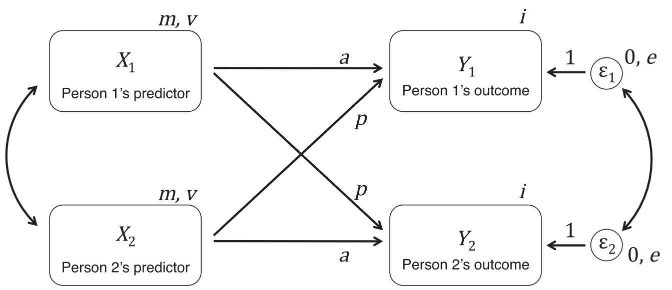

Fig. S1. Schematic representation of actor–partner interdependence model (APIM) for indistinguishable dyads with the following parameters: actor and partner effects (*a* and *p*), predictor mean and variance (*m* and *v*), outcome intercept (*i*), and residual variance (*e*). See refs. 24 and 25 for additional details on the APIM and its estimation using SEM.

Table S1. Results of three actor–partner interdependence models (APIMs) of T2 liking

T1 predictors of T2 liking	Reward system ROIs (mean of vmPFC and VS)			vmPFC ROI only			VS ROI only		
	β	Robust SE	*P*	β	Robust SE	*P*	β	Robust SE	*P*
Actor effects									
Neural activation	0.119	0.059	**0.044**	0.151	0.063	**0.016**	0.041	0.053	0.437
Initial liking	0.199	0.066	**0.003**	0.208	0.065	**0.001**	0.206	0.068	**0.003**
Partner effects									
Neural activation	0.167	0.059	**0.005**	0.155	0.061	**0.012**	0.134	0.058	**0.020**
Initial liking	0.159	0.059	**0.007**	0.162	0.057	**0.004**	0.172	0.061	**0.005**
(Intercept)	0.004	0.073	0.959	0.004	0.073	0.957	0.003	0.075	0.963

Results of three APIMs regressing T2 liking outcomes against T1 predictors—neural reward responses during the face-viewing task and a control variable for initial liking baselines—collected from oneself (actor) and the relevant other group member (partner). In other words, for baseline liking this means Anita's initial liking of Buddy and Buddy's initial liking of Anita. For the neural measure of reward value, this means Anita's neural valuation of Buddy and Buddy's neural valuation of Anita. Specifically, this predictor reflects neural activation parameter estimates extracted from a priori brain regions of interest (ROIs) underlying reward and valuation processes: the first model uses an aggregate measure of reward system activity (simple average of both ROIs' activations); the second model uses only the vmPFC ROI; and the third model uses only the VS ROI. All parameters are reported as standardized values. Bolded values denote $P < 0.05$, two-tailed.