

# **Gossip drives vicarious learning and facilitates social connection**

Jolly, Eshin and Chang, Luke J.

Computational Social Affective Neuroscience Laboratory  
Department of Psychological and Brain Sciences  
Dartmouth College  
Hanover, NH, USA 03755  
Corresponding author ([eshin.jolly@dartmouth.edu](mailto:eshin.jolly@dartmouth.edu); [@Eshjolly](#))

## SUMMARY

Complex language and communication is one of the unique hallmarks that distinguishes humans from most other animals. Interestingly, the overwhelming majority of our communication consists of *social topics* involving self-disclosure and discussions about others, broadly construed as *gossip*. Yet the precise social function of gossip remains poorly understood as research has been heavily influenced by folk intuitions narrowly casting gossip as baseless trash talk. Using a novel empirical paradigm that involves real interactions between a large sample of participants we provide evidence that gossip is a rich, multifaceted construct, that plays a critical role in vicarious learning and social bonding. We demonstrate how the visibility or lack thereof of others' behavior shifts conversational content between self-disclosure and discussions about others. Social information acquired through gossip aids in vicarious learning, directly influencing future behavior and impression formation. At the same time, conversants come to influence each other, form more similar impressions, and build robust social bonds. Consistent with prior work, gossip also helps promote cooperation in groups without a need for formal sanctioning mechanisms. Altogether these findings demonstrate the rich and diverse social functions and effects of this ubiquitous human behavior and lay the groundwork for future investigations.

**Keywords:** gossip, social-influence, social-bonding, vicarious-learning, cooperation

## INTRODUCTION

In a typical day, humans speak about 16,000 words <sup>1</sup> and at least 65% of these conversations involve discussing “social topics” <sup>2,3</sup>. These social topics primarily consist of self-disclosure (i.e. exchanges concerning the conversing parties) or discussions about others (i.e. concerning absent third parties) <sup>4,5</sup>. These types of exchanges have been long hypothesized to serve critical functions including information exchange, cultural learning, entertainment, influence, and social bonding <sup>6–9</sup>. Yet, despite the ubiquity of this behavior and its distinction from more general forms of communication (e.g. the exchange of facts, instruction or non-personal exchanges), our scientific understanding of it is severely limited. A key factor contributing to this stymied progress is the difficulty of establishing a consensus definition of *gossip* amongst researchers <sup>10</sup>.

Folk intuitions reflect the belief that gossip is limited to primarily negative evaluative commentary about absent individuals, and hence generally taboo <sup>11</sup>. This narrow definition has provided the theoretical basis for several empirical studies using economic games, in which participants make decisions that can be selfish or cooperative and are either directed to share evaluative information about others, or made aware that others can share evaluative information about them<sup>1</sup> <sup>12</sup>. These conditions are sufficient to motivate individuals to behave more cooperatively, even when compared to more direct sanctioning mechanisms such as costly punishment <sup>13,14</sup>. Evaluative commentary can also inform group members’ decisions to ostracize selfish individuals through partner selection, thereby improving outcomes for all remaining group members <sup>15</sup>. In general, participants are motivated to maintain a positive reputation amongst their social group because it attracts altruistic behavior from others <sup>16</sup>. For this reason, when gossip consists of discussions about absent individuals, it provides an exchange of reputational information and a cheap means of social control of behavior <sup>17–20</sup>.

---

<sup>1</sup> In reality other “players” are often simulated agents, confederates or preset actions displayed to participants

However, gossip has long been thought to be more than just a tool for enforcing cooperation via reputation management. Despite the overwhelming scientific focus on gossip's relationship with cooperative behavior, in daily life, "social policing" has been estimated to comprise a mere 5% of naturally occurring private conversations<sup>4</sup>. For example, in a recent study investigating the base rates of gossip in daily life using an Electronically Activated Recorder (EAR), Robbins and Karan<sup>21</sup> found that gossip comprised approximately 14% of people's daily conversations and was primarily neutral in content rather than positively or negatively valenced. This suggests that folk intuitions about gossip being negative fail to reflect the majority of real world conversational behavior. Scientific investigations that prematurely adopt such a narrow definition fail to capture the diverse social value gossip can provide<sup>22</sup> and can hamper scientific progress.

A broader characterization of gossip includes many more diverse social functions such as the opportunity to learn from the "triumphs and misadventures of people beyond one's immediate perceptual sphere"<sup>9</sup>. Gossip as a tool for *vicarious learning* can allow individuals to gain an understanding of the world and adjust their behavior appropriately even without firsthand experience. At the same time, gossip has been proposed to be a uniquely human form of "social grooming," analogous to physical grooming by other primates, a process by which individuals build trust and strong *social bonds*<sup>2,6</sup>. For example, "chit chat" or "idle talk" in which individuals self-disclose, provides an opportunity for rewarding social connection, both because self-disclosure is intrinsically rewarding<sup>23</sup> and because it provides an opportunity to establish trust and learn about others, and cultural norms<sup>6,9</sup>. Observational and self-report questionnaire studies on gossip in school and workplace relationships indicate that it can facilitate building and configuring social ties<sup>24,25</sup> and increase friendship and trust<sup>26</sup>. Individuals may also use gossip as a means to share and compare their *impressions* of other individuals, both to regulate and understand their own feelings as well as clarify acceptable normative behavior<sup>27</sup>. Evaluations of others can be seen as a form of implied communication, in which the topic of discussion is actually

the implied acceptability of behaviors (i.e. norm violations, morals)<sup>9,28</sup>. These implied discussions can help establish a shared reality<sup>29</sup> with consensus understanding of social norms and broader cultural conventions.

In the present work, we sought to broaden the empirical study of gossip beyond its role in reputation management by exploring: (a) the social circumstances that lead individuals to spontaneously engage in gossip; (b) its role in information transmission and vicarious learning and; (c) its role in impression formation, impression sharing, and social bonding. Critical to our approach, is the observation that gossip and rumor transmission appear to occur more often when social situations are ambiguous or information is withheld from individuals<sup>30,31</sup>. For example, outside the laboratory we rarely have full information as events unfold, nor do we have opportunities to directly observe everyone's behavior in our local professional and social communities. Gossip may aid in reducing uncertainty by providing an efficient way to quickly disseminate knowledge via vicarious learning, creating a shared consensus of how to interpret an event or others' actions, and also enhancing feelings of social support by facilitating social bonds.

To capture these social dynamics in an experimental setting, we designed a large online experiment in which individuals played one of four variants of a *live and interactive* 10-round repeated public-goods game in six-person groups<sup>32</sup>. We employed a between-subjects design to experimentally manipulate: (a) the amount of information available about other players' behavior and (b) whether or not it was possible to gossip via private communication with a partner throughout the duration of the game (Fig. 1). In the *Complete Information* condition, participants were able to see the actions made by all of their neighbors, while in the *Incomplete Information* condition, participants could only directly observe a subset of their neighbors. These information contexts were crossed with the ability to communicate by exchanging up to two short, private, free-form messages with another individual during each round of the game. Over the course of each game, repeated interactions occurred between the same network of participants, enabling

individuals to develop extended conversations, engage in long-term learning, as well as potentially develop social relationships.

Because situational factors (in addition to conversational content) have been proposed as one of the key features that distinguishes gossip from other forms of communication<sup>8</sup>, this design allowed us to test the prediction that the *mixture* of social topics (i.e. discussions about others vs chit-chat) would change based upon the visibility and shared knowledge of others' behavior. When participants are unable to observe the behaviors of others first-hand, we expected discussions about others to comprise a large proportion of social discussion topics. If gossip does indeed serve as a mechanism for vicarious learning, participants' *own behavior* and their *impressions of others* should reflect any knowledge they acquire via gossiping, specifically when such knowledge is unavailable first-hand. On the other hand, when all group members have full visibility of each other's actions, we expected discussion topics to involve more chit-chat or idle-talk. We also expected that conversants would influence *each others'* behavior to a greater degree than non-conversing participants and thus exhibit stronger social bonding via increased affinity for one another at the end of the game.

**[Figure 1 about here]**

## RESULTS

### Visibility of others' behavior changes the makeup of gossip

Consistent with the idea that situational factors play a key role in distinguishing types of gossip<sup>8</sup>, we found that games with incomplete information contained proportionally more spontaneous discussions about others' behavior relative to games with complete information (21.7% vs 14.9% of messages),  $b = 0.375$ ,  $z = 6.246$ ,  $p < .001$  (Fig. 2). Incomplete information games also contained more random messages (9.8% vs 7.4%)  $b = 0.284$ ,  $z = 2.085$ ,  $p = 0.014$ . However, complete information games comprised a wider array of discussion topics the largest of which was casual chit-chat (29.1% vs 26.8% of messages),  $b = 0.082$ ,  $z = 1.969$ ,  $p = 0.049$ . Participants also made more brief expressions, inquiries, discussed game mechanics, or forwent

communication altogether (blank messages) (all  $ps < 0.05$ ; Table S3). Affirmations and strategic discussions did not differ between conditions.

In other words, in circumstances where direct observation of others' was not possible, participants shifted the content of their discussion to focus more on how other individuals behaved while instead focusing on more self-relevant and casual topics when social visibility was not restricted. Similar to observations outside of the laboratory, these social discussion topics comprised the majority of all participants' conversations (~54%) <sup>2,21</sup>.

**[Figure 2 about here]**

### **Gossip facilitates vicarious learning in the absence of direction observation**

To test the prediction that gossip functions as a mechanism for *vicarious learning*, we examined how conversation impacted participants' behavior by evaluating how a focal participant's *future* behavior was influenced by each of their neighbor's *past* behavior when this behavior was not directly visible and contrasted these estimates between games with and without communication.

Despite a lack of direct observation (Fig. 3A), both distant and remote neighbors' contributions from past rounds were significantly more predictive of focal participants' contributions in future rounds in games where gossip was possible  $b = 0.10$ ,  $t(3273.76) = 2.77$ ,  $p = 0.006$ ,  $b = 0.07$ ,  $t(3466.83) = 2.60$ ,  $p = 0.009$ , respectively. Importantly, the influence of distant neighbors was unlikely to be indirectly mediated through immediate neighbors as our model explicitly controls for their behavior. Full model results appear in Table 1. Given the deliberate circular arrangement of each game's network, a focal participant's *unobservable* distant neighbors were the *observable* immediate neighbors of their remote neighbor (communication partner) (Fig. 1A). Therefore, the only way in which a focal participant could come to be influenced by individuals they could not directly observe (distant neighbors), was via information transmitted through another person (remote neighbor) who could observe those individuals.

Interestingly, we also found that vicarious learning influenced participants' *impressions* of

individuals they could not observe. Participants' affinity ratings for both distant and remote neighbors were consistently predicted by their unobservable behavior in games where gossip was possible,  $b = 0.38$ ,  $t(1699.10) = 4.53$ ,  $p < 0.001$ ,  $b = 0.30$ ,  $t(1997.40) = 4.48$ ,  $p < 0.001$ , respectively (Fig. 3C). Furthermore, impressions of these unobserved players were likely directly influenced by social discussions, as there was a significant increase in similarity<sup>2</sup> of affinity ratings between conversing individuals,  $M = -10.75$ ,  $p < 0.001$  (Fig. 3B). The magnitude of these effects appeared to reflect the strength of the bond developed between communication partners. Conversants who reported a stronger affinity for each other<sup>3</sup>, exhibited greater synchronization of their impressions of other players in their group,  $r = 0.20$ ,  $p = .002$ , (Fig. S3).

Taken together, these results indicate that discussions about other individuals provide a learning signal in situations where direct observation is not possible. When individuals cannot observe the members of their social group, they turn to others to vicariously learn about their social environment through dialogue. Information transmitted via gossiping was the only way in which any information about unobserved neighbors was available to participants. Participants not only adapted their behavior based on this information, but also adjusted their social impressions of other players by adopting their conversation partner's direct observations of other player's behavior in the game (Fig 3C middle panel).

**[Figure 3 about here]**

### **Gossip increases social connection and influence**

Because building social bonds has been proposed as another function of gossip <sup>2</sup>, we expected that conversants would influence each other to a greater degree when it was unnecessary to discuss the actions of others (i.e. complete information games). Consistent with

---

<sup>2</sup> Mean change in euclidean distance between ratings. Smaller values indicate *decreased* distance, i.e. *increased* similarity. Inference via permutation testing, see Methods for details.

<sup>3</sup> Correlation between affinity euclidean distance (Fig. 3B) and average affinity rating of each pair of communicators. Inference via permutation testing, see Methods for details.



this idea, we found that remote neighbors' past contributions were significantly more predictive of focal participants' future contributions in complete information games where gossip was possible relative to games where it was not  $b = 0.09$ ,  $t(3486.55) = 3.40$ ,  $p < 0.001$  (Fig. 4A). Conversely, immediate neighbors' past contributions were significantly less predictive in games where gossip was possible relative to when it was not,  $b = -0.08$ ,  $t(3361.67) = -2.26$ ,  $p = 0.024$  (Table 2 contains full model results). This effect likely reflects the developing bond between conversants and their increased attention to each others' behavior. Conversing participants reported the highest affinity for each other ( $M = 81.12$ ,  $SD = 30.58$ ) relative to all other group members (remote neighbors no-communication  $M = 57.93$ ,  $SD = 38.14$ ; immediate neighbors  $M = 64.63$ ,  $SD = 31.53$ ; distant neighbors  $M = 64.11$ ,  $SD = 32.18$ )  $t(205.22) = 5.05$ ,  $p < 0.001$ , Cohen's  $d = 0.67$ ,  $t(1900) = 7.86$ ,  $p < 0.001$ , Cohen's  $d = 0.53$ ,  $t(1900) = 8.13$ ,  $p < 0.001$ , Cohen's  $d = 0.54$ , respectively (Table S1; Fig. S1). In addition, conversing participants shared more similar impressions of others  $M = -9.71$ ,  $p < 0.008^1$  (Fig. 4B), which was amplified among those conversants who felt the most positive about each other,  $r = 0.28$ ,  $p < .001^2$  (Fig. S4). Furthermore, conversation moderated the degree to which their behavior in the game influenced their impressions of each other  $b = -0.22$ ,  $t(1535.74) = -3.42$ ,  $p = 0.004$  (Fig. 4C left panel; Fig. S2, Table S2). Participants' affinity for the remote neighbor was less influenced by how cooperatively they behaved in the game when they were able to communicate with each other.

These results speak to the impact that gossip can have on individuals' behaviors and impressions when social topics focus less on discussions about others. Specifically, this form of social dialogue facilitates social bonds that are robust to deviant behavior and facilitates a shared reality, in which conversants align their behavior and social impressions. This provides support for gossip as a mechanism for bringing individuals closer together via self-disclosure and learning directly about each other <sup>4</sup> as well as the broader social world.

**[Figure 4 about here]**

**The opportunity to gossip increases cooperative group behavior**

While the previous analyses focused on the dynamics of individual participant's behavior throughout the game, we also examined the impact of our experimental design on average group behavior. Consistent with prior public goods dilemmas<sup>33</sup>, average group contributions tended to decline over the course of each game  $b = -2.15$ ,  $t(155) = -7.75$ ,  $p < 0.001$ . However, games with the potential for gossip ( $b = -1.22$ ) exhibited less of a decline than games without ( $b = -3.08$ ),  $b = 1.86$ ,  $t(155) = 3.35$ ,  $p = 0.001$  (Fig. 5A). Participants contributed significantly more on average when they could communicate ( $M = 61.31$ ;  $SD = 6.26$ ) relative to when they could not ( $M = 51.20$ ;  $SD = 9.44$ ),  $b = 10.22$ ,  $t(155) = 2.52$ ,  $p = 0.013$ . We also found that participants contributed marginally less on average ( $M = 52.60$ ;  $SD = 6.54$ ) in the incomplete information condition relative to complete information ( $M = 59.74$ ;  $SD = 7.66$ ),  $b = 7.26$ ,  $t(155) = 1.79$ ,  $p = 0.075$ . Full results are available in Table S5. Importantly, we also observed a significant positive relationship between the amount of gossip received within a group and the average amount of group cooperation  $b = 29.36$ ,  $t(150.57) = 3.538$ ,  $p = .001$ . However, this did not change as a function of information visibility  $b = -25.02$ ,  $t(150.57) = -1.507$ ,  $p = 0.134$ . In other words, more gossip received by participants in a group, produced more cooperative behavior on average, independent of full or partial observability. See methods for model and analysis details ("group contribution behavior").

These results replicate and extend prior work demonstrating how communication, and specifically gossip, can provide an effective social means to sustain cooperation without the need for additional interventions<sup>13,15,34</sup>. Interestingly, we observe that this effect is not because *all* individuals cooperate more, but rather that a subset of individuals become more cooperative while others remain self-interested (bimodal distribution Fig. 5B Left; red distributions). This is consistent with the idea that not all individuals are equally motivated by reputational concerns, but rather may exhibit variations in their individual preferences<sup>27,35</sup>.

**[Figure 5 about here]**

## DISCUSSION

In the present work, we explored the complexity of gossip by focusing on: (a) the social circumstances that led individuals to spontaneously engage in gossip; (b) its role in information transmission and vicarious learning; (c) its role in impression formation, impression sharing, and social bonding; (d) and its influence on cooperative group behavior. A novel contribution of our approach was building a method to facilitate *real* social interactions with financial consequences, in which individuals were able to have largely unrestricted private dyadic conversations. Key to our investigation was the manipulation of information visibility, whether or not participants were privy to the actions of all or a subset of their group members, and the ability to communicate privately with another group member.

Across all games in which communication was possible, social topics constituted the majority of what individuals discussed, yet the makeup of these topics differed based on information visibility. Participants more frequently discussed others' behavior when it was not directly observable. Unlike the folk view of gossip as baseless "trash talk," these exchanges served a clear social function - vicarious learning. Participants adjusted their future behavior based on knowledge of others' unobserved past actions reported by their conversation partner. Participants also adjusted their social judgments, such that they felt more affinity for group members who behaved cooperatively and less affinity for those who behaved selfishly, despite being unable to directly observe these behaviors themselves. These findings provide direct evidence that situational factors influence what gossip ultimately looks like and demonstrate that gossip can provide a rich source of information to aid in navigating the social environment.

These findings complement and extend previous work on information transmission in the absence of direct observation<sup>12</sup>. Unlike previous work, our participants were not instructed or incentivized in any way to discuss particular topics or exchange specific messages. Rather, like real world conversations<sup>2,3,21</sup> gossip emerged *spontaneously* and influenced how participants behaved in the game despite a lack of mechanisms like ostracization or punishment<sup>13,15</sup>. One

interpretation of this finding is that discussing others' behavior indicates implied-communication whereby conversants talk about the *implied* acceptability of certain behaviors (e.g. "What they did was bad, right?") thereby facilitating the construction or maintenance of moral rules or social norms <sup>9</sup>. In this way, leveraging and checking-in with the experiences of others provided an efficient mechanism for participants to improve the quality of their future outcomes.

Gossip also played a key role in bringing individuals closer together. In particular, conversational content shifted more towards chit-chat and positive affirmation when each participant's behavior was visible to all group members. Social topics that consist of self-disclosures, constitute a different dimension of gossip than evaluations of absent others, and have been long thought to be the primary ingredients of social bonding <sup>6</sup>. Consistent with this idea, we found that (a) conversational partners influenced each other to a greater degree than other group members, (b) felt more positive about each other compared to other group members, and (c) came to share similar impressions of other players by the end of the game. Interestingly, these latter two effects were related. Participants who felt the most positively about each other also tended to have the most similar impressions of others. This alignment of impressions may provide a distinct mechanism by which gossip can bring individuals closer together. Prior work has suggested that telegraphing trust by exchanging personal information through gossip can aid in building alliances between individuals <sup>8,36,37</sup>. We speculate that participants established a sense of commonality with one another, creating a "shared reality" that served to influence each other's behavior and perspectives <sup>38</sup> while satisfying each other's inherent desire for social connection <sup>29,39</sup>. This idea is consistent with observational work demonstrating the strong relationship between workplace gossip and friendship over time <sup>26</sup>.

At the group level, the possibility of gossip also facilitated sustained cooperation. Replicating previous research, individuals tended to contribute more money when they had the opportunity to gossip with one another <sup>13,15,17</sup>. This led to a higher mean group contribution at the end of the game, but upon closer examination, we note that the mean contribution did not appear

to fully characterize the underlying behavioral dynamics over time. Without the possibility of gossip, groups gradually switch to investing all of their money in the beginning of the game to keeping it all at the end (Fig. 5B right; blue distributions), similar to the highly replicated unraveling of cooperation observed in traditional public goods games<sup>15,32</sup>. In games where gossip is possible, contribution behavior appeared to stabilize into a bimodal distribution consisting of extreme cooperators and self-interested players (Fig. 5B Left; red distributions). While it's possible that the cooperation would eventually unravel if participants played more rounds, another possibility is that communication may have been sufficient to disrupt the self-interested equilibrium in the game. This could be facilitated through the building of social bonds over time or alternatively through reputation management by providing a cheap form of social sanctioning<sup>19,40,41</sup>. While we cannot adjudicate between these possibilities with our design, this presents an interesting opportunity for future research. By varying the amount and channels of communication between group members (e.g. discussions between more than two players to form coalitions), it may be possible to stabilize this bimodal equilibrium and possibly lead more groups to adopt a maximal contribution strategy.

A limitation of the unconstrained messaging component of our task was the infeasibility of estimating the relationship between time-lagged behavior and communicative content at the *round* level. Because participants were not directed or restricted in what they could say, messages often changed topic across game rounds (e.g. chit-chat, strategy, nothing). This breaks the continuity required for predicting contribution behavior with time-lagged analyses. It also leads to large imbalances between messages of a specific type, on a per-round basis, making the interactions between contribution behavior on a given round and the kinds of messages received on that round, impossible or unreliable to compute. Additionally, because of the synchronous multiplayer nature of our experiment, participants were required to wait for inputs from *all* group members before the game advanced. Future work can address this shortcoming by permitting participants to have longer on-going conversations throughout the duration of an experiment,

independent of the actions that other participants take. Such a design could also facilitate more sophisticated analyses of richer conversational content.

Altogether, our findings provide a novel characterization of gossip and its social functions. Gossip is not a monolithic construct and is more complex than the narrow definition of baseless trash-talk reflected by our folk intuitions. Instead, the ability to glean knowledge about one's social world appears to be a key factor that changes the makeup of social topics in conversation. Entertainment, influence, and social bonding may emerge as consequences from gossip consisting more heavily of "idle-talk," "chit-chat," and self-disclosure, because the primary communicative purpose is connecting with another individual. This is consistent with the "social grooming hypothesis"<sup>6</sup>, the idea that communicative exchanges between individuals provide the building blocks of social bonds akin to physical grooming in other primate species. Our data demonstrate that unacquainted individuals can form robust social bonds through private discussions in a shared social context<sup>4</sup>. Interestingly, much like real friendships, gossip partners' affinity towards one another was robust to how each partner actually behaved, indicating that their relationship was based more on their communicative behavior rather than how they played the game.

Our work also demonstrates that when communication serves as a tool for vicarious learning, gossip may align more closely with our folk intuitions, i.e. evaluative discussions of other people. This is readily apparent with the advent of the term "whisper network" popularized by the #MeToo movement, whereby individuals shared private information about dangerous individuals to be avoided. This is consistent with a coalitional view of gossip<sup>24</sup>, in which bonds form between individuals who are engaged in private discourse with each other about an absent third-party. By sharing the experiences of others, gossip enables communities to grow and protect their members

---

<sup>4</sup> As an extreme example, one of our participants used the MTurk reviewing platform, Turktopticon ([turktopticon.ucsd.edu](http://turktopticon.ucsd.edu)) to try to reach out and send a positive message to their communication partner from the game, even after they had finished participating in our study.

from having negative first-hand experiences.

Given the incredible complexity of gossip, future research could benefit by moving beyond folk intuitions about its defamatory and taboo nature. Defamatory comments comprise a relatively small proportion of real world gossip<sup>21</sup> and there are many more features of gossip that contribute to social connection. Creating novel paradigms that involve unrestricted conversations between individuals, and using prospective prediction to test *how* these conversations influence behavior provides a promising approach for future research. Such an approach would also provide the ability to more formally study the independence and interdependence of the various aspects of gossip including impression sharing and vicarious learning. Investigating how humans use gossip as means to establish culture, build trust, learn vicariously, and influence one another is a vast and largely unexplored avenue that can grow our scientific understanding of this ubiquitous behavior.

**Acknowledgements.**

We would like to thank Emma Templeton, Jin Hyun Cheong, Antonia Hoidal, and Alec Smith for helpful feedback about this manuscript. We would also like to thank Andrew Mao and Lili Dworkin for software assistance during initial development of our experiment platform.

**Author contributions.**

E.J. and L.J.C designed the study. E.J. built the experiment platform, collected, and analyzed the data. E.J. and L.J.C. wrote the manuscript.

**Declaration of Interests**

This work is supported by a National Science Foundation CAREER award (1848370). The authors declare no competing interests.



## FIGURE TITLES AND LEGENDS

### Figure 1 | Experimental Design.

**A.** Network structure for each group from the perspective of a focal participant (the cow in this instance). Participants were always able to see the contributions of their immediate neighbors, but could only see contributions from all other players in complete information games. In games with communication, participants were able to exchange messages with a single other player. In incomplete information games, this communication partner acted as a remote neighbor because their visibility of others' contributions did not overlap with the focal participant; in an incomplete information game, the cow would see contributions from the bird and the lion, whereas the bee would instead see contributions from the cat and the pig. **B.** The temporal ordering of events each game round. Private message exchanges (blue) only occurred in games with communication.

### Figure 2 | Visibility of others' behavior changes the makeup of gossip.

Communicative content labels applied to exchanges in games with communication (see methods for details on how coding was performed). When direct observation of other participants was not possible (Incomplete Information), individuals discussed others' behavior significantly more relative to when observation was possible (Complete Information). On the other hand, when social visibility was not restricted, participants discussed a significantly broader range of topics including: casual chit-chat, brief expressions, game mechanics, and inquiries. Much like real-world measurements, casual chit-chat comprised the largest proportion of conversations across both experimental conditions. Data are represented as mean and 95% bootstrapped confidence intervals.  $*p < 0.05$

### Figure 3 | Gossip facilitates vicarious learning in the absence of direction observation

**A.** Time-lagged influence contrasts between incomplete information games with and without communication from the perspective of a focal participant. Each arrow represents the test-statistic on the contrast (communication versus no communication) between how much a specific neighbor's past contribution predicted a focal participant's future contribution. Warmer colors indicate a stronger degree of influence (higher parameter estimate) in games with communication and cooler colors represent a stronger degree of influence in games without communication for that type of neighbor. Despite a lack of direct observation, in games where gossip was possible, the past behaviors of a focal participant's distant and remote neighbors (faded cat, pig, and bee) were more predictive of their future behavior while the past behaviors of their immediate neighbors (bird, lion) were less predictive of their future behavior. **B.** Focal participants and their remote neighbors' affinity ratings of other players were more similar (decreased euclidean distance) in games where gossip was possible. **C.** The relationship between a focal participant's affinity toward each neighbor type neighbor and that neighbor's average contribution: left: remote (bee); middle distant (cat, pig); right immediate (bird, lion). Despite a lack of direct observation this relationship was stronger for games where gossip was possible for both remote (left red;  $b = 0.34$ ) and distant neighbors (middle red;  $b = 0.52$ ), relative to games where gossip was not possible: remote (left blue;  $b = 0.04$ ), distant (middle blue;  $b = 0.14$ ). Immediate neighbors were always visible and never communicated with and therefore demonstrated no difference between game contexts. All error bars and bands represent standard errors (SEM).

### Figure 4 | Gossip increases social connection and influence

**A.** Time-lagged influence contrasts between complete information games with and without communication from the perspective of a focal participant. Each arrow represents the test-statistic on the contrast (communication versus no communication) between how much a specific neighbor's past contribution predicted a focal participant's future contribution. Warmer colors indicate a stronger degree of influence

(higher parameter estimate) in games with communication and cooler colors represent a stronger degree of influence in games without communication for that type of neighbor. In games where gossip was possible, the past behavior of a focal participant's remote neighbor was significantly more predictive of their future behavior, while the past behavior of a focal participant's immediate neighbors were less predictive of their future behavior. **B.** Focal participants and their remote neighbors' affinity ratings of other players were more similar (decreased euclidean distance) in games where gossip was possible. **C.** The relationship between a focal participant's affinity toward their neighbor and their neighbor's average contribution. This relationship was stronger for remote neighbors in games without communication (left blue;  $b = 0.80$ ) and was moderated as a consequence of being able to communicate (left red;  $b = 0.58$ ). No differences were observed for distant or remote neighbors across game contexts. All error bars and bands represent standard error (SEM).

**Figure 5 | The opportunity to gossip increases cooperative group behavior**

**A.** Average group contributions across game rounds (left). Consistent with previous work contributions decline over time across all game types. However, games where gossip is possible (red) show less of decline than games where it is not (blue). These games also show higher cooperative behavior overall, when averaged across game rounds (right; collapsed over complete and incomplete information). No differences were observed between complete or incomplete information conditions or their interaction. All error bars and bands represent standard error (SEM). **B.** Distribution of contribution behavior over rounds displays a different pattern in games where gossip is or is not possible. Contribution behavior converges to a stable bimodal distribution when participants can communicate (left), yet displays a prominent zero contribution skew when they cannot (right).

## TABLES

**Table 1 | Time-lagged influence - Incomplete Information Games.**

Time-lagged mixed effects regression predicting participants' future contribution as a function of their neighbors' past contribution behavior, controlling for their past behavior in incomplete information games. Confidence intervals represent 95% assuming a quadratic log-likelihood surface (Wald method). Bolded results are significant at  $p < 0.05$ .

<b>Incomplete Information Games</b>	<i>b (CI)</i>	<i>se</i>	<i>t</i>	<i>df</i>	<i>p</i>
Participant contribution round <sub>t-1</sub>	0.29 (0.26 - 0.32)	0.01	21.20	3508.49	<b>&lt;.001</b>
Round	-1.02 (-1.32 - -0.71)	0.16	-6.52	434.47	<b>&lt;.001</b>
Gossip Possible	5.69 (2.44 - 8.94)	1.66	3.43	255.21	<b>.001</b>
Remote neighbor round <sub>t-1</sub>	0.06 (0.03 - 0.09)	0.01	4.30	3487.10	<b>&lt;.001</b>
Immediate neighbor round <sub>t-1</sub>	0.28 (0.24 - 0.31)	0.02	14.95	3344.56	<b>&lt;.001</b>
Distant neighbor round <sub>t-1</sub>	0.12 (0.8 - 0.15)	0.02	6.34	3333.10	<b>&lt;.001</b>
Gossip Possible x Remote neighbor round <sub>t-1</sub>	0.07 (0.02 - 0.12)	0.03	2.60	3466.83	<b>0.009</b>
Gossip Possible x Immediate neighbor round <sub>t-1</sub>	-0.07 (-0.14 - 0.002)	0.04	-1.89	2957.40	0.058
Gossip Possible x Distant neighbor round <sub>t-1</sub>	0.10 (0.03 - 0.17)	0.04	2.77	3273.76	<b>0.006</b>
Round x Gossip Possible	0.39 (-0.22 - 0.99)	0.31	1.25	433.38	0.213

**Table 2 | Time-lagged influence - Complete Information Games.**

Time-lagged mixed effects regression predicting participants' future contribution as a function of their neighbors' past contribution behavior, controlling for their past behavior in complete information games. Confidence intervals represent 95% assuming a quadratic log-likelihood surface (Wald method). Bolded results are significant at  $p < 0.05$ .

<b>Complete Information Games</b>	<i>b (CI)</i>	<i>se</i>	<i>t</i>	<i>df</i>	<i>p</i>
Participant contribution round <sub>t-1</sub>	0.28 (0.26 - 0.31)	0.01	19.92	3743.66	<b>&lt;.001</b>
Round	-1.06 (-1.37 - -0.75)	0.16	-6.66	385.70	<b>&lt;.001</b>
Gossip Possible	3.15 (0.10 - 6.28)	1.60	1.97	235.91	.050
Remote neighbor round <sub>t-1</sub>	0.17 (0.14 - 0.19)	0.01	11.75	3616.91	<b>&lt;.001</b>
Immediate neighbor round <sub>t-1</sub>	0.17 (0.13 - 0.21)	0.02	9.10	3481.63	<b>&lt;.001</b>
Distant neighbor round <sub>t-1</sub>	0.20 (0.17 - 0.24)	0.02	10.95	3502.14	<b>&lt;.001</b>
Gossip Possible x Remote neighbor round <sub>t-1</sub>	0.09 (0.04 - 0.15)	0.03	3.40	3486.55	<b>&lt;.001</b>
Gossip Possible x Immediate neighbor round <sub>t-1</sub>	-0.08 (-0.16 - -0.01)	0.04	-2.26	3361.67	<b>0.024</b>
Gossip Possible x Distant neighbor round <sub>t-1</sub>	-0.02 (-0.10 - 0.05)	0.04	-0.67	3326.56	0.502
Round x Gossip Possible	-0.74 (-1.36 - -0.11)	0.32	-2.32	385.52	<b>0.021</b>

## **STAR METHODS**

### **RESOURCE AVAILABILITY**

#### **Lead Contact**

Further information and requests for resources should be directed to and will be fulfilled by the lead contact Eshin Jolly ([eshin.jolly@dartmouth.edu](mailto:eshin.jolly@dartmouth.edu)).

#### **Materials Availability**

This study did not generate new materials or reagents.

#### **Data and Code Availability**

The custom online platform to run the experiment is available at:

[https://github.com/cosanlab/PGG\\_meteor](https://github.com/cosanlab/PGG_meteor). Data and analysis code that support the findings, and a preprint of this manuscript are available at: <https://psyarxiv.com/qau5s/>

## EXPERIMENTAL MODELS AND SUBJECT DETAILS

### Participants (main experiment)

2373 individuals were recruited from Amazon's Mechanical Turk marketplace<sup>42</sup> and consented to participation in accordance with the Committee for the Protection of Human Subjects (CPHS) at Dartmouth College. Because of the interactive nature of the experiment, it was critical to recruit groups of participants that fully understood the experiment, had no technical difficulties (i.e. were able to behave interactively in real-time), and did not leave the experiment prematurely. To ensure these criteria were met, we employed a rigorous vetting procedure as follows. Each participant was presented with instructions based on their random condition assignment, a required comprehension quiz, and a 10 minute waiting "room" during which time they waited to be matched with five other participants. In order for participants to enter this waiting period, they were given a maximum of two attempts to pass the required comprehension quiz and were otherwise prohibited from participating. Given both the synchronous nature of this experiment and the diminished experimenter control present when running online studies<sup>43,44</sup>, several sources of participant attrition existed for this experiment. Overall, 460 participants disconnected before attempting the comprehension quiz, 278 participants failed the comprehension quiz, 275 were unable to be matched with a complete group of participants within the allotted waiting period, and 406 experienced other connection or technical issues while playing resulting in failure to complete all 10 rounds of the experiment.

As a result, a smaller subset of 954 participants (382 females; 79 not indicated;  $M_{\text{age}} = 33.74$  years,  $SD = 9.80$  years) successfully completed the screening quiz, matched with other participants, and completed all rounds of the experiment and are therefore included in all reported analyses. Individuals participated in groups of six for the experiment, and were paid \$1.50 as a baseline payment and earned a bonus (\$4.50 max) based on the decisions made by the group ( $M = \$2.65$ ;  $SD = \$0.62$ ). Games lasted an average of 12 minutes, yielding an average total hourly pay-rate of \$20.75. Of these remaining individuals, 12.4% contained repeated Amazon Worker

Ids from real individuals. While this introduces a minimal amount of repeat participation, we do not believe that this substantially impacts our findings for two principal reasons. First, unlike non-synchronous or even dyadic experiments, any single individual within a group of six, has a limited ability to influence the behavior of all other group members. Second, previous literature has demonstrated that workers in the Mechanical Turk marketplace are well versed and highly experienced with related economic games (e.g. public goods, prisoner's dilemma, ultimatum game) especially compared to typical in lab subjects, and yet this does not diminish the replicability of research findings<sup>45,46</sup>. If anything, it might decrease the estimates of expected effects<sup>47</sup>. Further, our analyses were restricted to games that include complete data for 10 rounds of play. Unlike previous work, we believe this to be a more conservative approach than choosing random actions on behalf of disconnected players or assuming the role of a participant ourselves and taking preset actions as in some previous work<sup>13</sup>.

While formal power analyses were not conducted prior to data collection, the authors were sensitive to issues surrounding small sample sizes and underpowered inferences in psychological science. For this reason, to improve uncertainty estimates and achieve adequate power, data collection aimed to achieve a useable sample size between three to four times that of similar previous research ( $N = 954$ ) (Feinberg et al, 2014,  $N = 216$ ; Wu et al, 2016,  $N = 265$ ; Fehr & Gächter, 2002,  $N = 240$ ; Fowler & Christakis, 2010;  $N = 240$ ).

### **Participants (message labeling)**

To label communicative content from messages in the main experiment, an additional sample of 1,454 individuals was recruited from Amazon's Mechanical Turk marketplace during June of 2016 to read and apply pre-specified labels to each message sent during the main experiment. In order to ensure quality and consistency of labeling, individuals were provided with detailed instructions that included background context for the mechanics of a public-goods-game, as well as information about how to use each of 10 content labels. Individuals were also required

to pass a comprehension quiz that ensured they understood the meaning of each label type and were required to provide a complete set of labels for every message in a single game. As a result of this rigorous procedure, only 381 individuals (179 females; 2 not indicated;  $M_{\text{age}} = 33.09$  years,  $SD = 10.16$  years), successfully consented, passed our comprehension quiz, and provided a complete set of labels for a given game. Each game was therefore labeled by several different individuals ( $M = 4.94$ ,  $SD = 0.87$ ). Participants were paid a base payment (\$0.50) for accepting the HIT and attempting the comprehension quiz, and were paid a bonus (\$2.50) for completing the full labeling task.



## METHOD DETAILS

### Main experiment

**Overview.** Each group of six participants, played one of four variants of an iterated, 10 round public-goods style game. Unlike previous work, games did not employ between-round rematching or shuffling<sup>15,48</sup>. Because the focus of this experiment was on modeling real interactions between individuals, it was necessary to build custom software that enabled synchronous interactions between multiple participants. In order to achieve this, the experiment was built using the MeteorJS open-source framework in conjunction with TurkServer<sup>49</sup>, an open-source platform for creating web-based behavioral experiments and interfacing with Mturk. Due to the inevitability of connection and participant drop-out issues that arise during synchronous group experiments, all analyses were only conducted on games with 10 complete rounds of data and no participant dropouts<sup>50</sup>.

**Design.** Each participant was randomly assigned to one of four experimental conditions varying along two crossed dimensions: information (complete/incomplete), communication (gossip/no gossip). Participants were first presented with a series of instructions describing the basic rules of a public goods game<sup>32</sup> and asked to pass a comprehension quiz demonstrating they understood the rules of the game and the details of the condition in which they were participating. Only participants who passed the comprehension quiz within two attempts were permitted to participate. Games began when 6 eligible participants were found. After participants completed all 10 rounds, they rated their desire for repeat play (affinity) with every other participant in their group.

At the start of the game, participants were given animal avatars (used from: <https://github.com/niavlys/memoryKivy>) to help them more easily identify themselves and other individuals during the experiment. Each group of participants was organized as a static “ring network” and proceeded in one of four variants. In the complete information without gossip variant ( $N_{\text{Games}} = 41$ ;  $N_{\text{Participants}} = 246$ ), participants played a canonical public goods game as described

below. In the incomplete information without gossip variant ( $N_{\text{Games}} = 41$ ;  $N_{\text{Participants}} = 246$ ), participants were only able to see the contributions made by their immediate neighbors, although group earnings were computed based on all participants' contributions. In the complete information with gossip variant ( $N_{\text{Games}} = 37$ ;  $N_{\text{Participants}} = 222$ ), participants were able to send and receive two 140 character private messages to their remote neighbor once per round after seeing other individuals' contributions. Unlike previous research, participants were free to say anything they wanted (or nothing at all) and always communicated with the same individual across all rounds<sup>12,13,15</sup>. Critically, these messages were private dyadic communication observable only to communicators and not other group members. Finally, in the incomplete information with gossip variant ( $N_{\text{Games}} = 40$ ;  $N_{\text{Participants}} = 240$ ), participants were only able to see contributions made by their immediate neighbors, but were still able to privately communicate with their remote neighbor, despite being unable to see their behavior.

Therefore, from the perspective of a focal participant (cow) (Fig. 1A), each group member fell into one of three different types of neighbors: (a) immediate - those whose contribution behavior was always directly observed (bird and lion); (b) distant - those whose contribution behavior could only be observed during complete information games (cat and pig); (c) remote - those whose contribution behavior could only be observed during complete information games and with whom they communicated during games that permitted communication (bee). Using this group structure, remote neighbors (i.e. communication partners in games where gossip was possible) were set up such that their immediate and distant neighbors were reversed: the immediate neighbors of one participant were the distant neighbors of another, and vice versa. Upon each game's conclusion, participants rated their affinity for every other participant by indicating their desire for repeat play with each of these neighbors (Fig. 1B).

**Public Goods Game.** Each game consisted of 10 rounds with the same basic structure (Fig 1B): 1) participants received an initial endowment of 100 points and decided how many points to contribute to a group account versus keep for themselves, 2) participants saw the decisions made

by other individuals in the group based on their condition, 3) players saw the earnings of other individuals after contributions were summed, multiplied by 1.5 and evenly distributed to all players.

4) If a game included communication, participants sent and received two messages (one at a time) from another participant. At the end of the game a random round was selected to calculate participant bonuses using an exchange rate of 1 point = 2 cents.

### **Communicative Content Labeling**

Individuals were provided with a set of 10 labels that could be applied to each message in an exchange and were instructed to select as many labels as they wanted to categorize the communication content (one minimum). To facilitate this, all messages were laid out as a conversation in the order they were generated during the game for each communicating dyad from the main experiment. That is, each participant saw a 40 message exchange (10 rounds, 2 messages per individual) laid out together on a single page and labeled each message individually. At any point during this task, individuals were easily able to remind themselves of the meaning of each label by viewing a clickable reference. The following 10 labels were provided for individuals to use, along with a description of their meaning:

- 1) **Discuss Others** - When an individual is discussing information about other individuals.  
This may include how many points they contributed, or may include passing judgment on their behavior. Sometimes this may simply be the number of points contributed by another person with no other words.
- 2) **Inquiry** - When an individual is asking a question to the other person.
- 3) **Affirmation** - When an individual replies to a question or statement in a confirmatory way.
- 4) **Strategy** - When an individual is discussing game tactics such as: what they think the best action is in the game, what action they are planning to take, or what action the other individual ought to take.

- 5) **Chit-chat** - 'Getting-to-know-each-other' types of messages or playful social exchanges.

This might be things like finding out more about the other individual or trying to relate to them in some way.

- 6) **Game mechanics** - Messages remarking on some aspect of the game setup such as the number of rounds left, what the purpose of this game/experiment is, comments about the requester or experimenter and questions clarifying how to play (but not explicitly strategic questions).

- 7) **Nonsense** - Messages that seem like random strings of letters or numbers or seem to be completely random and meaningless.

- 8) **Brief expressions** - Brief one or two word messages that include things like emoticons, exclamations, laughing, abbreviations and chat slang.

- 9) **Random** - Messages that don't fit into one of the previous labels. We have found that most messages can be described using a combination of the previous labels, but for some messages that may not be true. If so please use this label.

- 10) **No Text** - Messages that are completely blank. Occasionally some individuals sent no message to the other person.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Overview

All data were organized and restructured using the python data analysis library (pandas)<sup>51</sup> and visualizations were made using plotnine<sup>52</sup>. All statistics were computed using the R statistical package<sup>53</sup> and scientific tools implemented in python<sup>54</sup>. All linear mixed models were fit using the lme4 package<sup>55</sup> and corrected marginal contrasts (i.e. minimizing family-wise-error rates) were computed using the lsmeans package<sup>56</sup> both implemented via Pymer4<sup>57</sup>. P-values for linear mixed-effects models were computed using the lmerTest package<sup>58</sup> via Satterthwaite approximation for degrees of freedom calculations, which has been demonstrated to produce reliable Type 1 error rates<sup>59</sup>.

### Behavioral analyses

***Time-lagged analyses.*** Time-lagged effects were estimated using linear mixed-effects models separately for games with complete and incomplete information. Each model estimated an individual's contribution on Round<sub>t+1</sub>, as a function of each of their neighbor's contributions at Round<sub>t</sub>, controlling for their own contribution at Round<sub>t</sub>, along with a contrast code for communication. That is, each model contained fixed-effects for centered game round, participants' centered contribution at Round<sub>t</sub>, a categorical contrast code for whether gossip was or was not possible, the centered contributions of each of three neighbors (immediate, distant, and remote) at Round<sub>t</sub>, along with interaction terms between neighbor contributions, round, and communication (see full model notation below). Prior to modeling, contributions for each pair of immediate and distant neighbors were averaged such that only three types of neighbors were entered into the model. To model variance across participants, random intercepts and slopes across game rounds, were estimated for each individual. More complex random-effects structures yielded unstable model estimates that failed to achieve convergence during estimation. Parameterized in this way, the interaction terms between specific neighbor contributions and

communication represented the contrast between games with and without gossip. These models were estimated in a spirit similar to that of <sup>48</sup>. Specifically we used the following specification:

$$\begin{aligned} \widehat{\text{Contribution}}_{s,t+1} = & (\beta^0 + \gamma_s^0) + \\ & (\beta^1 + \gamma_s^1) * \text{Round} + \\ & \beta^2 \text{Contribution}_{s,t} + \\ & \beta^3 \text{Gossip Possible} + \\ & \beta^4 \text{Comm Partner Contribution}_{s,t} + \\ & \beta^5 \text{Imm Neighbor Contribution}_{s,t} + \\ & \beta^6 \text{Dist Neighbor Contribution}_{s,t} + \\ & \beta^7 \text{Round} * \text{GossipPossible} + \\ & \beta^8 \text{Comm Partner Contribution} * \text{Gossip Possible} + \\ & \beta^9 \text{Imm Neighbor Contribution} * \text{Gossip Possible} + \\ & \beta^{10} \text{Dist Neighbor Contribution} * \text{Gossip Possible} + \\ & \epsilon_{t+1} + \\ & \epsilon_s \end{aligned} \tag{eq 1}$$

$$\epsilon_{t+1} \sim \mathcal{N}(0, \sigma_{t+1}^2) \tag{eq 2}$$

$$\epsilon_s \sim \mathcal{N}(0, \sigma_s^2) \tag{eq 3}$$

$$(\gamma_s^0, \gamma_s^1) \sim N \left( 0, \begin{bmatrix} \tau^0 & \rho^{01} \\ \rho^{10} & \tau^1 \end{bmatrix} \right) \tag{eq 4}$$

$$\tau^n = \text{var}(\gamma^n) \tag{eq 5}$$

$$\rho^{nm} = \text{cov}(\gamma^n, \gamma^m) \tag{eq 6}$$

Where  $n$  and  $m$  refer to index of the parameter (0 = intercept, 1 = slope)

Eq1 predicts the contribution made by participant  $s$  on round  $t + 1$  ( $\widehat{\text{Contribution}}_{s,t+1}$ ) using a time-lagged mixed-effects regression model with random intercepts and slopes for each subject.

$\gamma_s^0$  reflects the deviation of subject  $s$ 's intercept from the group fixed effect intercept  $\beta^0$ , which accounts for the fact that each participant differed in their mean contribution (i.e. a *random-intercept*).

$\gamma_s^1$  reflects an individual participant's *deviation* from the fixed effect  $\beta^1$ , which accounts for the fact that the *change* in each participant's contribution differed over time (game rounds) (i.e. a *random-slope*). These parameters are estimated by the model (i.e. *random-effects*) assuming a multivariate normal distribution with mean 0, and variance of random-intercepts, variance of random-slopes, and co-variance between them (eq 4).

$\beta^2$  controls for participant  $s$ 's contribution on the previous round  $t$ , i.e. their *past* behavior.

$\beta^3$  is a contrast code reflecting the *difference* between game types depending on the particular model being estimated, i.e. full information gossip possible - full information gossip not possible, or incomplete information gossip possible - incomplete information gossip not possible.

$\beta^4, \beta^5, \beta^6$  capture the degree each neighbor's *past* behavior in round  $r$  influenced participant  $s$ 's future behavior on round  $t + 1$ .

$\beta^7$  captures *differences* between how contributions change over time as a function of whether or not gossip was possible.

$\beta^8, \beta^9, \beta^{10}$  capture interactions between neighbor influence terms and game type, reflecting the *difference* in influence a particular neighbor had on participant  $s$ 's future behavior between games where gossip was or was not possible.

**Affinity.** Subsequent post-game affinity ratings were modeled in three ways. Prior to modeling, ratings for each pair of immediate and distant neighbors were averaged such that neighbor role was a factor containing only three levels (one for each neighbor type). First, to test for overall differences in affinity, a linear mixed effects model was estimated in which ratings were modeled as a function of orthogonally coded fixed effects of game type (four levels), neighbor type (three levels), and their interaction. To model variance across games and participants in each analysis, random intercepts were estimated for individual games and individual participants. More complex random-effects structures yielded unstable model estimates that failed to achieve convergence during estimation. Analysis of Variance tests were computed on each fixed effect to estimate overall differences between factor levels. To examine differences between specific neighbor types, least-squares pairwise comparisons were conducted and corrected for multiple-comparisons at a false-discovery rate of  $q = 0.05$ . A full list of these comparisons is available in Table S1.

Second, to test for interactions with behavior, a separate linear mixed effects model was estimated in which ratings were modeled as a function of fixed effects for centered average

individual contributions across all game rounds, game type, neighbor type, and the interaction between all three. To model variance across games and participants in each analysis, random intercepts were estimated for individual games and individual participants. Analysis of Variance tests were computed on each fixed effect to estimate overall differences between factor levels. Subsequently, specific least-squares pairwise comparisons were conducted, correcting for multiple-comparisons by controlling familywise error rates at  $\alpha = 0.05$ . A full list of these comparisons is available in Table S2.

To test for inter-individual similarity of affinity ratings, we computed the similarity between affinity ratings made by focal participants' and the same ratings made by their remote neighbors. Specifically, we computed the euclidean distance between vectors of length four for each pair of such participants. For focal participants, this vector included affinity ratings of their distant neighbors (left and right) and their immediate neighbors (left and right). For the matching remote neighbor, this vector included affinity ratings of their immediate neighbors (right and left) and their distant neighbors (right and left). We used this distance metric because (a) it reflects interpretable values on the original 100 point rating scale used by participants (b) it is calculable and interpretable even when participants rate all neighbors identically (i.e. if vectors have zero variance correlation distance cannot be computed). Distance scores were contrasted between games where communication was and was not possible separately for full information games and incomplete information games. Inference was performed via permuting group labels and recomputing means 5,000 times to build null distributions. We also examined the relationship between inter-individual similarity of affinity ratings and remote neighbors ratings of each other. Specifically, we correlated the mean affinity of each pair of remote neighbors with their inter-individual similarity ratings (reported in main text), as well as the absolute value difference of affinity of each pair of remote neighbors with their inter-individual similarity ratings. These analyses were performed separately for incomplete (Fig. S3) and complete information games (Fig S4) and all inference was performed via permuting each vector of inter-individual similarity



ratings and re-running the correlation 5,000 times to build null distributions.

**Group Contribution Behavior.** Group contributions were modeled in two ways. First, we estimated using a linear mixed effects model in which contributions were modeled as a function of fixed effects for centered game round, categorical contrast codes for gossip possible and information, and interactions between all three. To model variance across games and participants, random intercepts and round slopes were estimated for individual games and individual participants. A full list of model estimates can be found in Table S5.

Second we sought to test that groups in which participants received more gossip, were more cooperative on average (i.e. averaging over game rounds and individuals with a group). To perform this analysis we fit a multi-level model with random intercepts and slopes for each group predicting average contribution behavior (i.e. an individual's contributions averaged over time) from: (a) the group-centered average amount of gossip they received; (b) an indicator for full or incomplete information context; (c) their interaction. Group-centered gossip was computed by first calculating the proportion of all messages *received* by a participant that were labeled as gossip and then mean-centering within-group. Gossip messages were determined based upon content labeling performed by an independent sample as described in the methods.

### Message content analyses

**Annotation reliability.** To compute the reliability of annotations we took two approaches. First, we computed Krippendorff's alpha <sup>60</sup> across all messages and individuals using the approach outlined by Hayes & Krippendorff <sup>61</sup> implemented in the `simplendorff` Python package (<https://github.com/LightTag/simplendorff>). To perform inference testing on this statistic, we used a bootstrap procedure in which we randomly resampled individual raters with replacement 5,000 times and recomputed alpha for each bootstrapped sample. 95% confidence intervals were estimated from the distribution of 5,000 alpha statistics using percentile cutoffs. Krippendorff's

alpha suggested a significant, moderate level of inter-rater reliability across annotations:  $\alpha = 0.503$  [0.489 0.517].

This metric allows us to account for different numbers of raters and labels per message, but it does not consider that the *absence* of a message label may also be informative about its content. To take this into consideration, we also computed the pairwise Simple-Matching-Coefficient (Rand similarity) between annotator's labels, separately for each message and averaged these similarity scores for each message ( $M = 0.893$ ,  $SD = 0.071$ , [0.892 0.895]). Together these two metrics indicate a moderate to high levels of agreement across our raters.

**Comparing communicative content.** Given the reliability of our annotations, we calculated an index of communicative content for each message by counting how often each label was applied to any given message and converting this to a proportion, while aggregating across responses from all individuals. We then applied the label comprising the highest proportion as the label for a particular message. In this way, the label for each message reflects the *modal* label (that with the highest agreement used most often) across annotators. Finally, using these labels, we computed the proportion of each label occurrence across all messages and games and submitted these values to a log-linear mixed-effects model (multilevel Poisson model). This model contained orthogonal fixed-effects contrasts for information, labels, and their interaction and random intercepts per participant. To examine differences between specific labels, least-squares pairwise comparisons were conducted, correcting for multiple-comparisons by controlling false discovery rate (FDR) at  $q < 0.05$ . This allowed us to capture the multifaceted content that may comprise gossip and observe how this content changes based on situational factors.

## SUPPLEMENTAL ITEM TITLES AND LEGENDS

**Figure S1 | Post-game affinity.** *Related to Fig 3C and 4C. Mean differences between overall affinity ratings for each neighbor type in all games. Participants felt the most affinity towards their remote neighbors (communication partner) in games where gossip was possible, relative to all other neighbors across all game types. See Table S1 for estimates. Error bars represent standard errors (SEM).*

**Figure S2 | Post-game affinity and contribution behavior.** *Related to Fig 3C and 4C. Regression slope differences of the relationship between affinity and contribution across all game types. Higher values indicate that participants' affinity ratings were more correlated with a neighbor's behavior (contributions) during play. See Table S2 for estimates. Error bars represent standard errors (SEM).*

**Figure S3 | Impression Similarity - Incomplete Information.** *Related to Fig. 3B. Correlation between (i) the similarity (euclidean distance) of a participant and their communication partner's (remote neighbor) affinity for other players and; (ii) after (or absolute difference) affinity for each other. **X-axis:** Mean affinity (blue) is the average between a participant's affinity for their communication partner and their communication partner's affinity for them on a 100 point scale. Absolute difference is simply the absolute value of the difference between their ratings. **Y-axis:** The euclidean distance between a participant's affinity for all other players excluding their communication partner, and their communication partner's affinity for all other players excluding the participant. Conversants who rated other players (i.e. distant and immediate neighbors) more similarly (lower euclidean distance y-axis) felt more positively about each on average (orange x-axis) and felt more similar (lower absolute difference) about each other (blue, x-axis). Error bands represent bootstrapped 95% confidence intervals.*

**Figure S4 | Impression Similarity - Complete Information.** *Related to Fig. 4B. Correlation between (i) the similarity (euclidean distance) of a participant and their communication partner's (remote neighbor) affinity for other players and; (ii) after (or absolute difference) affinity for each other. **X-axis:** Mean affinity (blue) is the average between a participant's affinity for their communication partner and their communication partner's affinity for them on a 100 point scale. Absolute difference is simply the absolute value of the difference between their ratings. **Y-axis:** The euclidean distance between a participant's affinity for all other players excluding their communication partner, and their communication partner's affinity for all other players excluding the participant. Conversants who rated other players (i.e. distant and immediate neighbors) more similarly (lower euclidean distance y-axis) felt more positively about each on average (orange x-axis) and felt more similar (lower absolute difference) about each other (blue, x-axis). Error bands represent bootstrapped 95% confidence intervals.*

**Table S1. Post-game affinity.** *Related to Methods - Affinity. Least-squares pairwise comparisons of mean post-game affinity ratings. Results are corrected for multiple comparisons using FDR correction ( $q = 0.05$ ) and bolded results are those that survive correction. Only pairwise comparisons between different neighbor types within the same game or the same neighbor type across different games were made C = complete information; I = incomplete information; G = gossip possible; NG = gossip not possible*

**Table S2. Post-game affinity and contribution behavior.** *Related to Methods - Affinity. Least-squares pairwise comparisons of the regression slope between post-game affinity ratings and average contribution. Results are corrected for multiple comparisons using FDR correction ( $q = 0.05$ ) and bolded results are those that survive correction. Only pairwise comparisons between different neighbor types within the same game or the same neighbor type across different games were made. C = complete information; I = incomplete information; G = gossip possible; NG = gossip not possible*

**Table S3. Message topic frequency comparisons.** *Related to Fig 2. Least-squares pairwise comparisons of the frequency of discussing particular topics in games with Incomplete or Complete information. Estimates reflect the difference of Incomplete Information games > Completion Information games estimated using a mixed effect log-linear model on topic counts. Results are corrected for multiple comparisons by controlling the False-Discovery-Rate at  $q < 0.05$  and bolded results are those that survive correction.*

**Table S4. Time-lagged influence - Games where Gossip is possible.** *Related to Table 1 and Table 2. Time-lagged mixed effects regression predicting participants' future contribution as a function of their neighbors' past contribution behavior, controlling for their past behavior. Confidence intervals represent 95% assuming a quadratic log-likelihood surface (Wald method). Estimates reflect the contrast between Full Information games > Incomplete Information games. Complementary to other models, the interaction between Distant Neighbor and Information is not significant suggesting that distant neighbors had similar amounts of influence on participants' behavior when gossip was possible, even when they could not be seen. Bolded results are significant at  $p < 0.05$ .*

**Table S5. Group Cooperation.** *Related to Methods - Group Contribution Behavior. Mixed effects regression estimates for group contributions as a function of round, gossip and information. Confidence intervals represent 95% assuming a quadratic log-likelihood surface (Wald method). Bolded results are significant at  $p < 0.05$ .*

## REFERENCES

1. Mehl, M.R., Vazire, S., Ramírez-Esparza, N., Slatcher, R.B., and Pennebaker, J.W. (2007). Are women really more talkative than men? *Science* 317, 82.
2. Dunbar, R.I. (2004). Gossip in evolutionary perspective. *Rev. Gen. Psychol.* 8, 100–110.
3. Dahmardeh, M., and Dunbar, R.I.M. (2017). What Shall We Talk about in Farsi? *Human Nature* 28, 423–433.
4. Dunbar, R.I., Marriott, A., and Duncan, N.D. (1997). Human conversational behavior. *Hum. Nat.* 8, 231–246.
5. Emler, N. (1994). Gossip, reputation, and social adaptation. In Good gossip, R. F. Goodman and A. Ben-Ze'ev, eds. (University Press of Kansas), pp. 117–138.
6. Dunbar, R.I. (1998). *Grooming, gossip, and the evolution of language* (Harvard University Press).
7. Stirling, R.B. (1956). Some psychological mechanisms operative in gossip. *Soc. Forces*.
8. Foster, E.K. (2004). Research on gossip: Taxonomy, methods, and future directions. *Rev. Gen. Psychol.* 8, 78–99.
9. Baumeister, R.F., Zhang, L., and Vohs, K.D. (2004). Gossip as cultural learning. *Rev. Gen. Psychol.* 8, 111–121.
10. Wert, S.R., and Salovey, P. (2004). Introduction to the Special Issue on Gossip. *Rev. Gen. Psychol.* 8, 76–77.
11. Emler, N. (1990). A Social Psychology of Reputation. *European Review of Social Psychology* 1, 171–193.
12. Sommerfeld, R.D., Krambeck, H.-J., Semmann, D., and Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17435–17440.
13. Wu, J., Balliet, D., and Van Lange, P.A.M. (2016). Gossip Versus Punishment: The Efficiency of Reputation to Promote and Maintain Cooperation. *Sci. Rep.*, 1–8.
14. Wu, J., Balliet, D., and Van Lange, P.A.M. (2015). When Does Gossip Promote Generosity? Indirect Reciprocity Under the Shadow of the Future. *Soc. Psychol. Personal. Sci.* 6, 923–930.
15. Feinberg, M., Willer, R., and Schultz, M. (2014). Gossip and Ostracism Promote Cooperation in Groups. *Psychol. Sci.* 25, 656–664.
16. Milinski, M. (2016). Reputation, a universal currency for human social interactions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371, 20150100.
17. Beersma, B., and Van Kleef, G.A. (2011). How the Grapevine Keeps You in Line: Gossip Increases Contributions to the Group. *Soc. Psychol. Personal. Sci.*

18. Feinberg, M., Willer, R., Stellar, J., and Keltner, D. (2012). The virtues of gossip: Reputational information sharing as prosocial behavior. *J. Pers. Soc. Psychol.* 102, 1015–1030.
19. Piazza, J., and Bering, J.M. (2008/5). Concerns about reputation via gossip promote generous allocations in an economic game. *Evol. Hum. Behav.* 29, 172–178.
20. Giardini, F., and Conte, R. (2012). Gossip for social control in natural and artificial societies. *Simulation* 88, 18–32.
21. Robbins, M.L., and Karan, A. (2019). Who Gossips and How in Everyday Life? *Soc. Psychol. Personal. Sci.*, 1948550619837000.
22. Bloom, P. (2004). Postscript to the special issue on gossip. *Rev. Gen. Psychol.* 8, 138–140.
23. Tamir, D.I., and Mitchell, J.P. (2012). Disclosing information about the self is intrinsically rewarding. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8038–8043.
24. Wittek, R., and Wielers, R. (1998). Gossip in Organizations. *Comput. Math. Organ. Theory* 4, 189–204.
25. Shaw, A.K., Tsvetkova, M., and Daneshvar, R. (2011). The effect of gossip on social networks. *Complexity* 16, 39–47.
26. Ellwardt, L., Steglich, C., and Wittek, R. (2012). The co-evolution of gossip and friendship in workplace social networks. *Soc. Networks* 34, 623–633.
27. Wert, S.R., and Salovey, P. (2004). A social comparison account of gossip. *Rev. Gen. Psychol.* 8, 122–137.
28. Rosnow, R.L. (2001). Rumor and gossip in interpersonal interaction and beyond: A social exchange perspective. In *Behaving badly: Aversive behaviors in interpersonal relationships*, (pp. R. M. Kowalski, ed. (American Psychological Association, xiii), pp. 203–232.
29. Rossignac-Milon, M., Bolger, N., Zee, K.S., Boothby, E.J., and Higgins, E.T. (2020). Merged minds: Generalized shared reality in dyadic relationships. *J. Pers. Soc. Psychol.*
30. Rosnow, R.L. (1991). Inside rumor: A personal journey. *American Psychologist* 46, 484–496.
31. Schachter, S., and Burdick, H. (1955). A field experiment on rumor transmission and distortion. *J. Abnorm. Soc. Psychol.* 50, 363–371.
32. Fehr, E., and Gächter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140.
33. Chaudhuri, A. (2010). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Exp. Econ.* 14, 47–83.
34. Balliet, D. (2010). Communication and Cooperation in Social Dilemmas: A Meta-Analytic Review. *J. Conflict Resolut.* 54, 39–57.
35. Nevo, O., Nevo, B., and Derech-Zehavi, A. (1994). The tendency of gossip as a psychological disposition: Constructing a measure and validating it. In *Good gossip*, R. F.

Goodman and A. Ben-Ze'ev, eds. (University Press of Kansas), pp. 180–189.

36. Hannerz, U. (1967). Gossip, networks and culture in a black American ghetto. *Ethnos* 32, 35–60.
37. Merry, S.E. (1997). Rethinking gossip and scandal. *Reputation: Studies in the voluntary elicitation of good conduct* 47.
38. Echterhoff, G., Higgins, E.T., and Levine, J.M. (2009). Shared Reality: Experiencing commonality with others' inner states about the world. *Perspect. Psychol. Sci.* 4.
39. Baumeister, R.F., and Leary, M.R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117, 497.
40. Feinberg, M., Cheng, J.T., and Willer, R. (2012). Gossip as an effective and low-cost form of punishment. *Behav. Brain Sci.* 35, 25.
41. Milinski, M., Semmann, D., and Krambeck, H.-J. (2002). Reputation helps solve the "tragedy of the commons." *Nature* 415, 424–426.
42. Buhrmester, M., Kwang, T., and Gosling, S.D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspect. Psychol. Sci.* 6, 3–5.
43. Rand, D.G., Arbesman, S., and Christakis, N.A. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proc. Natl. Acad. Sci. U. S. A.* 108, 19193–19198.
44. Paolacci, G., Chandler, J., and Ipeirotis, P.G. (2010). Running Experiments on Amazon Mechanical Turk. *Judgement and Decision Making* 5.
45. Rand, D.G. (2012). The promise of Mechanical Turk: how online labor markets can help theorists run behavioral experiments. *J. Theor. Biol.* 299, 172–179.
46. Suri, S., and Watts, D.J. (2011). Cooperation and contagion in web-based, networked public goods experiments. *PLoS One* 6, e16836.
47. Rand, D.G., Peysakhovich, A., Kraft-Todd, G.T., Newman, G.E., Wurzbacher, O., Nowak, M.A., and Greene, J.D. (2014). Social heuristics shape intuitive cooperation. *Nat. Commun.* 5, 3677.
48. Fowler, J.H., and Christakis, N.A. (2010). Cooperative behavior cascades in human social networks. *Proceedings of the National Academy of Sciences* 107, 5334–5338.
49. Parkes, D.C., Mao, A., Chen, Y., Gajos, K.Z., Procaccia, A., and Zhang, H. (2012). Turkserver: Enabling synchronous and longitudinal online experiments. In *Proceedings of the Fourth Workshop on Human Computation (HCOMP'12)* (AAAI Press).
50. Thomas, K.A., and Clifford, S. (2017). Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Comput. Human Behav.* 77, 184–197.
51. McKinney, W. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, pp. 51–56.
52. Kibirige, H., Lamp, G., Katins, J., O., A., gdowding, Funnell, T., matthias-k, Arnfred, J.,

- Blanchard, D., Chiang, E., et al. (2018). has2k1/plotnine: v0.5.1.
53. R Core Team (2013). R: A Language and Environment for Statistical Computing.
  54. Jones, E., Oliphant, T., and Peterson, P. (2001--). {SciPy}: Open source scientific tools for {Python}.
  55. Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1--7. 2014. Institute for Statistics and Mathematics of WU website. <http://CRAN.R-project.org/package=lme4>. Accessed March 18.
  56. Lenth, R.V. (2016). Least-squares means: the R Package lsmeans. J. Stat. Softw. 69, 1–33.
  57. Jolly, E. (2018). Pymer4: Connecting R and Python for Linear Mixed Modeling. JOSS 3, 862.
  58. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest package: tests in linear mixed effects models. J. Stat. Softw. 82.
  59. Luke, S.G. (2017). Evaluating significance in linear mixed-effects models in R. Behav. Res. Methods 49, 1494–1502.
  60. Krippendorff, K. (2011). Agreement and Information in the Reliability of Coding. Commun. Methods Meas. 5, 93–112.
  61. Hayes, A.F., and Krippendorff, K. (2007). Answering the Call for a Standard Reliability Measure for Coding Data. Commun. Methods Meas. 1, 77–89.



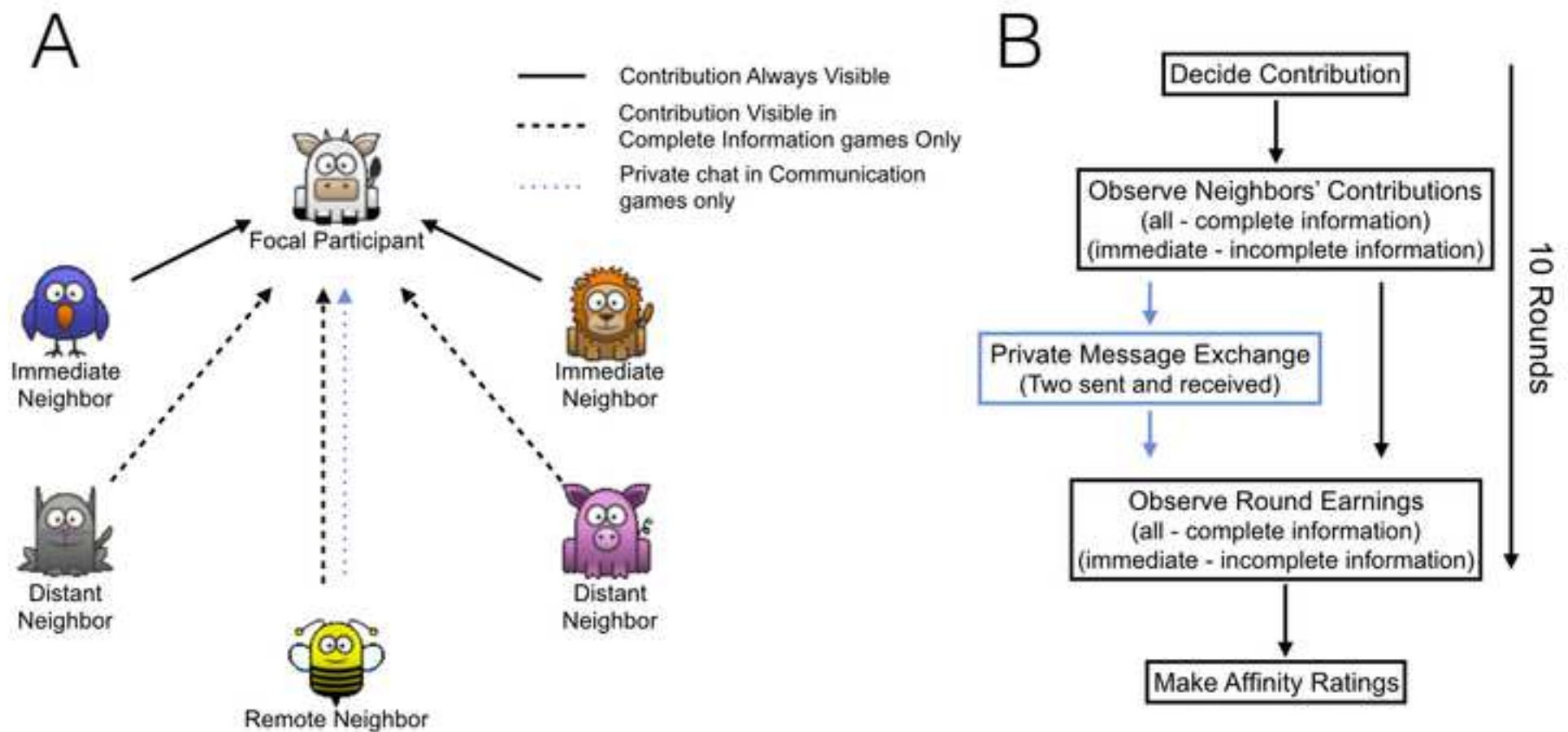


Figure2

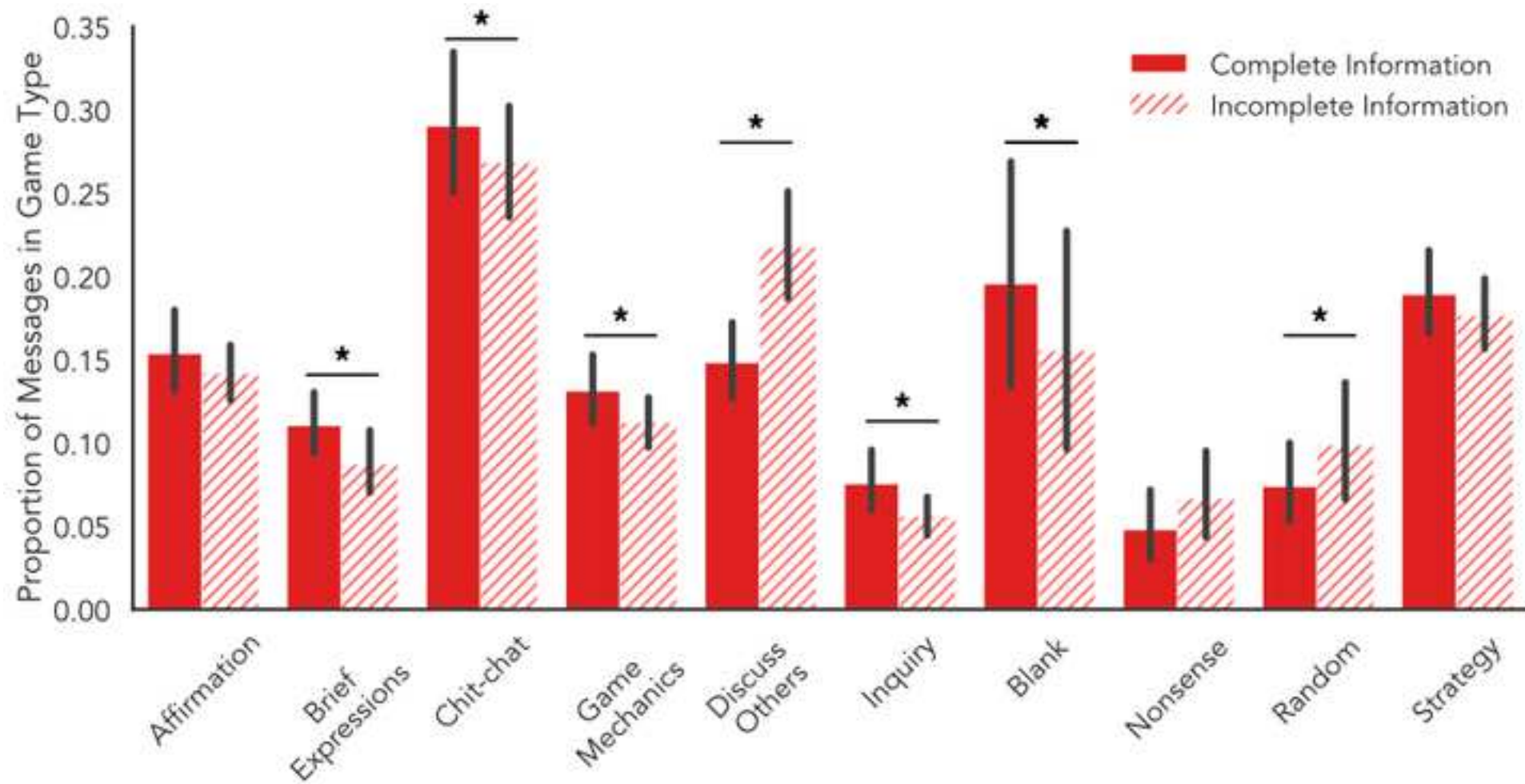
[Click here to access/download;Figure;Fig2.tiff](#)

Figure3

[Click here to access/download;Figure;Fig3.tiff](#)

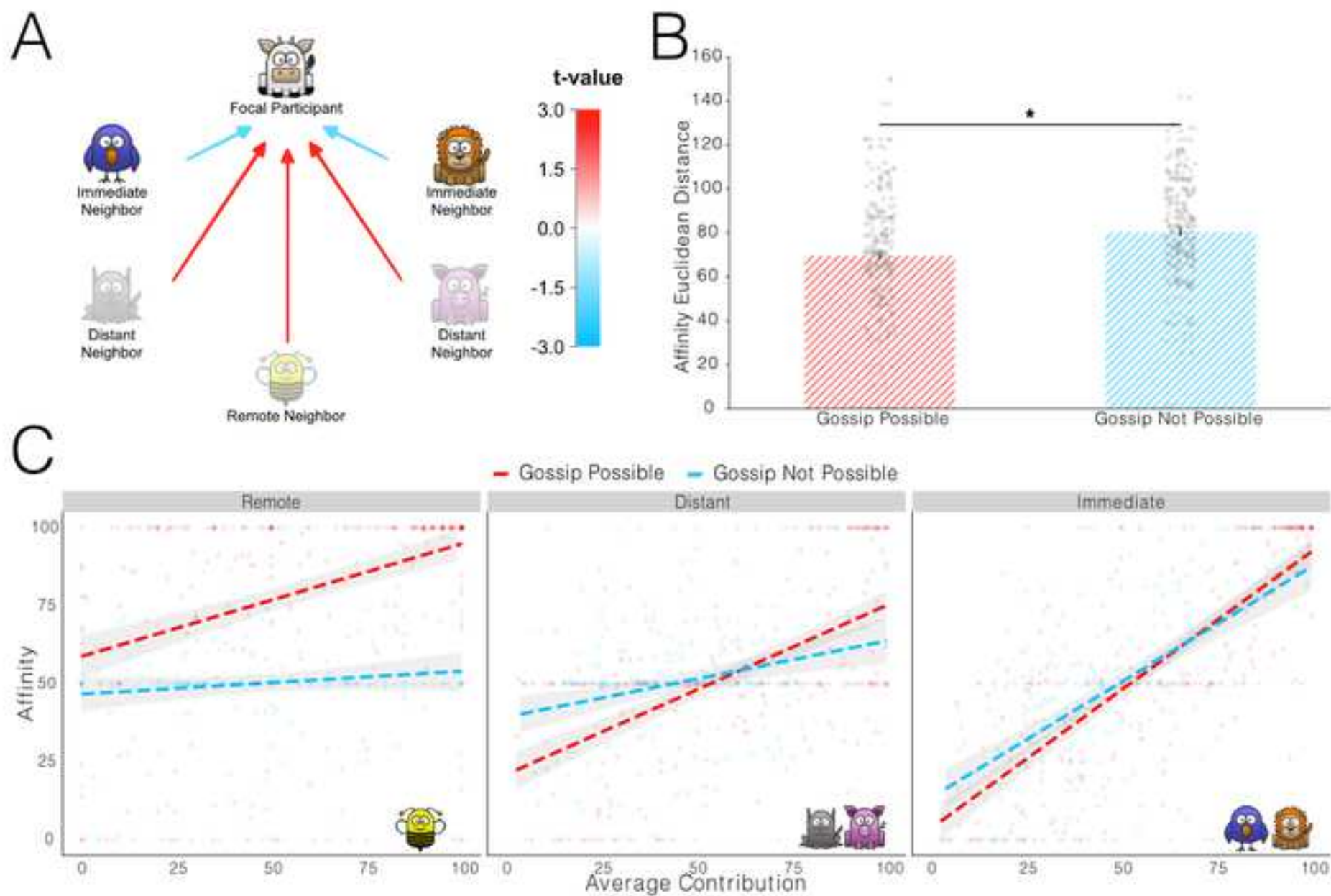
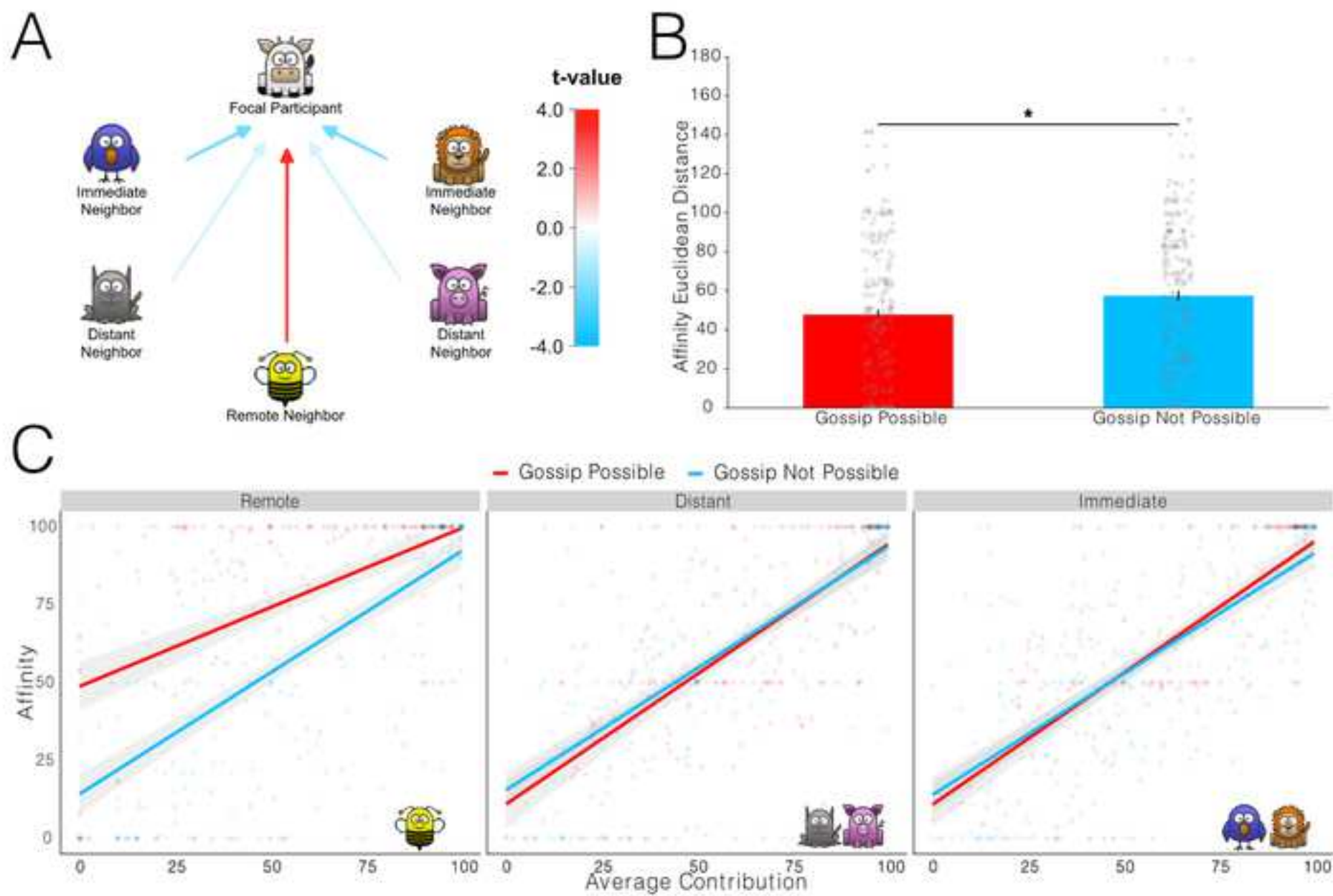
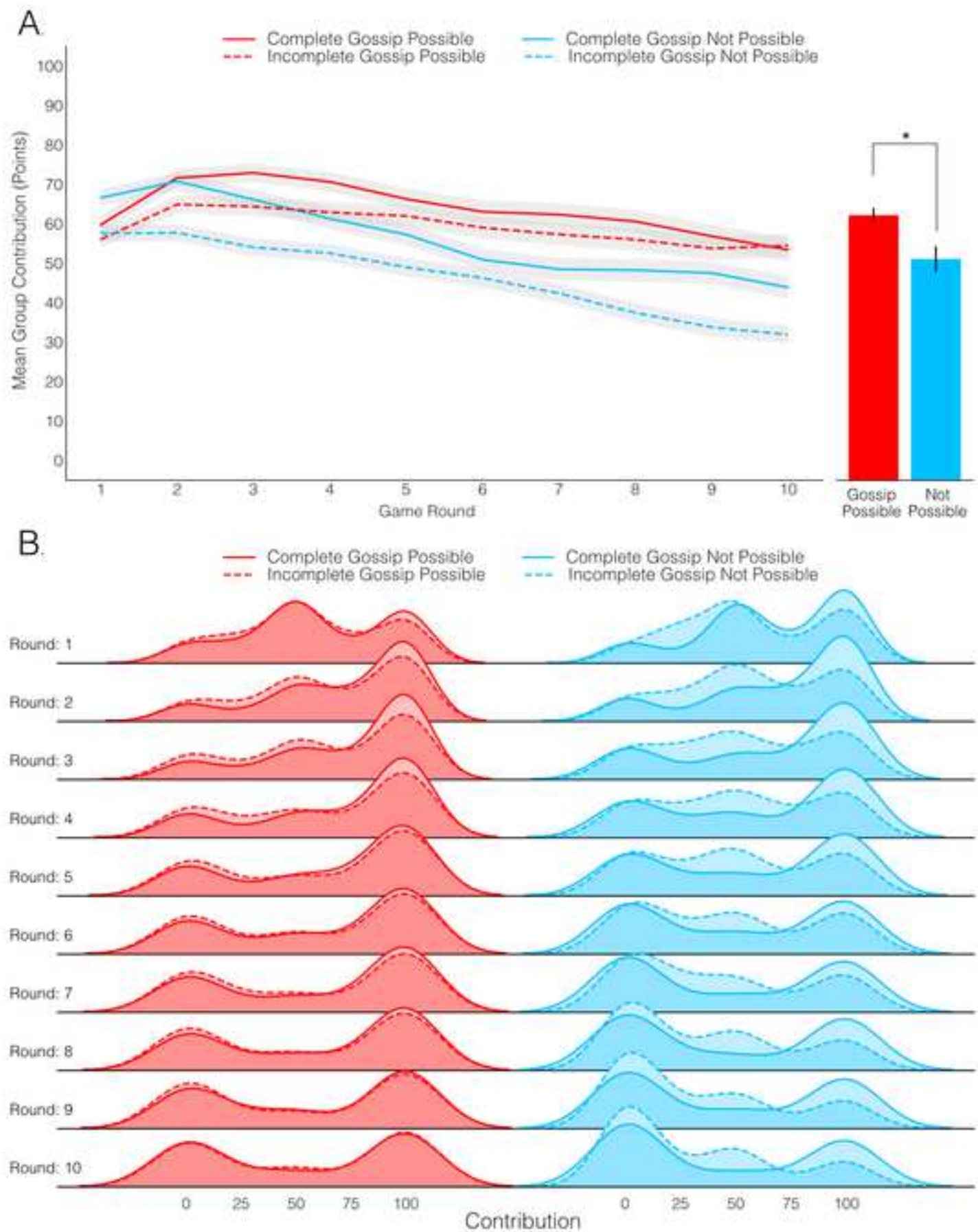


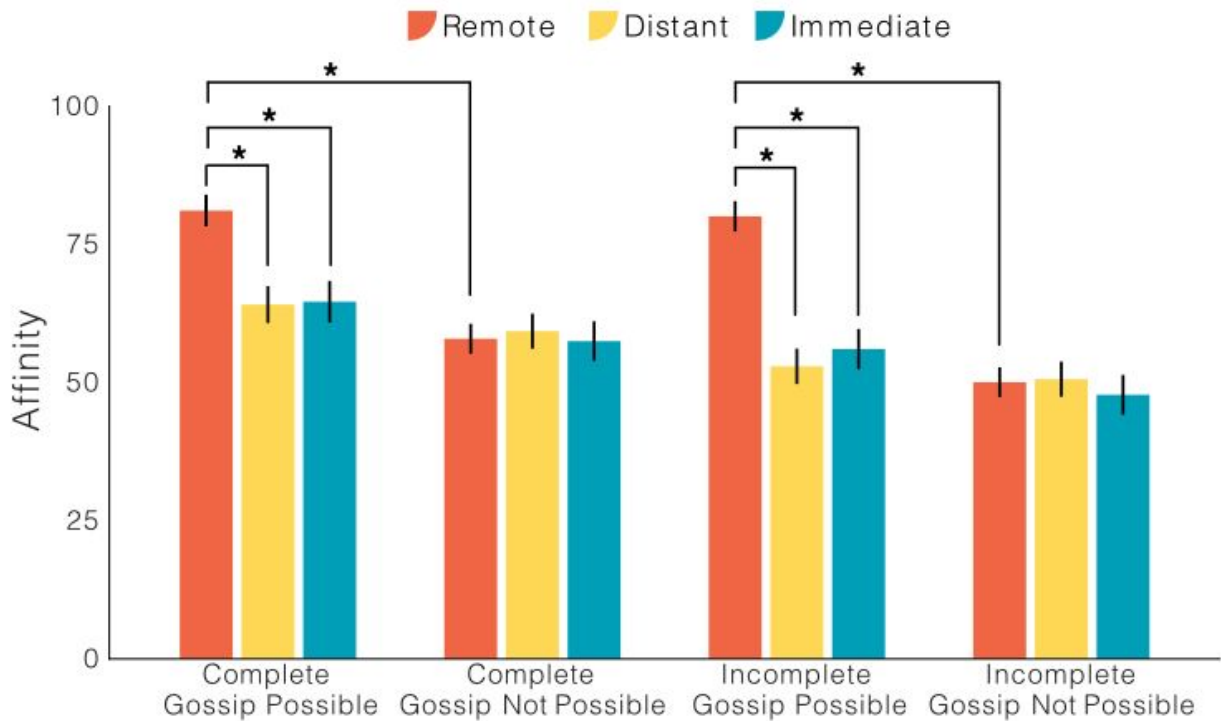
Figure4

[Click here to access/download;Figure;Fig4.tiff](#)

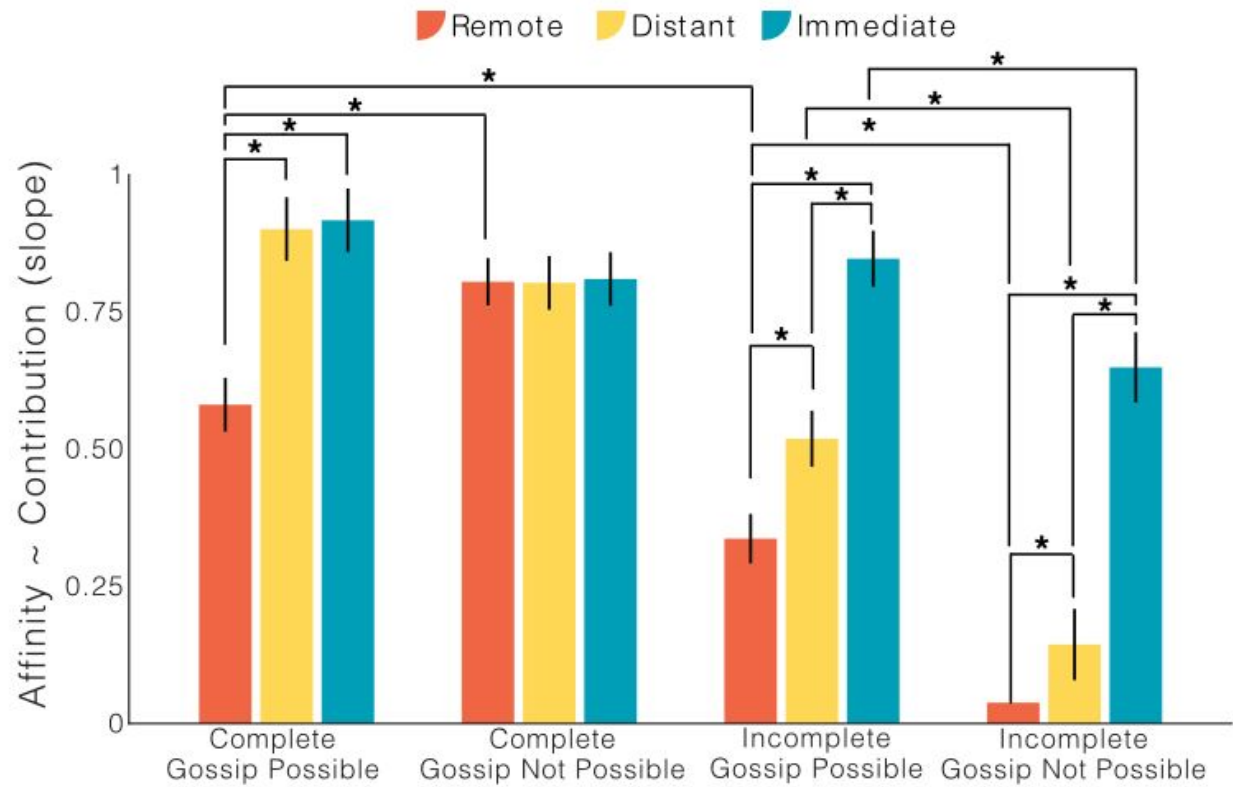




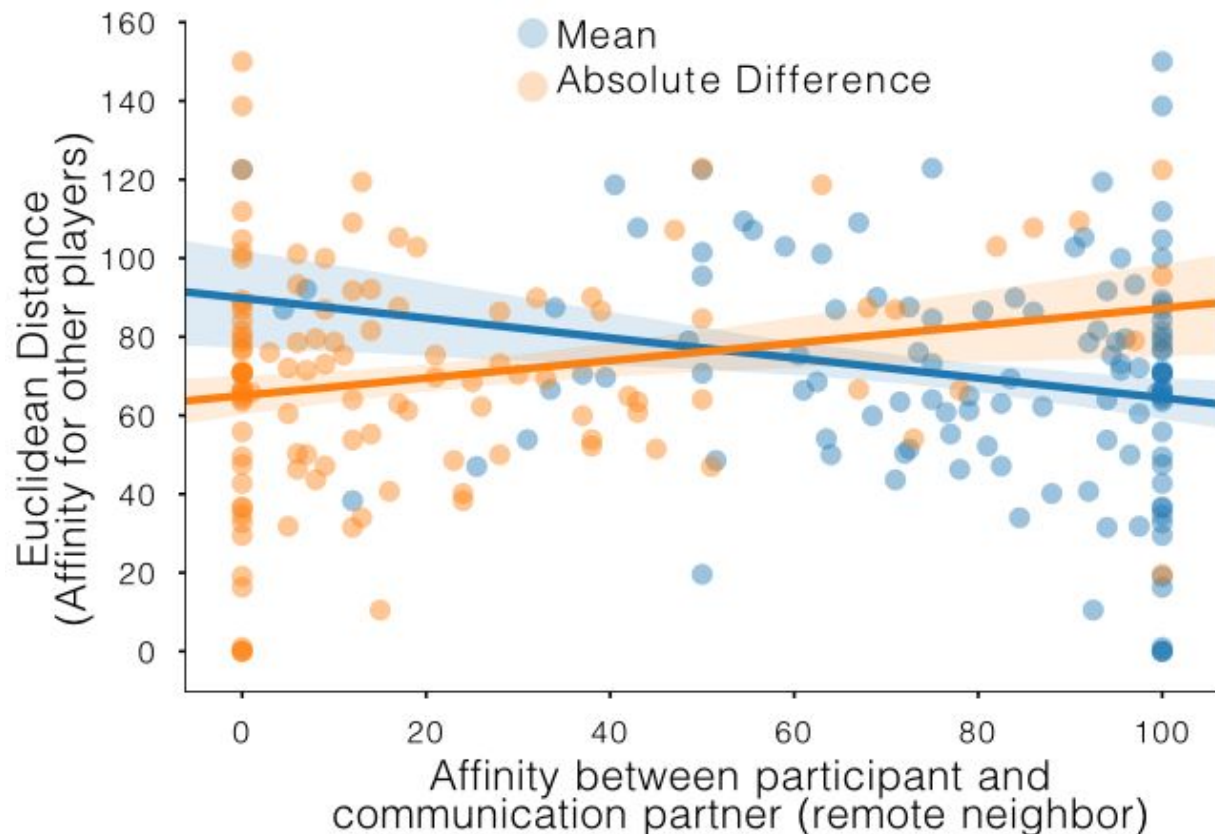
**Figure S1 | Post-game affinity.** Related to Fig 3C and 4C. Mean differences between overall affinity ratings for each neighbor type in all games. Participants felt the most affinity towards their remote neighbors (communication partner) in games where gossip was possible, relative to all other neighbors across all game types. See Table S1 for estimates. Error bars represent standard errors (SEM).



**Figure S2 | Post-game affinity and contribution behavior.** Related to Fig 3C and 4C. Regression slope differences of the relationship between affinity and contribution across all game types. Higher values indicate that participants' affinity ratings were more correlated with a neighbor's behavior (contributions) during play. See Table S2 for estimates. Error bars represent standard errors (SEM).

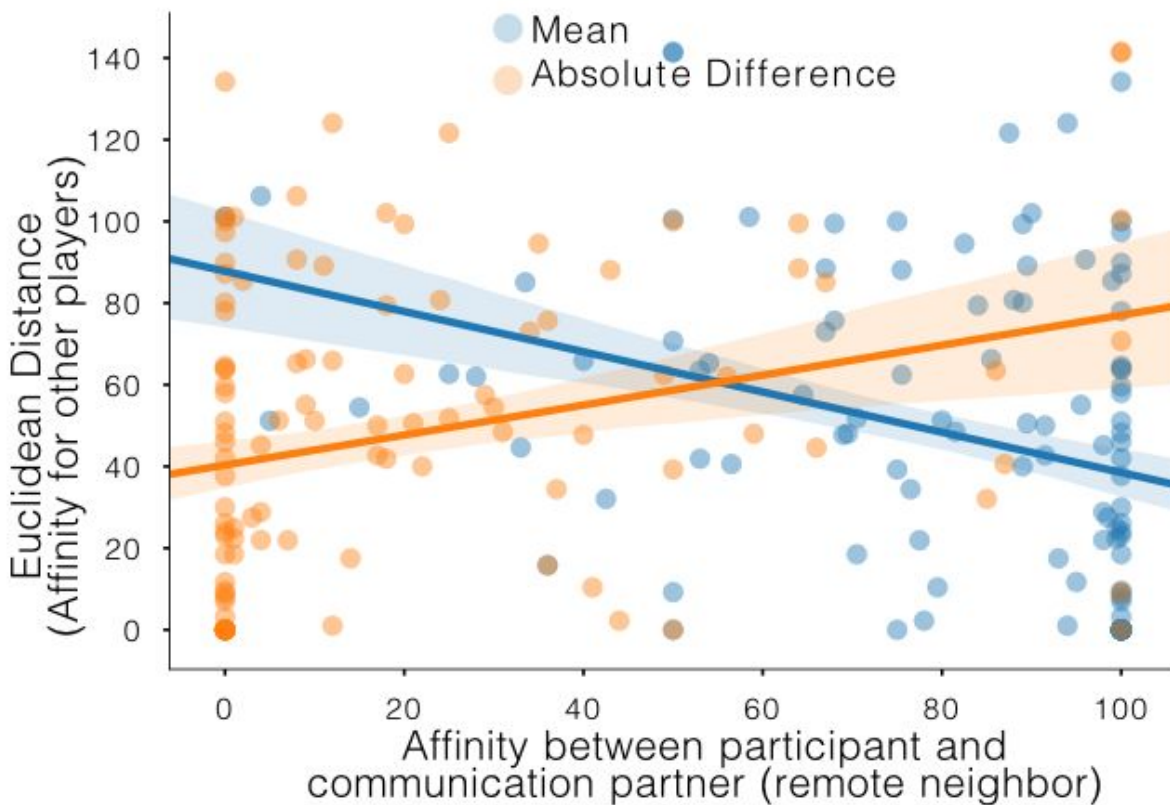


**Figure S3 | Impression Similarity - Incomplete Information.** Related to Fig. 3B. Correlation between (i) the similarity (euclidean distance) of a participant and their communication partner's (remote neighbor) affinity for *other* players and; (ii) after (or absolute difference) affinity for each other. **X-axis:** Mean affinity (blue) is the average between a participant's affinity for their communication partner and their communication partner's affinity for them on a 100 point scale. Absolute difference is simply the absolute value of the difference between their ratings. **Y-axis:** The euclidean distance between a participant's affinity for all other players excluding their communication partner, and their communication partner's affinity for all other players excluding the participant. Conversants who rated other players (i.e. distant and immediate neighbors) more similarly (lower euclidean distance y-axis) felt more positively about each on average (orange x-axis) and felt more similar (lower absolute difference) about each other (blue, x-axis). Error bands represent bootstrapped 95% confidence intervals.





**Figure S4 | Impression Similarity - Complete Information.** Related to Fig. 4B. Correlation between (i) the similarity (euclidean distance) of a participant and their communication partner's (remote neighbor) affinity for *other* players and; (ii) after (or absolute difference) affinity for each other. **X-axis:** Mean affinity (blue) is the average between a participant's affinity for their communication partner and their communication partner's affinity for them on a 100 point scale. Absolute difference is simply the absolute value of the difference between their ratings. **Y-axis:** The euclidean distance between a participant's affinity for all other players excluding their communication partner, and their communication partner's affinity for all other players excluding the participant. Conversants who rated other players (i.e. distant and immediate neighbors) more similarly (lower euclidean distance y-axis) felt more positively about each on average (orange x-axis) and felt more similar (lower absolute difference) about each other (blue, x-axis). Error bands represent bootstrapped 95% confidence intervals.



**Table S1. Post-game affinity.** Related to Methods - Affinity. Least-squares pairwise comparisons of mean post-game affinity ratings. Results are corrected for multiple comparisons using FDR correction ( $q = 0.05$ ) and bolded results are those that survive correction. Only pairwise comparisons between different neighbor types within the same game or the same neighbor type across different games were made

C = complete information; I = incomplete information; G = gossip possible; NG = gossip not possible

<i>Neighbor Comparison (game type)</i>	<i>Mean (SD)</i>		<i>Mean diff (CI)</i>	<i>se</i>	<i>t</i>	<i>df</i>	<i>p</i>
Remote - Immediate (C/G - C/G)	81.12 (30.58)	64.63 (31.53)	16.50 (11.59 - 21.40)	2.09	7.89	1900.06	<b>&lt;.001</b>
Remote - Distant (C/G - C/G)	81.12 (30.58)	64.11 (32.18)	17.00 (12.10 - 21.91)	2.09	8.13	1900.06	<b>&lt;.001</b>
Immediate - Distant (C/G - C/G)	64.63 (31.53)	64.11 (32.18)	0.51 (-4.39 - 5.42)	2.09	0.24	1900.06	0.810
Remote - Immediate (C/NG - C/NG)	57.93 (38.14)	57.49 (32.87)	0.44 (-4.22 - -5.10)	1.99	0.22	1900.06	0.844
Remote - Distant (C/NG - C/NG)	57.93 (38.14)	59.30 (32.72)	-1.37 (-6.04 - 3.29)	1.99	-0.69	1900.06	0.752
Immediate - Distant (C/NG - C/NG)	57.49 (32.87)	59.30 (32.72)	-1.81 (-6.47 - 2.85)	1.99	-0.911	1900.06	0.752
Remote - Immediate (I/G - I/G)	80.09 (29.43)	56.06 (33.41)	24.03 (19.31 - 28.74)	2.01	11.94	1900.06	<b>&lt;.001</b>
Remote - Distant (I/G - I/G)	80.09 (29.43)	52.95 (27.11)	27.14 (22.42 - 31.86)	2.01	13.49	1900.06	<b>&lt;.001</b>
Immediate - Distant (I/G - I/G)	56.06 (33.41)	52.95 (27.11)	3.11 (-1.60 - 7.84)	2.01	1.55	1900.06	0.128
Remote - Immediate (I/NG - I/NG)	50.08 (22.81)	47.79 (28.31)	2.29 (-2.37 - 6.96)	1.99	1.16	1900.06	0.452
Remote - Distant (I/NG - I/NG)	50.08 (22.81)	50.62 (22.21)	-0.54 (-5.20 - 4.12)	1.99	-0.27	1900.06	0.792
Immediate - Distant (I/NG - I/NG)	47.79 (28.31)	50.62 (22.21)	-2.83 (-7.49 - 1.83)	1.99	-1.43	1900.06	0.452
Remote - Remote (C/G - C/NG)	81.12 (30.58)	57.93 (38.14)	23.19 (11.30 - 30.08)	4.59	5.05	205.22	<b>&lt;.001</b>
Remote - Remote (I/G - I/NG)	80.09 (29.43)	50.08 (22.81)	30.01 (18.36 - 41.66)	4.50	6.67	205.22	<b>&lt;.001</b>

Remote - Remote (C/G - I/G)	81.12 (30.58)	80.09 (29.42)	1.03 (-9.56 - 11.63)	3.97	0.26	162.41	0.795
Immediate - Immediate (C/G - C/NG)	64.63 (31.53)	57.49 (32.87)	7.13 (-6.81 - 21.08)	5.22	1.37	155.86	0.208
Immediate - Immediate (I/G - I/NG)	56.06 (33.41)	47.79 (28.31)	8.27 (-5.39 - 21.94)	5.12	1.62	155.86	0.162
Immediate - Immediate (C/G - I/G)	64.63 (31.53)	56.06 (33.41)	8.56 (-5.46 - 22.59)	5.25	1.63	155.86	0.162
Distant - Distant (C/G - C/NG)	64.11 (32.18)	59.30 (32.72)	4.81 (-7.51 - 17.13)	4.61	1.04	159.50	0.358
Distant - Distant (I/G - I/NG)	52.95 (27.11)	50.62 (22.21)	2.33 (-9.75 - 14.40)	4.52	0.52	159.50	0.608
Distant - Distant (C/G - I/G)	64.11 (32.18)	52.95 (27.11)	11.17 (-1.23 - 23.56)	4.64	2.41	159.50	0.052

**Table S2. Post-game affinity and contribution behavior.** Related to Methods - Affinity.

Least-squares pairwise comparisons of the regression slope between post-game affinity ratings and average contribution. Results are corrected for multiple comparisons using FDR correction ( $q = 0.05$ ) and bolded results are those that survive correction. Only pairwise comparisons between different neighbor types within the same game or the same neighbor type across different games were made.

C = complete information; I = incomplete information; G = gossip possible; NG = gossip not possible

<i>Neighbor Comparison (game type)</i>	<i>b (se)</i>		<i>b diff (CI)</i>	<i>se</i>	<i>t</i>	<i>df</i>	<i>p</i>
Remote - Immediate (C/G - C/G)	0.58 (0.05)	0.92 (0.06)	-0.34 (-0.49 - -0.19)	0.06	-5.29	2145.65	<b>&lt;.001</b>
Remote - Distant (C/G - C/G)	0.58 (0.05)	0.90 (0.06)	-0.32 (-0.47 - -0.17)	0.06	-5.03	2143.89	<b>&lt;.001</b>
Immediate - Distant (C/G - C/G)	0.92 (0.06)	0.90 (0.06)	0.02 (-0.14 - 0.17)	0.07	0.26	1954.65	0.798
Remote - Immediate (C/NG - C/NG)	0.80 (0.04)	0.81 (0.05)	-0.005 (-0.13 - 0.12)	0.05	-0.09	2071.39	0.978
Remote - Distant (C/NG - C/NG)	0.80 (0.04)	0.80 (0.05)	0.001 (-0.12 - 0.13)	0.05	0.03	2049.14	0.978
Immediate - Distant (C/NG - C/NG)	0.81 (0.05)	0.80 (0.05)	0.006 (-0.13 - 0.14)	0.06	0.11	2016.66	0.978
Remote - Immediate (I/G - I/G)	0.34 (0.05)	0.85 (0.05)	-0.51 (-0.64 - -0.38)	0.06	-9.06	2090.27	<b>&lt;.001</b>
Remote - Distant (I/G - I/G)	0.34 (0.05)	0.52 (0.05)	-0.18 (-0.31 - -0.05)	0.06	-3.26	2048.33	<b>.001</b>
Immediate - Distant (I/G - I/G)	0.85 (0.05)	0.52 (0.05)	0.33 (0.19 - 0.47)	0.06	5.61	1995.38	<b>&lt;.001</b>
Remote - Immediate (I/NG - I/NG)	0.04 (0.05)	0.65 (0.06)	-0.61 (-0.79 - -0.44)	0.07	-8.17	2335.18	<b>&lt;.001</b>
Remote - Distant (I/NG - I/NG)	0.04 (0.05)	0.14 (0.06)	-0.11 (-0.28 - 0.07)	0.07	-1.44	2198.93	0.149
Immediate - Distant (I/NG - I/NG)	0.65 (0.06)	0.14 (0.06)	0.51 (0.32 - 0.69)	0.08	6.30	2132.69	<b>&lt;.001</b>
Remote - Remote (C/G - C/NG)	0.58 (0.05)	0.80 (0.04)	-0.22 (-0.39 - -0.06)	0.07	-3.42	1535.74	<b>.001</b>
Remote - Remote	0.34	0.04	0.30	0.07	4.48	1997.40	<b>&lt;.001</b>

(I/G - I/NG)	(0.05)	(0.05)	(0.13 - 0.47)				
Remote - Remote (C/G - I/G)	0.58 (0.05)	0.34 (0.05)	0.244 (0.07 - 0.42)	0.07	3.66	1517.59	<b>&lt;.001</b>
Immediate - Immediate (C/G - C/NG)	0.92 (0.06)	0.81 (0.05)	0.11 (-0.09 - 0.31)	0.08	1.42	1270.36	0.233
Immediate - Immediate (I/G - I/NG)	0.85 (0.05)	0.65 (0.06)	0.20 (-0.02 - 0.42)	0.08	2.41	1722.69	<b>0.048</b>
Immediate - Immediate (C/G - I/G)	0.92 (0.06)	0.85 (0.05)	0.07 (-0.13 - 0.28)	0.08	0.92	1246.37	0.432
Distant - Distant (C/G - C/NG)	0.90 (0.06)	0.80 (0.05)	0.10 (-0.10 - 0.30)	0.08	1.28	1265.29	0.200
Distant - Distant (I/G - I/NG)	0.52 (0.05)	0.14 (0.06)	0.38 (0.16 - 0.59)	0.08	4.53	1699.08	<b>&lt;.001</b>
Distant - Distant (C/G - I/G)	0.90 (0.06)	0.52 (0.05)	0.38 (0.18 - 0.59)	0.08	4.93	1236.10	<b>&lt;.001</b>

**Table S3. Message topic frequency comparisons.** Related to Fig 2. Least-squares pairwise comparisons of the frequency of discussing particular topics in games with Incomplete or Complete information. Estimates reflect the *difference* of Incomplete Information games > Completion Information games estimated using a mixed effect log-linear model on topic counts. Results are corrected for multiple comparisons by controlling the False-Discovery-Rate at  $q < 0.05$  and bolded results are those that survive correction.

Label	Estimate	2.5_ci	97.5_ci	SE	Z-stat	P-val
Discuss Others	0.375	0.257	0.492	0.06	6.246	<b>&lt; .001</b>
Strategy	-0.073	-0.172	0.026	0.05	-1.448	0.148
Affirmation	-0.086	-0.198	0.026	0.057	-1.499	0.134
Chit-chat	-0.082	-0.164	-0.0	0.042	-1.969	<b>0.049</b>
Game Mechanics	-0.163	-0.294	-0.032	0.067	-2.441	<b>0.015</b>
Brief Expressions	-0.249	-0.401	-0.098	0.077	-3.222	<b>0.001</b>
Inquiry	-0.314	-0.533	-0.095	0.112	-2.814	<b>0.005</b>
Nonsense	0.322	-0.114	0.758	0.222	1.447	0.148
Random	0.284	0.017	0.55	0.136	2.085	<b>0.037</b>
Blank	-0.25	-0.45	-0.05	0.102	-2.449	0.014

**Table S4. Time-lagged influence - Games where Gossip is possible.** Related to Table 1 and Table 2. Time-lagged mixed effects regression predicting participants' future contribution as a function of their neighbors' past contribution behavior, controlling for their past behavior. Confidence intervals represent 95% assuming a quadratic log-likelihood surface (Wald method). Estimates reflect the contrast between Full Information games > Incomplete Information games. Complementary to other models, the interaction between Distant Neighbor and Information is not significant suggesting that distant neighbors had similar amounts of influence on participants' behavior when gossip was possible, even when they could not be seen. Bolded results are significant at  $p < 0.05$ .

<b>Games where Gossip is possible</b>	<i>b (CI)</i>	<i>se</i>	<i>t</i>	<i>df</i>	<i>p</i>
Participant contribution round <sub>t-1</sub>	0.265 (0.24 - 0.29)	0.01	18.63	3748.73	<b>&lt;.001</b>
Round	-1.14 (-1.44 - -0.84)	0.15	-7.39	354.97	<b>&lt;.001</b>
Information	0.57 (-2.66 - 3.79)	1.65	0.35	242.64	0.730
Remote neighbor round <sub>t-1</sub>	0.15 (0.13 - 0.18)	0.01	10.88	3698.59	<b>&lt;.001</b>
Immediate neighbor round <sub>t-1</sub>	0.19 (0.15 - 0.23)	0.02	9.92	3548.41	<b>&lt;.001</b>
Distant neighbor round <sub>t-1</sub>	0.19 (0.15 - 0.22)	0.02	9.73	3568.97	<b>&lt;.001</b>
Information x Remote neighbor round <sub>t-1</sub>	0.11 (0.06 - 0.17)	0.06	4.06	3422.35	<b>&lt;.001</b>
Information x Immediate neighbor round <sub>t-1</sub>	-0.13 (-0.21 - -0.06)	0.04	-3.47	3450.01	<b>0.001</b>
Information x Distant neighbor round <sub>t-1</sub>	0.02 (-0.06 - 0.09)	0.04	0.49	3467.05	0.625
Round x Information	-0.63 (-1.23 - -0.02)	0.31	-2.04	354.78	<b>0.043</b>

**Table S5. Group Cooperation.** Related to Methods - Group Contribution Behavior. Mixed effects regression estimates for *group contributions* as a function of round, gossip and information. Confidence intervals represent 95% assuming a quadratic log-likelihood surface (Wald method). Bolded results are significant at  $p < 0.05$ .

	<i>b (CI)</i>	<i>se</i>	<i>t</i>	<i>df</i>	<i>p</i>
Round	-2.15 (-2.69 - -1.61)	0.28	-7.75	155	<b>&lt;.001</b>
Gossip Possible	10.22 (2.28 - 18.17)	4.05	2.52	155	<b>0.013</b>
Information	7.25 (-0.69 - 15.20)	4.05	1.79	155	0.075
Gossip Possible x Information	-5.17 (-21.06 - 10.71)	8.10	-0.64	155	0.524
Round x Gossip Possible	1.86 (0.77 - 2.95)	.056	3.35	155	<b>0.001</b>
Round x Information	-0.25 (-1.34 - 0.84)	.056	-0.45	155	0.653
Round x Gossip Possible x Information	-0.66 (-2.84 - 1.51)	1.11	-0.60	155	0.551