Norm talk and human cooperation: Can we talk ourselves into cooperation?

Abstract

Norm talk is verbal communication that explicitly states or implicitly implies a social norm. To investigate its ability to shape cultural dynamics, two types of norm talk were examined: injunction, which explicitly states what should be done, and gossip, which implies a norm by stating an action approved or disapproved of by the communicator. In two experiments, participants engaged in norm talk in repeated public goods games. Norm talk was found to help sustain cooperation relative to the control condition; immediately after every norm talk opportunity, cooperation spiked, followed by a gradual decline. Despite the macro-level uniformity in their effects on cooperation, evidence suggests different micro-level mechanisms for the cooperation-enhancing effects of injunction and gossip. A third study confirmed that both injunction and gossip sustain cooperation by making salient the norm of cooperation, but injunction also effects mutual verification of the communicated norm, whereas gossip emphasizes its reputational implications by linking cooperation to status conferral and noncooperation to reputational damage. A fourth experiment provided additional evidence that norm talk was superior to the promise of conditional cooperation in sustaining cooperation. Implications of the findings for cultural dynamics are discussed in terms of how feelings of shared morality, language-based interpersonal communication, and ritualization of norm communication contribute to social regulation.

Keywords: Public goods dilemma; social dilemma; communication; gossip; norms; cultural dynamics

Norm talk and human cooperation: Can we talk ourselves into cooperation?

Norms are said to be one of the most, if not *the* most, cited of all explanations of human social behavior in the social sciences (Gibbs, 1968). In social psychology too, the concept of norms has gained prominence in recent times (e.g., Gelfand & Jackson, 2016; Morris, Hong, Chiu, & Liu, 2015; Zou & Leung, 2015). Of all the variants of the norm concept (for an attempt to catalog it, see Gibbs, 1965; also see Anderson & Dunning, 2014), the present article is concerned with what Cialdini, Reno, and Kallgren (1990; Cialdini, Kallgren, & Reno, 1991) called an *injunctive* norm – that is, a prescriptive or proscriptive norm, or what should or should not be done. From health behavior (e.g., Reid & Aiken, 2013) to environmental behavior (e.g., Cialdini et al., 1990; Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007) and beyond (e.g., sexual behavior; van de Bongardt, Reitz, Sandfort, & Deković, 2015), injunctive norms have been found to influence intentions and behaviors in a wide range of social contexts (for a meta-analysis, see Manning, 2009).

Whereas it is possible for people to adopt these injunctive norms through observation or self-reflection, we believe that the socionormative influence often occurs through the communication of an injunctive norm, what we call *norm talk*. We define norm talk as utterances (i.e., verbal behaviors) that explicitly state or implicitly imply an injunctive norm about what should or should not be done. When someone says "should" or "ought to", the speaker is likely engaged in an explicit form of norm talk, whereas when someone praises or condemns a course of action the speaker is likely engaged in an implicit form of norm talk by implying an injunctive norm through that (dis)approval. George Homans ([1961] 1974), a sociologist of renown, even went as far as *defining* a norm as a "*statement* made by a number of members of a group….that the members *ought to* behave in a certain way in certain circumstances (p. 46, italics added)." To him, a social norm *is* norm talk, and presumably,

norm talk is the primary means by which injunctive norms are communicated in human societies.

It is an open question whether people's fleeting norm talk has any socio-cultural implications. When one individual engages in norm talk to another individual, does it influence the latter's action? When norm talk circulates within a local social network, does it increase the likelihood of norm-consistent action in the network? Cialdini et al.'s (1990, 1991) focus theory suggests that the answer to these questions should be "yes". According to focus theory, an injunctive norm influences action when it is salient in the actor's mind. Norm talk, by definition, increases the salience of injunctive norms, which means that it should influence action. Even though a single instance of norm talk may only have a minute effect, if it happens with some regularity its cumulative socio-cultural consequences may be far reaching. From a cultural dynamics perspective, everyday interpersonal conversations may have significant consequences when they accumulate across many social situations over time. For example, interpersonal communications may have cumulative effects on the formation, maintenance, and transformation of cultural stereotypes and intergroup relationships over time (e.g., Clark & Kashima, 2007; Hunzaker, 2014; Kashima, 2000; Lyons & Kashima, 2003; Schaller, Conway, & Tanchuk, 2002; Schaller & Conway, 1999; see Kashima, 2013, for a recent review). Indeed, interpersonal communication through social networks may be a critical micro-level mechanism that can shape societal outcomes (e.g., DiFonzo et al., 2013; DiFonzo et al., 2014; E. R. Smith, 2014).

If norm talk matters, it is most likely to matter in the provision of public goods (i.e., goods that are freely available to everyone, such as public parks, clean air, and stable climate). Typically, the provision of these goods involves a social dilemma, where the formation and maintenance of collective goods crucially depends on individuals' cooperation, but there is a strong incentive for the individuals to free ride (for reviews, see Dawes, 1980;

Kollock, 1998; Komorita & Parks, 1995; Parks, Joireman, & van Lange, 2013; Van Lange, Joireman, Parks, & Van Dijk, 2013). If interpersonal norm talk about cooperation – an individual talking about an injunctive norm of cooperation to another individual in their social network – discourages free riding, then this may be one micro-level process by which humans can overcome the social dilemma at the heart of public goods provision. Because language-based communication is a uniquely human psychological tool (e.g., Holtgraves & Kashima, 2008), which may socially regulate collective action (E. A. Smith, 2010), we investigated whether such a uniquely human *communicative regulation of collective action* is possible in public goods game experiments.

Norm Talk in Social Dilemmas

A social dilemma situation is prototypically modeled by a public goods game, where everyone is better off if everyone contributes to the provision of public goods than if no one does; however, individuals are always better off when they free ride than when they cooperate (e.g., Chaudhuri, 2010; Ledyard, 1995). In these situations, where a *Homo economicus* – what Nobel laureate Eleanor Ostrom (2000) called a rational egoist – would never cooperate, an injunctive norm of cooperation appears to be a critical factor in the maintenance of human cooperation (e.g., Biel & Thøgersen, 2007; Chudek & Henrich, 2011; Fehr & Fischbacher, 2004a). Behavioral game theoretic research has shown that people often do cooperate in one shot public goods experiments (for reviews, see Chaudhuri, 2010; Ledyard, 1995), and when someone fails to do so, people are willing to incur a cost to themselves to sanction the free rider (e.g., Fehr & Fischbacher, 2004b; Fehr & Gachter, 2002). Given that injunctive normativity drives the sanctioning of norm violations (Brauer & Chaurand, 2010), it is reasonable to suggest that the injunctive norm of cooperation is shared among those who punish free riders in public goods experiments (for a relational explanation of norm enforcement see Horne, 2001).

Evidence that norm talk, specifically, may matter comes from studies on group discussion in public goods games. When a group of individuals repeatedly engage in a public goods game for a finite number of times, the level of cooperation typically declines over time (Ledyard, 1995). However, one of the factors that can slow or even stop this decline is a group discussion (e.g., Isaac & Walker, 1988; Ledyard, 1995; Ostrom, 2000), and there is some suggestive evidence that norm talk in the group discussion may be partly responsible for sustaining this cooperation. First, it is well established that when all members of a group participate in a discussion the very act of discussing the dilemma situation can increase cooperation relative to the control condition in which no communication takes place (e.g., Dawes, McTavish, & Shaklee, 1977; Kerr & Kaufmann-Gilliarnd, 1994; Orbell, van de Kragt, & Dawes, 1988). Whether communication occurs face-to-face or online (e.g., Jensen, Farnham, Drucker, & Kollock, 1999), via open communication or pre-set messages (e.g., Chen, 1996; Chen & Komorita, 1994), meta-analyses have consistently found a robust effect of communication on cooperation (Balliet, 2009; Sally, 1995). E. A. Smith (2010) noted that language appears to enable humans to resolve social dilemmas more easily than other species.

Second, a number of researchers have made informal observations about the strongly normative undertone of the experimental participants' discussions in social dilemma experiments (e.g., Cason & Khan, 1999; Dawes et al., 1977). In a more systematic exploration, Bonacich (1972, 1976) reported the words most frequently used in these discussions included normatively evaluative terms like "cheat," "greed," "honest" or "dishonest". The content of these face-to-face discussions appears so normative in the injunctive sense that Bicchieri and Lev-On (2007, p. 163) summarized it this way: "Discussants emphasize not only the mutual gains obtained from cooperation, but also its appropriateness and normative appeal." Whether this normative undertone can be decomposed into more specific norms such as norms of commitment or equity, as Kerr

(1995) argued, it seems likely that there exists the normative talk of cooperation, that is, "We *should* cooperate in social dilemmas" (for examples of normative written messages, see Baum, Paciotti, Richerson, Lubell, & McElreath, 2012) or "You *should not* take advantage of your team members by shirking" (Fehr & Fischbacher, 2004a, p. 185).

Whereas this suggests that norm talk accompanies sustained cooperation, it does not necessarily mean that interpersonal norm talk itself produces cooperation. First, in explaining the benefits of communication, theorists have proposed a number of different mechanisms. Although many theorists have argued that communication affects cooperation because of social norms (e.g., Bicchieri, 2002; Kerr, 1995), the details of the arguments vary. For example, Kerr and Kaufmann-Gilliarnd (1994) have argued that people often make a public pledge to cooperate in such group discussions, and it is the norm of commitment and consistency – i.e., "now that I've said I'd cooperate, I should cooperate" – that drives cooperation (also see Kerr, Garst, Lewandowski, & Harris, 1997). In a related vein, Bicchieri (2002) suggested that making a pledge to cooperate is backed up by the norm of promise keeping, which supports cooperation. In contrast, Bouas and Komorita (1996) argued that this norm-cooperation relationship is mediated by perceived consensus to cooperate, rather than commitment. As Bicchieri (2002) noted, perceived consensus is not inconsistent with the norm argument, and in fact may serve to strengthen the normative effect. Indeed, Kusumi, Hirayama, and Kashima (2017) found that perceived consensus strengthened the effect of a perceived norm on people's behavioral intention.

There is also a perspective that argues that social identification may be the psychological process that mediates the effect of communication on cooperation (e.g., Orbell et al., 1988). However, there are empirical findings that are somewhat difficult to explain from the identification perspective. For instance, Kerr and Kaufmann-Gilliarnd (1994) argued that if the discussion effect is due to group identification, group discussion should increase

the value of the collective outcome (i.e., the more the group gains as a whole, the more valuable), and that this effect should be stronger if an individual is more efficacious in producing the collectively beneficial outcome. Their experimental results were opposite to this prediction. The effect of group discussion was stronger when an individual was *less* efficacious in producing the collectively beneficial outcome. In another instance, Bouas and Komorita (1996) had two conditions in which university students had a discussion prior to playing a social dilemma game; in one condition, they discussed the game, but in the other condition, they discussed the possibility of their university's tuition hike. In both conditions, the participants had similar levels of group identification; however, the discuss-dilemma condition had a far higher level of cooperation than the discuss-tuition condition. Furthermore, in both Kerr and Kaufman-Gilliland (1994) and Bouas and Kororita (1996), measured group identification could not account for the group discussion effect on cooperation (see Bicchieri, 2002, for further detailed discussion). Nonetheless, it is important to note that this alternative explanation for the discussion effect on cooperation exists.

The critical reason that the existing evidence cannot speak to the role of interpersonal norm talk in cooperation is that much of the evidence about communication effects on cooperation comes from *group discussion*, in which all members of the group communicate with each other before or during the social dilemma experiment. In many real life situations, however, communications do not occur simultaneously within a group as a whole, but occur only piecemeal, interpersonally between individual members suspended in social networks. The experiment by Kinukawa, Saijo, and Une (2000) suggests that interpersonal communications have weaker effects on cooperation than group communications, and in fact, when the interpersonal communication was limited to two individuals, it had no effect on cooperation. They structured the communication network within a six-person group to create four different conditions: a no communication control, a condition in which each member

could interpersonally communicate only with one other member, a condition in which each member could interpersonally communicate with two other members, and a condition in which all members communicated with each other as a group. The level of cooperation increased as the number of communication partners increased, with the group discussion condition producing the highest level of cooperation. It is important to note that the condition in which a group member had a discussion with only one other partner was no better than the no communication control; furthermore, this communication occurred face to face, and there is no guarantee that this discussion involved norm talk.

Summary. All in all, the literature is inconclusive about the causal efficacy of interpersonal norm talk in affecting cooperation in repeated public goods games. Although there is evidence that norm talk and cooperation co-occur, there is a variety of interpretations of this empirical finding. After all, norm talk is "cheap" in that there is often no binding obligation for the speaker or recipients to abide by the espoused norm with little or no material cost to those who violate the stated norm. If there is no incentive to cooperate when someone suggests it, either the speaker or the recipient may choose to not cooperate. In addition, there is no direct evidence that links norm talk to cooperation. That group discussions enhance cooperation, and that group discussions often contain norm talk do not guarantee that the instance of a norm talk increases cooperation. In fact, the existing studies have not shown a direct correlation between the occurrence of norm talk and cooperation. Finally, it is an open question whether *interpersonal*, as opposed to group, norm talk can facilitate cooperation at all. In the present studies, we investigate whether interpersonal norm talk within a structured social network can sustain cooperation in a repeated public goods game.

Norm Talk as a Speech Act

In addition to the first research question of whether norm talk sustains cooperation, we investigate the micro-level dynamics of how norm talk may be able to do so. In particular, norm talk can take at least two different forms, which have different implications for subsequent cultural dynamics. One is an *explicit* form, that is, an *injunction* about what should or should not be done. If someone says, "We should cooperate", it explicitly communicates an injunctive norm of cooperation, and is explicitly about the action that the utterance refers to. This type of statement can be considered to be a speech act (e.g., Holtgraves, 2002; Levinson, 1983). In Searle's (1969) typology, it may be regarded as a directive because it directs the recipient of the message to perform a certain action, in this case, cooperation. Secondarily, it also directs the speaker him or herself to perform the same action because the subject of this utterance, "We", refers at a minimum to the sender of the statement as well as its receiver. In this sense, explicit norm talk of cooperation may be regarded primarily as a directive, and secondarily, as a pledge, to cooperate.

The other type of norm talk is *implicit*. It is a statement that implies that people should cooperate without saying so explicitly. In particular, gossip is one of the most prominent forms of implicit norm talk. Gossip is defined broadly as a communication of reputational information about an absent other (e.g., Baumeister, Zhang, & Vohs, 2004; Dunbar, 2004; Foster, 2004; Sommerfeld, Krambeck, Semmann, & Milinski, 2007). Therefore it can easily occur as a part of the *interpersonal* communication within a small group, because other group members are absent from the communication, but reputational information about them is generally known and relevant. Gossip typically involves a triadic relationship among the sender, the receiver, and the target of gossip (e.g., Peters & Kashima, 2007, 2015), and in communicating positive (negative) gossip to the receiver, the sender can convey approval (or, of course, disapproval) of the target or the target's behavior.

In the context of social dilemma, if someone sends a message to another saying that a target person has defected, it is gossip. While such gossip need not necessarily indicate the sender's approval or disapproval of the target's action, it frequently does. Approval can be conveyed directly (e.g., "He did not contribute and I disapprove"), indirectly through labels and valenced language (e.g., "He is a free rider") (Torsvik, Molander, Tjøtta, & Kobbeltvedt, 2011), or indirectly through nonverbal or contextual factors (e.g., a display of disappointment). When gossip directly or indirectly signifies approval of a target's behavior, it can be seen as an implicit form of norm talk that communicates the injunctive norm that one should cooperate, very much in line with the argument that gossip conveys cultural norms (e.g., Baumeister et al., 2004; Gluckman, 1963). Whatever the gossiper's intent, the recipient of the gossip may learn or infer the gossiper's attitude towards the target's behavior (i.e., approval or disapproval); therefore, the primary information conveyed is reputational. However, the gossiper's inferred attitude may then be generalized to the group to infer the injunctive norm held in the group (see Table 1 for a schematic contrast).

Table 1. Injunction and gossip as two forms of norm talk

	Form	Information		Norm Salience	Reputational System Salience	Norm Verifiability
		Explicit	Implication			
Injunction	Explicit	Norm	Pledge	High	Low	High
Gossip	Implicit	Reputation	Norm	Low	High	Low

Thus, the two forms of norm talk, injunction and gossip, reflect two aspects of Cialdini et al.'s (1990) injunctive norm. Specifically, they have characterized injunctive norms as having "the *ought* meaning of social norms (p. 1015; italics in the original)" and also as referring to "rules or beliefs as to what constitutes morally *approved and disapproved* conduct (p. 1015; italics added)." Injunction explicitly conveys the ought meaning, whereas

gossip can explicitly convey approval and disapproval, and in so doing, imply the underlying norm.

Social psychological functions of injunction and gossip. If we are correct in conceptualizing injunction and gossip as norm talk, then they should have one social psychological function in common. Both injunction and gossip should make the injunctive norm salient in the sender's and receiver's minds. As Cialdini et al.'s (1990, 1991) focus theory suggests, norms influence social action when they are salient and cognitively accessible to the actors. Cialdini et al. (1990; Study 1) provided empirical support. Observing a confederate littering in an otherwise clean parking lot made the injunctive norm of non-littering salient, and therefore a large majority did not litter (only 6% did); however, the likelihood of littering was greater (14%) when people did not observe a littering confederate (i.e., norm was not salient). Likewise, by making the injunctive norm of cooperation salient, norm talk can encourage the actors to cooperate in social dilemmas. To be sure, a norm salience effect may not be long lasting if it is akin to conceptual priming effects. Nevertheless, the momentary salience of the norm of cooperation, activated by norm talk, may momentarily affect cooperation in social dilemma situations. When these effects are repeated over time, norm talk may have a significant cumulative effect.

The common normative impact of injunction and gossip notwithstanding, the distinction between the two is also important because they have potentially different social psychological functions. First, injunction and gossip may differ in their capacity to make the underlying norm salient (Table 1). Injunction explicitly states the norm, and hence activates the norm in both the sender and receiver of that norm talk. In contrast, gossip may not make the injunctive norm as salient as injunction. Therefore, the norm salience effect of injunction may be stronger than that of gossip.

Second, gossip differs from injunction in that gossip contains reputational information whereas injunction does not. Thus, gossip can enhance or damage the target person's reputation, but injunction does not have this consequence directly. To the extent that good reputation can confer status (e.g., Willer, 2009) and foster further cooperation from others (e.g., Barclay, 2004; Wedekind & Milinski, 2000), reputational information can act as reward or punishment. In this sense, gossip can provide what French and Raven (1959) called coercive or reward power. Consistent with this reasoning, gossip has been found to foster cooperation in economic bargaining games (e.g., Beersma & Van Kleef, 2011; Feinberg, Willer, Stellar, & Keltner, 2012; Piazza & Bering, 2008; Sommerfeld et al., 2007) and a public goods game in particular (Feinberg, Willer, & Schultz, 2014; we will return to this point later). Simply engaging in gossip makes one privy to examples of sanctioning and status conferral, first making salient *the entire reputational system*, and consequently one's own reputation (Table 1).

Whereas injunctions and gossip are both "cheap talk" in economics' parlance – that is they are typically non-binding – they differ in the enforceability of the reputational costs. Injunction incurs no direct cost to one's reputation if its recipient defies the injunction, but defying the injunctive norm implied by gossip means one will receive similar sanctions as the gossip target has received. Depending on the type of sanctions available, this may involve ostracism (e.g., Feinberg et al, 2014), other forms of punishment (e.g., non-cooperation afterwards), or simply a negative verbal evaluation. If gossip involves a negative evaluation, the target is devalued in the gossiper's and the gossip recipient's minds, and imagining that one is devalued in someone's mind may be an enough punishment even in the absence of a tangible negative sanction.

Finally, injunction and gossip differ in the extent to which the response made by the receiver has an implication for the normative status of the injunctive norm. Norm talk

typically occurs in interpersonal interactions. The receiver of norm talk can verify the

normative status of the injunctive norm by agreeing with the sender, or the receiver can

disagree with the sender, thus undermining its normative status. However, this is more likely

for injunction. When the sender sends an injunction message, the receiver has a chance to

accept it or not accept it. If it is accepted, the sender and receiver intersubjectively verify the

injunctive norm. In so doing, they form a shared reality (e.g., Echterhoff, Higgins, & Levine,

2009; Hardin & Higgins, 1996; Higgins, 1992), and establish an intersubjective culture (Chiu,

Gelfand, Yamagishi, Shteynberg, & Wan, 2010). If an injunction is not accepted by the

receiver, the non-acceptance implies that the receiver is less likely to follow the norm than

otherwise, thus undermining its normative status. In contrast, if gossip is or is not accepted, it

is primarily the reputational information that is accepted or not accepted directly, but not the

norm *per se*. Like the gossip itself, gossip's acceptance or rejection can infer an injunctive

norm, but does not explicitly declare one. Consequently, injunction and gossip may differ in

what may be called *norm verifiability* (Table 1). The injunctive norm is intersubjectively

verifiable if it is conveyed by injunction, but less so if it is conveyed by gossip. Figure 1

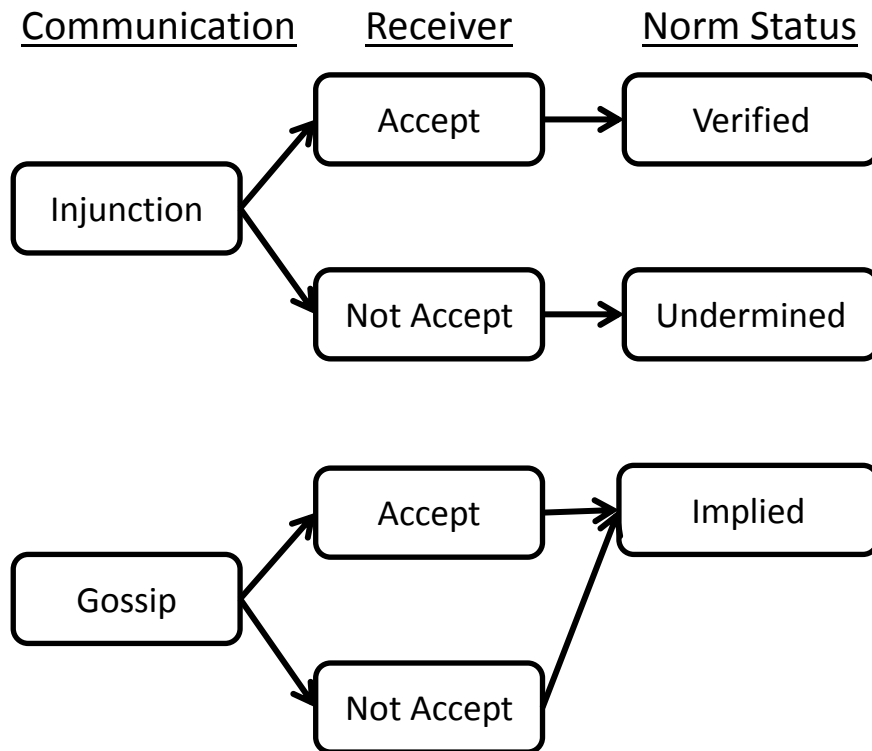schematically summarizes these points.

Figure 1. Injunction and gossip have different patterns of norm verifiability.

Summary. Explicit and implicit forms of norm talk, injunction and gossip, are both speech acts – producing social consequences by saying something. They both communicate and make salient an injunctive norm, but injunction does so explicitly, whereas gossip, when explicitly communicating the approval or disapproval of a conduct, implies the norm underlying this normative judgment. Both injunction and gossip should maintain cooperation in social dilemmas by activating the injunctive norm in the communicators' minds, but potentially through different processes. Injunction may do so by activating the injunctive norm of cooperation, whereas gossip may do so by activating the norm to a lesser extent, but making salient reputational sanctions. Finally, injunction and gossip have different implications in intersubjective verifiability of the status of the injunctive norm *qua* moral norm. The norm of cooperation can be mutually verified when injunction is accepted, but

gossip is less amenable to mutual verification – whether a gossip message is accepted or not does not directly speak to the normative status of the norm of cooperation. Thus, if norm talk sustains cooperation in social dilemmas, injunction and gossip can present two communicative mechanisms by which this occurs.

Present Studies

This paper summarizes the results of four experiments. Experiments 1 and 2 investigated the capacity of norm talk to sustain cooperation in public goods games; Experiments 3 and 4 investigated the intersubjective mechanisms through which different forms of norm talk have their effects. In Experiments 1, 2 and 4, participants took part in a repeated public goods game in four-person groups. Experiment 3 presented participants with scenarios based on this set up. The game was run online using purpose-built software that provided participants with an accurate summary of their own behavior as well as the aggregate behavior of their group (in the gossip conditions, participants were also presented with a summary of the behavior of a specific group member in communication rounds; e.g., "Person B contributed 1 token in the last 5 turns"). The experimental manipulation varied participants' ability to engage in norm talk. Participants in the control condition were not able to communicate with one another. In contrast, participants in the norm talk conditions were able to exchange norm talk with another specific player after every five rounds. Participants selected their norm talk from a menu of statements that could include injunctions (e.g., "We should each contribute a lot to the group"), gossip messages (e.g., "Person B contributed very little and I disapprove") or both as a function of study and condition. Participants responded to the norm talk that was sent to them by selecting from another menu of statements (e.g., "I strongly agree with you"). This paradigm has elements in common with previous work (notably that examining the functions of gossip; see Feinberg et al., 2014; Sommerfeld et al.,

2007, and Ahn, Esarey & Scholz, 2009); we examine these commonalities as well as the unique features of the present experimental paradigm in the General Discussion.

The present studies were designed to answer four main questions. First, we tested the hypothesis that mere interpersonal norm talk – injunction and gossip – is sufficient to sustain cooperation in repeated public goods games. Both injunction and gossip would make the injunctive norm salient, which then should increase cooperation at the time of norm talk. Second, we examined the possibility that the different types of norm talk sustain cooperation through somewhat different micro-level mechanisms. In particular, we surmised that norm talk *receivers' responses* have a more important role in verifying the norm for injunction than for gossip. Injunctions sent to receivers would increase their cooperation especially when they accept the injunctions because it signals the sender's and the receiver's mutual verification of the status of the injunctive norm *qua* norm. In contrast, gossip sent to receivers should increase their cooperation regardless of their acceptance. This is because gossip enforces its normative implication with the reputational threat or reward that it carries. Third, we also explored whether the *sending* of injunction or gossip would predict the *sender's cooperation*. This should enable us to examine what senders mean when they engage in norm talk. If norm talk predicts the sender's cooperation, it can be interpreted as a pledge to cooperate; if it does not, its intended meaning is less clear. Fourth, we directly examined the intended meaning of norm talk and replies from the perspective of the sender, receiver, and a third-party observer to uncover the mechanisms involved. Norm talk could be primarily interpreted as a pledge, or norm talk with reply could be primarily interpreted as mutual agreement to cooperate. However, we argue and then demonstrate that norm talk primarily increases the salience of the injunctive norm and only secondarily signals a pledge or mutual agreement to cooperate.

Experiment 1

Experiment 1 used the setup in Shank, Kashima, Saber, Gale, and Kirley (2015) to examine the effect of norm talk on cooperation in a repeated public goods game. Participants in the norm talk conditions were allowed to send a message containing an injunction, gossip, or both to another member of the group; if a message was sent, the receiver was given a chance to respond, by accepting or rejecting it. In the control conditions, no communication was allowed. The norm talk factor was crossed with a benefit factor in which the size of a public goods payoff was set at high or low levels, so as to explore whether the effectiveness of norm talk differs as a function of the public goods benefit.

Method

Participants. Recruited through Amazon Mechanical Turk, participants residing in the USA with records of reliable study participation participated in the experiment on their own devices. They were told that study sessions occurred once per hour and were directed to a webpage running custom software. For the ten minutes prior to each hour they were allowed to login and proceeded through pages including a study information page, a consent form, instructions on the primary decision-making task (the public goods task), a quiz about those instructions, instructions on message-sending rounds, and finally a virtual waiting room where participants were told to wait until an experiment began. Participants that reached the waiting room remained until the top of the hour, when the software automatically began forming groups of four out of participants in the waiting room, delegating excess participants to smaller groups not used in this study. They then participated in 31 rounds of public goods dilemma decisions interspersed with six communication rounds in the norm talk conditions. The number of public goods or communication rounds was not told to participants in advance. After this, participants filled out demographic and other questionnaires and were provided a code to submit their completion on Mechanical Turk. Participants were paid a

base rate of 2.50 USD, plus a bonus based on the cumulative points earned from the public goods dilemma. An entire session typically lasted from 20 to 40 minutes.

Because our experimental manipulations and analysis occurs at the level of the group, we conducted a power analysis on the required group sample size (Snijders, 2005). To obtain power = .80 for a within-factor repeated-measure design with six key measurements (the six cooperation rounds following norm talk) with a medium-strong effect size (d = .65).. The power analysis indicated 8 groups were required per condition, but due to our varied analyses, we conservatively predetermined to collect a minimum of 12 complete sessions per condition. The experimental platform was designed to pool all potential participants signed up in an online "waiting room," to randomly construct as many groups of specified size (4 in the present case) as possible at a time, and start the experimental procedure. Anyone who has waited longer than five minutes was given a chance to participate in a different experiment, left the present study, and therefore could not be included in the study. In addition, due to the nature of the interactive group, participants who miss three decisions (i.e., a majority of a block of five decisions; see below for details about missing a decision) were deemed unreliable, and were also surmised that they may bias the group processes. For these reasons, we set the program to drop these participants while the rest of the group continue to play the public goods game; however, we predetermined that these data should be treated as incomplete and therefore should not be included in our data. Because of the possibility of incomplete sessions, we conducted more sessions than were needed, meaning we slightly exceeded our minimum session numbers. When a target was reached, all complete data were kept and analysed; no further data collection took place thereafter. The final sample consisted of 212 participants in 53 four-person groups.

Procedure. Participants were given the identical instructions prior to reaching the virtual waiting room, and then were pseudorandomly assigned, by arrival order, to the

experimental conditions when the software formed groups from the participants in the waiting room. The instructions indicated that everyone in one's group had to decide whether to contribute a renewed endowment (100 points) to the group (cooperate) or keep it for oneself (defect) each round. All members' contributions to the group were then summed and multiplied by a parameter called the MPCR (marginal per capita return) to determine the number of points each player individually received in that round from the group. Therefore, if a player kept the endowment, he or she obtained the sum of the endowment and the return from all the members' contributions. In contrast, a player who contributed the endowment received only the return from all the members' contributions. On every dilemma round all of the possible outcomes were presented in a table. Based on pilot tests, we set the level of MPCR to .3 (low benefit) and .4 (high benefit). After the instructions, participants had to correctly answer a four-question quiz about contribution decisions using an example table before proceeding.

The instructions also contained information about how participants may be able to send and receive messages during the experiment (see below for further details), so that everyone was familiar with the messages that they may be able to send and receive, and the interface (drop-down boxes) that they may use, regardless of whether they were in the norm talk conditions or not. However, they were explicitly told that in some conditions they would be asked to use this functionality, whereas in other conditions, they would not. This was necessary to keep the salience of injunction and gossip constant in all conditions, so that we can gauge the impact of actual norm talk, not just the potential of it, on cooperation.

Once they were assigned to a group in a condition, participants began to play a public goods game, and for each round, they were shown the table of possible outcomes given their and others' decisions. There was also a timer, set to 1 minute in the first rounds and 30 seconds in all other rounds in which they had to make their decision to keep or contribute.

When everyone in a group had made their decisions or time had run out, the program advanced to the next screen. Regardless of whether the next screen was a public goods or communication round, it would display the results from the previous public goods round in terms of the number of contributors, the participant's decision, the points they had earned, and their updated cumulative point total. Note that no information about any other individual player's decision was given. There were altogether 31 rounds of the public goods game.

The timers on each round ensured that individuals who did not make a selection would not hold up the entire group. When the time expired, the software advanced to the next screen. Because of the interactive nature of the public goods dilemma and communications, participants who missed public good decisions were treated as choosing to cooperate, and missed messages and replies were treated as choosing not to communicate in terms of their effect on the participants and in terms of analysis. If any participant in a session missed three decisions, that participant was dropped by the software and that session was incomplete and not included in the dataset. The missed decisions in complete sessions therefore were a maximum of two per participant, and were only excluded from analysis when they were the dependent variable.

Norm talk. In the norm talk conditions, six communication rounds occurred, placed after every five public goods rounds. Each communication round consisted of a message and reply stage, presented on two separate screens. For the message stage, participants were given information about one group member's contribution over the past five rounds and allowed to send a message to another participant. For the reply stage, participants then received a message from another group member and they were allowed to reply. Replies were then displayed prior to the beginning of the next public goods game round. All communication rounds had 30 second timers.

Because the study involved 31 public goods rounds, sixth and final communication round was presented between rounds 30 and 31, enabling us to examine the effect of norm talk on the final dilemma round. On each of the six communication rounds, every participant took the role of a sender, receiver, and target. For a particular message, the sender, receiver, and target were always different people. See Figure 2 for a schematic representation of these roles. For example, on one round, A would communicate to B about C, B would communicate to C about D, C would communicate to D about A, and D would communicate to A about B. These were structured so that they always formed an inclusive ring, meaning that there was never reciprocal communication in the same round (e.g., A communicating to B while B communicated to A). Over the course of the six communication rounds, these inclusive rings ensured that each sender had every receiver by target combination exactly once. From A's perspective, A communicate about B once to C and once to D, about C once to B and once to D, and about D once to B and once to C, and received messages that were similarly distributed (see Supplementary Table S1 for distribution pattern; supplementary information is provided to allow for exact replication of these methods and for details of specific outcomes). Senders and receivers knew who they were sending and receiving from and the sender initially knew who the target was and could communicate that to the receiver. Participants were only identified by the letters A, B, C, and D.
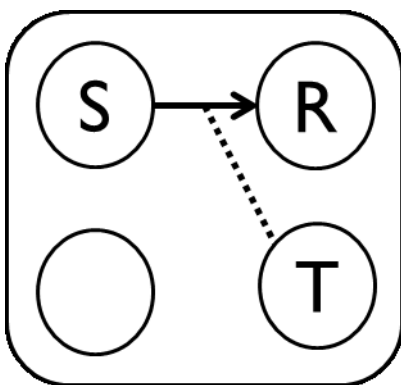

Figure 2. A Schematic Presentation of the Norm Talk Structure

Note. S = Sender; R = Receiver; T = Target. S has information about T, and sends a message to R; R replies to S.

Norm talk components. Participants were allowed to communicate a range of preset messages from dropdown menus that combined two message components (Table S2). The first was a gossip component (shown with B as the target as an example). We gave participants two different types of messages they could send. One described the target's behavior, and the other, the target's disposition and only one of those two types could be selected. A behavioral message could be selected from five options with varying degrees of prosociality, from "Person B contributed very little" to "Person B contributed very much". Similarly, a dispositional message could be selected from five options from "B is a very unhelpful person" to "B is a very helpful person". The gossip component included a supplementary gossip approval component with five options from "I strongly disapprove" to "I strongly approve". This supplementary component was only enabled when a gossip message was selected. Both the gossip message and gossip approval components could vary and only their alignment produces the sort of norm talk that conveys explicit reputational information with an implicit norm. Therefore, approving of prosocial behavior and disapproving of self-interested behavior were both coded 1, with all other combinations of the gossip message and gossip approval (e.g., approving of selfish behavior, not approving nor disapproving, not sending gossip) were coded as 0 (See Table S3 for more details).

The second was an injunction component, which could be selected independently from the gossip component (i.e., they were separate drop-down boxes). Again, we gave participants two types of message options: a more concrete and more abstract injunction (Table S2). They could select only one type. The five concrete injunctions varied from "We should keep a lot for ourselves" to "We should contribute a lot to the group", and the five abstract injunctions varied from "We should focus completely on being individuals" to "We

should focus completely on being one group". These were coded from -2 to +2, with larger numbers represented more prosociality (Table S4). All of the components that a given sender selected (i.e., gossip – with or without approval – or injunction, or both) were joined together and sent to the receiver. When a sender failed to choose either gossip or injunction components, the receiver was informed that the sender had not sent a message.

If a message was sent, the receiver had an opportunity to reply. Based on pilot testing, we included five agreement messages from "I strongly disagree with you" to "I strongly agree with you", and three evaluative messages: "That is terrible", "Okay", and "That is excellent" (Table S6). Intermediate evaluative messages were considered, but pilot testing revealed they were rarely used and so we did not include them in the study proper. If the receiver chose not to reply, the sender was informed of this. All replies were coded from -2 to +2, with higher values indicating a more positive agreement or evaluation (Table S6). No other manipulations or measures were present in this study.

Results

A total of 212 participants in 53 usable groups were collected with at least 12 usable groups per condition. Participants only failed to make a public goods decision on 28 of the 6572 (0.4%) trials and these were reasonably evenly distributed among conditions (Table S7).

Norm talk sustains cooperation over time. Figure 3 presents the mean levels of cooperation over time in all conditions. In the control conditions, there are clear trends of declining cooperation over time, whereas in the norm talk conditions, there is a cyclical pattern in which the cooperation level "spikes" after every norm talk, which occurred after every five rounds, i.e., spikes on round 6, 11, 16, 21, 26, 31.
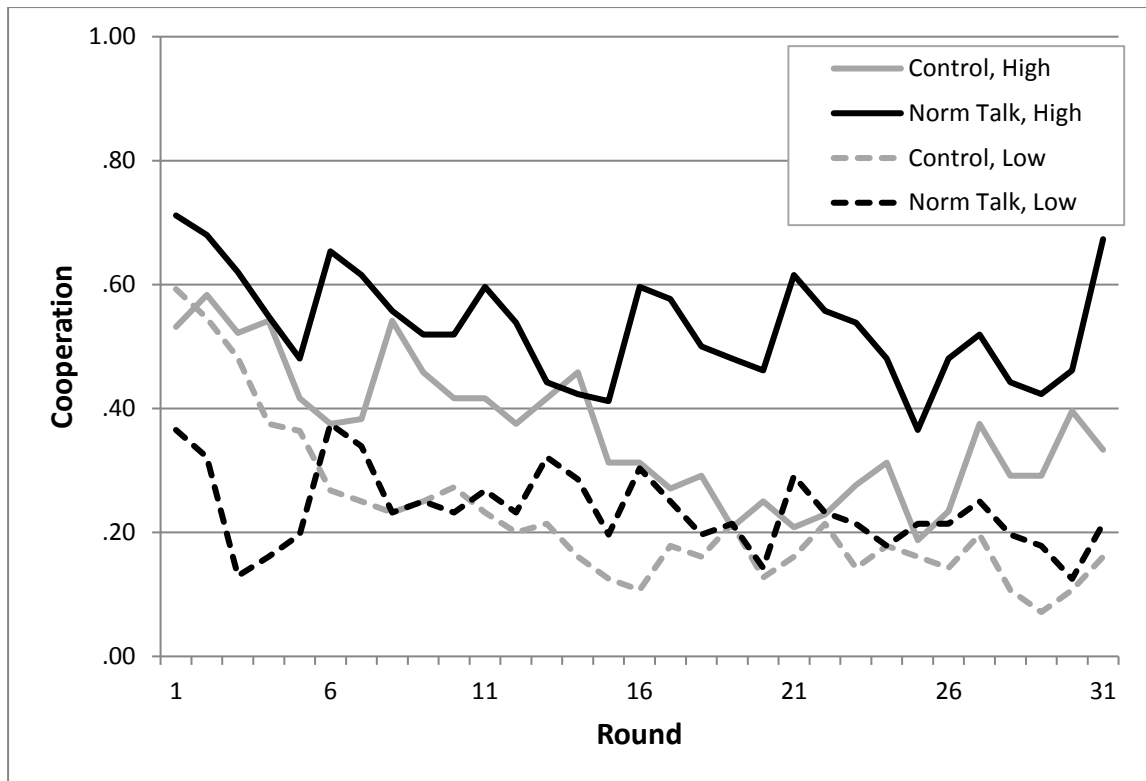
Figure 3: Mean cooperation levels in Experiment 1.

To determine if norm talk helps sustain cooperation over time, we fit a Mixed Model Logistic Regression model to cooperation decisions. The predictors were the experimental variables, Norm Talk (norm talk vs. control), Benefit (high vs. low), Time (linear and quadratic trends), and their relevant interactions (Table 2). Note that individuals are nested under their group, and time is a repeated measures factor. There is a significant decline in cooperation over time as indicated by a negative linear trend of time, but the rate of decrease slows as indicated by the positive quadratic trend of time. The interaction effect of norm talk with time reveals that norm talk increases cooperation over time (Coef = .045, $p < .001$) and high benefit has a marginal positive effect on cooperation over time (Coef = .020, $p < .1$). Consistent with our expectations, norm talk can maintain cooperation over time.

Table 2: General Linear Mixed Model Logistic Regression predicting cooperation in Experiment 1.

	Coef.	Odds Ratio

Intercept	-.303	.739
Benefit	.845	2.329
Norm Talk	-.406	.666
Time	-.137***	.872
Time^2	.002***	1.002
Benefit X Norm Talk	1.050	2.858
Benefit X Time	.020†	1.020
Norm Talk X Time	.045***	1.046
Benefit X Norm Talk X Time	-.022	.978
Group Intercept	2.822	
% Correctly Predicted	78.1	
N	6544	

Note: *** p < .001; ** p < .01; * p < .05; † p < .1.

We noted the spiking pattern of cooperation in the norm talk conditions. Each time it occurs, a sudden jump is followed by a gradual decline until the next norm talk occurrence, where a sudden jump occurs again. To investigate this pattern statistically, we conducted two sets of follow-up analyses. First, we compared the levels of cooperation immediately before and immediately after each norm talk round. There was a statistically significant increase in cooperation in both the low, from .18 to .28, t(669) = 2.874, p =.004, and high conditions, from .45 to .60, t(621)=3.848, p ≤ .001. Second, we fit a set of Mixed Model Logistic Regression model with dummy variables representing the five rounds (immediately after norm talk, two rounds after norm talk, etc.; five rounds after norm talk was the reference). Again, individuals were nested under groups. Table 3 indicates that the increased cooperation is maintained the first three turns after norm talk, relative to the reference point of the fifth round after norm talk.

The analyses verify the spiking pattern we noted. Next, we examine the micro-level dynamics that produce this spike in an effort to understand how norm talk may boost the level of cooperation, at least momentarily.

Table 3: General Linear Mixed Model Logistic Regression predicting cooperation in the norm talk conditions only for rounds 5 to 31.

	Coef.	Odds Ratio
Intercept	-.952	.386
Benefit	1.532†	4.625
Time	-.074*	.929
Time^2	.001	1.001
Benefit X Time	.018	1.018
Round After Norm Talk [a]		
+1	.792***	2.207
+2	.599***	1.821
+3	.332*	1.393
+4	.212	1.236
Group Intercept	4.904	
Percent Correctly Predicted	80.6	
N	2806	

Note: *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .1$. [a] +5 rounds after norm talk is the reference

Norm talk: What does it communicate and what are the consequences?

Participants in the norm talk conditions sent messages a majority of the time (513/648; 79.2%). When they sent messages, they almost always used both injunction and gossip (490; 95.5%); When a message was sent, the receiver almost always replied (474; 92.4%). When a gossip message was sent, people also included the approval-disapproval component (482/512; 94.1%), resulting in 314 (61.2%) of the total messages being coded as prosocial gossip (i.e., approving of cooperation and disapproving of defection). See Supplementary Information (Table S8) for a detailed breakdown of the message content.

Truth, gossip, injunction, and replies. The gossip message was overwhelmingly truthful with a .89 correlation ($p \leq .001$) between the valence of the sender's gossip message and the target's cooperation level that was presented to the sender (also see Sommerfeld et al., 2007). When both gossip and injunctive messages were sent, the prosociality of the gossip

(1 versus 0) correlated with the injunction at $r=.55$ (p ≤ .001), suggesting that, although correlated, gossip and injunction are not always consistent. Specifically, the most extreme prosocial injunctions (i.e., those coded +2) were frequently accompanied by prosocial gossip (205/235; 87.2%) and the extremely self-interested injunctions were infrequently accompanied by prosocial gossip (20/96; 20.8%), yet the less extreme and neutral injunctions (i.e., coded -1, 0, +1) were about as likely to be accompanied by prosocial gossip as not (85/159 prosocial gossip; 53.5%). Positive replies were moderately correlated with prosocial gossip (.31, p ≤ .001) and with injunctions (.40, p ≤ .001).

Table 4: General Linear Mixed Model Logistic Regression predicting sender and receiver cooperation after norm talk in Experiment 1

	Senders		Receivers	
	Coef.	Odds Ratio	Coef.	Odds Ratio
Intercept	-1.189	.305	-1.349†	.259
Controls				
High Benefit	1.153	3.167	.626	1.870
Time	-.097	.908	-.058	.943
Time^2	.002	1.002	.001	1.001
High X Time	.035	1.036	.029	1.029
Prior turn individual cooperation	2.772***	16.000	2.981***	19.701
Prior turn others cooperation	.126	1.135	-.258	.773
Information about target	-.163	.850		
Sender Norm Talk				
Prosocial Gossip	.054	1.056	.048	1.049
Prosocial Injunction	.499***	1.647	.175	1.192
Receiver Response				
Positive Reply	.337†	1.401	-.298	.743
Prosocial Gossip x Positive Reply	-.162	.850	.379	1.461
Prosocial Injunction x Positive Reply	-.127	.881	.333***	1.395
Group Intercept	1.941		1.487	
Percent Correctly Predicted	84.4		81.4	
N	513		513	

Note: *** p < .001; ** p < .01; * p < .05; † p < .1.

Receiver behavior: Replies to norm talk and cooperation. We first focus on the receivers' cooperation. We expected that the receivers' acceptance of an injunction would be

critical in predicting cooperation. It is when the receivers accept the senders' injunctions that they would cooperate. We fit a Mixed Model Logistic Regression to receiver's next-round cooperation using sender's norm talk and the receiver's response as predictors and including only the cases where norm talk was sent. We also controlled for the relevant experimental and other variables endogenous to the evolving levels of cooperation in the groups (i.e., the receiver's cooperation in the previous turn, the cooperation of other members of the group). Table 4's right columns reports the results. Consistent with our expectation, the receiver tends to cooperate when he or she receives an injunctive message *and* accepts it (i.e., prosocial injunction x positive reply interaction).

Sender behavior: Norm talk and cooperation. We then explored the relationship between norm talk and cooperation focusing on the senders. Table 4's left columns reports the result of a Mixed Model Logistic Regression analysis of sender's immediate cooperation on the basis of sender's gossip and injunction message content. We controlled for potential confounds as described in the receiver analysis above, as well as the information that the sender had received about the target's cooperation over the last five rounds. This analysis showed that senders who have cooperated before, and sent a prosocial injunction, were more likely to cooperate in the next turn. Thus, norm talk can be interpreted as pledges. In this instance, the receiver's reply appears to have little effect if any.

Discussion

Consistent with our expectation, norm talk helps sustain cooperation in repeated public goods games. Although cooperation declines over time as previous studies have shown (e.g., Ledyard, 1995), norm talk helped to arrest this decline. These findings are similar to Feinberg et al.'s (2014) gossip only condition, in that gossip alone can arrest the decline of cooperation. However, we have extended their finding by showing that gossip does not have

to be shared by everyone in a group in order for it to be effective. In our experiment, a sender sent gossip to only one receiver at a time within a social network.

The subsequent micro-level analyses supported our hypothesis that the exchange of injunctive norm talk – the combination of the sender's injunctions and the receivers' acceptances – drives up the cooperation levels of both the sender and receiver immediately following it. The senders tend to cooperate after sending their injunctive messages, and the receivers who respond positively to these tend to cooperate as well. Interestingly, it is the sending and replying to injunctions, rather than gossip, that explains the spike in cooperation following the norm talk. Nonetheless, this effect dissipates on the average after two or three rounds. This spiking pattern is in line with Cialdini et al.'s focus theory (1990, 1991), which suggests that norm salience drives normative conduct. If the salience of the norm declines across rounds, its effect on cooperation can be expected to dissipate. There may be other mechanisms that could account for this effect; we will explore this possibility more fully in the general discussion.

The fact that the norm talk effect lasted up to the third round after the norm talk occurred is less consistent with an alternative mechanistic account. According to this account, norm talk could work by establishing a temporary mutual agreement between the sender and receiver to conditionally cooperate (vs. a norm of cooperation). At minimum, these findings suggest that an agreement to cooperate in the next round is not all that is going on. It is also important to note that in this paradigm (which does not permit participants to sanction each other for breaking an agreement) a mutual agreement to cooperate must nevertheless be backed up by the norm of promise keeping (Bicchieri, 2002) or commitment and consistency (Kerr & Kaufman-Gilliland, 1994). In other words, agreement to cooperate may be interpretable as an expression of the more specific norm of promise keeping, which can be

understood as secondary to the broader norm of cooperation. We will revisit this issue in Experiments 3 and 4.

Our findings that gossip tends to communicate accurate information about the target person, and is likely to be a moral and prosocial act, corroborates previous work (e.g., Feinberg et al., 2012; Peters & Kashima, 2015). At the same time, we did not observe gossip-specific effects on cooperation, which is somewhat surprising in light of previous findings to the contrary (Feinberg et al., 2014). At this point, we surmised that there may be two possible reasons why gossip did not have a direct effect on senders' or receivers' cooperation. One possibility is that broadcast gossip as used in Feinberg et al. (2014) may enhance cooperation, but interpersonal gossip used in the current research has a weaker effect. Another possibility is that when gossip is combined with injunction, gossip may be regarded as an example or evidence that is used in support of the injunction. In this case, the senders' and receivers' thoughts may be directed primarily to the injunction itself, and only secondarily to the gossip. As a result, both cognitive and communicative processes may be mostly "about" injunctions. We examined these possibilities in Experiment 2.

Experiment 2

In order to examine the difference between injunction and gossip further, and to replicate and shed light on the process when injunction and gossip messages are combined, we factorially manipulated the capabilities to send injunction and gossip in Experiment 2. There were four conditions altogether: in the control condition, no norm talk was allowed; in the injunction only condition, only injunction was allowed; in the gossip only condition, only gossip was allowed; and in the combined condition, both injunction and gossip were allowed. These conditions all used an MPCR of .4, replicating the high benefit conditions in Experiment 1.

As in Experiment 1, we expected that cooperation would be sustained longer in the norm talk conditions than in the control condition, and that this is due to the capacity of norm talk to encourage and sustain cooperation in the short term, so that cooperation would "spike" following the communication. In addition, this design allowed us to examine the processes involving injunction and gossip separately without interference from each other. In particular, we wished to examine the role of receiver replies in the sustenance of cooperation. Our analysis suggests that receiver replies should play a critical role in sustaining cooperation in the injunction only condition, such that it is when the receivers accept the injunctions that they will cooperate. However, they should play a minor role if any in the gossip only condition. By comparing the combined condition to these injunction and gossip only conditions, we should be able to diagnose to what degree the ability of norm talk to sustain cooperation depends on inclusion of both the injunction and gossip components.

Method

Experiment 2 methods were identical to Experiment 1 with the following exceptions. In the instructions, participants were shown all of the messages, and told that they may be allowed to send none, some, or all of the messages. In the gossip only condition, participants were only presented with the dropdown boxes for gossip and approval of gossip. In the injunction only condition, participants did not receive any information about a target's previous behavior, they were also only presented with one dropdown box for the injunction message. Reply options were identical to Experiment 1.

Results

A total of 208 participants, different from those in Experiment 1, in 52 usable groups were collected (as before, we aimed to collect a minimum of 12 complete groups per

condition). Only 30 of the 6448 (0.5%) public goods selections were missing and these were fairly evenly distributed among conditions (Table S7).

Norm talk again sustains cooperation over time. Figure 4 presents average percentage of cooperative responses in the four conditions. Generally replicating Experiment 1, the level of cooperation declined over time, most discernibly in the no communication control condition. However, the spikes found in Experiment 1 are again observable. The level of cooperation in the rounds immediately after norm talk picks up, followed by a gradual decrease, and this cyclical pattern continues in the norm talk conditions. Nevertheless, the effect of norm talk appears to differ as a function of the type of norm talk permitted.
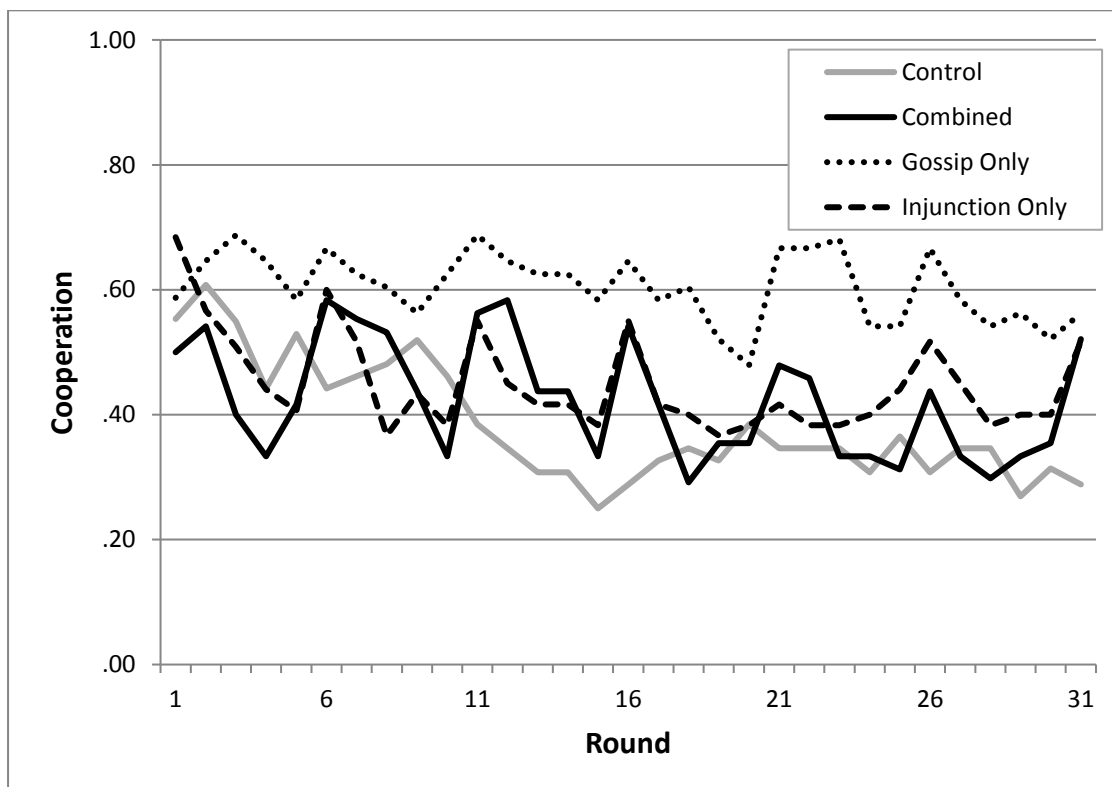

Figure 4: Mean cooperation levels in Experiment 2.

Mixed Model Logistic Regression models were fit to cooperation decisions (Table 5). Replicating Experiment 1, a linear trend of time was negative and significant, showing a gradually declining cooperation level; however, a quadratic trend was also significant, suggesting that the rate of decline eased over time. Most importantly, injunction and gossip interacted with time. Injunction and gossip can arrest the declining level of cooperation, thus helping to sustain cooperation over time. In particular, the gossip only condition showed the highest level of cooperation.

Table 5: General Linear Mixed Model Logistic Regression predicting cooperation in Experiment 2.

	Coef.	Odds Ratio
Intercept	.118	1.125
Injunction	.601	1.824
Gossip	1.370	3.934
Time	-0.90***	.914
Time^2	.001***	1.001
Injunction X Gossip	-1.848	.157
Injunction X Time	.019†	1.019
Gossip X Time	.030**	1.030
Injunction X Gossip X Time	-.029*	.972
Group Intercept	4.597	
% Correctly Predicted	75.7	
N	6418	

Note: *** p < .001; ** p < .01; * p < .05; † p < .1.

Next, we fit a Mixed Model Logistic Regression to cooperation decisions in rounds 6 to 31 in the norm talk conditions; dummy variables representing the one to four rounds after norm talk were also included (Table 6). This time, one and two rounds after norm talk showed higher levels of cooperation relative to the reference of five rounds after the communication. Nevertheless, the pattern was very similar to that found in Experiment 1.

Table 6: General Linear Mixed Model Logistic Regression predicting cooperation (Norm talk conditions, round 6 to 31 only) in Experiment 2

	Coef.	Odds Ratio
Intercept	1.262	3.532
Injunction	-1.195	.303
Gossip	.092	1.096
Time	-.045	.956
Time^2	.001†	1.001
Injunction X Time	-.014	.986
Gossip X Time	-.018	.982
Round After Norm Talk [a]		
+1	.848***	2.335
+2	.482***	1.620
+3	.162	1.176
+4	.120	1.127
Group Intercept	6.189	
Percent Correctly Predicted	77.7	
N	4051	

Note: *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .1$.

Norm talk: What does it communicate and what are the consequences? The content of Experiment 2's messages had similar patterns to Experiment 1. Participants in the norm talk conditions sent messages a majority of the time (804/936; 85.9%). When a message was sent, the receiver almost always replied (761; 94.7%). When a gossip message was allowed and sent, the approval-disapproval component was usually included with it (469/508; 92.3%). When a message was sent in gossip-only or combined conditions, 64.1% (327/ 510) of them were coded as prosocial gossip (see Table S9 for more details).

Truth, gossip, injunction, and replies. Gossip message valence was highly correlated to the information the sender received about the target's cooperation (.84, $p \le .001$) again indicating a lack of malicious false gossip. In the combined condition when both gossip and injunction were included in the same message, the prosociality of the gossip (0 or 1) and

injunction messages were moderately correlated (.32, p ≤ .001). Similar to Experiment 1, extremely prosocial injunctions (i.e., coded +2) were frequently accompanied by prosocial gossip (88/115; 76.5%); however, extremely individualistic (i.e., coded -2) and moderate (i.e., coded -1, 0, +1) injunctions were mixed as to their inclusion of prosocial gossip (10/24; 41.7% and 45/91; 49.5%, respectively). Mirroring Experiment 1, positive replies were moderately correlated with both prosocial gossip (.28, p ≤ .001, 480 cases) and prosocial injunction (.43, p ≤ .001, 503 cases).

Injunction only condition. Table 7 reports the results of the General Linear Mixed Model analyses for the injunction only condition. First, focusing on the receivers, receiving an injunction *and* accepting it increased their immediate cooperation. Consistent with our expectation, the receivers' receipt and acceptance of an injunction was a significant predictor of their cooperation. Next, looking at the senders, their own norm talk did not predict their cooperation, but it was their receipt of the receivers' acceptance that prompted their immediate cooperation. This difference from Experiment 1 highlights the importance of the receivers' acceptance as a mechanism that sustained cooperation when injunction is the only form of norm talk.

Table 7: General Linear Mixed Model Logistic Regression predicting sender and receiver cooperation after norm talk in the injunction only condition in Experiment 2

	Senders		Receivers	
	Coef.	Odds Ratio	Coef.	Odds Ratio
Intercept	-.802	.448	-1.699*	.183
Controls				
Time	-.087	.917	.008	1.008
Time^2	.001	1.001	-.001	.999
Prior turn individual cooperation	2.497***	12.144	2.578***	13.170
Prior turn others cooperation	.822**	2.276	.507*	1.660
Information about target	.004	1.004		
Senders				
Prosocial Injunction	.250	1.284	.467**	1.595
Receivers				
Positive Reply	.598**	1.819	.131	1.140

	Senders		Receivers	
	Coef.	Odds Ratio	Coef.	Odds Ratio
Prosocial Injunction X Positive Reply	-.157	.855	.363**	1.438
Group Intercept		.225		.297
Percent Correctly Predicted		83.7		83.3
N		294		294

Note: *** p < .001; ** p < .01; * p < .05; † p < .1.

Gossip only condition. In the gossip only condition, again, consistent with our expectation, receivers cooperated more when they received gossip, but cooperation was not contingent on how they replied (Table 8). Regardless of whether they accepted gossip or not, the sheer receipt of the gossip message was sufficient to instigate cooperation. The effectiveness of gossip in sustaining cooperation does not depend on how receivers respond to the gossip. Intriguingly, senders' sending prosocial gossip did not enhance their likelihood of immediate cooperation.

Table 8: General Linear Mixed Model Logistic Regression predicting sender and receiver cooperation after norm talk in the gossip only condition in Experiment 2

	Senders		Receivers	
	Coef.	Odds Ratio	Coef.	Odds Ratio
Intercept	-.254	.776	-1.238	.290
Controls				
Time	.042	1.042	.092	1.097
Time^2	-.002	.998	-.003	.997
Prior turn individual cooperation	2.281***	9.785	2.446***	11.547
Prior turn others cooperation	.003	1.003	-.133	.875
Information about target	-.125	.882		
Senders				
Prosocial Gossip	.072	1.075	.951*	2.587
Receivers				
Positive Reply	-.180	.835	-.270	.764
Prosocial Gossip X Positive Reply	.512	1.668	.294	1.342
Group Intercept	1.325		1.059	
Percent Correctly Predicted	83.8		83.4	
N	259		259	

Note: *** p < .001; ** p < .01; * p < .05; † p < .1.

Combined condition. This is a condition that closely replicates Experiment 1 in which both injunction and gossip can be communicated (Table 9). The findings are generally in line with the results from Experiment 1. The receiver's acceptance of the sender's injunction predicted the receiver's cooperation immediately afterwards. Thus, the injunction-acceptance exchange appears to drive the cooperation of the receiver. Somewhat different from Experiment 1, we found a gossip x reply interaction effect on receivers' cooperation, suggesting that the receiver's acceptance of gossip messages too was a booster for the receiver cooperation. Again replicating Experiment 1, a sender's sending an injunction predicted the sender's cooperation immediately following the norm talk.

Table 9: General Linear Mixed Model Logistic Regression predicting sender and receiver cooperation after norm talk in the combined condition in Experiment 2

	Senders		Receivers	
	Coef.	Odds Ratio	Coef.	Odds Ratio
Intercept	.900	2.458	.305	1.357
Controls				
Time	-.137	.872	-.084	.919
Time^2	.003	1.003	.001	1.001
Prior turn individual cooperation	2.091***	8.091	2.301***	9.981
Prior turn others cooperation	-.322	.725	-.181	.835
Information about target	-.124	.884		
Senders				
Prosocial Gossip	.265	1.303	-.290	.748
Prosocial Injunction	.419**	1.521	-.062	.940
Receivers				
Positive Reply	.393†	1.480	-.584*	.558
Prosocial Gossip X Positive Reply	.011	1.011	.643*	1.902
Prosocial Injunction X Positive Reply	-.201†	.818	.504***	1.655
Group Intercept	.947		.220	
Percent Correctly Predicted	77.7		76.9	
N	251		251	

Note: *** p < .001; ** p < .01; * p < .05; † p < .1.

Discussion

As expected, norm talk can help sustain cooperation by boosting cooperation when it occurs. Replicating the results in Experiment 1, when participants could send injunction, gossip, or both, they could talk themselves into cooperation by increasing their levels of cooperation although the booster effect was relatively short-lived. Somewhat surprisingly, the level of cooperation was higher in the gossip only condition relative to the other norm talk conditions. We will return to this finding later.

The detailed analyses of micro-level norm talk showed that the mechanisms for sustaining cooperation subtly differ for injunction and gossip. For injunctions, the receiver's response was, as expected, critical. The "pure" injunctions made by the senders do not predict their cooperation in the injunction only condition without their receivers' acceptance. Likewise, receivers were more likely to cooperate when they accepted the senders' injunction. In the absence of receiver acceptance, the sender and the receiver were less likely to cooperate. Thus, mutual norm verification, i.e., the sender's message making salient the norm of cooperation and the receiver's agreement verifying the norm, was an important determinant of the effectiveness of pure injunctions. As we noted in Experiment 1, it is also possible that a receiver's acceptance is interpreted as a pledge for mutual conditional cooperation (i.e., "I'll cooperate if you cooperate."). We address this possibility in Experiment 3.

In contrast, for gossip, the receiver's response did not predict cooperation. Pure gossip can sustain cooperation as long as it is sent – Beersma and Van Kleef (2011) may be right in emphasizing the amount of gossip circulated in a group as a critical factor that effects cooperation. It is important to note that the booster effect of pure gossip *cannot* be interpreted in terms of a mutual commitment for conditional cooperation. We suggest that gossip acts as a reminder of the injunctive norm and the reputational system that need to be monitored in social settings.

Within this spectrum, the role of norm talk was somewhat more complex in the combined condition where both injunction and gossip are available. The senders of an injunction, which often accompanied gossip, tended to cooperate after sending it, and their receipt of the receivers' acceptance did not have a consistent effect (in both Experiment 1 and 2, they were only marginally significant). As expected, the receivers of an injunction tended to cooperate when they accepted the injunction in both Experiments 1 and 2. In Experiment 2, however, the receivers' acceptance of gossip also indicated their cooperation too, though this was not observed in Experiment 1. In all, consistent with our expectation, the receivers' acceptance of a norm talk message plays an important role in the combined condition too; however, the senders who sent injunctive messages were also likely to cooperate.

Overall, the exchange of sender injunction and receiver acceptance appears to be a critical mechanism that drives cooperation when injunctions are sent. In both the injunction only and combined conditions, receivers' cooperation was predicted by their receipt and acceptance of injunctions. In contrast, the effectiveness of gossip in increasing cooperation does not appear to depend on the receiver's acceptance. We examine the meanings behind the exchanging of these norm talk messages in Experiment 3.

Experiment 3

Experiments 1 and 2 provided evidence that norm talk – injunction, gossip, or their combination – can sustain cooperation in repeated public goods games by making the injunctive norm of cooperation salient. Nevertheless, these studies have not provided direct evidence for our claims about the mechanisms. Our first aim here is to test our claim that injunction increases the salience of the cooperative norm, and that the receiver's positive response verifies it. Our second aim is to explore an alternative mechanism, which is that the injunction-acceptance exchange is a mutual commitment to conditionally cooperate. In other words, it may be that the sender and the receiver are not so much agreeing that it is a good

thing (i.e., normatively right) to cooperate, as saying to each other, "if you cooperate, I'll cooperate." Note that while our reasoning does allow for the possibility that injunction can invoke a cooperation pledge, we expect such a pledge to be less salient than the norm and its verification. Our third aim is to test our claim that gossip simultaneously increases the salience of the underlying injunctive norm and the reputational system.

To address these issues, Experiment 3 examined participants' interpretation of norm talk consisting of injunction and/or gossip. Specifically, they were asked to indicate the extent to which each instance of norm talk implied (1) an injunctive norm of cooperation, (2) an intent to conditionally cooperate, and (3) reputational concern. Norm talk took the form of prosocial messages (e.g., "We should each contribute a lot to the group."), and we compared ratings for these messages with two controls: non-prosocial messages (e.g., "We should each keep some for ourselves and contribute some to the group.") and no message. We also asked participants to interpret a complete sender – receiver exchange.

We hypothesize that prosocial injunction messages (with or without gossip) would be more likely interpreted as a reminder of the injunctive norm of cooperation than an intention to conditionally cooperate. Further, we hypothesize that receivers who agree with these prosocial messages would be more likely seen to be verifying the norm than conveying an intention to conditionally cooperate. We also hypothesize that pure gossip messages would be interpreted as a reminder of both the injunctive norm of cooperation and the reputational system.

Method

Participants. We estimated the number of participants required to obtain power = .80 under the assumption that one of the critical hypotheses required an F-test to compare at most five conditions ($df_{hyp} = 4$) and at least three conditions (to be described in detail below) with a medium to strong effect size ($d \approx .65$). According to Murphy and Myors (2004), this requires

df_{error} of approximately 120, or 20~30 per condition. To err on the conservative side, we aimed to collect data between approximately 40 and 50 per condition, because we expected a fairly large attrition rate in a sample recruited from Amazon Mechanical Turk based on our previous experience. In the end, 282 participants completed the study; no further data collection was undertaken. A catch question was included in the section about participants' interpretations of the situation in which no message was received. It asked 'If you read this item, please respond "Strong No"'. 61 participants who responded other than "Strong No" were deemed not to be paying sufficient attention to this task, and therefore removed from further analyses. In the end, 193 participants (92 men and 101 women) were retained (see Supplementary Information for the same analyses including those removed. We found results largely consistent with those reported here).

Design and Procedure. Participants were told that the purpose of the study was to gain insight into the thinking of participants who had taken part in a previous experiment on decision making. To familiarise them with the previous experiment, participants were provided with a simplified version of the norm talk instructions. They were also given the opportunity to play two rounds of the public goods game (with pre-set feedback) and to complete two communication rounds. In the communication rounds they were asked to use the drop-down menus described in the previous experiments to construct one message about a player who had contributed on all of the previous five rounds and a second message about a player who had contributed on none of these rounds.

Participants were then asked to take the receiver role and they were presented with different types of norm talk as a function of their allocation to one of five different conditions. As can be seen from Table 10, there was one *injunction only* condition. There were also two *gossip only* conditions and two *combined gossip and injunction* conditions, reflecting the fact that prosocial gossip can be framed positively or negatively: "Person C

contributed very much, and I strongly approve" or "Person A contributed very little, and I strongly disapprove".

Table 10. Sender messages presented to participants

	Prosocial Message	Non-prosocial Message
Injunction	We should each contribute a lot to this group.	We should each keep some for ourselves and contribute some to the group.
Gossip (Positive Frame)	Person C contributed very much and I strongly approve.	Person C contributed very much and I neither approve nor disapprove.
Gossip (Negative Frame)	Person C contributed very little and I strongly disapprove.	Person C contributed very little and I neither approve nor disapprove.
Combined (Positive Frame)	Person C contributed very much and I strongly approve. We should each contribute a lot to this group.	Person C contributed very much and I neither approve nor disapprove. We should each keep some for ourselves and contribute some to the group.
Combined (Negative Frame)	Person C contributed very little and I strongly disapprove. We should each contribute a lot to the group.	Person C contributed very little and I neither approve nor disapprove. We should keep some for ourselves and give some to the group.

In each condition, participants were asked to report their interpretation of three different messages: *prosocial*, *non-prosocial*, and no message. They were also asked to indicate how they would respond to the prosocial and non-prosocial messages, and describe what they meant by their response. Interpretations were elicited with the following statement: "By sending this message, [the person is] basically saying…". Participants were then asked to respond to five items on 5-point scales (where +2 = strong Yes, +1 = yes, 0 = neither, -1 = no and -2 = strong no). Two items measured an *injunctive norm*, "I think people should contribute" and "It is a good thing to contribute" (they were averaged due to good reliability; $\alpha > .77$). Two items tapped into *reputational concern*, "I don't like people who don't contribute" and "I wouldn't like to do this task with people who don't contribute" (they were

averaged because their reliability was acceptable, α > .69). One item measured *intention to conditionally cooperate*, "I will contribute if you contribute." The order of items was randomized across presentations.

After completing the sender message and receiver response ratings, participants were presented with a norm talk-agreement exchange consisting of a prosocial message and the response, "I strongly agree with you." They were then asked to rate their interpretation of the exchange in terms of norm verification ("They agree that people should contribute") and mutual commitment ("They promise each other that they will contribute"). Participants responded to several additional items that we do not describe here. They were removed from analyses because they were either irrelevant to the hypothesis or because our item analyses found that the factor structure of the reduced set of items was clearer and less ambiguous than the full set. Please see Supplementary Information (Analyses for Experiment 3, Interpretation Ratings) for further detail and evidence that, with one exception, our results hold if we repeat our analyses with the full set of items and entire sample (i.e., including those who failed the manipulation check)[1]. After completing one more section, which is unrelated to this experiment, participants were asked about their demographics, debriefed, and thanked for their participation.

Results

  Preliminary Analyses. We first examined whether there is a meaningful difference between positively and negatively framed messages with a view to aggregating across them. We used t-test to compare the interpretation ratings for gossip and combined messages separately (this resulted in 31 comparisons per condition). Only three comparisons were significant, all other |t| < 1.952. Specifically, in the combined condition, the positively framed

---

[1] When the analyses are conducted with the entire sample and full scales, the significant interaction between message prosociality and norm talk type on ratings of reputational concern no longer achieves significance, although the descriptive pattern of results is consistent.

message was rated as providing a stronger signal of the norm of cooperation, M = 1.63 vs. 1.17, t(77) = 2.11, p = .038, and a weaker signal of reputational issues, M = .55 vs. 1.07, t(77) = 2.46, p. = .016, than the negatively framed message. In the gossip condition too, the positively framed message was rated as providing a weaker signal of reputational issues, M = .49 vs. 1.11, t(72) = 2.59, p = .022. Because there are only a few significant comparisons (and approximately as many as would be expected based on chance), and the differences that were observed were of quantity rather than quality (i.e., they all fell on the same side of the neutral scale point), we aggregated across frame. Further support for this decision was provided by a cluster analysis using the mean ratings on the items for each of the five norm talk types. The Between-Groups Linkage algorithm showed a clear pattern of two clusters: one cluster for both the positively and negatively framed gossip messages, and the other cluster for the injunction and positively and negatively framed combined messages. As a result, the analyses that we report below relate to three types of norm talk: injunction, gossip, and combined messages.

Injunction signals norms more than conditional cooperation. To test our hypothesis that injunction messages (with or without gossip) are interpreted more as a reminder of the injunctive norm than an offer of conditional cooperation, we first compared injunctive norm and conditional cooperation ratings. We examined these ratings for prosocial and control messages (where control ratings were the average for non-prosocial and no messages), and explored the difference between injunction and combined messages, using a mixed-design three-way ANOVA on ratings as the dependent variable with interpretation type (injunctive norm vs. conditional cooperation) and prosociality of message (prosocial vs. control) as within-participant factors and message type (injunction vs. combined) as a between-participant factor. There was a large effect of interpretation type, F(1,117) = 65.93, η_p^2 = .36, p = .000, where the injunctive norm interpretation (M = .81, SE = .06) was

endorsed more than the conditional cooperation interpretation (M = .16, SE = .07). This was qualified by an interpretation type x prosociality interaction, $F(1,117) = 46.12$, $\eta_p^2 = .28$, p = .000, indicating that the norm interpretation was far more strongly endorsed for the prosocial relative to the control message, M = 1.50 vs. .13, than conditional cooperation interpretation, M = .43 vs. -.11.

There was an effect of prosociality, $F(1,117) = 97.26$, $\eta_p^2 = .49$, p = .000, and a two-way effect of prosociality and message type, $F(1,117) = 5.65$, $\eta_p^2 = .05$, p = .011, such that the contrast between prosocial and control messages was greater for the injunction (M = 1.08 vs. -.11) than for the combined messages (M = .86 vs. .13), but there was no three-way, $F(1,117) = 1.60$, $\eta_p^2 = .01$, p. = .208, or other effect. Table 11 reports the mean ratings. Clearly, prosocial injunction and combined messages were interpreted as a reminder of the injunctive norm of cooperation; both means were significantly greater than the neutral point of zero (1.57 and 1.42). Participants saw prosocial injunction or combined messages as an offer of conditional cooperation to some extent. Both means were greater than the neutral point (.58 and .29), suggesting that injunctions connote conditional cooperation; however, they were lower than those for injunctive normative interpretations as shown by the paired sample t-tests (Table 11).

Table 11: Mean ratings of the prosocial message and third party interpretations.

Message Interpretation	Injunction		Combined		Gossip	
	Prosocial	Non	Prosocial	Non	Prosocial	Non
Injunctive Norm	1.57**	.05	1.42**	.21**	1.53**	-.09
Conditional Cooperation	.58**	-.27**	.29*	.06	.23	-.29**
Reputational Concern	.71**	.16	.83**	-.22**	.98**	-.34**
t-test comparing injunctive norm and conditional cooperation	5.21**		7.88**		9.94**	
Third Party Interpretation						
Norm Verification	1.63**		1.59**		1.62**	
Conditional Cooperation	1.05*		1.10**		.28*	

t-test comparing norm verification and conditional cooperation	3.80**	4.85**	9.27**

Note: t-test of H_0: $\mu = 0$, ** $p < .01$; * $p < .05$. For independent sample t-tests, ** $p < .01$; * $p < .05$. Ratings vary from -2 (Strong No), 0 (Neither), to +2 (Strong Yes).

Second, we examined the meaning of the receivers' response. Recall that, after they were presented with a sender's message, the participants chose their response as a receiver (i.e., whether they agreed with the message or not) and then rated the meaning that they intended to convey with the response. We coded the response in terms of receiver agreement, that is, the degree to which the response agreed with the sender message in the exact same manner as in Experiments 1 and 2. We then correlated this index with the extent to which participants meant to verify the underlying injunctive norm (i.e., norm verification, or agreeing with the injunctive norm) and the extent to which participants mean to communicate their intent to conditionally cooperate (i.e., mutual conditional cooperation, or agreeing that they would cooperate if the sender cooperates).

Table 12 reports the correlations. When receivers agreed to an injunction or a combined message, they clearly implied norm verification. Receiver agreement highly correlated with norm verification for injunction (.82), and also for combined messages (.60). It is noteworthy that receiver agreement also correlated with conditional cooperation for injunction (.46), suggesting that receivers may convey an intent to conditionally cooperate when they agreed with an injunction. However, receiver agreement appears to imply norm verification more than conditional cooperation. The correlation of receiver agreement with norm verification was significantly greater than that with conditional cooperation (Hoerger, 2013; Steiger, 1980). Furthermore, the partial correlation of receiver agreement with norm verification remained significant when conditional cooperation was controlled, but the reverse was not true – the partial correlation of receiver agreement with conditional cooperation became nonsignificant when norm verification was controlled, suggesting that

the receiver agreement-conditional cooperation correlation was a by-product of norm verification. There is no evidence that receiver agreement with a combined prosocial message implied an intent to conditionally cooperate. Receiver agreement did not correlate with conditional cooperation for combined messages (.18).

Table 12: Correlation of receiver agreement with norm verification and conditional cooperation.

Interpretation	Injunction	Combined	Gossip
Norm Verification (rnv)	.817**	.603**	.657**
Conditional Cooperation (rcc)	.462**	.179	.165
Comparison between rnv and rcc	3.56**	3.62**	4.55**
Partial correlation with Norm Verification controlling for Conditional Cooperation (prnv.cc)	.760**	.586**	.658**
Partial correlation with Conditional Cooperation controlling for Norm Verification (prcc.nv)	.050	-.009	-.173

Note. ** p. < .01.

Finally, we examined how participants interpreted from a third party perspective an exchange that involved a sender's prosocial norm talk and a receiver's agreement with it. As a third party, participants interpreted norm talk-agreement exchanges more as norm verification than mutual commitment to cooperate. The mean ratings for norm verification were higher than those for mutual commitment to cooperate in all conditions (see Table 11). To verify this, we conducted a mixed-design ANOVA on third party ratings with interpretation type (norm verification vs. conditional cooperation) as a within-participant factor and message type (injunction vs. combined) as a between-subject factor. The only significant effect was due to interpretation type, $F(1,117) = 35.65$, $\eta_p^2 = .234$, p = .000. The norm verification interpretation was endorsed more (M = 1.61, SE = .062) than the conditional cooperation interpretation (M = 1.08, SE = .097). There was some indication that these exchanges implied a degree of mutual commitment to cooperate (see Table 11); the mean ratings of conditional cooperation were above the midpoint. However, norm

verification was rated to be more pertinent than mutual commitment in all conditions (Table 11).

Gossip is a reminder of norm and reputation. We theorized that gossip explicitly speaks to a particular individual's reputation, but in so doing implicitly implies the underlying injunctive norm. So, first, we expected that gossip would be interpreted as implying the injunctive norm of cooperation, potentially less so than injunction and combined messages. This was indeed the case. The injunctive normative interpretation conferred to prosocial gossip was rated at 1.53, significantly above the neutral midpoint of zero (Table 11). Furthermore, this was significantly greater than the control, $t(74) = 15.30$, $p < .001$. We examined whether gossip was interpreted differently from injunction and combined messages by conducting a repeated measures ANOVA on normative interpretation with message prosociality (prosocial vs. control) as a within-subject factor and norm talk type (injunction vs. combined vs. gossip) as a between-participant factor. There was a strong effect of prosociality, $F(1,190) = 442.70$, $\eta_p^2 = .700$, $p = .000$, showing that all prosocial norm talk was normatively interpreted ($M = 1.51$, $SE = .06$) compared to the control ($M = .06$, $SE = .05$). Overall, an inspection of the means suggests that gossip appears to be interpreted as a reminder of the injunctive norm at a level similar to pure injunction or combined messages (Table 11). However, there was a significant interaction between prosociality and norm talk type, $F(2, 180) = 3.93$, $\eta_p^2 = .040$, $p = .021$, which suggests that, if anything, relative to the control, prosocial gossip implicates the injunctive norm (prosocial – control = 1.62) as much as prosocial injunction (prosocial – control = 1.53), and even more than prosocial combined messages (prosocial – control = 1.22).

In addition, we suggested that gossip would be interpreted as implicating reputational concerns more than injunction. This expectation was also supported (for means, see Table 11). When we conducted an analogous ANOVA with reputational concern as the dependent

variable, there was a main effect of message prosociality, $F(1,190) = 158.09$, $\eta_p^2 = .454$, p = .00; however, this was moderated by norm talk type, $F(2,190) = 7.10$, $\eta_p^2 = .070$, p = .00. Relative to the control, prosocial gossip was seen to implicate reputational concerns (prosocial – control = 1.32) more than prosocial injunction (prosocial – control = .56). Prosocial combined messages (prosocial – control = 1.05) were perceived as being between pure injunction and pure gossip, but perhaps more akin to gossip. This makes sense because combined messages contained gossip information as well.

Finally, we examined what the receiver's agreement with gossip may mean. Table 12 reports the correlation of the receiver's agreement with norm verification and also conditional cooperation. Clearly, agreeing to gossip implies norm verification (r = .657, p < .001), but not conditional cooperation (r = .165, n.s.). Also, the correlation of receiver agreement with norm verification is marginally weaker for gossip (r = .657) than for injunction (r = .817), z = 1.78, p = .075. Put differently, disagreement with pure injunction can undermine the normative status of the injunction marginally more than disagreement with pure gossip, making gossip less vulnerable to a challenge than pure injunction. In this regard, combined messages appear to be similar (r = .603) to gossip.

Discussion

Experiment 3 was designed to examine three main issues. First, we sought and obtained evidence for our claim that injunction and combined norm talk remind its receivers of the injunctive norm of cooperation, and the receiver's agreement with this type of norm talk verifies the status of the injunction as a norm. As expected, injunction and combined norm talk was likely to be interpreted as an assertion of the norm of cooperation, the receiver's agreement to it was meant to be an expression of its verification, and the norm talk-agreement exchange was seen as a mutual verification of the norm of cooperation.

Second, we wished to examine the rival hypothesis that a norm talk-agreement exchange may be interpreted as an agreement to cooperate with each other conditional on the other party's cooperation. The participant ratings suggested that the participants also interpreted norm talk as potentially indicating the message sender's intention to conditionally cooperate. However, this interpretation was secondary to the first injunctive normative interpretation. The level of endorsement was always lower for conditional cooperation interpretation than for injunctive normative interpretation.

As well, as we noted earlier, even if injunction and combined messages imply a pledge to cooperate and an agreement to conditionally cooperate, the fact that this mutual agreement is carried out may be due to the norm of promise keeping (Bicchieri, 2002). This is because, in the present experiments, there was no tangible sanction associated with breaking a promise. In the context of social dilemmas, the injunctive norm of cooperation may also activate the more specific injunctive norm of promise keeping.

Third, we sought and found evidence for our claim that gossip conveys reputational implications, but at the same time acts as a reminder of the injunctive norm underlying the reputational judgments. As expected, gossip was seen to implicate reputations more than the pure injunction or combined messages. It was somewhat surprising that the injunctive normative interpretation was as strongly endorsed for gossip as for injunction and combined message. Apparently, without explicitly saying so, gossip implicates the norm of cooperation as much as its explicit statements, at least in the context of public goods dilemma games.

The results of Experiment 3 may shed some light on the finding that the level of cooperation was highest in the gossip only condition in Experiment 2, even higher than in the injunction only or combined conditions. First of all, gossip reminds its receivers of *both* the norm *and* the reputational system, more so than pure injunctions or combined messages. Therefore, gossip may be a powerful "cloak and dagger", firstly, holding out a silent threat of

reputational damage or a social reward of status gain, and secondly, simultaneously upholding the injunctive norm of cooperation and shielding it from challenges. Nonetheless, it remains to be seen whether gossip always outperforms injunction in sustaining cooperation in social dilemmas.

Experiment 4

Experiment 4 aimed to examine participants' interpretations of messages, and message-reply exchange in an actual public goods game, using the paradigm described in Experiments 1 and 2. This experiment also included a non-norm talk communication condition, in which a sender could make a pledge to contribute (i.e., conditional cooperation). Experiment 4 therefore included three conditions that varied the types of messages that were exchanged. In the injunction only and gossip only conditions, the messages and replies were the same as those in Experiment 2. In the conditional cooperation condition, participants were given a chance to send a message that promises a conditional cooperation, "I will definitely contribute if you agree to contribute. Do you agree to contribute?" and to reply by agreeing or disagreeing.

In this way, we aimed to answer three main questions. First, we examined whether norm talk and promises of conditional cooperation are equally effective at sustaining cooperation. We do this because Experiment 3 found that while norm talk more strongly implies the norm of cooperation than conditional cooperation, it is also interpreted as a conditional cooperation pledge. Second, we sought to replicate our Experiment 3 findings about receivers' interpretations of senders' messages. Specifically, we expected that messages would be interpreted as connoting injunctive norms to a greater degree in the injunction and gossip conditions than in the conditional cooperation condition. We also expected that the reputational issues would be rated most important in the gossip condition and least important in the conditional cooperation condition. Third, we examined whether

participants' interpretations for their message-reply exchanges – as norm verification, reputational concerns, or conditional cooperation – predicted their cooperation. In light of our Experiment 3 findings that agreement in both the injunction and gossip conditions were interpreted as implying norm verification, we expected norm verification to predict cooperation in both the injunction and gossip conditions, reputational concern to predict cooperation in the gossip condition, and conditional cooperation to predict cooperation in the conditional cooperation condition. However, in light of the results of Experiment 3 that norm talk also connotes conditional cooperation to some degree, we thought that conditional cooperation may also predict cooperation in the injunction and gossip conditions.

Method

The methods were identical to Experiment 2 with the following exceptions. First, we included the Experiment 3 message rating tasks into each communication round. This meant that upon receiving the message, and before replying, receivers were asked to interpret the message. Additionally, at the end of the communication round, senders and receivers were asked to rate their own interpretations of their message-reply exchange. This substantially increased the steps required to complete each communication round; to manage participant demands, we reduced the number of game and communication rounds (to 16 and 3, respectively). Second, rather than being presented with all of the messages, the sender saw and chose the message from one of the three sets of messages according to the condition. Messages for gossip and injunction conditions were identical to Experiment 2. In the conditional cooperation condition the messages were promises of conditional cooperation such as "*I will definitely contribute if you agree to contribute. Do you agree to contribute?*" (see Table S5 for details). Third, in all conditions, receivers could choose a reply from "*I strongly disagree with you*" to "*I strongly agree with you*". Fourth, we increased the

minimum number of groups per condition to 20 as differences in interpretation would require more statistical power to detect compared to differences in cooperation only.

Results and Discussion

A total of 256 participants in 64 usable groups were collected with at least 21 complete groups per condition. Only 8 of the 4096 (0.19%) public goods selections were missing and these were evenly distributed among conditions (Table S7).

Norm talk sustains cooperation over time. Figure 5 depicts the average percentage of cooperation in the three conditions. From this, while cooperation shows the expected temporal decline to some extent, it is not as marked as in Experiments 1 and 2 presumably due fewer decision rounds (16 as opposed to 31). Levels of cooperation appear to be comparable across the three conditions, but they appear to be somewhat higher in the norm talk conditions than in the conditional cooperation condition. To formally test these observations, Mixed Model Logistic Regression models were fit to cooperation decisions (Table 13). This revealed a significant negative quadratic trend of time, suggesting a decline towards the end of the time series, and a marginally significant effect of injunction, such that levels of cooperation in the injunction condition were somewhat higher than those in the conditional cooperation condition. The results imply that norm talk is at least as effective as a promise of conditional cooperation in sustaining cooperation, with a suggestion that injunction may be even more effective.
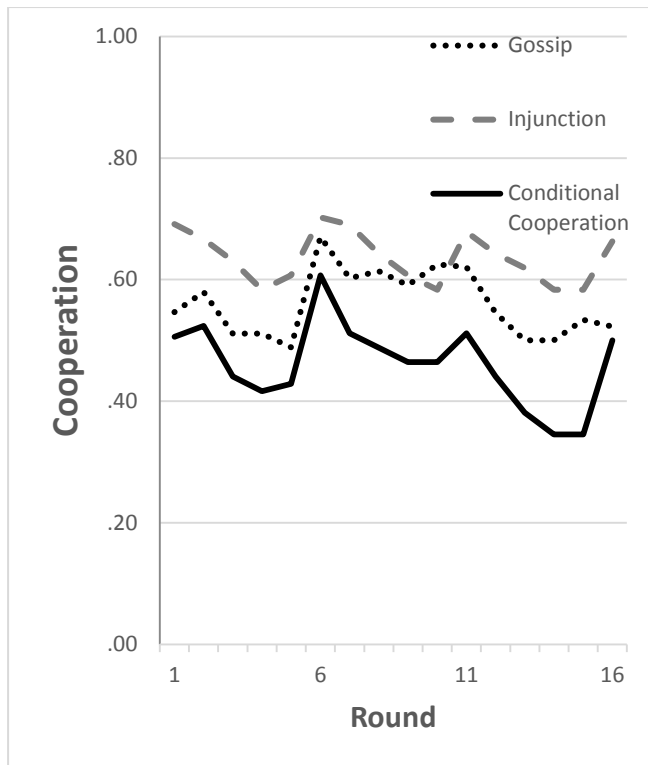
Figure 5. Mean cooperation levels in Experiment 4

Table 13: General Linear Mixed Model Logistic Regression predicting cooperation in Experiment 4.

	Coef.	Odds Ratio
Intercept	.003	1.003
Injunction (dummy)	1.199†	3.317
Gossip (dummy)	.387	1.473
Time	.041	1.042
Time^2	-.006**	.994
Injunction X Time	.029	1.030
Gossip X Time	.028	1.029
Group Intercept	4.810	
% Correctly Predicted	76.4%	1
N	4088	

Note: *** $p < .001$; ** $p < .01$; * $p < .05$; † $p = .098$. Reference for dummy coding = conditional cooperation condition.

Next, we fit a Mixed Model Logistic Regression predicting round 6 to 16 cooperation

decisions in all conditions, with dummy variables representing the rounds after

communication round (Table 14). There was a significant declining trend (negative linear trend of time). The round immediately after communication showed significantly higher levels of cooperation relative to the reference. Although the other effects were not significant, the estimated coefficients suggest that the effect of communication was greatest immediately following it and gradually waned over the subsequent rounds.

Table 14: General Linear Mixed Model Logistic Regression predicting cooperation for rounds 6 to 16.

	Coef.	Odds Ratio
Intercept	.967	2.629
Injunction (dummy)	.979	2.662
Gossip (dummy)	.698	2.009
Time	-.128	.880
Time^2	.001	1.001
Injunction X Time	.061	1.063
Gossip X Time	.002	1.002
Round After Norm Talk [a]		
+1	.490**	1.632
+2	.087	1.091
+3	-.052	.949
+4	-.164	.849
Group Intercept		
Percent Correctly Predicted	80.3	
N	2814	

Note: *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .1$. [a] +5 rounds after norm talk is the reference

Norm talk. The patterns of norm talk were similar to Experiments 1 and 2. Participants sent messages most of the time 765/768 (99.6%): 251/252 (99.6%) in the injunction condition, 262/264 (99.2%) in the gossip condition, and 252/252 (100%) in the conditional cooperation condition. When a message was sent, the receiver replied most of the time 747/768 (97.2%): 247/252 (98.0%) in the injunction condition, 256/264 (97.0%) in the gossip condition, and 242/252 (96.0%) in the conditional cooperation condition. A majority

of the messages was prosocial, 509/765 (66.5%): 186/251 (74.1%) in the injunction condition, 183/262 (69.9%) in the gossip condition, and 157/252 (62.3%) in the conditional cooperation condition. See Table S16 for complete counts of norm talk content.

Norm talk is interpreted as an assertion of an injunctive norm. The ratings of message interpretations were first analyzed using an ANOVA with message prosociality (sender's message was prosocial vs. not prosocial), condition (injunction, gossip, and conditional cooperation), and interpretation (injunctive norm, reputational concern, and conditional cooperation) with the last being a within-participant factor. Message prosociality had a main effect, $F(1,741) = 33.21$, $p < .001$, and this was not qualified by any higher order interaction effects, all $F < 1.54$. This suggests that prosocial messages were interpreted to have more prosocial implications ($M = .85$, $SD = .98$) than non-prosocial messages ($M = .38$, $SD = .99$). There were also main effects of condition and interpretation, $F(2,741) = 4.13$, $p = .016$, and $F(2,740) = 87.95$, Wilks' $\Lambda = .808$, $p < .001$, respectively, although they were qualified by a significant condition x interpretation interaction, $F(4,1480) = 4.51$, Wilks' $\Lambda = .976$, p. $= .001$. Table 15 shows the mean ratings for the interaction.

This reveals that the ratings of message interpretations were higher for injunctive norm than for conditional cooperation in all conditions, although this difference was smallest in the conditional cooperation condition. This is verified by a significant interaction due to interpretation type (injunctive norm vs. conditional cooperation) and condition, $F(2,744) = 6.96$, Wilks' $\Lambda = .982$, $p < .01$. It is interesting to note that even an offer to conditionally cooperate (I'll cooperate if you cooperate) is seen to connote the injunctive norm of cooperation in public goods games (albeit less than in the injunction and gossip conditions, $F(2,744) = 12.36$, $p < .001$; injunction vs. conditional cooperation, $t(486) = 4.31$, $p < .001$, gossip vs. conditional cooperation, $t(498) = 4.10$, $p < .001$). This analysis also revealed that the conditional cooperation condition evoked less reputational concern than the other two

conditions, $F(2,744) = 6.47$, $p < .01$ (injunction vs. conditional cooperation, $t(487) = 2.70$, $p < .01$, gossip vs. conditional cooperation, $t(498) = 3.49$, $p < .001$; unexpectedly, ratings in the injunction and gossip conditions were comparable, $t(503) = .67$, $p = .50$).

Table 15. Mean ratings of the message interpretations in each condition

	Injunction	Gossip	Conditional Cooperation
Message Interpretation			
Injunctive Norm	1.13	1.07	.67
Conditional Cooperation	.84	.64	.52
Reputational Concern	.54	.60	.30
t-test comparing injunctive norm and conditional cooperation	5.38**	7.60**	3.25**

Note: ** $p < .01$; * $p < .05$; † $p < .1$; all means are significantly greater than 0 (neutral point) at $p < .001$ by independent t-tests.

Norm verification precipitates cooperation. We investigated which interpretation of the participants' message-reply exchanges – norm verification, conditional cooperation, and reputational concern – predicted their cooperation with a General Linear Mixed Model Logistic Regression analysis while controlling for the same predictors as in Experiments 1 and 2. Table 16 reports the coefficient for each interpretation in the relevant conditions for senders and receivers separately. The results of the full model are reported in Supplementary Information (Table S17).

Table 16. Coefficient for the interpretations of message-reply exchange by senders and receivers in predicting their cooperation in the subsequent round

	Injunction		Gossip		Conditional Cooperation	
	Sender	Receiver	Sender	Receiver	Sender	Receiver
Norm Verification	.28†	.47**	.63**	.46**	.22	.15
Conditional Cooperation	.12	.74**	.40*	.60**	.53**	.19
Reputational Concern	.01	.15	.33*	.21	-.08	.01

Note: ** $p < .01$; * $p < .05$; † $p < .1$

Consistent with our expectation, norm verification was an important driver of cooperation for both senders and receivers in both the injunction and gossip conditions. In contrast, norm verification did not predict cooperation in the conditional cooperation condition. Again, consistent with our expectation, reputational concern predicted senders' cooperation in the gossip condition, but not in the injunction condition; reputational concerns did not predict cooperation in the conditional cooperation condition. As expected, conditional cooperation interpretation – a message-reply exchange interpreted as a promise of conditional cooperation – predicted the message senders' cooperation in the conditional cooperation.

However, somewhat unexpectedly, conditional cooperation did not predict the message receivers' cooperation. This finding is particularly intriguing because conditional cooperation interpretation predicted receivers' cooperation in both the injunction and gossip conditions. To put it differently, the receivers' interpretation that their communication implied a promise to cooperate had a binding effect on the message receivers to cooperate only in the norm talk conditions. This may be because they saw verification of the norm of cooperation as implying a verification of the norm of promise keeping as well. As we discussed in Experiment 3, a promise of conditional cooperation does not, in and of itself, ensure people would cooperate. A promise to cooperate can ensure cooperation only when it is backed up by the norm of promise keeping. In the conditional cooperation condition, the norm of promise keeping may not have been as salient. In fact, there is even an incentive for people to "cheap talk" their communication partners into cooperation while choosing to defect themselves. This is because a defector's payoff increases in the public goods game if there is one additional cooperator. It would be important to investigate the role potentially played by the norm of promise keeping in the sustenance of conditional cooperation.

General Discussion

The present research has shown that language-based interpersonal norm talk within a social network can sustain cooperation in social dilemma situations as modelled by repeated public goods games. The ability to use a complex language is a uniquely human capacity. Whether human language has evolved to solve the collective action problem in human society (Dunbar, 1996, 2004), it appears to be an effective tool by which to do so. Even with the very limited capacities to exchange norm talk and replies that we used in the current experiments, it can act to activate the injunctive norm of cooperation and sustain cooperation with or without the power of reputational information. Both explicit and implicit norm talk – injunction and gossip – can make salient the injunctive norm of cooperation as well as a more specific norm of promise keeping. Consistent with focus theory (Cialdinini et al., 1990, 1991), norm talk can arrest declining cooperation. Such communicative regulation of collective action is not only possible but may also be highly beneficial in sustaining human cooperation. This is because it does not require any additional tangible resources other than the cost of maintaining the communication channels. Although this cost is not negligible, norm talk seems to represent a relatively low cost alternative to punishment and other forms of social regulation of collective action.

Nevertheless, there are subtly different aspects to the mechanisms by which norm talk sustain cooperation depending on their types. On the one hand, injunction maintains cooperation by the interactants' mutual verification of the injunctive norm. When both the sender and receiver of an injunction say that they should cooperate, this injunction-acceptance exchange verifies the injunctive norm of cooperation. There appears to be an element of a *de facto* "contract" that they both contribute to the public goods. As shown in Experiment 3, however, the *de facto* contract element is secondary to the mutual verification of the injunctive norm of cooperation, and perhaps backed by the norm of promise keeping as a more specific norm implied by the norm of cooperation. Consistent with this interpretation,

Experiment 4 showed that both norm verification and conditional cooperation underpinned cooperation. By cumulating these dyadic verifications of the norm and tacit agreements to cooperate within a group, people can arrest the declining trend towards non-cooperation. On the other hand, gossip, when communicated, acts as a reminder of the norm and reputational implications to its receivers. Both Experiments 3 and 4 showed the importance of reputational concern in addition to norm verification as a significant driver of cooperation in the gossip condition. Presumably if it is positive gossip, it reminds its receiver of the benefits of social status and approval, whereas if it is negative it illustrates to the receiver the disapproval heaped on those who deviate from the norm. In everyday life, this approval is linked to tangible rewards and this disapproval to social sanctions such as punishment and ostracism.

Different types of norm talk appear to send subtly different signals. An injunction message with gossip appears to signal the sender's intention to cooperate, or it can be taken as a pledge to cooperate (i.e., in Experiment 1 and the combined condition in Experiment 2); gossip sent alone says little about the sender's intention; and an injunction sent alone seems to mean that the sender awaits the receiver's consent, rather than announcing a unilateral decision to contribute to the public goods – presumably for fear of being a sucker who ends up benefiting a free rider. Injunctive norm talk may be taken as a unilateral pledge to cooperate only when it accompanies gossip, or reputational information about a concrete individual.

One interpretation is that gossip may have a different meaning when it accompanies injunctions. Gossip sent alone is a reminder of reputational implications of a particular action as well as of the injunctive norm. The results are mixed as to whether gossip's dual "cloak and dagger" will consistently produce higher cooperation (i.e., Experiment 2 vs Experiment 4). However, gossip sent with an injunction acts as an exemplification of the injunctive norm: that is, the message says, "Here is someone who has violated the norm, it is to be

disapproved, and we should all adhere to the norm." Our Experiment 3 corroborates this interpretation. Combined messages were interpreted more similarly to injunctions than to gossip. Thus fortified, the sender of the injunctive statement may feel certain about its moral righteousness and simply carries out its implication in action, and thus the receiver's explicit acceptance is unnecessary. It may be this feeling of sharedness, which Echterhoff et al. (2009) called shared reality, that makes the difference, in this instance, the anticipation of shared morality.

In contrast, an offer to conditionally cooperate was not more effective than norm talk in sustaining cooperation in Experiment 4. We found that conditional cooperation can in and of itself maintain cooperation to some extent; people tend to cooperate after making a promise to cooperate. Interestingly, norm talk communications are also interpreted to entail a promise of conditional cooperation as shown in Experiments 3 and 4; however, norm talk appears to entail additional implications – a reminder of the general norm of cooperation plus the specific norm of promise keeping. It was in the norm talk conditions where these norms are salient that conditional cooperation was an especially powerful driver of cooperation in message receivers' decisions in the social dilemma situations. In the conditional cooperation condition, where communications had relatively weak normative connotations, promise of cooperation did not predict cooperation. Although this last finding was unexpected, it nonetheless highlights the importance of injunctive norms.

Finally, we observed in Experiments 1, 2, and 4, the consistent pattern of spike and decline in cooperation following interpersonal communication. Whereas the spike may be explained by the norm salience in line with Cialdini et al.'s focus theory (1990, 1991), the decline may be somewhat more difficult to explain purely by norm salience. In fact, the decline mimics the typical pattern of cooperation in repeated public goods games (e.g., Fischbacher, Gächter, & Fehr, 2001; Isaac & Walker, 1988), and Fischbacher et al. explained

this as follows. At the start of a game, people's expectations that others would cooperate are often somewhat higher than the actual level of cooperation; as they continue to play the game repeatedly, their expectation that others would cooperate is often violated and their lowered expectation drives down the cooperation level. Likewise, immediately following norm talk, the norm salience may increase participants' levels of expectations that others would cooperate. However, their expectations may be unrealistically high – their observation that the actual cooperation does not match their expectations may drive down their cooperation subsequently as the effect of the norm of cooperation dissipates until the next round of communication. The cyclical pattern of spike and decline may be explained in terms of the cycle of norm salience-raised expectation-expectation violation. Future research should investigate the possibility that norm talk may affect cooperation through expectations of others' cooperation.

How is the Current Experimental Paradigm Different from Previous Paradigms?

Although the present research has some similarities with the previous work, it differs from it in several respects. First, in most research on communication and social dilemmas, participants were usually allowed to engage in unhindered group discussions where everyone could communicate with everyone else at the same time (for an exception, see Kinukawa et al., 2000); in contrast, in the current studies, participants could only concurrently send a message to one other participant, and reply to another participant's message. As well, the current communication structure significantly diverges from that examined by Feinberg et al. (2014). In their gossip condition, each participant had the information about all other members of the group, selected one member about whom to write a gossip note, and the note was given to all players of the game in which this target player was to play in the next round. In essence, Feinberg et al.'s gossip was broadcast to all members of the gossip target's group.

In contrast, gossip in our studies was interpersonally communicated to only one person in the group at a time.

In this regard, our studies are closest to Sommerfeld et al.'s (2007) and Ahn, Esarey, and Scholz's (2009) experimental paradigms. However, in Sommerfeld et al.'s experiments, participants did not play public goods game, but played a dyadic indirect reciprocity game in which a participant was paired with two other players, and decided whether or not to give some money. If one is said to have given more money to others, one is likely to receive more money from others. This game does not present the stark social dilemma of contributing to the public goods at one's own cost. As well, one's reputation is likely to be directly linked to the amount of money one receives because the giving is directed towards a particular individual in this experiment.

Ahn et al.'s local information condition is potentially closer to our experimental condition; their participants were allowed to seek reputational information about another player from other players when they chose their interaction partners in two-person prisoners' dilemma games. They found an elevated level of cooperation in this condition compared to the control. Even this condition, however, diverges significantly from the norm talk conditions in the current paradigm in that participants played a dyadic game rather than a public goods game within a group, and more importantly, they could select their partners based on the reputational information. In this sense, Ahn et al.'s experimental condition resembled Feinberg et al.'s (2014) gossip with ostracism condition, where participants who received gossip had a chance to ostracize a player with whom they did not wish to play the game. By choosing an interaction partner as in Ahn et al. or excluding a player from the game as in Feinberg et al.'s gossip with ostracism condition, these experiments allowed participants to make an individually targeted partner choice decision; the possibility for partner choice is a significant factor that is known to increase cooperation in and of itself (e.g., Rand et al.,

2009). In contrast, interpersonal norm talk in our experiments was the only factor that differed from the control conditions.

Conclusion

When prevailing behaviors (i.e., descriptively normative behaviors) are undesirable from a moral or practical viewpoint, injunctive norms can play a significant role in socially down regulating those behaviors (e.g., Cialdini et al., 2006; Schultz et al., 2007). Whereas descriptive norms may be learnable by observing others' behaviors in the immediate setting (e.g., Cialdini et al., 1990) or in social networks (e.g., Kashima, Wilson, Lusher, Pearson, & Pearson, 2013), injunctive norms often have to be learned and made salient by language-based communications (e.g., Lapinski & Rimal, 2005). The present research has shown that such language-based interpersonal communications – norm talk – can help resolve social dilemmas by sustaining cooperation in the provision of public goods. Although it is often assumed that injunctive norms may regulate social behaviors only when accompanied by the threat of sanctions (e.g., Bendor & Swistak, 2001), norm talk can arrest the declining level of cooperation at least in the short-term through the verification of the norms of cooperation and promise keeping, but also through the reminder of potential reputational damage or status conferral in the case of gossip.

The present results suggest that communicative regulation of collective action in social dilemmas is possible, but also that normative communication may need to be *ritualized*. That is to say, the injunctive norm of cooperation may need to be communicated and made salient periodically with relatively short intervals. As Rossano (2012) noted, rituals may be a critical mechanism by which social norms are communicated and maintained. In many large-scale societies, low intensity rituals (i.e., rituals that do not involve severe physical pain and high levels of physiological arousal) are performed frequently (Atkinson &

Whitehouse, 2011; Collins, 2004). Frequent ritualized norm talk (e.g., Sunday sermons, conversations around the water cooler, Facebook comments) may be a significant cultural artifact that helps sustain cooperation. Whether this mechanism works under all circumstances is an open question, however. In the present experiments, the level of benefit that individuals receive from the provision of public goods was set at a moderate level (i.e., MPCR = .3 or .4), but not extremely low. It is quite possible that norm talk cannot sustain cooperation at much lower levels of MPCR (e.g., .1 or .2). Likewise, if MPCR is much higher, norm talk may not need to be repeated as often or may not be even necessary. Future research may further investigate the interaction of incentive structure and cultural processes such as norm talk.

References

Ahn, T., Esarey, J., & Scholz, J. T. (2009). Reputation and cooperation in voluntary exchanges: Comparing local and central institutions. *Journal of Politics, 71*(2), 398-413.

Anderson, J. E., & Dunning, D. (2014). Behavioral norms: Variants and their identification. *Social and Personality Psychology Compass, 8*(12), 721–738.

Atkinson, Q. D., & Whitehouse, H. (2011). The cultural morphospace of ritual form. *Evolution and Human Behavior, 32*(1), 50-62. doi:10.1016/j.evolhumbehav.2010.09.002

Balliet, D. (2009). Communication and Cooperation in Social Dilemmas: A Meta-Analytic Review. *Journal of Conflict Resolution, 54*(1), 39-57. doi:10.1177/0022002709352443

Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the "tragedy of the commons". *Evolution and Human Behavior, 25*, 209-220.

Baum, W. M., Paciotti, B., Richerson, P., Lubell, M., & McElreath, R. (2012). coperation due to cultural norms, not individual reputation. *Behav Processes, 91*, 90-93.

Baumeister, R. F., Zhang, L., & Vohs, K. D. (2004). Gossip as cultural learning. *Review of General Psychology, 8*(2), 111-121. doi:10.1037/1089-2680.8.2.111

Beersma, B., & Van Kleef, G. A. (2011). How the grapevine keeps you in line: Gossip increases contributions to the group. *Social Psychological and Personality Science*, 1948550611405073.

Bendor, J., & Swistak, P. (2001). The evolution of norms. *American Journal of Sociology, 106*(6), 1493-1545.

Bicchieri, C. (2002). Covenants without Swords: Group Identity, Norms, and Communication in Social Dilemmas. *Rationality and Society, 14*(2), 192-228. doi:10.1177/1043463102014002003

Bicchieri, C., & Lev-On, A. (2007). Computer-mediated communicaiton and cooperation in social dilemmas: An experimental analysis. *Politics, Philosophy, Economics, 6*(2), 139-168.

Biel, A., & Thøgersen, J. (2007). Activation of social norms in social dilemmas: A review of the evidence and reflections on the implications for environmental behaviour. *Journal of Economic Psychology, 28*(1), 93-112. doi:10.1016/j.joep.2006.03.003

Bonacich, P. (1972). Norms and Cohesion as Adaptive Responses to Potential Conflict: An Experimental Study. *Sociometry, 35*(3), 357-375. doi:10.2307/2786500

Bonacich, P. (1976). Secrecy and Solidarity. *Sociometry, 39*(3), 200-208. doi:10.2307/2786513

Bouas, K. S., & Komorita, S. S. (1996). Group discussion and cooperation in social dilemmas. *Personality and Social Psychology Bulletin, 22*(11), 1144-1150.

Brauer, M., & Chaurand, N. (2010). Descriptive norms, prescriptive norms, and social control: An intercultural comparison of people's reactions to uncivil behaviors. *European Journal of Social Psychology, 40*(3), 490-499. doi:10.1002/ejsp.640

Cason, T. N., & Khan, F. U. (1999). A laboratory study of voluntary public goods provision with imperfect monitoring and communication. *Journal of Development Economics, 58*(2), 533-552. doi:http://dx.doi.org/10.1016/S0304-3878(98)00124-2

Chaudhuri, A. (2010). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics, 14*(1), 47-83. doi:10.1007/s10683-010-9257-1

Chen, X.-P. (1996). The group-based binding pledge as a solution to public goods problems. *Organizational Behavior and Human Decision Processes, 66*(2), 192-202.

Chen, X.-P., & Komorita, S. S. (1994). The effects of communication and commitment in a public goods social dilemma. *Organizational Behavior and Human Decision Processes, 60*, 367-386.

Chiu, C. Y., Gelfand, M. J., Yamagishi, T., Shteynberg, G., & Wan, C. (2010). Intersubjective Culture: The Role of Intersubjective Perceptions in Cross-Cultural Research. *Perspectives on Psychological Science, 5*(4), 482-493. doi:10.1177/1745691610375562

Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences, 15*(5), 218-226. doi:10.1016/j.tics.2011.03.003

Cialdini, R. B., Demaine, L. J., Sagarin, B. J., Barrett, D. W., Rhoads, K., & Winter, P. L. (2006). Managing social norms for persuasive impact. *Social Influence, 1*(1), 3-15.

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology, 24*, 201-234.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Rrecycling the concept of norms to reduce littering in public places. . *Journal of Personality and Social Psychology, 58*(6), 1015-1026.

Clark, A. E., & Kashima, Y. (2007). Stereotypes help people connect with others in the community: A situated functional analysis of the stereotype consistency bias in communication. *Journal of Personality and Social Psychology, 93*(6), 1028-1039. doi:10.1037/0022-3514.93.6.1028

Collins, R. (2004). *Interaction Ritual Chains*. Princeton: Princeton University.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology, 31*(1), 169-193. doi:10.1146/annurev.ps.31.020180.001125

Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology, 35*(1), 1-11.

DiFonzo, N., Bourgeois, M. J., Suls, J. M., Homan, C., Stupak, N., Brooks, B., . . . Bordia, P. (2013). Rumor clustering, consensus, and polarization: Dynamic social impact and self-organization of hearsay. *Journal of Experimental Social Psychology, 49*(3), 378-399. doi:10.1016/j.jesp.2012.12.010

DiFonzo, N., Suls, J., Beckstead, J. W., Bourgeois, M. J., Homan, C. M., Brougher, S., . . . Terpstra-Schwab, N. (2014). Network structure moderates intergroup differentiation of stereotyped rumors. *Social Cognition, 32*(5), 409-448. doi:10.1521/soco.2014.32.5.409

Dunbar, R. I. M. (1996). *Grooming, gossip, and the evolution of language*. Boston, MA: Harvard University Press.

Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology, 8*(2), 100-110. doi:10.1037/1089-2680.8.2.100

Echterhoff, G., Higgins, E. T., & Levine, J. M. (2009). Shared reality experiencing commonality with others' inner states about the world. *Perspectives on Psychological Science, 4*(5), 496-521. doi:10.1111/j.1745-6924.2009.01161.x

Fehr, E., & Fischbacher, U. (2004a). Social norms and human cooperation. *Trends in Cognitive Sciences, 8*(4), 185-190. doi:10.1016/j.tics.2004.02.007

Fehr, E., & Fischbacher, U. (2004b). Third-party punishment and social norms. *Evolution and Human Behavior, 25*(2), 63-87. doi:10.1016/s1090-5138(04)00005-4

Fehr, E., & Gachter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137-140. doi:10.1038/415137a

Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science, 25*(3), 656-664.

Feinberg, M., Willer, R., Stellar, J., & Keltner, D. (2012). The virtues of gossip: reputational information sharing as prosocial behavior. *Journal of Personality and Social Psychology, 102*(5), 1015.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters, 71*(3), 397-404.

Foster, E. K. (2004). Research on gossip: Taxonomy, methods, and future directions. *Review of General Psychology, 8*(2), 78-99. doi:10.1037/1089-2680.8.2.78

French, J. R. P., Jr., & Raven, B. H. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power.* (pp. 150-167). Ann Arbor, MI: Institute for Social Research.

Gelfand, M., & Jackson, J. C. (2016). From one mind to many: The emerging science of cultural norms. *Current Opinion in Psychology*.

Gibbs, J. (1965). The problem of definition and classification. *American Journal of Sociology, 70*(5), 586-594.

Gibbs, J. (1968). The study of norms. In D. Sills (Ed.), *International Encyclopedia of the Social Sciences* (Vol. 11). New York: Macmillan.

Gluckman, M. (1963). Gossip and scandal. *Current Anthropology, 4*(3), 307-316.

Hardin, C. D., & Higgins, E. T. (1996). Shared reality: How social verification makes the subjective objective. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: The interpersonal context* (Vol. 3, pp. 28-84). New York, NY: Guilford Press.

Higgins, E. T. (1992). Achieving "shread reality" in the communication game: A social action that creates meaning. . *Journal of Language and Social Psychology, 11*, 107-131.

Hoerger, M. (2013). ZH: An updated version of Steiger's Z and web-based calculator for testing the statistical significance of the difference between dependent correlations.

Holtgraves, T. M. (2002). *Language as social action: Social psychology and language use*. Mahwah, NJ: Lawrence Erlbaum.

Holtgraves, T. M., & Kashima, Y. (2008). Language, meaning, and social cognition. *Pers Soc Psychol Rev, 12*(1), 73-94. doi:10.1177/1088868307309605

Homans, G. C. ([1961] 1974). *Social Behavior: Its Elementary Forms*. New York: Harcourt Brace Jovanovich.

Horne, C. (2001). The enforcement of norms: Group cohesion and meta-norms. *Social Psychology Quarterly*, 253-266.

Hunzaker, M. B. F. (2014). Making Sense of Misfortune: Cultural Schemas, Victim Redefinition, and the Perpetuation of Stereotypes. *Social Psychology Quarterly, 77*(2), 166-184. doi:10.1177/0190272514521219

Isaac, R. M., & Walker, J. M. (1988). Communication and free-riding behavior: The voluntary contribution mechanism. *Economic Inquiry, 26*(4), 585-608. doi:10.1111/j.1465-7295.1988.tb01519.x

Jensen, C., Farnham, S. D., Drucker, S. M., & Kollock, P. (1999). *The effect of communication modality on cooperation in online environments*. Technical Report MSR-TR-99-75. Retrieved from https://www.researchgate.net/profile/Shelly_Farnham/publication/2404191_The_Effect_of_Communication_Modality_on_Cooperation_in_Online_Environments/links/00b4952853f904a1fb000000.pdf

Kashima, Y. (2000). Maintaining cultural stereotypes in the serial reproduction of narratives. *Personality and Social Psychology Bulletin, 26*(5), 594-604.

Kashima, Y. (2013). Cultural dynamics of intergroup relations: How communications can shape intergroup reality. In C. Stangor & C. S. Crandall (Eds.), *Stereotyping and prejudice* (pp. 119-150). New York: Psychology Press.

Kashima, Y., Wilson, S., Lusher, D., Pearson, L. J., & Pearson, C. (2013). The acquisition of perceived descriptive norms as social category learning in social networks. *Social Networks, 35*, 711-719. doi:10.1016/j.socnet.2013.06.002

Kerr, N. L. (1995). Norms in social dilemmas. In D. A. Schroeder (Ed.), *Social dilemmas: Perspectives on individuals and groups* (pp. 31-). Westport, CT: Praeger.

Kerr, N. L., Garst, J., Lewandowski, D. A., & Harris, S. E. (1997). That still, small voice: Commitment to cooperate as an internalized versus a social norm. *Personality and Social Psychology Bulletin, 23*(12), 1300-1311.

Kerr, N. L., & Kaufmann-Gilliarnd. (1994). Communication, commitment, and cooperation in social dilemmas. *Journal of Personality and Social Psychology, 66*(3), 513-529.

Kinukawa, S., Saijo, T., & Une, M. (2000). Partial Communication in a Voluntary-Contribution-Mechanism Experiment. *Pacific Economic Review, 5*(3), 411-428. doi:10.1111/1468-0106.00113

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology, 24*(1), 183-214. doi:10.1146/annurev.soc.24.1.183

Komorita, S. S., & Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology, 46*(183-207).

Kusumi, T., Hirayama, R., & Kashima, Y. (2017). Risk perception and risk talk: The case of the Fukushima Daiichi nuclear radiation risk. *Risk Analysis, 37*(12), 2305-2320.

Lapinski, M. K., & Rimal, R. N. (2005). An explication of social norms. *Communication Theory, 15*(2), 127-147.

Ledyard, J. (1995). Public goods: a survey of experimental research. In J. Kagel & A. Roth (Eds.), *Handbook of experimental economics* (Vol. 111-194). Princeton: Princeton University Press.

Levinson, S. C. (1983). *Pragmatics*. Cambridge, United Kingdom: Cambridge University Press.

Lyons, A., & Kashima, Y. (2003). How are stereotypes maintained through communication? The influence of stereotype sharedness. *Journal of Personality and Social Psychology, 85*(6), 989-1005. doi:10.1037/0022-3514.85.6.989

Manning, M. (2009). The effects of subjective norms on behaviour in the theory of planned behaviour - A meta-analysis. *British Journal of Social Psychology, 48*, 649-705.

Morris, M. W., Hong, Y.-y., Chiu, C.-y., & Liu, Z. (2015). Normology: Integrating insights about social norms to understand cultural dynamics. *Organizational Behavior and Human Decision Processes, 129*, 1-13. doi:10.1016/j.obhdp.2015.03.001

Murphy, K. R., & Myors, B. (2004). Statistical power analysis: A simple and general model for traditional and modern hypothesis tests. Mahwah, NJ: Lawrence Erlbaum.

Orbell, J. M., van de Kragt, A. J. C., & Dawes, R. M. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology, 54*(5), 811-819.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives, 14*(3), 137-158.

Parks, C. D., Joireman, J., & van Lange, P. A. M. (2013). Cooperation, trust, and antagonism: How public goods are promoted. *Psychological Science in the Public Interest, 14*(3), 119-165. doi:10.1177/1529100612474436

Peters, K., & Kashima, Y. (2007). From social talk to social action: Shaping the social triad with emotion sharing. *Journal of Personality and Social Psychology, 93*(5), 780-797. doi:10.1037/0022-3514.93.5.780

Peters, K., & Kashima, Y. (2015). Bad habit or social good? How perceptions of gossiper morality are related to gossip content. *European Journal of Social Psychology, 45*, 784-798.

Piazza, J., & Bering, J. M. (2008). Concerns about reputation via gossip promote generous allocationsin an economic game. *Evolution and Human Behavior, 29*, 172-178.

Rand, D. G., Pfeiffer, T., Dreber, A., Sheketoff, R. W., Wernerfelt, N. C., & Benkler, Y. (2009). Dynamic remodeling of in-group bias during the 2008 presidential election. *Proc Natl Acad Sci U S A, 106*(15), 6187-6191. doi:10.1073/pnas.0811552106

Reid, A. E., & Aiken, L. S. (2013). Correcting injunctive norm misperceptions motivates behavior change: A randomized controlled sun protection intervention. *Health Psychology, 32*(5), 551-560.

Rossano, M. J. (2012). The essential role of ritual in the transmission and reinforcement of social norms. *Psychological Bulletin, 138*(3), 529-549. doi:10.1037/a0027038

Sally, D. (1995). Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992. *Rationality and Society, 7*(1), 58-92. doi:10.1177/1043463195007001004

Schaller, M., Conway, L. G., III, & Tanchuk, T. L. (2002). Selective pressures on the once and future contents of ethnic stereotypes: Effects of the communicability of traits. *Journal of Personality and Social Psychology, 82*(6), 861-877. doi:10.1037//0022-3514.82.6.861

Schaller, M., & Conway, L. G. I. (1999). Influence of impression-management goals on the emerging contents of group stereotypes: Support for a social-evolutionary process. *Personality and Social Psychology Bulletin, 25*(7), 819-833. doi:10.1177/0146167299025007005

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. . *Psychological Science, 18*(5), 429-434.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge, United Kingdom: Cambridge University Press.

Shank, D. B., Kashima, Y., Saber, S., Gale, T., & Kirley, M. (2015). Dilemma of dilemmas: how collective and individual perspectives can clarify the size dilemma in voluntary linear public goods dilemmas. *PLoS ONE, 10*(3), e0120379. doi:10.1371/journal.pone.0120379

Smith, E. A. (2010). Communication and collective action: language and the evolution of human cooperation. *Evolution and Human Behavior, 31*(4), 231-245. doi:10.1016/j.evolhumbehav.2010.03.001

Smith, E. R. (2014). Evil acts and malicious gossip: a multiagent model of the effects of gossip in socially distributed person perception. *Pers Soc Psychol Rev, 18*(4), 311-325. doi:10.1177/1088868314530515

Snijders, T. A. B. (2005). Power and Sample Size in Multilevel Linear Models. In B. S. Everitt & D. C. Howell (Eds.), Encyclopedia of Statistics in Behavioral Science (Vol. 3, pp. 1570-1573). Chichester: Wiley.

Sommerfeld, R. D., Krambeck, H. J., Semmann, D., & Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proc Natl Acad Sci U S A, 104*(44), 17435-17440. doi:10.1073/pnas.0704598104

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245-251.

Torsvik, G., Molander, A., Tjøtta, S., & Kobbeltvedt, T. (2011). Anticipated discussion and cooperation in a social dilemma. *Rationality and Society, 23*(2), 199-216.

van de Bongardt, D., Reitz, E., Sandfort, T., & Deković, M. (2015). A meta-analysis of the relations between three types of peer norms and adolescent sexual behavior. *Personality and Social Psychology Review, 19*(3), 203-234.

Van Lange, P. A. M., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes, 120*(2), 125-141. doi:10.1016/j.obhdp.2012.11.003

Wedekind, C., & Milinski, M. (2000). Cooperation through Image Scoring in Humans. *Science, 288*(5467), 850-852.

Willer, R. (2009). Groups Reward Individual Sacrifice: The Status Solution to the Collective Action Problem. *American Sociological Review, 74*(1), 23-43.

Zou, X., & Leung, A. K. y. (2015). Special Issue: Intersubjective Norms. *Journal of Cross-Cultural Psychology, 46*(10), 1238-1371.