Neural detection of socially valued community members

Sylvia A. Morelli[a,1,2], Yuan Chang Leong[b,1], Ryan W. Carlson[b], Monica Kullar[b], and Jamil Zaki[b]

[a]Department of Psychology, University of Illinois at Chicago, Chicago, IL 60607; and [b]Department of Psychology, Stanford University, Stanford, CA 94305

Edited by Tor D. Wager, University of Colorado, Boulder, CO, and accepted by Editorial Board Member Michael S. Gazzaniga June 21, 2018 (received for review July 21, 2017)

As people form social groups, they benefit from being able to detect socially valuable community members—individuals who act prosocially, support others, and form strong relationships. Multidisciplinary evidence demonstrates that people indeed track others' social value, but the mechanisms through which such detection occurs remain unclear. Here, we combine social network and neuroimaging analyses to examine this process. We mapped social networks in two freshman dormitories ($n = 97$), identifying how often individuals were nominated as socially valuable (i.e., sources of friendship, empathy, and support) by their peers. Next, we scanned a subset of dorm members ("perceivers"; $n = 50$) as they passively viewed photos of their dormmates ("targets"). Perceiver brain activity in regions associated with mentalizing and value computation differentiated between highly valued targets and other community members but did not differentiate between targets with middle versus low levels of social value. Cross-validation analysis revealed that brain activity from novel perceivers could be used to accurately predict whether targets viewed by those perceivers were high in social value or not. These results held even after controlling for perceivers' own ratings of closeness to targets, and even though perceivers were not directed to focus on targets' social value. Overall, these findings demonstrate that individuals spontaneously monitor people identified as sources of strong connection in the broader community.

social networks | social value | fMRI | prediction | mentalizing

Communities allow people to cooperate and support each other, bolstering their collective and individual well-being. One key way in which groups maximize collective benefit is by rewarding prosocial individuals—for instance, through direct and indirect reciprocity. Such "social selection" is likely a crucial driver for the evolution of prosocial behaviors, and similar processes promote and sustain prosociality in economic games (1–6). Social selection, in turn, requires group members to first detect socially valuable peers, who are generous, trustworthy, and supportive to the community at large. In addition to benefiting the group, such detection can also aid individuals. Socially valuable others provide high-quality support and minimize others' stress (7–10). They also tend to be connected to other community members, and close relationships with them can offer a gateway to additional social resources (11, 12).

Despite the importance of social value detection, it remains unclear how capably people detect socially valuable others in dynamic, real-world communities, or the mechanisms through which this detection takes place. Here, we use a combination of social network analysis and neuroimaging to explore the possibility that—even absent explicit instructions to do so—individuals track their peers' social value. We focused on individuals undergoing the transition to college. During this period, individuals are separated from their previous social networks (i.e., family and high school friends) while facing increasing academic demands (13, 14). They also rapidly build new communities, and students who quickly form close relationships on campus exhibit improved adjustment during the first year of college and beyond (15–18).

We recruited newly matriculated college students from two freshman-only dormitories at Stanford University ($n = 97$) (12). In the second week of the academic year, we asked participants to nominate dorm members in response to eight prompts: for instance, identifying dorm members they viewed as socially supportive, positive, and empathic. We then identified "hubs" in each dorm: individuals who received unusually high numbers of nominations. In a second phase of the study, we scanned a subset of 50 students ("perceivers") using fMRI while they viewed photos of their fellow dorm members ("targets").

During this task, participants were not instructed to evaluate targets in any way. Nonetheless, we predicted that targets' social value would be reflected in perceivers' brain activity. Past work suggests that when individuals encounter popular individuals from their networks, they engage brain regions, including medial prefrontal cortex, temporoparietal junction, and ventral striatum (19, 20). These regions are broadly associated with mentalizing—considering the internal states of other people—and with value computation. These regions are also preferentially engaged by salient social targets, such as ingroup members, suggesting that popular individuals likewise take on motivational relevance in social networks.

Here, we build on that work in several ways. First, we controlled for perceivers' own relationship to each target when

Significance

To form successful communities, people must be able to detect socially valued individuals: people who are generous, supportive, and well-connected. Here, we provide evidence that people accomplish this detection by monitoring how the broader community views individuals. We used social network analysis to identify highly socially valued individuals in two college dormitories. We then scanned dorm residents using fMRI as they passively viewed pictures of dormmates. Activity in brain systems related to mentalizing and reward increased when people viewed highly valued, versus less valued, dormmates—even when controlling for individuals' own impressions of their dormmates. These data suggest that people robustly monitor peers' social value, potentially allowing them to efficiently locate high-quality social ties.

Author contributions: S.A.M., R.W.C., M.K., and J.Z. designed research; S.A.M., R.W.C., and M.K. performed research; S.A.M., Y.C.L., R.W.C., and M.K. analyzed data; and S.A.M., Y.C.L., and J.Z. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. T.D.W. is a guest editor invited by the Editorial Board.

Data deposition: All data and code are located in a GitHub Repository: https://github.com/esclabUIC/NetworkFMRI. All activation maps for neuroimaging results are stored in NeuroVault: https://neurovault.org/collections/2715/.

[1]S.A.M. and Y.C.L. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: smorelli@uic.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1712811115/-/DCSupplemental.

Published online July 23, 2018.

isolating brain activity related to targets' social value. This allowed us to test the prediction that individuals identify hub individuals using community-level information, above and beyond their own idiosyncratic experience with targets. Second, we tested whether targets' social value could be predicted based on perceivers' brain activity. We first isolated brain regions—at the group level—in which activity tracked the number of nominations targets received (hereafter: targets' "hub index"). We then ran a series of leave-one-out cross-validation prediction models, using activity in mentalizing- and reward-related brain structures to differentiate between targets with high, medium, and low hub index. In the first model, we used overall activity in these regions to predict targets' hub category. In the second, we used multivariate patterns within these regions to make similar predictions. In both models, we iteratively trained an algorithm to differentiate between hub categories using data from 49 perceivers. We then tested whether we could use brain activity in the fiftieth perceiver to predict targets' hub category. Critically, the accuracy of each model provided an objective measure of how closely dorm members were tracking social value hubs in their community.

Finally, this approach allowed us to test different predictions about the nature of social value detection. One possibility is that group members retain a linear internal model of social value: drawing similar distinctions between low, medium, and high value peers. Another potentially more efficient strategy would be for individuals to specifically monitor peers of high social value but not differentiate between those with medium and low social value. Some evidence from social network science suggests that people indeed track unusually popular group members (21), but this type of monitoring has never been examined at the neural level. By marrying levels of analysis, we generated a test of the mechanisms underlying value detection in new communities.

Results

Quantifying Social Value in Dorm Communities. A factor analysis revealed that nominations in response to our eight prompts cohered into a single factor. Therefore, we computed a hub index to represent each dorm member's score on this factor, simultaneously capturing the number of unique connections for each individual and the strength of these connections. More specifically, each person's hub index was calculated by multiplying the factor loading for each prompt with the total number of people who nominated that individual when given that prompt (*SI Appendix*, Table S1) and then averaging across all prompts (i.e., weighted average). Hub index was positively skewed (dorm A skewness = 0.71; dorm B skewness = 0.91), with long tails on the right side (Fig. 1, *Top*), suggesting that a handful of individuals emerge as hubs in the dorm community.

Brain Activity Tracking Social Value. Our neuroimaging analysis focused on the hypothesis that perceivers monitor how targets are perceived by the broader dorm community, irrespective of how perceivers themselves think about or interact with targets. To capture these community-level perceptions, we counted the number of ties directed to each target from the entire dorm, for each of the eight questions. However, we did not include any direct ties from the perceiver to the target in these totals. We then computed a weighted average of these community nominations.

Despite the exclusion of the perceiver's nominations, it is possible that brain activity tracking group nominations could still reflect shared variance with a perceiver's idiosyncratic evaluations. In our sample, correlations between individual and group nominations across targets were low (average within-subject r = 0.13). Nonetheless, we controlled for the perceiver's nominations of targets and their self-reported closeness to targets in all analyses. This further ensured that resulting brain activity reflected community-wide perceptions of targets' status as a

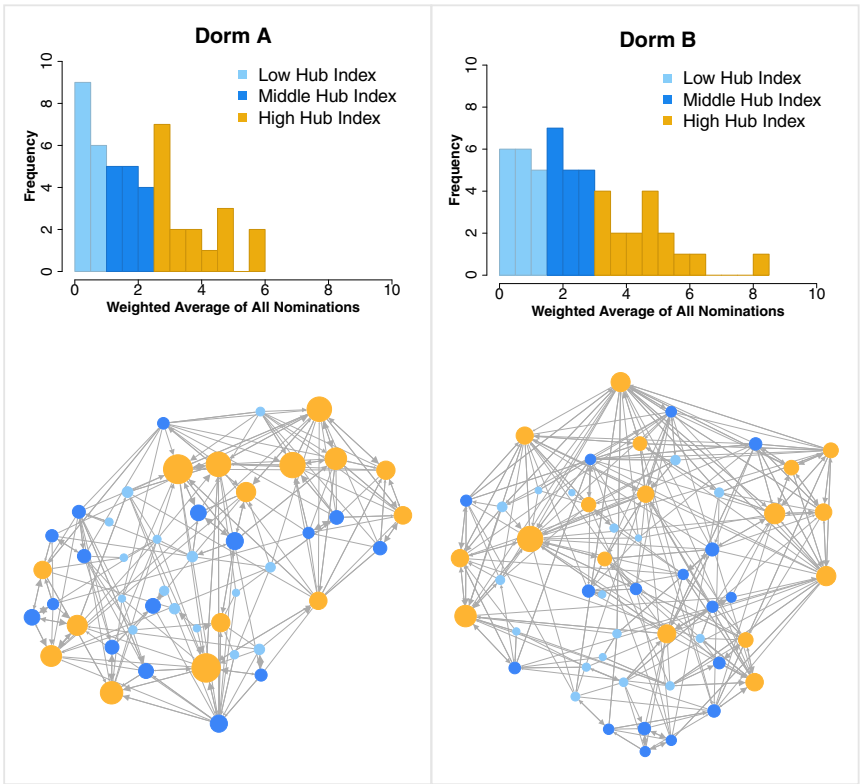

Fig. 1. (*Top*) The distribution of hub index in each dorm is divided into low (light blue), middle (dark blue), and high (orange) hub categories. (*Bottom*) Graphs of the social network in each dorm. Larger nodes indicate higher numbers of nominations received from the dorm. The darkest, thickest arrows indicate that an individual was nominated for all eight prompts whereas the lightest, smallest arrows suggest that an individual was only nominated for one prompt.

social value hub, irrespective of how the perceiver in the scanner viewed targets.

We conducted a whole-brain parametric analysis to isolate brain regions in which activity increased as a function of targets' hub index. This analysis revealed activity in regions related to (*i*) mentalizing (22), including medial prefrontal cortex (MPFC), temporal poles, and precuneus (*SI Appendix*, Table S2), and (*ii*) value processing (23), including left striatum.

Neural Prediction of Social Value. The previous group-level analyses suggest that perceivers passively tracked their dorm members' hub index as assessed by the broader community, in regions associated with mentalizing and value computation. Our second analytic approach focused on using brain activity from a given perceiver to predict whether or not targets they encountered were viewed by the community as social connection hubs. To do so, we divided each dorm into three groups: individuals with low, middle, and high hub index (hereafter, "hub category") (Fig. 1). Dividing dorms into terciles afforded sufficient power for subsequent prediction analyses, while also allowing the high hub category to capture individuals on the right tail of the distribution who were viewed as having unusually high levels of social value.

As expected, individuals in the high hub category maintained strong ties and were connected to several individuals in the community. If these hub individuals were nominated by a member of their community, they were typically nominated for 3.8 out of the 8 questions (SD = 0.65). Across all eight questions, they were also nominated by the highest number of unique individuals (*SI Appendix*, Fig. S5). In dorm A, individuals in the high hub category had more unique ties [mean (M) = 9.33, SD = 2.06] than people in the middle [M = 5.06, SD = 1.61; $t(29) = 6.46$, $P < 0.001$] and low hub categories [M = 1.53, SD = 1.55; $t(28) = 11.72$, $P < 0.001$]. In dorm B, individuals in the high hub category (M = 11.76, SD = 3.67) had more unique ties than people in the middle [M = 5.65, SD = 2.15; $t(32) = 5.94$, $P < 0.001$] and low hub categories [M = 2.76, SD = 1.52; $t(32) = 9.35$, $P < 0.001$].

We then averaged perceivers' brain activity in response to targets in each hub category and applied two leave-one-out cross-validation models to the resulting data. In each, we iteratively trained an algorithm to differentiate between targets in low, middle, and high hub categories using activation data from each subset of 49 perceivers. We then tested whether we could accurately predict the hub category of targets viewed by the fiftieth perceiver, based on that perceiver's brain activity alone.

To assess the accuracy of these models, we used two different metrics: (*i*) average within-subject correlation and (*ii*) forced-choice classification accuracy. First, we tested if targets' predicted hub category (i.e., low, middle, or high hub index) linearly increased with the targets' actual hub category in the held-out participant. Second, we examined forced-choice classification accuracy by testing the model on every pairwise comparison of hub category (i.e., low vs. middle, middle vs. high, low vs. high hub category) and assessed the model's accuracy in predicting which data came from a higher hub category.

Univariate prediction. We first assessed our ability to predict targets' hub category from the mean activity of the entire mentalizing network. Using the online metaanalytical tool Neurosynth, we identified a set of brain regions related to mentalizing. This map included MPFC, posterior medial cortex (PMC), temporoparietal junction (TPJ), and temporal poles (TP) (*SI Appendix*, Fig. S7). We trained a linear regression model to predict targets' hub category from average activity across this network in each subset of 49 perceivers and used regression weights drawn from this training to predict targets' hub category in the held-out participant. On average, the within-subject correlation between the predicted and actual levels of target hub category was moderate (*SI Appendix*, Table S3) (mean r = 0.337, SE = 0.100, P = 0.002). (Significance was assessed using a nonparametric permutation test. We repeated the analysis 1,000 times with randomly shuffled labels for low, middle, and high hub categories within each participant, generating a null distribution of r values. P values were computed by comparing the r value obtained from the unshuffled data with the corresponding null distribution.)

We also performed a forced-choice classification, testing whether the predicted hub category of targets matched their actual categories in the held-out perceiver. To make this prediction, the model used average activity in the mentalizing network for different target categories (i.e., low vs. middle, middle vs. high, and low vs. high hub category) in the other 49 perceivers. This analysis correctly distinguished between high and middle hub categories in 72% of held-out participants (36 out of 50, $P < 0.05$, binomial test) and between high and low hub categories in 70% of held-out participants (35 out of 50, $P < 0.05$, binomial test) but did not distinguish between middle and low hub categories (24 out of 50, $P = 0.887$, binomial test) (Fig. 2). These results held when we repeated prediction analysis using data from single regions within the mentalizing network (Table 1 and *SI Appendix*, Fig. S7). Taken together, our results demonstrate that activity across the mentalizing network indeed tracks

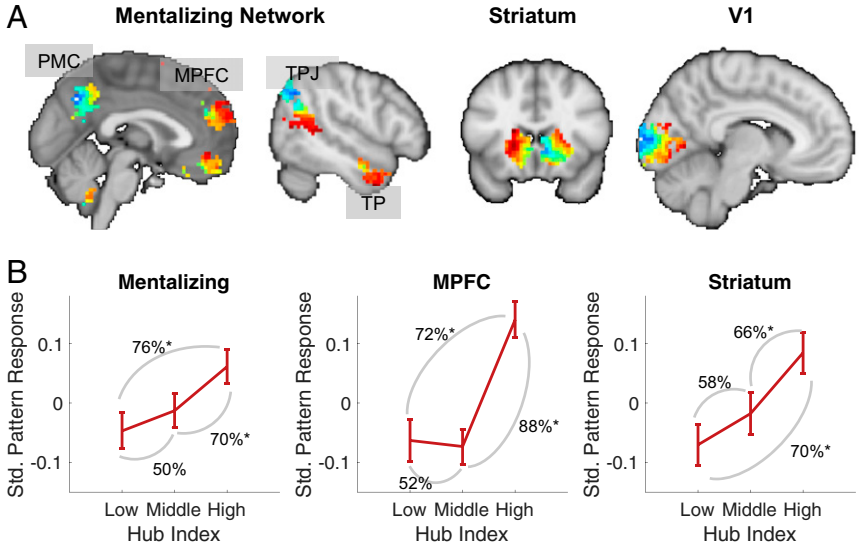

Fig. 2. Multivariate prediction of social value hubs. (*A*) Patterns of weights associated with viewing targets with higher hub index. (*B*) Average standardized pattern response to hub index in the mentalizing network, MPFC, and striatum. The standardized pattern response was computed between the cross-validated pattern weights learned by the multivariate model and the activation map associated with each hub category. This correlation corresponds to the standardized pattern response reflecting the extent to which a given map is similar to the pattern associated with increasing hub status. The percentage indicates forced-choice classification between each pairwise comparison. Error bars indicate SEM across participants. *$P < 0.05$ for a two-sided binomial test. L, left; MPFC, medial prefrontal cortex; PMC, posterior medial cortex; R, right; TP, temporal poles; TPJ, temporoparietal junction.

Table 1. Forced-choice classification accuracy between different hub categories from mean ROI activity

ROI	Low vs. mid	Mid vs. high	Low vs. high
Mentalizing Network	0.48 (0.07)	0.72 (0.06)*	0.70 (0.07)*
MPFC	0.48 (0.07)	0.70 (0.07)*	0.64 (0.07)
PMC	0.52 (0.07)	0.64 (0.07)	0.72 (0.06)*
R TP	0.52 (0.07)	0.72 (0.06)*	0.68 (0.07)*
L TP	0.46 (0.07)	0.72 (0.06)*	0.70 (0.07)*
R TPJ	0.52 (0.07)	0.64 (0.07)	0.68 (0.07)*
L TPJ	0.48 (0.07)	0.70 (0.07)*	0.66 (0.07)*
Striatum	0.40 (0.07)	0.62 (0.07)	0.58 (0.07)
V1	0.48 (0.07)	0.60 (0.07)	0.60 (0.07)

Numbers represent percentage of accuracy. Parentheses denote the standard error of the mean. L, left; MPFC, medial prefrontal cortex; PMC, posterior medial cortex; R, right; ROI, region of interest; TP, temporal poles; TPJ, temporoparietal junction.
*$P < 0.05$ for a two-sided binomial test.

community members' hub category robustly enough to predict hub category in new participants. It also provides further evidence that people monitor the difference between hubs and nonhubs but fail to differentiate between people who form middle versus low numbers of close ties.

Based on prior work (19), we hypothesized that average activity in perceiver brain regions associated with value computation would also track targets' hub category. To test this prediction, we used a reward-based functional localizer (*SI Appendix*, Fig. S3) to identify a cluster of activity spanning bilateral ventral and dorsal striatum (*SI Appendix*, Fig. S7). We then repeated prediction analyses using average activity in this region of interest (ROI). This regression model failed to linearly predict targets' hub category (*SI Appendix*, Table S3) or reach above chance levels for forced-choice accuracy (*SI Appendix*, Fig. S7).

Finally, we tested whether targets' hub category is reflected in brain regions not canonically associated with either mentalizing or reward processing. We repeated the above analyses using mean activity from the primary visual cortex (V1) (*SI Appendix*, Fig. S7) as identified using NeuroSynth. We found that prediction accuracy was no different from chance for within-participant correlation (*SI Appendix*, Table S3) and forced-choice accuracy (*SI Appendix*, Fig. S7).

Multivariate prediction. Our next analysis took advantage of multivariate patterns in perceivers' brain activity to further predict targets' hub category. This approach builds on work using multiregion brain "signatures" to predict personal experiences of emotion or pain (24, 25), but with key differences. Unlike past work, this analysis used signatures of brain activity to predict perceptions of others (i.e., targets). It further moved beyond examining neural responses related to a person's personal relationship with each target (26) and instead isolated patterns of brain activity that track perceptions in the broader community.

We trained a least absolute shrinkage and selection operator principal component regression (LASSO-PCR) algorithm to identify multivariate patterns of activity that tracked targets' hub category. These analyses focused specifically on patterns (*i*) across the entire mentalizing network, (*ii*) within each mentalizing-related region (MPFC, PMC, TPJ, and TP), (*iii*) in value-encoding regions (striatum), and (*iv*) in a control region (primary visual cortex). We again used a leave-one-participant-out procedure to train the algorithm on each set of 49 perceivers as they viewed targets who varied in hub index and then tested this pattern (Fig. 2) in the held-out perceiver.

When our model employed neural patterns spanning perceivers' mentalizing network, model predictions of targets' hub category significantly correlated with their actual hub category (mean $r = 0.343$ SE = 0.09, $P = 0.003$) (*SI Appendix*, Table S3). Consistent with univariate analyses, forced-choice classification accuracy correctly distinguished between high and middle hub categories in 76% (38 out of 50, $P < 0.05$, binomial test) of held-out participants and between high and low hub categories in 70% (35 out of 50, $P < 0.05$, binomial test) of held-out participants but did not distinguish between middle and low hub categories (25 out of 50, $P = 1$, binomial test) (Fig. 2). Further, patterns of activity within each mentalizing-related region showed similar levels of prediction of accuracy as the entire network (Fig. 2) (Table 2 and *SI Appendix*, Fig. S8).

Multivariate patterns in the striatum also produced model predictions of hub index that significantly correlated with targets' actual hub index (mean $r = 0.298$, SE = 0.09, $P = 0.005$) (*SI Appendix*, Table S3). An algorithm trained on striatal patterns also distinguished between faces with low vs. high (70% forced-choice accuracy, 35 out of 50, $P < 0.05$, binomial test) as well as middle vs. high hub category (66% forced-choice accuracy, 33 out of 50, $P < 0.05$, binomial test), but not between faces with low vs. middle hub category (58% forced-choice accuracy, 29 out of 50, $P = 0.32$, binomial test) (Fig. 2). These results stand in contrast to the univariate analyses, which failed to predict hub category from mean activity in the striatum. As with the univariate models, prediction accuracy was no different from chance when the algorithm was applied to activity in the primary visual cortex (Table 2 and *SI Appendix*, Fig. S8).

Univariate vs. multivariate prediction. Lastly, we examined whether average levels of neural activity versus fine-grained patterns in each ROI would more accurately predict targets' hub category. For each perceiver, we calculated the root mean squared error (RMSE)—a metric for how close the targets' actual hub categories were to targets' predicted hub categories—for both types of models. We then conducted a paired-samples *t* test for each ROI to determine if RMSE was significantly different for univariate versus multivariate prediction (*SI Appendix*, Table S6). Multivariate RMSE was numerically lower—suggesting higher accuracy—for all regions of interest, but these differences were only significant in the MPFC [$t(49) = -2.779$, $P = 0.008$] and striatum [$t(49) = -2.055$, $P = 0.045$] (see also *SI Appendix*, Fig. S9). Although these differences are significant, they should be interpreted with caution because the effect is weak.

Taken together, our results suggest that viewing photos of individuals with high hub index is associated with an increase in activity across the mentalizing network and that this increase can be used to predict when participants are viewing faces of hubs.

Table 2. Forced-choice classification accuracy between different hub categories from multivoxel activity

ROI	Low vs. mid	Mid vs. high	Low vs. high
Mentalizing Network	0.50 (0.07)	0.70 (0.07)*	0.76 (0.06)*
MPFC	0.52 (0.07)	0.88 (0.05)*	0.72 (0.06)*
PMC	0.56 (0.07)	0.66 (0.07)*	0.68 (0.07)*
R TP	0.40 (0.07)	0.70 (0.06)*	0.64 (0.07)
L TP	0.52 (0.07)	0.72 (0.06)*	0.68 (0.07)*
R TPJ	0.46 (0.07)	0.68 (0.07)*	0.72 (0.07)*
L TPJ	0.52 (0.07)	0.62 (0.07)	0.66 (0.06)*
Striatum	0.58 (0.07)	0.66 (0.07)*	0.70 (0.07)*
V1	0.50 (0.07)	0.54 (0.07)	0.58 (0.07)

Numbers represent percentage of accuracy. Parentheses denote the standard error of the mean. L, left; MPFC, medial prefrontal cortex; PMC, posterior medial cortex; R, right; ROI, region of interest; TP, temporal poles; TPJ, temporoparietal junction.
*$P < 0.05$ for a two-sided binomial test.

Discussion

To encourage prosociality and maximally benefit from social connections, individuals must be able to detect socially valuable others. Correspondingly, these detection skills develop early and are—to some degree—evolutionarily preserved. For instance, children (27, 28) and nonhuman primates (29) choose to repeat interactions with individuals who have acted cooperatively in the past. Adults also automatically evaluate others' trustworthiness, cooperative nature, and warmth in social exchanges (30–32). Here, we provide evidence about how this detection occurs in emerging real-world networks. Freshman dorm residents engaged brain systems associated with mentalizing and value computation when they viewed peers nominated as supportive, empathic, and well-connected. This activity was robust enough that, based on individuals' brain activity alone, we could accurately predict whether or not dormmates they encountered were viewed as hubs by the broader community.

Several features of these findings stand out. First, brain activity tracked hubs even though perceivers were not directed to evaluate dorm members' role in the community when viewing their faces. This suggests that people identify hub individuals spontaneously (cf. ref. 20) and might thus efficiently use this information to identify socially valuable targets. Second, individuals' brain activity tracked hubs even when removing and controlling for perceivers' relationship with these individuals. In other words, dorm members track individuals who are viewed throughout their community as helpful and positive social forces, irrespective of how perceivers personally related to that person (21, 33).

There are numerous ways in which the ability to detect socially valuable peers could serve perceivers. Empathic individuals provide high-quality social support that can buffer their friends during times of heightened stress and vulnerability. And these individuals also tend to be connected to numerous other community members, meaning that forming relationships with them is likely to provide access to other social resources. Several motivational factors, including targets' status or the likelihood of forming close connections, drive perceivers' tendency to mentalize with them (34). Here, we do not examine the motives that drive perceivers to detect high-value community members, nor do we examine the consequences of such detection for perceivers' later well-being or social connectedness. It would be useful, however, to leverage social network and neuroimaging approaches to explore whether some perceivers are better than others at detecting hub individuals, what drives accurate detection of hubs, and whether savvy detectors benefit as a result.

Lastly, by comparing univariate and multivariate analyses, we gained insight into how social value hubs are represented in the brain. In general, average activation and multivariate patterns in mentalizing and reward-related regions proved to be equally predictive of when individuals were viewing hubs (or not). The only exceptions were in the MPFC and striatum, where multivariate patterns outperformed average activity. Thus, these regions may contain more fine-grained patterns that encode additional information about community standing whereas other regions may have a more uniform, graded response to hubs. It will be worthwhile to attempt to replicate these findings, verifying that multivariate patterns in these regions more accurately predict social value hubs.

Our predictive analyses also shed light on an intriguing pattern in the perception of social value. Both univariate activity and multivariate patterns in perceivers' brains linearly predicted the social value of targets they encountered, but forced-choice analysis revealed that this effect was driven by hub individuals who received unusual numbers of nominations. This insight extends prior work combining social network science and neuroimaging. Two studies have demonstrated that brain activity in perceivers tracks targets' popularity and network centrality (19, 20). Here, we demonstrate that, in the domain of social value, this relationship is nonlinear. Rather than monitoring social value as a continuous variable, community members experience a "pop out" effect, tracking only especially valuable peers. This could represent an efficient metric for detecting trustworthy and supportive relationship partners, rather than expending significant cognitive resources to monitor all network members equally (35).

In sum, we identify a potential mechanism for detecting high-quality social connections and provide insight into what neural systems track individuals who build numerous strong ties in their community. These findings can, in the future, inform how individuals accrue mental and physical health benefits from their broader community.

Materials and Methods

Participants. As part of a larger study on social networks (12), newly matriculated college students were recruited from two freshman-only dormitories at Stanford University. Ninety-seven freshmen for the larger study and 52 students who were eligible and willing to participate in the fMRI scanning session provided informed consent. Eligibility for scanning required that participants be right-handed, free of ferrous metal, not actively taking psychoactive medications, and have no history of neurological problems. Two participants were excluded due to excessive signal dropout in their functional and anatomical scans.

The final sample consisted of 50 participants (25 males): 26 participants in dorm A and 24 participants in dorm B with a mean age of 18.08 y (SD = 0.27) and 18% East Asian, 10% Black or African American, 32% White or Caucasian, 12% Hispanic or Latino/a, 6% South Asian, and 22% Mixed Race or Other. Our final sample was representative of the larger sample of 97 participants (52 males): 46 participants in dorm A and 51 participants in dorm B with a mean age of 18.11 y (SD = 0.32) and 19% East Asian, 6% Black or African American, 34% White or Caucasian, 11% Hispanic or Latino/a, 6% South Asian, and 24% Mixed Race or Other.

Procedure. During the second week of the academic quarter, 97 participants completed social network nominations in an online survey. During the third week of the quarter, an experimenter photographed each participant. Participants were asked to pose with a neutral expression and forward eye gaze. Over the next 6 wk, a subset of 50 students participated in fMRI scanning sessions. Approximately 5 d before their scan, participants completed prescan ratings that assessed their perceptions of other individuals in their dorm. During the scanning session, participants started with the face-viewing task (described below) followed by a functional localizer task for reward. They then completed four additional tasks related to social network perceptions, prosociality, and empathy (not analyzed here). As part of the larger study, participants also completed trait and daily surveys, as well as providing saliva samples. All experiments involving human subjects were approved by the Stanford University Institutional Review Board.

Social Network Nominations. To assess social value to the broader dorm community, we asked participants to nominate up to eight people in their dormitory in response to each of eight prompts (in the listed order): (*i*) "Who are your closest friends?", (*ii*) "Whom do you spend the most time with?", (*iii*) "Whom have you asked for advice about your social life?", (*iv*) "Who do you turn to when something bad happens?", (*v*) "Whom do you share good news with?", (*vi*) "Who makes you feel supported and cared for?", (*vii*) "Who is the most empathetic?", and (*viii*) "Who usually makes you feel positive (e.g., happy, enthusiastic)?".

Face-Viewing Task. The face-viewing task was modified from a study by Zerubavel et al. (19). Participants were scanned as they viewed 30 photographs of dorm members. For each participant, we applied a face selection algorithm to pick a unique set of 30 faces that maximized how much targets varied in closeness to the participant. Photos were cropped to only include targets' faces and necks, converted to grayscale, and adjusted to have equal luminance. Participants were not instructed to do anything while viewing the faces, except press a button when they saw an occasional red dot (i.e., an attention check) instead of a face. In a rapid event-related design, these 30 photos were randomly presented two times for 1 s each (*SI Appendix*, Fig. S3), for a total of 60 face presentations. Twelve red dots were randomly intermixed with these faces and displayed for 1 s each. A fixation cross was displayed during the interstimulus intervals (ISIs) for a jitter of 1 to 8 s (mean duration = 3.5 s).

Analyses for Face-Viewing Task. For the face-viewing task, first-level effects were estimated for two models using the general linear model. The first model was a whole-brain parametric analysis aimed at identifying brain regions in which activity increased as a function of targets' hub index. The first regressor represented the average hemodynamic response across all 60 presentations to dorm members' photos (i.e., all photos > baseline). To control for the perceivers' personal relationship with the target, the second regressor was a parametric modulator modeling the hemodynamic response that linearly varied with perceiver's self-reported ratings of closeness to the target collected before the scan. The third regressor modeled the hemodynamic response that varied parametrically with targets' hub index (i.e., excluding the perceiver's own ties to targets). Due to serial orthogonalization of parametric modulators, variance associated with the third regressor would reflect blood oxygen level-dependent activity correlated with network-level perceptions of each target, after removing and controlling for perceivers' personal relationship with these individuals. Therefore, first-level contrasts were only computed for the parametric effect of targets' hub index in model 1.

Model 2 was a whole-brain analysis conducted to produce statistical *t* maps for leave-one-out cross-validation models. In model 2, the first three regressors modeled the average hemodynamic response for targets in the following three categories: (*i*) low hub index, (*ii*) middle hub index, and (*iii*) high hub index. For the 30 targets selected specifically for each perceiver, the high hub category included targets in the 67th percentile or higher on hub index. The middle hub category included targets in the 33rd percentile or higher and lower than the 67th percentile on hub index. The low hub category included all targets below the 33rd percentile on hub index. To control for the perceivers' personal relationship with the target, we included a fourth regressor that modeled the hemodynamic response that varied parametrically with prescan ratings of closeness. Then, we computed three first-level contrasts comparing each hub category to baseline (i.e., low hub category > baseline, middle hub category > baseline, high hub category > baseline).

Both models included additional nuisance covariates: attention check trials, six motion parameters from image realignment, and regressors modeling time points where in-brain global signal change exceeded 2.5 SDs of the mean global signal change or where estimated motion exceeded 0.5-mm translation or 0.5° rotation. All events were modeled as a boxcar spanning their duration and convolved with a canonical (double-gamma) hemodynamic response function. The time series was high pass-filtered using a cutoff period of 128 s. Serial autocorrelations were modeled as an AR(1) process.

For the parametric analysis, random effects analyses of the group were computed using the contrast images generated for each participant. For whole-brain group-level analyses, all images were thresholded using the cluster_correct function in bspmview (https://doi.org/10.5281/zenodo.595175) with a cluster-defining threshold of $P < 0.001$, followed by a cluster-level correction at a family-wise error of 0.05. For leave-one-out cross-validation analyses, we used multivoxel pattern analysis at the whole-brain level, as well as constraining the search space to specific regions of interest related to mentalizing and reward processing. For visualization of results, thresholded results were surface rendered using bspmview (www.bobspunt.com/bspmview/).

Univariate Prediction. For each of the three *t* maps associated with low, middle, and high hub categories (see model 2 above), we averaged the *t* values in each ROI for each participant. A linear model was then trained to predict the hub category of a particular set of faces given the average *t* statistic of a given ROI. A separate model was run for each ROI. (The *t* statistic maps are affected by the amount of data collected. We could have trained the models on beta estimates, which are less affected by the amount of data. Training the models on *t* statistic maps, however, penalizes voxels with high variability, which would improve the model's ability to pick up meaningful signal in the maps.) To avoid overfitting the data, we followed a leave-one-participant-out cross-validation procedure. That is, we trained the models on the average *t* values of all but one participant and assessed prediction accuracy in the held-out participant.

Multivariate Prediction. For each participant, we extracted and vectorized the *t* values associated with each hub category in a given ROI. We then applied a LASSO-PCR algorithm to predict hub category from the vectors of *t* values (implemented with CANlab Core Tools package: canlab.github.io/CanlabCore/). As a dimensionality reduction step, we first applied a principle components analysis (PCA) on the vectors of *t* values. We retained the top number of components that explained 35% of the variance in the data. As the PCA was done separately for each ROI, this procedure allowed us to take a data-driven approach to determine the number of components to retain for each ROI. The retained components were then used to predict the hub category using least squares regression with L-1 regularization (LASSO), which encourages sparse regression coefficients by shrinking them toward zero. For visualization purposes, the regression coefficients were then back-projected into voxels in 3D Montreal Neurological Institute space. As was the case with ROI prediction, we followed a leave-one-participant-out cross-validation procedure to evaluate the algorithm's prediction accuracy.

ACKNOWLEDGMENTS. We thank Rucha Makati and Desmond Ong for help with data collection and analysis, as well as Matthew Jackson for advice and guidance on social network analyses. This work was funded by National Institute of Mental Health Grants R21 MH104464 (to J.Z.) and F32 MH098504 (to S.A.M.).

1. Nesse RM (2007) Runaway social selection for displays of partner value and altruism. *Biol Theory* 2:143–155.
2. Simon HA (1990) A mechanism for social selection and successful altruism. *Science* 250:1665–1668.
3. Trivers R (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46:35–57.
4. Wedekind C, Milinski M (2000) Cooperation through image scoring in humans. *Science* 288:850–852.
5. Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437:1291–1298.
6. Feinberg M, Willer R, Schultz M (2014) Gossip and ostracism promote cooperation in groups. *Psychol Sci* 25:656–664.
7. Morelli SA, Lee IA, Arnn ME, Zaki J (2015) Emotional and instrumental support provision interact to predict well-being. *Emotion* 15:484–493.
8. Canevello A, Crocker J (2010) Creating good relationships: Responsiveness, relationship quality, and interpersonal goals. *J Pers Soc Psychol* 99:78–106.
9. Reis HT, Gable SL (2015) Responsiveness. *Curr Opin Psychol* 1:67–71.
10. Sorkin D, Rook KS, Lu JL (2002) Loneliness, lack of emotional support, lack of companionship, and the likelihood of having a heart condition in an elderly sample. *Ann Behav Med* 24:290–298.
11. Kardos P, et al. (2017) Empathic people have more friends: Empathic abilities predict social network size and position in social network predicts empathic efforts. *Soc Networks* 50:1–5.
12. Morelli SA, et al. (2017) Empathy and well-being correlate with centrality in different social networks. *Proc Natl Acad Sci USA* 114:9843–9847.
13. Pratt MW, et al. (2000) Facilitating the transition to university: Evaluation of a social support discussion intervention program. *J Coll Student Dev* 41:427.
14. Hicks T, Heastie S (2008) High school to college transition: A profile of the stressors, physical and psychological health issues that affect the first-year on-campus college student. *J Cult Divers* 15:143–147.
15. Friedlander LJ, et al. (2007) Social support, self-esteem, and stress as predictors of adjustment to university among first-year undergraduates. *J Coll Student Dev* 48:259–274.
16. Swenson LM, Nordstrom A, Hiester M (2008) The role of peer relationships in adjustment to college. *J Coll Student Dev* 49:551–567.
17. Gray R, et al. (2013) Examining social adjustment to college in the age of social media: Factors influencing successful transitions and persistence. *Comput Educ* 67:193–207.
18. Lee C, et al. (2014) A closer look at self-esteem, perceived social support, and coping strategy: A prospective study of depressive symptomatology across the transition to college. *J Soc Clin Psychol* 33:560–585.
19. Zerubavel N, Bearman PS, Weber J, Ochsner KN (2015) Neural mechanisms tracking popularity in real-world social networks. *Proc Natl Acad Sci USA* 112:15072–15077.
20. Parkinson C, Kleinbaum AM, Wheatley T (2017) Spontaneous neural encoding of social network position. *Nat Hum Behav* 1:0072.
21. Banerjee A, Chandrasekhar A, Duflo E, Jackson MO (2017) *Gossip: Identifying Central Individuals in a Social Network*. arXiv:1406.2293v6. Preprint, posted May 8, 2017.
22. Van Overwalle F, Baetens K (2009) Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *Neuroimage* 48:564–584.
23. Delgado MR (2007) Reward-related responses in the human striatum. *Ann N Y Acad Sci* 1104:70–88.
24. Chang LJ, Gianaros PJ, Manuck SB, Krishnan A, Wager TD (2015) A sensitive and specific neural signature for picture-induced negative affect. *PLoS Biol* 13:e1002180.
25. Wager TD, et al. (2013) An fMRI-based neurologic signature of physical pain. *N Engl J Med* 368:1388–1397.
26. Eisenberger NI, et al. (2011) Attachment figures activate a safety signal-related neural region and reduce pain experience. *Proc Natl Acad Sci USA* 108:11721–11726.
27. House BR, et al. (2012) The ontogeny of human prosociality: Behavioral experiments with children aged 3 to 8. *Evol Hum Behav* 33:291–308.
28. House B, et al. (2013) The development of contingent reciprocity in children. *Evol Hum Behav* 34:86–93.
29. Melis AP, Hare B, Tomasello M (2006) Engineering cooperation in chimpanzees: Tolerance constraints on cooperation. *Anim Behav* 72:275–286.
30. Everett JA, Pizarro DA, Crockett MJ (2016) Inference of trustworthiness from intuitive moral judgments. *J Exp Psychol Gen* 145:772–787.
31. Fiske ST, Cuddy AJ, Glick P (2007) Universal dimensions of social cognition: Warmth and competence. *Trends Cogn Sci* 11:77–83.
32. Rilling J, et al. (2002) A neural basis for social cooperation. *Neuron* 35:395–405.
33. Casciaro T (1998) Seeing things clearly: Social structure, personality, and accuracy in social network perception. *Soc Networks* 20:331–351.
34. Zaki J (2014) Empathy: A motivated account. *Psychol Bull* 140:1608–1647.
35. Meyer ML, Spunt RP, Berkman ET, Taylor SE, Lieberman MD (2012) Evidence for social working memory from a parametric functional MRI study. *Proc Natl Acad Sci USA* 109:1883–1888.