

Psychological Adaptations for Assessing Gossip Veracity

Nicole H. Hess

Institute for Theoretical Biology, Humboldt University, Berlin
Max Planck Institute for Human Development, Berlin

Edward H. Hagen

Institute for Theoretical Biology, Humboldt University, Berlin

Evolutionary models of human cooperation are increasingly emphasizing the role of reputation and the requisite truthful “gossiping” about reputation-relevant behavior. If resources were allocated among individuals according to their reputations, competition for resources via competition for “good” reputations would have created incentives for exaggerated or deceptive gossip about oneself and one’s competitors in ancestral societies. Correspondingly, humans should have psychological adaptations to assess gossip veracity. Using social psychological methods, we explored cues of gossip veracity in four experiments. We found that simple reiteration increased gossip veracity, but only for those who found the gossip relatively uninteresting. Multiple sources of gossip increased its veracity, as did the independence of those sources. Information that suggested alternative, benign interpretations of gossip decreased its veracity. Competition between a gossipier and her target decreased gossip veracity. These results provide preliminary evidence for psychological adaptations for assessing gossip veracity, mechanisms that might be used to assess veracity in other domains involving social exchange of information.

KEY WORDS: Cooperation; Gossip; Language; Reputation; Signaling theory

REPUTATION, COOPERATION, AND GOSSIP

Reputation plays a central role in recent evolutionary models of human cooperation, deepening the theoretical connections between cooperation and language. In the indirect reciprocity theories (Alexander 1987; Leimar and Hammerstein 2001;

Received October 17, 2004; revisions requested November 30, 2004; revised version received July 1, 2005; accepted July 19, 2005.

Address all correspondence to Nicole Hess, Institute for Theoretical Biology, Humboldt University, Invalidenstr. 43, 10115 Berlin, Germany. Email: n.hess@biologie.hu-berlin.de

Human Nature, Fall 2006, Vol. 17, No. 3, pp. 337–354.

1045-6767/98/\$6.00 = .15

Mohtashemi and Mui 2003; Nowak and Sigmund 1998; Panchanathan and Boyd 2003), benefits are provided to an individual based on information about his or her past contributions to others in the group—generous individuals are rewarded by receiving benefits from group members. In the “health insurance” theories (Gurven et al. 2000; Sugiyama and Chacon 2000), individuals increase the likelihood that they will be taken care of when ill or injured by generously providing benefits to group members when they are well. In the “show-off” or “costly signaling” theories (Gintis et al. 2001; Hawkes 1991; Smith and Bliege Bird 2000), individuals engage in behavior, such as big-game hunting, that signals their quality as mates or social partners, and consequently reap valuable mating or social benefits. Reputation can also play an important role in reciprocal altruism, whereby individuals benefit from learning whether future social partners previously defected or cooperated with others (e.g., Cox, Sluckin, and Steele 1999; Enquist and Leimar 1993; Pollock and Dugatkin 1992). In most of these models, reputations are based on direct observations of behaviors, but the proponents of these models suggest that in the real world reputations are often based on the spread of information about these behaviors via personal conversations, that is, via gossip.

Consistent with reputation-based models of cooperation, in the small, kin-based societies that typified the ancestral environment, reputation frequently regulates access to scarce resources like food, mates, protection, and care (e.g., Chagnon 1988; Gurven et al. 2000; Hawkes 1991, 1993; Hawkes, O’Connell, and Blurton Jones 2001; Marlowe 1999; Patton 2000; Smith and Bliege Bird 2000; Sosis 2000; Sugiyama and Chacon 2000). Reputation has also been shown to predict resource contributions in experimental economics games (e.g., Milinski et al. 2001, 2002; Wedekind and Milinski 2000).

In light of this recent theoretical and empirical work, which strongly suggests that reputation impacts one’s access to social partners and the resources they provide, truthful “gossiping” about reputation-relevant topics, far from being a trivial pastime, is instead a foundation of human cooperation.

LIES, EXAGGERATIONS, AND MISREPRESENTATIONS

A signal is “an action or structure that increases the fitness of an individual by altering the behaviour of other organisms detecting it, and that has characteristics that have evolved because they have that effect” (Maynard-Smith and Harper 1995). Crucially, signals evolve because they provide fitness benefits to the signaler, not to the recipient (Dawkins and Krebs 1978). In the presence of conflict, signals can therefore contain a substantial dose of deception. Mimicry and crypsis, for example, are extremely common in vertebrates, arthropods and opisthobranch gastropods (Starrett 1993). Deceptive signals select for ever better signal discrimination on the part of signal receivers, which, in turn, selects for more effective deceptive signaling, and so on.

Language is a “cheap” signaling system that could only have evolved if commu-

nication were mostly honest in ancestral environments, implying that language evolved in a social context involving cooperation, shared interests, coordination, and/or repeated interactions with opportunities to punish deception (Lachmann, Szamado, and Bergstrom 2001; Silk, Kaldor, and Boyd 2000). Speech should be honest when “on average, the incentive to the signaler to misrepresent the state of the world [is] outweighed by the incentive not to do so” (Lachmann et al. 2001:13189). Lying can sometimes confer social benefits, however, such as misleading competitors, extracting additional resources from social partners, or avoiding punishment for proscribed behavior. In these circumstances, speech will often be dishonest.

Reputation is one domain in which there are particularly strong incentives “to misrepresent the state of the world.” The role of reputation in resource competition among humans creates incentives for signalers to deceive receivers about the social world in ways that benefit the signaler, such as promoting good reputations for kin and allies, and derogating reputations of competitors. Gossip—honest or deceptive—may be one strategy by which individuals compete for scarce resources by using information to damage their opponents’ reputations and improve their own (Barkow 1992; Buss and Dedden 1990; Emler 1990; Hess and Hagen 2006; McAndrew and Milenkovic 2002; Paine 1967; Radin 1927). In one of the few laboratory studies of gossip, McAndrew and Milenkovic (2002) found that student participants were more likely to pass on negative information about strangers and powerful others (professors), but were more likely to share positive information about friends and relatives.

OTHER FUNCTIONS OF GOSSIP

Like proponents of reputation-based theories of cooperation, we use the word “gossip” as a convenient short-hand for personal conversations about reputation-relevant behavior, a definition that is close to one influential view of gossip as strategic “information management” (Paine 1967). This is not, however, our view of gossip in general. We agree with other gossip researchers that gossip, broadly construed, probably has many other functions, including maintaining the unity, morals, and values of social groups (e.g., Gluckman 1963), cultural learning (e.g., Baumeister et al. 2004), social bonding (e.g., Dunbar 1996, 2004), social comparison (e.g., Wert and Salovey 2004), and norm learning and enforcement (e.g., Wilson et al. 2000). We would go further than most gossip theorists, in fact, and endorse Bloom’s (2004) view that “gossip” is probably not a natural category and has no single function. It is a culturally constructed category whose meaning has evolved significantly over the past ten centuries (see the Oxford English Dictionary, 1989 edition, OED Online). We believe the most useful scientific definition of gossip is *a personal conversation on social topics*. The information exchanged, truthful or deceptive, can benefit the sender, the receiver, or both. Our view is thus closest to Barkow’s (1992) view of gossip as information that had important implications for individuals’

fitness-relevant social strategies in the environment of evolutionary adaptedness (EEA).

But, to simplify discussion of the issues addressed in this paper, which were motivated primarily by the literature on reputation and cooperation, we narrowly operationalize gossip as information relevant to reputation. In particular, we are not proposing or testing any theory of the function of gossip. Our experimental results, described below, are relevant to any theory of gossip that is consistent with the solid empirical evidence that gossip is, in part, used aggressively and competitively to manipulate reputations.

GOSSIP AND AGGRESSION

The empirical evidence is overwhelming that gossip (in our narrow sense) is used aggressively and competitively. Researchers have recently devoted considerable attention to nonphysical aggression, variously termed relational, social, or indirect aggression. In each of these forms of aggression, individuals aggress against others using gossip and other nonphysical tactics. These forms of aggression have been documented in more than 60 studies on four continents (see Archer 2004 and Archer and Coyne 2005 for reviews; see Goodwin 1990a, 1990b for a very detailed ethnographic account of gossip in conflicts among children). In a study of college women, for example, Holland and Eisenhart (1990) present “Rosalind’s” account of what one woman did in order to attract Rosalind’s boyfriend; this account clearly illustrates the use of apparently false gossip to attack a competitor’s reputation in order to obtain a valued resource—a mate:

That girl would do anything in her power to spite me. . . . She’s always trying to get something against me. . . . [The authors ask what the woman would do.] Well, to start off she likes [my boyfriend]. And she’ll tell him things [lies] about me. . . . And she’ll come over to [my neighbor’s] room. You can hear right through the walls. She’ll even open the door . . . and she’ll strike up a conversation about me. She calls me every name in the book . . . trying to provoke me into fighting her . . . and trying to make [my boyfriend] think that I’m lying to him. . . . [She’ll be] telling him that some [other] man paged me . . . or came and picked me up . . . [when] no guy called me that morning . . . or picked me up (1990:114; brackets and omissions in the original).

NOISE

In addition to deception, unintentional errors (noise) can occur during communication. Signal errors can be costly to both senders and receivers. In reciprocal altruism, for instance, signaling cooperation is essential to receiving future benefits because individuals who defect lose out on future cooperation. If a social partner misread one’s cooperative intentions as defection, however, both the individual and his or her partner would suffer by failing to capitalize on the benefits of cooperation. Wu and Axelrod (1995) and Boerlijst et al. (1997) have explored strategies for

overcoming the costs of noisy signals. The contrite tit-for-tat strategy, for example, enables quick recovery from error by preventing player A from defecting when player B defects from A's initial unintended defection. Errors of perception and implementation are also important in models of indirect reciprocity (e.g., Leimar and Hammerstein 2001; Panchanathan and Boyd 2003). The costly consequences of misperceiving signals implies that adaptations for signaling should produce signals with a high signal-to-noise ratio, and adaptations for signal reception should effectively distinguish signals from noise.

PREVIOUS RESEARCH AND PREDICTIONS FOR THE CURRENT EXPERIMENTS

If gossip were occasionally deceptive or erroneous in the EEA, then there was an important selection pressure for the evolution of psychological mechanisms to discriminate truthful and accurate signals from deceptive and erroneous signals. These mechanisms, if they exist, should attend to cues that correlate with signal veracity. Although mechanisms to evaluate signal veracity could be used in many domains of information transfer, here we emphasize their use in evaluating social information relevant to reputation because, as we argued above, there are strong incentives to manipulate this kind of information. "Gossip" in our narrow sense thus provides a focused domain to explore cognitive mechanisms that evaluate signal veracity, mechanisms that might be used much more broadly to evaluate information received from others.

One important cue of veracity should be the number of times the same information is received from others. Random errors can be introduced into a signal during transmission and reception, but the probability that the *same* random error has been introduced into separate transmissions and receptions of the same signal decreases rapidly with the number of signal transmissions and receptions. Previous studies have shown that repeated statements are more believed than unrepeated statements, whether or not the statements are actually true (e.g., Bacon 1979; Hasher, Goldstein, and Toppino 1977; Hertwig, Gigerenzer, and Hoffrage 1997), a phenomenon referred to as the "truth" or "reiteration" effect. This effect has been demonstrated for "trivia"—possible factual statements about the non-social world, such as "the Eiffel Tower is 986 feet tall." We were skeptical that this effect would work for gossip statements, however, because unlike for trivia statements there are incentives to manipulate gossip statements. Humans should therefore have evolved to resist fairly simple manipulations of gossip. If so, the mere repetition of the same gossip from the same source or an unspecified number of sources (in the reiteration paradigm, the number of sources is not specified) should not have a strong, positive impact on believability. In experiment I, we test whether the reiteration effect works for gossip statements. We predicted the effect would be very small and perhaps nonexistent.

In addition to noise, information can be in error because an informant misinterpreted an event. If an informant conveys not the raw facts but his or her interpreta-

tion of the facts, this interpretation could be wrong. Additional facts that imply alternative interpretations of an event might lead a receiver to discount an interpretation of the signaler, and/or to make an alternative interpretation of the event. In experiment II, we test whether such additional facts undermine the believability of a piece of gossip.

Unlike mere repetition of a signal from the same, potentially deceptive source, receiving the same information from multiple, independent sources does decrease the probability of deception for several reasons. First, if an individual manipulates information to serve his or her own individual interest, then the probability that other signalers share exactly the same interest, and thus an incentive to manipulate their signals in exactly the same way, will decrease as the number of signalers increases. Second, deception imposes a severe coordination problem on multiple deceivers. Whereas there is only one truth, there are many different deceptions that can serve the same end. Multiple sources who share exactly the same interest in deception could each generate radically different stories serving this interest. Without sophisticated, costly coordination, it would be very difficult to keep their stories straight. Third, lying invites punishment (e.g., Silk, et al. 2000). Multiple signalers colluding to deceive would also all have to be willing to risk punishment for lying, a likelihood that declines as the number of signalers increases. Fourth, people are more likely to pass on a rumor the stronger their confidence is in its truth (Rosnow, Yost, and Esposito 1986). In contrast to simple repetition, receiving information from explicitly identified multiple sources should be a reliable cue of veracity for gossip statements. In experiment II, we test whether an increasing number of sources for a piece of gossip increases its believability.

Because deceivers can share interests and coordinate stories, receivers should attend to cues that multiple sources of gossip are independent—that is, one gossiper did not simply receive the information from another, and, importantly, one gossiper does not have an obvious incentive to collude with another gossiper. In experiment III, we test whether increasing independence among gossipers increases the believability of gossip. Finally, because competition between a gossiper and her target increases the incentive for the gossiper to manipulate her story, such competition should decrease the believability of gossip, a prediction we test in experiment IV.

Although there is a massive literature on the psychophysical cues that improve the detection of individual lies (see Vrij 2000 for a recent review), only a handful of experimental studies have examined the social and contextual cues hypothesized to affect the believability of gossip. Kuttler, Parker, and La Greca (2002) found that children rated gossip as less likely to be true when it was difficult to hear the source of the gossip (in this case, because the school bus was loud). They found that children rated firsthand information as more believable than secondhand information, a hypothesis we test with adults in experiment III. They also found that gossipers who presented conflicting stories were less believed than a single gossiper, and that children were less likely to believe gossipers who were unreliable due to, for example, ulterior motives.

Wilson and colleagues (2000) presented participants with four versions of a scenario in which a teaching assistant (TA) learns that student may have cheated on an exam. The TA either (1) sees him do it, (2) hears the information from one trusted friend who saw him do it, (3) hears the information from one unknown student who saw him do it, or (4) hears the information from two unknown students who saw him do it. Participants were then asked to rate the believability of the information. Study participants ranked the believability of accounts as follows: TA eyewitness > two unknown students > one trusted friend of the TA > one unknown student (only the first comparison was significant, perhaps owing to insufficient sample size). The stimuli in this experiment contained a mix of cues of multiple sources, credibility, authority, ulterior motives, and source independence. But, two unknown students were believed more than one unknown student, and the TA eyewitness, with presumably no ulterior motive, was believed the most, more than students and more than the hearsay of the TA's "friend." In experiments II, III, and IV we separately test the effects of multiple sources, source dependence, and ulterior motives.

Harkins and Petty (1987) varied attributes of the members of a committee that presented a non-gossip argument to participants. They found that participants' opinions were more persuaded when the committee members were independent rather than dependent sources. They also found that when committee members held dissimilar perspectives, they were more believed. We investigate how the relationship among multiple sources affects gossip believability in experiment III, wherein "friends" (people who are likely to share motives, information, and/or opinions) convey a piece of gossip.

EXPERIMENTS

In experiments I–III, UCSB undergraduates received packets containing one or more experiments. Participants were randomly assigned to experimental conditions, and counterbalancing was used to prevent the conditions of one experiment from varying systematically with those of another. Experiments II and III used similar stimuli, so these stimuli never appeared together in the same packet. Ages ranged from 18 to 55, with 96% \leq 25. In experiment IV, German students were randomly assigned to experimental conditions. Their ages ranged from 18 to 38, with 94% \leq 30. Parametric tests were used when variables met assumptions of these tests, for example, normality, homogeneity of variance, similar sample sizes, and independence; otherwise, nonparametric tests were used. Statistics were computed using SPSS 10 for Windows and R 1.91.

Experiment I: The Reiteration Effect and Gossip

Study Rationale and Predictions. Repetition is one of the most widely studied cues of believability, but it has not been tested with gossip statements. In contrast to previous studies, we predicted the reiteration effect for gossip would be small and

perhaps nonexistent because humans should have evolved to resist being manipulated by the simple repetition of information in the domain of potentially competitive gossip.

Method. Because we wished to show that any reiteration effect was, at best, small, we needed to ensure that we had adequate power to detect a small effect if one existed. Based on the standard deviations of previous studies, we estimated that in our study a small effect size (Cohen's $d = 0.2$) would be equivalent to a 1.4-point difference in the believability scores of repeated vs. unrepeated statements (see below). A power analysis indicated that a sample size of 165 would be able to detect this difference with a power of .9 at an alpha of .05. We therefore recruited more than 165 participants (52 male and 135 female, $N = 187$) for this two-condition, within-subjects experiment. Participants were asked to read and rate a set of gossip statements. Because it was impossible to acquire real gossip about the important members of each subject's actual social environment, we used fictional but plausible gossip statements about celebrities. Reasoning about celebrities could be a by-product of a psychology that evolved for reasoning about members of one's in-group (Barkow 1992). The content of the gossip statements dealt with the health, relationship status, mate value, career, and reproductive decisions of relatively young female celebrities who were popular when the studies were done. For example: "Jennifer Lopez said that she's sexier and more popular than Jewel" and "Heather Locklear refuses to allow her child to take music lessons, despite husband rocker Richie Sambora's wishes." Sixteen total gossip statements were used.

Participants were first asked to read eight of the gossip statements and rate each on a ten-point Likert scale ranging from "extremely uninteresting" to "extremely interesting." This rating task was originally intended to ensure that participants would actually read each gossip statement (though later these "interestingness" data were used in analyses). Participants then spent approximately ten minutes completing tasks related to other experiments. Finally, participants were asked to read sixteen more gossip statements and rate each for truthfulness on a ten-point Likert scale ranging from "definitely false" to "definitely true"; eight of the sixteen were novel statements (UNREPEATED) and eight were the same statements that they had previously ranked for interestingness (REPEATED). Some of the statements may have been perceived as truer because of the character of the celebrity in question, because of recent media exposure, or because of participants' biases about the celebrity. To control for these possible content effects, for each subject a unique, random selection of eight statements from the pool of sixteen statements was repeated; the remaining eight were unrepeated. Statements were also randomly ordered within each rating task for each subject.

Results. We calculated each subject's cumulative truth score for REPEATED and their cumulative truth score for UNREPEATED. TRUTHDIFF (REPEATED - UNREPEATED) was a measure of the reiteration effect for each participant. The possible range was between 8 and 80 for each condition. See Table 1 for descriptive statistics. Unless otherwise stated, all p -values are two-tailed.

Table 1. Descriptive Statistics for Variables in Experiment I

Variable	Range	Mean	s.d.
AGE	18–55	21.5	3.96
REPEATED	8–80	44.25	9.42
UNREPEATED	8–68	42.78	8.40
TRUTHDIFF	–26 to 51	1.47	8.30
INTEREST	8–73	27.28	13.87

Distributions of variables deviated from normal, so we used nonparametric tests. Similar to previous results with trivia statements, a Wilcoxon signed rank test showed that the difference in truth ratings between repeated and unrepeated statements was small (mean TRUTHDIFF = 1.47, s.d. = 8.30, $d = 0.18$) but significant, $Z = -1.98$, $p = .048$; repeated statements were more believed than unrepeated statements.

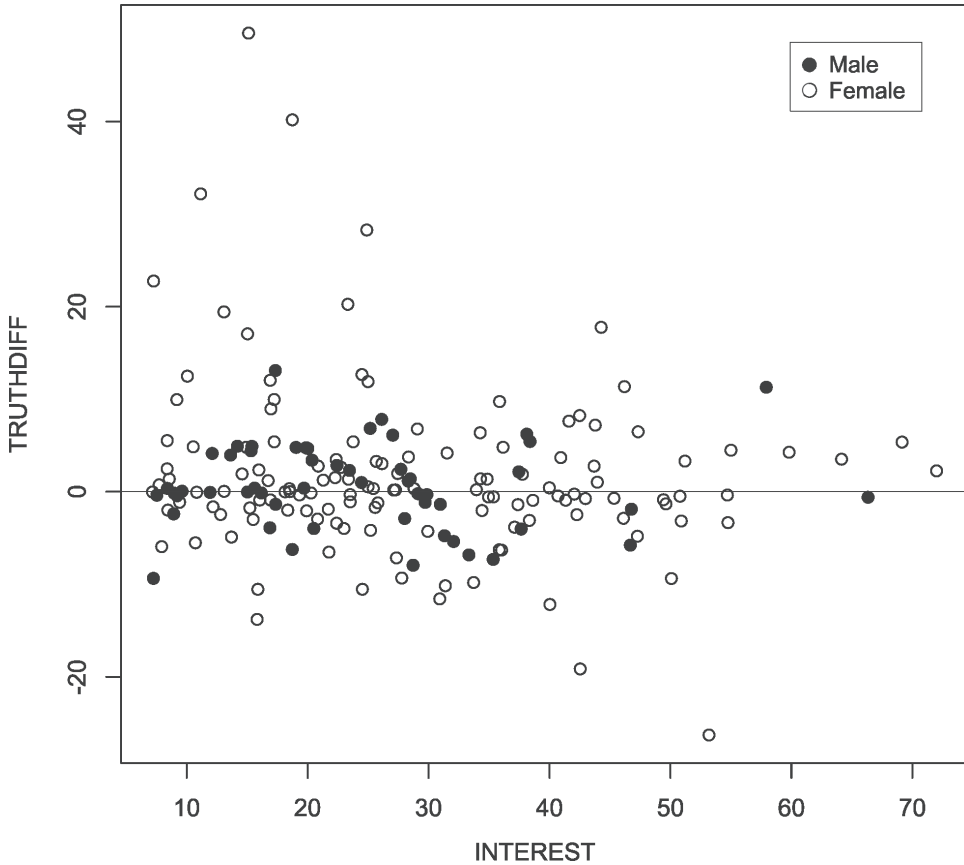
Because the gossip statements were mostly about female celebrities, we checked for sex differences in the results for this experiment. Females showed a slight positive bias in TRUTHDIFF (mean = 1.84) relative to males (mean = 0.52). This difference was not significant according to a Mann-Whitney test ($Z = -0.27$, $p = .79$). Female variance ($\sigma^2 = 87.6$) in TRUTHDIFF, however, was significantly greater than the male variance ($\sigma^2 = 21.6$), $F = 4.06$, $p = .007$.

To further explore the results, we examined a scatter plot of TRUTHDIFF against a measure of participants' INTEREST in these gossip statements (Figure 1). Though we had not originally intended to analyze the interest data, there was a noteworthy result. As can easily be seen in Figure 1, the entire reiteration effect was due to a small number of participants (all female) who had little interest in the gossip statements but whose truth ratings were strongly positively affected by repetition. We performed a median-split on the data with respect to interest in gossip statements (median = 25) and then tested for the reiteration effect in both subsamples.

For participants with low interest scores, the reiteration effect was more than double that of the entire sample, and it was significantly greater than zero (mean TRUTHDIFF_{low interest} = 3.21, s.d. = 9.53, $Z = -3.12$, $p = .002$). For participants whose interest scores were greater than the median, there was a *negative*, albeit not significant, effect (mean TRUTHDIFF_{high interest} = -0.45, s.d. = 6.27, $Z = -0.40$, $p = .69$). Eight participants (all female) were outliers, having mean TRUTHDIFF scores that were more than two standard deviations above or below the mean (i.e., scores greater than 18.13 or less than -15.17). When these eight participants were removed from the sample, mean TRUTHDIFF dropped by more than half, from 1.47 to 0.70, and was no longer significantly greater than zero, $Z = -1.55$, $p = .12$. The six outliers whose truth ratings were strongly *positively* affected by repetition all had interest scores that were less than the median of 25.

Discussion. Although we predicted no reiteration effect for gossip statements, there was a small reiteration effect, but only for participants who were relatively

Figure 1. Difference in truth ratings for repeated vs. non-repeated gossip statements as a function of expressed interest in those statements. Those individuals who expressed low interest in the statements were particularly vulnerable to the repetition effect. A small amount of jitter was added to reveal overlapping data points.



uninterested in the gossip. When these outlying participants were removed, no significant effect remained. It is possible that uninterested participants were reacting to the gossip statements as if they were trivia statements.

Experiment II: Alternatives, Multiple Sources, and Believability

Study Rationale and Predictions. In contrast to our prediction for simple reiteration, we predicted that receiving gossip from explicitly identified, distinct, multiple sources would have a strong, positive effect on believability because this cue of veracity is more difficult to manipulate. We also tested our prediction that additional facts can undermine a particular interpretation of gossip, decreasing its believability.

Method. In this between-subjects experiment, 346 participants (203 women and 143 men) read one of eleven scenarios that described one or more individuals gossiping that they had seen a female coworker making sexual advances toward (“hitting on”) her boss at a party. All participants read this basic scenario:

Pretend that you are an employee at a company that is having a large office holiday party. You are at the party, the company’s president is there, and so is Cathy, who is another employee of the company. [*] You decide to leave the party early. The next day at work, another employee who was at the party, Linda, tells you that she saw Cathy hitting on the company president at the party after you left. [†]

The dependent variable BELIEVABILITY was the response to the question, “How likely is it that Cathy was really hitting on the company president?” on a 1–9 Likert scale from “It definitely did NOT happen” to “It definitely DID happen.” Replacing the first brackets [*] was a combination of zero or more of the following pieces of information:

- A: Cathy is a very talkative and friendly person.
- B: You notice the president is talking to a lot of people that night.
- C: The party takes place one night after work in December.

Inclusion of statements A and B were predicted to decrease the believability of the gossip by suggesting more benign interpretations of any interaction between Cathy and the president (if Cathy and/or the president were generally talkative people, their talkativeness might have been misinterpreted as “hitting”). To test the alternative hypothesis that simply increasing the information content of the basic scenario would have an impact on believability, statement C was included.

Replacing the second brackets [†] were zero to three of the following sentences in which additional people gossiped that they saw the coworker hitting on her boss:

- 2: Then another employee who was at the party, Steve, says he saw Cathy hitting on the president.
- 3: Later, Rachel tells you she saw Cathy hitting on the president.
- 4: Then you hear from Rob that he saw Cathy hitting on the president at the party.

Participants in the one-gossiper condition read no additional sentence; the two-gossiper condition added sentence 2; the three-gossiper condition added 2 and 3; the four-gossiper condition added 2, 3, and 4. We predicted that increasing the number of independent sources would increase the gossip’s believability. Unlike in experiment I, here the repetition of gossip was explicitly associated with different, named individuals, each of whom claims to have witnessed the interaction. In order to put the impact of multiple sources to the strictest test, we only added additional gossipers to the scenarios that included all the information suggesting benign alternatives.

Results: Benign Alternatives. We found no significant differences in BELIEVABILITY between the scenario with no extra pieces of information (BASIC, mean =

Table 2. Description of Each Condition in Experiment II

Condition	Extra Information	Number of Gossip Sources	Participants in Each Condition	Combined Conditions
BASIC	None	1	31	NONE
A	A	1	31	ALTERNATIVE
B	B	1	30	ALTERNATIVE
C	C	1	33	NONE
AB	A and B	1	34	ALTERNATIVE
AC	A and C	1	34	ALTERNATIVE
BC	B and C	1	31	ALTERNATIVE
ABC	A, B, and C	1	31	ALTERNATIVE
ABC2	A, B, and C	2	30	
ABC3	A, B, and C	3	32	
ABC4	A, B, and C	4	29	

4.90) and the scenario that only included statement C (mean = 5.27), $t_{61,9} = -0.99$, $p = .33$. We therefore combined these two conditions into the NONE condition (Table 2). We also found no significant differences in the BELIEVABILITY of any scenarios that included A or B and only one gossip, $F_{5, 185} = 0.099$, $p = .99$ (means ranged from 4.63 to 4.81), so we combined all scenarios with at least one of A or B, but only one source of gossip, into an ALTERNATIVE condition.

As predicted, mean BELIEVABILITY in the ALTERNATIVE condition (mean = 4.71, s.d. = 1.37) was slightly but significantly lower than the mean in the NONE condition (mean = 5.09, s.d. = 1.50), $t_{253} = 1.91$, $p = .029$ (one-tailed).

Results: Multiple Sources. We used a linear model with polynomial contrasts to test whether multiple independent sources of gossip increased BELIEVABILITY (only conditions ABC, ABC2, ABC3 and ABC4 were included in this test). Error variance was homogenous across conditions, $F_{3, 118} = .125$. As predicted, there was a large effect. BELIEVABILITY increased monotonically with the number of sources (ABC: mean = 4.81, s.d. = 1.56, $n = 31$; ABC2: mean = 5.40, s.d. = 1.81, $n = 30$; ABC3: mean = 6.69, s.d. = 1.23, $n = 32$; ABC4: mean = 6.97, s.d. = 1.48, $n = 29$). The mean BELIEVABILITY score of those who heard from four gossipers (6.97) was 45% higher than those who heard from one gossip (4.81). The effect size for ABC vs. ABC2, $d = 0.35$, was about double that for simple reiteration ($d = 0.18$). For ABC vs. ABC4, the effect size, $d = 1.43$, was very large. We found no significant sex differences. See Table 3, model 1 for parameters.

Discussion. In contrast to the findings for the simple reiteration of gossip tested in experiment I, experiment II found that when multiple sources of gossip are identified as distinct individuals, there is a strong, unambiguous increase in believability. Experiment II also found that information suggesting benign alternatives for negative gossip reduced gossip believability.

Table 3. Believability as a Function of Source Number and Independence

Model	Variable	Estimate	S.E.	<i>t</i>	<i>p</i>
1 (Experiment II)	Intercept	5.97	0.14	43.1	<.001
	NUM (linear)*	1.74	0.28	6.3	<.001
2 (Experiment III)	Intercept	5.54	0.17	32.39	<.001
	SOURCE (linear)*	1.01	0.30	3.41	<.001

Model 1: BELIEVABILITY as a function of number of independent gossip sources. Residual standard error = 1.53 on 120 degrees of freedom, $R^2 = 0.25$, Adj. $R^2 = 0.24$, $F_{1,120} = 39.11$, $p < .001$.

Model 2: BELIEVABILITY as a function of SOURCE. Residual standard error = 1.69 on 96 degrees of freedom, $R^2 = .108$, Adj. $R^2 = .099$, $F_{1,96} = 11.66$, $p = 0.0009$.

*Polynomial contrasts; higher-order terms were not significant.

Experiment III: Source Dependence and Gossip Believability

Study Rationale and Predictions. Although multiple sources are a cue of veracity, if these sources are not independent of one another, then perceived gossip veracity should be reduced. We predicted that multiple sources who were independent would be more believed than multiple sources who were dependent, because if one source heard the information from the other (dependent), there is no decrease in the probability that an error occurred. Friendship between sources makes deceptive collusion or slanted gossip possible because one friend might benefit by supporting the other friend's story. We therefore predicted that independent sources would be believed more than friends. Friends should be more believed than fully dependent sources, however, because the *possibility* of collusion (friends) is still more reliable than *actual* collusion (dependent); although friends might be intending to deceive, so, too, could the original source of gossip in the dependent condition.

Method. In this between-subjects experiment, 99 participants read this basic scenario:

Pretend that you are an employee at a company that is having a large office holiday party. Linda and Cathy, employees of the company, and their boss are also at the party. Linda is competing with Cathy for a promotion that will be decided on by their boss. The next day at work, Linda tells you that she saw Cathy hitting on their boss at the party.

The dependent measure was each participant's response to the question, "How likely is it that Cathy was really hitting on their boss?" Participants circled a number on a Likert scale from one ("It definitely did NOT happen") to ten ("It definitely DID happen").

The independent variable was the relationship between the two SOURCES of gossip, one of whom was Linda. This manipulation involved adding one of three sentences to the scenario:

Independent: Later that day, another employee also tells you that she saw Cathy hitting on their boss.

Friends: Later that day, another employee who is Linda's close friend also tells you that she saw Cathy hitting on their boss.

Dependent: Later that day, another employee tells you that Linda said she saw Cathy hitting on their boss.

Results. As predicted, BELIEVABILITY increased monotonically from dependent (mean = 4.82, s.d. = 1.91, $n = 34$) to friends (mean = 5.55, s.d. = 1.54, $n = 33$) to fully independent sources (mean = 6.26, s.d. = 1.61, $n = 32$). A polynomial contrast analysis indicated that the linear trend was significant. There were no significant sex differences. See Table 3, model 2.

Discussion. Experiment II showed that multiple sources of gossip increased perceived veracity. This experiment (III) showed that the believability of multiple sources was moderated by their degree of independence.

Experiment IV: Competition and Gossip Believability

Study Rationale and Predictions. Perceived gossip veracity should be moderated by motivations to manipulate gossip, a prediction we test here.

Method. Eighty female and 71 male German students (mean age = 24.8, s.d. = 3.47) read the following stimuli in German (the German version was translated from the English, and then back-translated to check for accuracy):

Please read the following scenario, and imagine yourself in it. You are an employee in an office where you work with Frau Hauser and Frau Schreiber. [*] One day, you run into Frau Hauser at a café during lunch. Frau Hauser tells you that Frau Schreiber was recently arrested for shoplifting.

The dependent measure was each participant's response to the following question, "How much do you believe Frau Hauser?" on a 10-point Likert scale that ranged from one ("I do NOT believe Frau Hauser") to ten ("I DO believe Frau Hauser"). COMPETITION was the independent variable, manipulated by replacing [*] with the following sentence "They are competing with each other to become the new office manager." We predicted participants in the COMPETITION condition would believe Frau Hauser less because her ulterior motive increased the probability of deception or manipulation.

Results. As predicted, there was a main effect for COMPETITION. Students in the non-competition condition believed the gossip significantly more (mean = 5.76, s.d. = 2.33) than students in the competition condition (mean = 4.86, s.d. = 2.49, $F = 4.9$, $p = .028$). There was a marginally significant interaction between COMPETITION and SEX ($F = 3.18$, $p = .077$). Women were more affected by competition (mean = 5.98 vs. 4.41) than were men (mean = 5.48 vs. 5.32).

GENERAL DISCUSSION

The results of these four experiments strongly suggest that the quantity, independence, and motives of information sources are used to evaluate veracity in the restricted domain of reputation-relevant gossip, and that participants were generally resistant to simple manipulations like mere repetition. Although repetition slightly increased the believability of gossip statements, the effect was due entirely to a small number of participants (3%) whose interest in the gossip was low. Multiple, identified sources of gossip, however, strongly increased perceived gossip veracity. Information reported by multiple, identified sources is more likely to be true because (1) the probability of error is lower and/or (2) it is less likely that several individuals would use the same deceptive story. Alternatively, participants' believability ratings for multiple sources might have increased owing to a reiteration effect of unknown origin (though the weak results of experiment I reduce the likelihood of this explanation). The independence of sources was also found to have a strong impact on veracity, whereas information suggesting benign alternative interpretations of a somewhat questionable social interaction slightly but significantly decreased the believability of negative gossip about that interaction. Finally, participants were less likely to believe gossip when there was competition between the gossiper and the target of the gossip. This effect seemed to be larger among women, perhaps because the stimulus involved an incident between women. Our theory predicted no such sex effect, however, even if the incident involved members of the opposite sex.

Our experiments suffer from three main limitations. First, almost all participants were college students age 18 to 30; the results may not generalize to other segments of the population. Second, these studies were conducted in Western populations. We are proposing that people possess a set of innate psychological adaptations to evaluate gossip. Cross-cultural universality is a necessary (but not sufficient) condition for complex adaptations, so this proposal requires considerable cross-cultural validation. Finally, the principal limitation is that our experiments do not clarify *why* the cues we examined are salient. They could indicate a lower possibility of random error, a lower possibility of deception, or other reasons. Future studies should attempt to discriminate between mechanisms that screen for errors versus those that screen for deception.

CONCLUSION

Current models of reputation-based cooperation suggest that errors are an important parameter. Although for language to evolve, most speech must have been honest in the EEA (Lachmann et al. 2001; Silk et al. 2000), this does not imply that deception and error were absent. If error and deception were frequent enough, there would have been a selection pressure for adaptations to assess the believability of language signals. We and others have proposed that cues of information believability should include the number and relationships of sources, and motives to deceive.

The results reported here support this view and provide preliminary evidence for psychological adaptations to assess gossip veracity. Such adaptations would play an important role in the evolution of reputation-based cooperation and could be used more broadly to evaluate information received from others.

Support for Hess provided by the LIFE Program, Max Planck Institute for Human Development. Support for Hagen provided by ECAgents, Contract No. 001940. We thank Peter Hammerstein, Don Symons, members of the ITB at Humboldt, several anonymous reviewers for helpful discussions and comments, and Gregor Caregnato of MPI for running experiment IV.

Nicole Hess is a postdoctoral fellow at the Institute for Theoretical Biology, Berlin. She specializes in evolutionary approaches to gossip and female coalitions. Edward Hagen is a research scientist at the Institute for Theoretical Biology, Berlin.

REFERENCES

- Alexander, R. D.
1987 *The Biology of Moral Systems*. New York: Aldine de Gruyter.
- Archer, J.
2004 Sex Differences in Aggression in Real-World Settings: A Meta-analytic Review. *Review of General Psychology* 8:291–322.
- Archer, J., and S. M. Coyne
2005 An Integrated Review of Indirect, Relational, and Social Aggression. *Personality and Social Psychology Review* 9:212–230.
- Bacon, F. T.
1979 Credibility for Repeated Statements: Memory for Trivia. *Journal of Experimental Psychology* 5:241–252.
- Barkow, J. H.
1992 Beneath New Culture Is Old Psychology: Gossip and Social Stratification. In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, J. H. Barkow, L. Cosmides, and J. Tooby, eds. Pp. 627–637. New York: Oxford University Press.
- Baumeister, R. F., K. D. Vohs, and L. Zhang
2004 Gossip as Cultural Learning. *Review of General Psychology* 8:111–121.
- Bloom, P.
2004 Postscript to the Special Issue on Gossip. *Review of General Psychology* 8:138–140.
- Boerlijst, M. C., A. Nowak, and K. Sigmund
1997 The Logic of Contrition. *Journal of Theoretical Biology* 185:281–293.
- Buss, D. M., and L. Dedden
1990 Derogation of Competitors. *Journal of Social and Personal Relationships* 7:395–422.
- Chagnon, N. A.
1988 Life Histories, Blood Revenge, and Warfare in a Tribal Population. *Science* 239:985–992.
- Cox, S. J., T. J. Sluckin, and J. Steele
1999 Group Size, Memory, and Interaction Rate in the Evolution of Cooperation. *Current Anthropology* 40:369–377.
- Dawkins, R., and J. R. Krebs
1978 Animal Signals: Information or Manipulation? In *Behavioral Ecology: An Evolutionary Approach*, J. R. Krebs and N. B. Davies, eds. Pp. 282–309. North Scituate, MA: Sinauer Associates.
- Dunbar, R. I. M.
1996 *Grooming, Gossip, and the Evolution of Language*. Cambridge: Harvard University Press.
2004 Gossip in Evolutionary Perspective. *Review of General Psychology* 8:100–110.
- Emler, N.
1990 A Social Psychology of Reputation. In *European Review of Social Psychology*, vol. 1, W. Stroebe and M. Hewstone, eds. Pp. 173–193. Chichester, UK: Wiley and Sons.

- Enquist, M., and O. Leimar
1993 The Evolution of Cooperation in Mobile Organisms. *Animal Behaviour* 45:747–757.
- Gintis, H., E. A. Smith, and S. Bowles
2001 Costly Signaling and Cooperation. *Journal of Theoretical Biology* 213:103–119.
- Gluckman, M.
1963 Gossip and Scandal. *Current Anthropology* 4:307–316.
- Goodwin, M. H.
1990a Tactical Uses of Stories: Participation Frameworks within Girls' and Boys' Disputes. *Discourse Processes* 13:52–71.
1990b *He-Said-She-Said: Talk as Social Organization among Black Children*. Bloomington: Indiana University Press.
- Curven M., W. Allen-Arave, K. Hill, and M. Hurtado
2000 "It's a Wonderful Life": Signaling Generosity among the Ache of Paraguay. *Evolution and Human Behavior* 21:263–282.
- Harkins, S. G. and R. E. Petty
1987 Information Utility and the Multiple Source Effect. *Journal of Personality and Social Psychology* 52:260–268.
- Hasher, L., D. Goldstein, and T. Toppino
1977 Frequency and the Conference of Referential Validity. *Journal of Verbal Learning and Verbal Behavior* 16:107–112.
- Hawkes, K.
1991 Showing off: Tests of an Hypothesis about Men's Foraging Goals. *Ethology and Sociobiology* 12:29–54.
1993 Why Hunter-Gatherers Work: An Ancient Version of the Problem of Public Goods. *Current Anthropology* 34:341–361.
- Hawkes, K., J. F. O'Connell, and N. G. Blurton-Jones
2001 Hadza Meat Sharing. *Evolution and Human Behavior* 22:113–142.
- Hertwig, R., G. Gigerenzer, and U. Hoffrage
1997 The Reiteration Effect in Hindsight Bias. *Psychological Review* 104:194–202.
- Hess, N. H., and E. H. Hagen
2006 Sex Differences in Indirect Aggression: Psychological Evidence from Young Adults. *Evolution and Human Behavior* 27:231–245.
- Holland, D. C., and M. A. Eisenhart
1990 *Educated in Romance: Women, Achievement, and College Culture*. Chicago: University of Chicago Press.
- Kuttler, A. F., J. G. Parker, and A. M. La Greca
2002 Developmental and Gender Differences in Preadolescents' Judgments of the Veracity of Gossip. *Merrill-Palmer Quarterly* 48:105–132.
- Lachmann, M., S. Szamado, and C. T. Bergstrom
2001 Cost and Conflict in Animal Signals and Human Language. *Proceedings of the National Academy of Sciences* 98:13189–13194.
- Leimar, O., and P. Hammerstein
2001 Evolution of Cooperation through Indirect Reciprocity. *Proceedings of the Royal Society of London B* 268:745–753.
- Marlowe, F.
1999 Showoffs or Providers: The Parenting Effort of Hadza Men. *Evolution and Human Behavior* 20:391–404.
- Maynard-Smith, J., and D. G. C. Harper
1995 Animal Signals: Models and Terminology. *Journal of Theoretical Biology* 117:305–311.
- McAndrew, F. T., and M. A. Milenkovic
2002 Of Tabloids and Family Secrets: The Evolutionary Psychology of Gossip. *Journal of Applied Social Psychology* 32:1064–1082.
- Milinski, M., D. Semmann, T. C. M. Bakker, and H.-J. Krambeck
2001 Cooperation through Indirect Reciprocity: Image Scoring or Standing Strategy? *Proceedings of the Royal Society of London B* 268:2495–2501.

- Milinski M., D. Semmann, and H.-J. Krambeck
2002 Reputation Helps Solve the "Tragedy of the Commons." *Nature* 415:424–426.
- Mohtashemi, M., and L. Mui
2003 Evolution of Indirect Reciprocity by Social Information: The Role of Trust and Reputation in Evolution of Altruism. *Journal of Theoretical Biology* 223:523–531.
- Nowak, M. A., and K. Sigmund
1998 Evolution of Indirect Reciprocity by Image Scoring. *Nature* 393:573–577.
- Paine, R.
1967 What Is Gossip About? An Alternative Hypothesis. *Man* 2:278–285.
- Panchanathan, K., and R. Boyd
2003 A Tale of Two Defectors: The Importance of Standing for Evolution of Indirect Reciprocity. *Journal of Theoretical Biology* 224:115–126.
- Patton, J. Q.
2000 Reciprocal Altruism and Warfare: A Case from the Ecuadorian Amazon. In *Adaptation and Human Behavior: An Anthropological Perspective*, L. Cronk, N. A. Chagnon, and W. Irons, eds. Pp. 417–436. New York: Aldine de Gruyter.
- Pollock, G, and L. A. Dugatkin
1992 Reciprocity and the Emergence of Reputation. *Journal of Theoretical Biology* 159:25–37.
- Radin P.
1927 *Primitive Man as Philosopher*. New York: Appleton.
- Rosnow, R. L., J. H. Yost, and J. L. Esposito
1986 Belief in Rumor and Likelihood of Rumor Transmission. *Language & Communication* 6:189–194.
- Silk, J. B., E. Kaldor, and R. Boyd
2000 Cheap Talk When Interests Conflict. *Animal Behaviour* 59:423–432.
- Smith, E. A., and R. L. Bliege Bird
2000 Turtle Hunting and Tombstone Opening: Public Generosity as Costly Signaling. *Evolution and Human Behavior* 21:245–261.
- Sosis, R.
2000 Costly Signaling and Torch Fishing on Ifaluk Atoll. *Evolution and Human Behavior* 21:223–244.
- Starrett, A.
1993 Adaptive Resemblance: A Unifying Concept for Mimicry and Crypsis. *Biological Journal of the Linnean Society* 48:299–317.
- Sugiyama, L. S., and R. Chacon
2000 Effects of Illness and Injury among the Yora and Shiwiar. In *Human Behavior and Adaptation: An Anthropological Perspective*, L. Cronk, N. A. Chagnon, and W. Irons, eds. Pp. 371–395. New York: Aldine de Gruyter.
- Vrij, A.
2000 *Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice*. Chichester, England: Wiley.
- Wedekind, C., and M. Milinski
2000 Cooperation through Image Scoring in Humans. *Science* 288:850–852.
- Wert, S. R., and P. Salovey
2004 A Social Comparison Account of Gossip. *Review of General Psychology* 8:122–137.
- Wilson, D. S., C. Wilczynski, A. Wells, and L. Weiser
2000 Gossip and Other Aspects of Language as Group-Level Adaptations. In *The Evolution of Cognition*, C. Heyes and L. Huber, eds. Pp. 347–365. Cambridge: MIT Press.
- Wu, J., and R. Axelrod
1995 How to Cope with Noise in the Iterated Prisoner's Dilemma. *Journal of Conflict Resolution* 39:183–189.