Article

RATIONALITY AND SOCIETY

Rationality and Society
2015, Vol. 27(3) 309–333
© The Author(s) 2015
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1043463115593109
rss.sagepub.com
SAGE

Lifting the veil of ignorance: An experiment on the contagiousness of norm violations

Andreas Diekmann*
ETH Zurich, Switzerland

Wojtek Przepiorka*
Utrecht University/ICS, The Netherlands

Heiko Rauhut*
University of Zurich, Switzerland

Abstract

Norm violations can be contagious. Previous research analyzed two mechanisms of why knowledge about others' norm violations triggers its spread: (1) actors lower their subjective beliefs about the probability or severity of punishment or (2) they condition their compliance on others' compliance. While earlier field studies could hardly disentangle both effects, we use a laboratory experiment which eliminates any punishment threat. Subjects (n = 466) can throw a die and are paid according to their reported number. Our design rules out any possibility of personal identification so that subjects can lie about their thrown number and claim inflated payoffs without risking detection. The aggregate distribution of reported payoffs allows the estimation of the extent to which the honesty norm is violated. We compare two treatment conditions in which subjects are informed about lying behavior of

*Andreas Diekmann can be reached at: diekmann@soz.gess.ethz.ch, Wojtek Przepiorka can be reached at: W.Przepiorka@uu.nl, and Heiko Rauhut can be reached at: heiko.rauhut@uzh.ch.

Corresponding author:
Andreas Diekmann, Chair of Sociology, ETH Zurich, CLU D 4, Clausiusstrasse 50, CH-8092 Zurich, Switzerland.
Email: diekmann@soz.gess.ethz.ch

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

others with a control condition without information feedback. Observations from a subsequent die throw reveal that knowledge about liars triggers the spread of lying compared to the control condition. Results from a follow-up experiment show that this effect is moderated by subjects' beliefs about the prevalence of norm violations of others. Our results demonstrate the contagiousness of norm violations, where actors imitate norm violations of others under the exclusion of strategic motives.

Keywords
Cheating, conditional cooperation, experimental sociology, Heinrich Popitz, social influence, social norms

Introduction

Ignorance can act as a protective barrier to the spreading of norm violations. This hypothesis has its roots in the writings of the German sociologist Heinrich Popitz (1968). Popitz argued that once people learned about the full extent of other peoples' norm violations, they would adjust their behavior accordingly. Thus, assuming that people underestimate the frequency of infringements and social norm violations, Popitz expected, in his own words, a "preventive effect of ignorance." This hypothesis is a precursor to the social-psychological focus theory (Cialdini et al., 1990; see also Gino et al., 2009). Observations of others' norm violations may change actors' beliefs about the appropriateness of their own actions. In this sense, "descriptive norms," that is, what most people do, affect "injunctive norms," that is, what most people approve of. To put it simply: if others break a certain norm, it is fine if you break it too. We call this effect *conditional norm compliance.* More precisely, we mean with conditional norm compliance the motivation of individuals to adhere to norms if they believe that others do so as well and to violate them if they believe that others violate them.[1]

A similar notion has been raised by the so-called broken windows theory, which emphasizes the contagiousness of norm violations (Wilson and Kelling, 1982). The paradigmatic example is the observation of broken windows in a neighborhood or other signs of disorder, which eventually trigger the spread of norm violations.

There is a growing body of empirical research testing the hypothesis of the contagiousness of norm violations, an example of which is the series of field experiments by Keizer et al. (2008). These experiments tested whether information about a certain kind of disorder triggered the spread of other kinds of disorder. The authors were able to show that graffiti caused people to litter, illegal parking caused illegal trespassing, and a large number of unreturned supermarket shopping carts caused littering. Graffiti and litter

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

also caused people to steal money from sealed envelopes sticking out of mailboxes. In another set of field experiments, Cialdini et al. (1990) showed that people have a higher propensity to litter if they observe others littering and if the environment is already littered.

While there is some empirical evidence for the contagiousness of norm violations, the causal mechanisms driving this effect are unclear. Several factors can be identified, which are rooted in different schools of thought. First, the effect could be explained by *rational-choice theory*. Observing others' norm violations may cause people to update their beliefs regarding their cost–benefit analysis of norm violations (Allingham and Sandmo, 1972; Becker, 1968; Bikhchandani et al., 1998; Groeber and Rauhut, 2010). Signs of disorder could therefore imply that norm violations are hardly ever detected and, if so, only mildly punished. Moreover, norm violations in the public sphere, such as littering, concern a collective good. When the collective good is already destroyed, people might reason that further violations do not harm. Both factors affect the cost–benefit ratio. Second, the effect could be explained by the motivation of conditional norm compliance, as is proposed by psychological focus theory. Descriptive norms, or what other people do, affect injunctive norms, the motivation to comply with or to violate a social norm. The contagiousness effect is expected even if sanctions are absent or the perception of the likelihood and severity of sanctions is unrelated to the descriptive norm.

Our objective is to demonstrate that conditional norm compliance is sufficient to trigger the effects of ignorance on norm compliance. With our experimental design, we will test the hypothesis of a "pure" effect of conditional norm compliance. While other studies, like the recent one by Keizer et al. (2008) or those setting forth the broken windows theory (Sampson and Raudenbush, 2004; Wilson and Kelling, 1982), had to advance the additional assumption that punishment of observed norm violations is simply uncommon and therefore negligible, we can rule out the punishment argument not only with theoretical assumptions but also with experimental design. In what follows, we discuss examples, case studies, and previous research regarding the contagiousness of norm violations. Then, we introduce our experimental design, present empirical results, and discuss our findings.

Contagiousness of norm violations

Imagine you knew that most inhabitants of your city evaded taxes, cheated on their partners, dodged paying the fares for public transport, lied about their age at dates, and crossed against red lights. While we may intuitively think that a high detection rate of norm violations maintained social order, this very knowledge of others' norm violations may lower the willingness to comply with norms. Loosely based on the aphorism "what the eye does not

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

see, the heart does not grieve over," ignorance of norm violations can have a preventive effect. In what follows, we use the term *ignorance hypothesis* to refer to the preventive effect of not knowing about others' norm violations.

We all have every-day evidence about the contagious effects of norm violations from scenarios in which norm violations are visible. If many people start to cross against the red light at a crosswalk, others join in. If many cars in a street park in clearways, the sideway and other no-parking zones quickly become jammed of cars. Seeing a littered beach makes people more likely to litter themselves, which gradually pollutes the beach. However, many norm violations are not directly visible but conducted in private. Two-timing, tax evasion, consumption of pornography, visits of prostitutes, or alcohol abuse are behaviors which are often concealed from others. Some of which are revealed, but others remain covered. Seneca was probably one of the first making the observation that ignorance of these invisible norm violations can act as a protective barrier to the spread of norm violations. About 2000 years ago, he wrote in his "Moral Essays":

> In that state in which men are rarely punished a sympathy for uprightness is formed, and encouragement is given to this virtue as to a common good. Let a state think itself blameless, and it will be so; its anger against those who depart from the general sobriety will be greater if it sees that they are few. Believe me, it is dangerous to show a state in how great a majority evil men are. (Seneca, 1928 [first 63], xxii. 3–xxiv. 1)[2]

Nearly 2000 years later, Popitz (1968) suggested a more detailed account specifying the *ignorance hypotheses* by three impossibility statements:

> Impossible is the complete transparency of all norm-relevant behaviors in society, a normative system which could cope with the detection of all norm violations, and a punishment system which would retain its protective function if all norm violations were sanctioned. (p. 18, translated by the authors)

Popitz (1968) illustrates the ignorance hypotheses by a thought experiment from Thackeray's (1869) classical novel. Consider what would happen if every norm violation was actually detected:

> Just picture to yourself everybody who does wrong being found out, and punished accordingly. Fancy all the boys in all the schools being whipped; and then the assistants, and then the headmaster […] Fancy the provost marshal being tied up, having previously superintended the correction of the whole army. […] The butchery is too horrible. The hand drops powerless, appalled at the quantity of birch which it must cut and brandish. I am glad we are not all found out. (Thackeray, 1869, quoted in Popitz, 1968)

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

The impossibility of complete transparency of normative behavior builds the foundation for two macro-sociological hypotheses: (1) if the complete extent of norm violations was known, norm violations would gradually increase and the normative system would collapse and (2) if all norm violations were sanctioned, the punishment system would collapse. The counter-intuitive point here is that the political goal to uncover undetected norm violations actually counteracts the stability of the norm.

In this regard, Popitz' theory arrives at very different conclusions than the broken windows theory. Although both theories share the assumption of the contagiousness of norm violations, the broken windows theory mainly focuses on the visibility of transgressions in the public sphere while the ignorance hypothesis directs the attention to invisible norm violations in the private sphere. The broken windows theory emphasizes the macro-sociological consequences of a low tolerance level and is therefore often referred to by advocates of law and order (although this policy does not follow automatically from the theory). In contrast, Popitz' theory argues against over-regulation and a more cautious persecution of transgressions.[3]

It is possible to reconstruct the macro-sociological implication of the ignorance hypothesis by micro-level mechanisms. Again, its main premise is that actors do not exactly know if and to what extent other members of society violate a particular norm—be it black labor, corruption, smuggling, shoplifting, "deviant" sexual practices, or adultery: "There is in general a large discrepancy between the actual and the subjectively perceived extent of norm violations" (Popitz, 1968: 15). That is, ignorance typically implies underestimation of the extent of norm violations in the population, which lowers actors' propensity for norm violations. "Many social forces work towards generating a relatively favorable representation of norm compliance in society, which may also be described as the general requirement for the 'functioning' of a particular social norm in society" (Popitz, 1968: 15). Thus, if the actual extent of norm violations was known, the legitimacy of the norm would deteriorate.

The aforementioned theoretical reasoning can be illustrated by an example from the history of the German Democratic Republic (GDR). The socialist state discouraged to watch Western television channels and political campaigns were launched against households lacking this kind of compliance. The crux was that households consuming Western television were identifiable by the direction of their antennas upon their roofs. Walter Ulbricht, the then leader of the East Germany Communist Party, realized that this lifted the veil of ignorance and stated "the class enemy is sitting on the roof" (Spiegel, 1980: 42). He tried reverting the situation by encouraging the East German youth organization "Freie Deutsche Jugend" to tear down television antennas oriented to receive Western German channels

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

(Marks, 1983: 50). Too late—the otherwise unknown extent of Western television consumption was already revealed to everybody. In fact, this knowledge eventually triggered the collapse of the prohibitive norm against Western television consumption in 1973 (Spiegel, 1980: 44).

Furthermore, lifting the veil of ignorance may even trigger normative change as the Kinsey et al. (1948, 1953) report about sexual behaviors in the United States exemplified:

> In sum, Kinsey was the major factor in changing attitudes about sex in the twentieth century. [...] He changed the nature of sexual studies, forced a reexamination of public attitudes toward sex, challenged the medical and psychiatric establishment to reassess its own views, influenced both the feminist movement and the gay and lesbian movement, and built a library and an institution devoted to sex research. His reputation continues to grow, and he has become one of the legends of the twentieth century. (Bullough, 1998: 131)

According to our argument, the veil of ignorance of various sexual practices was lifted. The extent of homosexual behavior, masturbation, oral sex, and other practices became public. It became apparent that these behaviors were more widespread than previously thought, which triggered their subsequent spread and eventually contributed to a fundamental change in norms and values in people's sex lives. In what follows, we test the ignorance hypothesis in a simple laboratory experiment.

Design of the experiment

The question as to whether ignorance has a preventive effect is an empirical question. One problem with traditional research designs such as surveys or official crime records, however, is that the actual extent of norm violations is not only unknown to the population but also to the researcher. Generally, there is undercoverage of norm violations so that the actual rate has to be estimated by making a number of influential assumptions that are often hard to validate. In contrast, laboratory experiments enable measuring the complete rate of norm violations in a very direct way.

In laboratory experiments, however, it may be that the true rate of norm compliance is overestimated because subjects feel observed and may react overly norm adherent. These considerations led us to the implementation of a laboratory experiment that eliminated the possibility of identifying which particular subject committed a norm violation. Our goal was to establish maximal anonymity on the individual level by keeping the possibility to correctly infer the extent of norm violations from aggregated data.

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

We chose the dice experiment of Fischbacher and Föllmi-Heusi (2013) as basic design of our study. Later on, we realized that the complete absence of a sanctioning threat was a desirable side effect of the design. The experiment enabled subjects to commit a violation of the honesty norm. Subjects threw a die in a completely anonymous setting, meaning that they were alone and not observed at all. Then, they had to report their number in complete privacy by entering it in a computer system located in an isolated box. Subjects received cash in Swiss Francs according to the number they reported. The only exception was the number six, which corresponded to zero earnings (see Figure 8 in Appendix 3 for the exact wording of instructions). All subjects knew that they were the only ones knowing the true number they actually threw. Therefore, subjects could distort the truth and report a number that yielded them higher monetary earnings. This meant that subjects who did not throw a "five" faced an internal conflict between maximizing their own payoff by reporting a higher number and adhering to the honesty norm.[4] The implementation of anonymity among subjects and between subjects and experimenter eliminated any possibility to detect liars. This guaranteed that our experimental design avoided any confounding with punishment threats, neither material nor social ones. In addition, our experiment avoided any confounding with strategic reasoning. Individual lying did not affect the payoffs of other subjects.

Of course, our experiment did not allow for the observation of individual norm violations, but the point of the whole procedure was to provide individuals with maximal anonymity for a valid elicitation of behaviors. Nonetheless, we could estimate the extent of norm violations at the group level. If the honesty norm was consistently followed by all subjects, we would observe a rate of about one-sixth for every reported dot. Therefore, the extent of norm violations can be measured by computing the difference between theoretical and observed rates for each dot.[5]

According to the ignorance hypothesis, knowledge about norm violations in the population should trigger subsequently more norm violations. Therefore, we informed subjects about the distribution of reported numbers after the first throw. In this way, we uncovered the dark field of norm violations. We implemented two ways of communicating the empirical distributions of first throws. In the information condition 1 (info 1), we presented the distribution of the dice experiment of Fischbacher and Föllmi-Heusi (2013). Thus, we explained the subjects that the distribution they saw was an empirical distribution from 389 subjects who came from the same universities and previously participated in the same experiment (see Figure 9 in Appendix 3 for experimental instructions and the displayed distribution in info 1). The distribution has a clear pattern. The displayed over-reporting of the numbers "four" and "five" and the implied violation of the honesty norm is striking.

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

In the information condition 2 (info 2), we reported the distribution of the group which was currently playing the game (see Figure 10).

The advantage of the first information condition is that the distribution represents behaviors of almost 400 subjects. Furthermore, it does not vary over different experimental sessions. The disadvantage is that subjects may consider the distribution as “external,” because it stems from a group of different subjects. This is different for the “internal” implementation of info 2, which represented the behaviors of the very subjects who participated in the same experimental session. Here, however, the dice throws naturally generated a great deal of randomness. This means that the generated group distribution in each session was always a different distribution generated from a different group of subjects. Because both methods are complementary in their advantages and disadvantages, they balance each other and enable robustness checks of our empirical results.

Furthermore, we implemented a control treatment (info 0) in which we did not present the distribution of first throws. After the first throws and the presentation of the respective distribution (respectively, its absence in the control group), we implemented another round in which subjects were asked to throw the die again. We paid subjects for both throws similarly. Hence, the maximum payment was 10 Swiss Francs. Table 1 summarizes our design.

The instructions made it clear that subjects were allowed to test their die. Thus, they were allowed to throw the die as many times as they wished. However, only the first throw counted. This rule was stated very explicitly. This setup made it possible for subjects to ensure that they were not deceived by fraud dice. Furthermore, this setup implicitly provided self-justifications for lying, and therefore made it easier for subjects to violate the honesty norm in favor of their self-interest. Subjects who were unlucky with their first throw may have continued to throw the die and tempted to report another, more favorable number which occurred later on. Potentially, the inhibition threshold for illegitimately reporting an actual number of their series may be lower compared to directly entering a fraudulent number (see Shalvi et al., 2011, for an empirical confirmation of this argument).

Furthermore, subjects were asked to enter their earnings in addition to their thrown numbers. This allowed testing as to whether they understood the rules of the game. The devaluation of the number six was on purpose. This design feature should increase the propensity for norm violations. Since in most board games, the number “six” is a desirable outcome, subjects may feel particularly frustrated by a “six.”

The dice experiments were conducted at the end of four unrelated other experiments. The subjects were recruited from ETH Zurich and University of Zurich during May 2009 and May 2010. The experiments were

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

Table 1. The experimental design.

Control group (info 0)	O		O
"External" distribution of many others (info 1)	O	X_1	O
"Internal" distribution of own session (info 2)	O	X_2	O

O: observation; X: intervention.

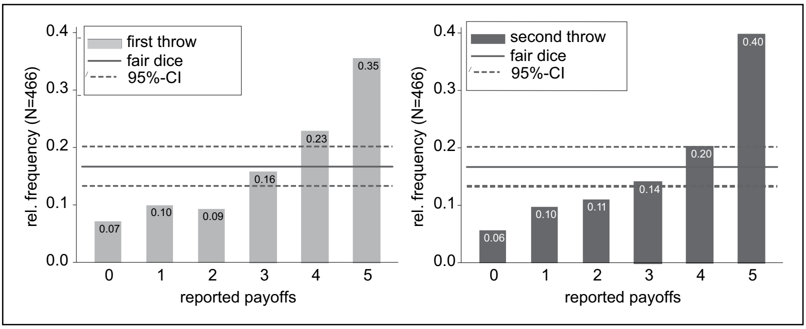

Figure 1. Distribution of reported payoffs from the first and second throws (averaged over all treatments).

conducted using the software z-Tree (Fischbacher, 2007). Thirty sessions were conducted, each of which with 14–16 subjects. In total, 466 subjects participated in the experiment. Subjects were male in 63% of cases and 23 years old on average (standard deviation (*s.d.*) = 3.2). After the experiment, a small questionnaire was administered eliciting socio-demographic information such as income, education of the parents, religious affiliation, and religiousness.

Results

Violation of the honesty norm

Is the honesty norm violated at all? The null hypothesis states that each number is rolled with a probability of 1/6 (i.e. 16.7%). Even in the first throw, there is an apparent discrepancy between what we see and what we would observe if everybody was honest (see Figure 1). The relative frequency of the highest payoff is 35% in the first and 40% in the second throw (averaged over all treatments). The expected probability of throwing twice a five is 1/36 (2.8%). Despite this, 20% report this occurrence, which is over

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

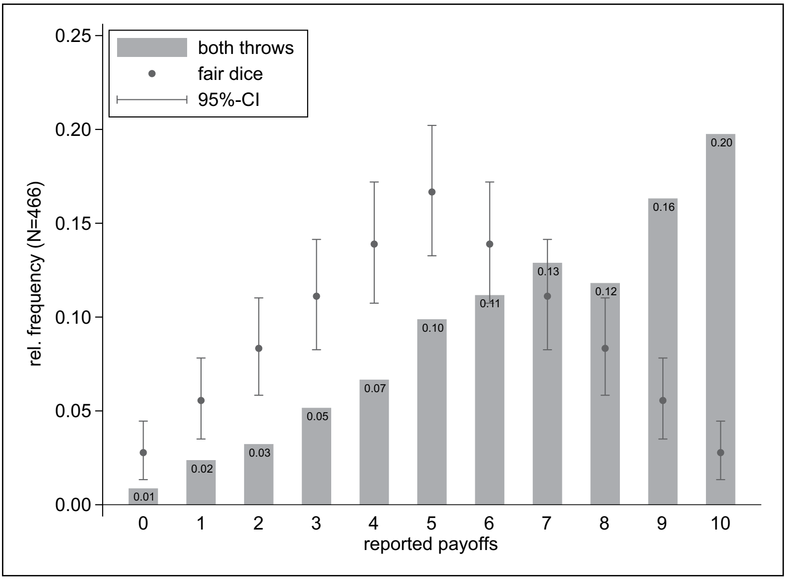

Figure 2. Distribution of reported cumulated payoffs from both throws (averaged over all treatments).

seven times as much as what we would observe in an honest population (Figure 2). In contrast, an "honest six" (with zero earnings) is only reported by 7% in the first and by 6% in the second throw. Moreover, a "super-honest" double six is reported in only 1% of the cases (for the frequency distribution of payoff combinations, see Figure 7, Appendix 2).

Interestingly, there is even fraud below the maximum. Apparently, people make compromises between their compliance with the honesty norm and their self-interest (see also Fischbacher and Föllmi-Heusi, 2013). Figure 2 shows the distribution of the sum of payoffs from both throws. The number nine occurs in 16% of the cases, although its expectation value would be 5.6%. Possibly, a combination of four and five is frequently reported, because it seems less suspicious than a double five.

Testing the ignorance hypothesis

While the previous analyses demonstrate that a substantial fraction of the population violates the honesty norm and claim more money than they are entitled to, the question remains as to whether lying behavior is even more widespread if people are informed about the lying behavior of others.

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

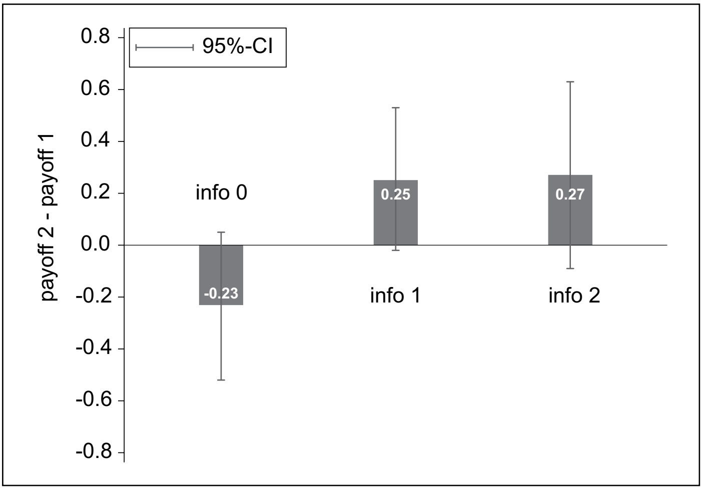

Figure 3. Difference between first and second reported payoffs by experimental condition.

Hence, we compare the extent of norm violations across the three experimental treatments. Figure 3 shows the differences between first and second reported payoffs in the control condition (info 0), in the condition with information about the external large group (info 1) and in the condition with the information about the internal small group (info 2). Although not significant, we observe a considerable decline in mean payoffs in the control group (Figure 3). By inspection of our data, we do not know whether this is a trend or random variation.[6] Therefore, we tested both: the increase through information feedback compared to the control group (differences in differences) and the payoff increase in treatment groups versus the null hypothesis of no effect.

A comparison between both information conditions and the control condition yields significant differences (analysis of variance (ANOVA); $F(2, 29)=4.90$; $p=0.015$; see Appendix 1; Table 3, ordinary least squares (OLS 1) for further details). Note that there is no significant difference between the treatment conditions (info 1 vs. info 2), suggesting that the kind of information feedback is less important than the fact of information feedback. In contrast to a comparison of the treatment conditions with the control condition, it is also possible to test whether the differences between the mean reported payoffs in the first and second throws are different from 0 in all three experimental conditions. This is tested by a linear regression model without intercept (see Appendix 1; Table 3, OLS 2). A joint test yields that

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

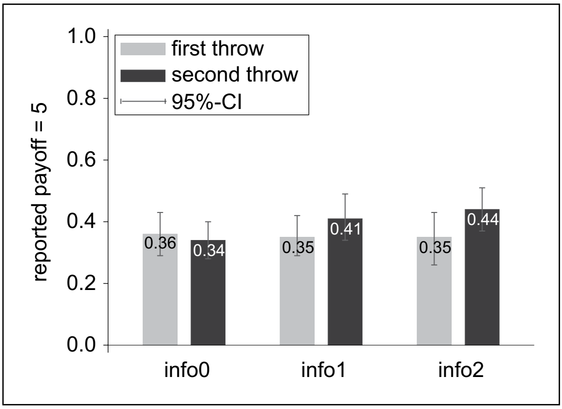

Figure 4. Fraction of reported maximum payoff (throwing a "five") by experimental condition.

the treatment differences are significantly different from 0 (ANOVA: $F(3, 29)=3.47$; $p=0.029$). However, each separate treatment effect is not significantly different from 0. The p-values for the difference of the control condition is $p=0.10$, for treatment condition 1, $p=0.07$, and for treatment condition 2 $p=0.14$. However, the effects of both treatment conditions are in line with the hypothesis and support the conclusions from the linear regression model with intercept, which yielded significant differences in mean outcomes between each treatment condition and the control.

Figure 4 shows that the rate of subjects who claim the highest payoff in the first and the second throws of the control condition is almost similar. However, this rate increases in both experimental conditions. Hence, the rate of norm violations increases if people are informed about the extent of norm violations in an external group (info 1) or in their own group (info 2). We can confirm the statistical significance of this finding by logistic regression models, using the probability of a reported five as the outcome variable and the experimental condition as the predictor. (Note that confidence intervals in Figure 4 are computed from this logistic regression; see Appendix 1; Table 3.)

Session variability

Our second treatment condition (info 2) offers a more specific test of the ignorance hypothesis. In this condition, subjects have seen different rates of reported first payoffs, because information feedback was based on the

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

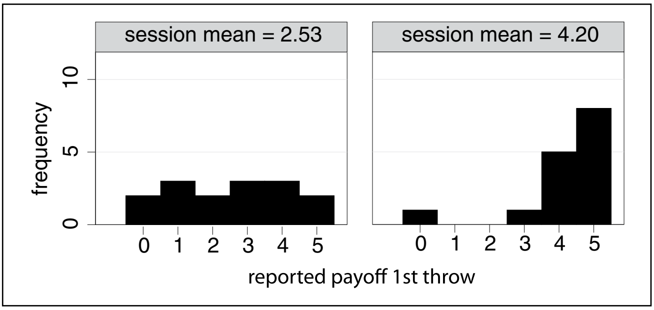

Figure 5. Illustration of the two sessions with the smallest and largest session average of reported first payoffs.

specific sessions subjects were participating in.[7] Sessions varied substantially regarding the distribution of reported first payoffs. This variation may have partly been due to different propensities for lying and partly to pure randomness in throws. We can exploit session variability in lying as an indicator for the extent of revealed lying. In some sessions, an almost even distribution indicated a small extent of lying, while in others a strongly skewed distribution indicated a large extent of lying (see Figure 5 for the two most extreme sessions).

Our more specific ignorance hypothesis states that a larger extent of revealed lying in first throws triggers more subsequent lying in second throws. We used the session mean as an indicator of revealed lying. We regressed session means of reported first payoffs on individuals' propensity to report a five in the second throw (Table 2, model 1) and on the reported mean payoff in the second throw (Table 2, model 2). Both regressions only refer to data from information treatment 2, where information about distributions varied.

The more specific ignorance hypothesis has to be rejected. The analysis in Table 2 reveals that session variability in information feedback regarding initial lying does not have statistically significant effects on subsequent lying. We conducted a number of additional robustness checks of the null finding (see Online Appendix for details). Our robustness checks consisted of different operationalizations of revealed lying behavior, such as using the number of reported fives in a session instead of session means. We also tried different statistical specifications of the hypotheses by taking different models, such as Poisson regressions. Furthermore, we ensured by simulation scenarios that the variability in sessions would have been sufficient to elicit significant findings and to make sure that our models are not affected by

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

Table 2. Regression models quantifying whether a higher indicator of lying about first throws in sessions (measured by session means) increases the number of reported fives (model 1) and reported payoffs (model 2) in second throws.

	(1)	(2)
	Five in second throw	Reported payoff second throw
	Only info 2 treatment	Only info 2 treatment
Session mean first throw	−0.067 (−0.17)	0.029 (0.085)
Intercept	−0.012 (−0.0089)	3.58** (3.01)
N	150	150

** $p < 0.01$, t statistics in parentheses.

statistical artifacts like the so-called "regression to the mean." All these analyses indicate that the more specific ignorance hypothesis has to be rejected.

Overestimation

So far we had assumed implicitly that subjects underestimated the extent of norm violations. This assumption is also an integral part of the ignorance hypothesis. However, it may occur that the population studied is heterogeneous, that is, there are underestimators and overestimators (see Groeber and Rauhut, 2010; Opp, 2012; Schultz et al., 2007). In accordance with the ignorance hypothesis, information about others' behavior will drive underestimators' norm violations upward. In contrast, the level of overestimators' norm violations is expected to move downward. Schultz et al. (2007) present evidence on household energy consumption. After having been given information about their neighbors' electricity consumption, overconsumers' electricity usage declined and underconsumers' increased. An extension of our study by one of the authors (Rauhut, 2013) did not corroborate the ignorance hypothesis for the whole sample. However, after controlling for subjects' beliefs about how many others in the population complied with the honesty norm, the ignorance hypothesis could be confirmed for underestimators (Figure 6): lying increased in underestimators who were given an estimate of others' norm violations (relative to underestimators who received no information about others' behavior). The opposite effect emerged in the group of overestimators: lying decreased in overestimators who were given an estimate of others' norm violations (relative to overestimators who received no information). The conclusion from all this evidence is that the testability of the ignorance hypothesis crucially depends on the

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

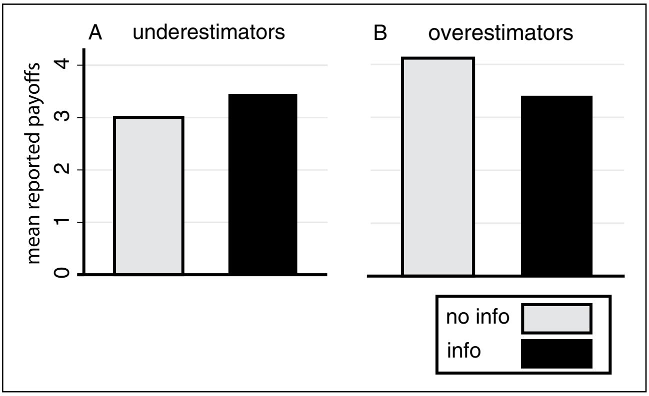

Figure 6. Interaction between beliefs and lying dynamics in dice experiments by Rauhut (2013). The figure displays mean reported payoff claims without (gray) and with (black) information feedback. Panel A shows differences in reported payoffs between control and information treatments for "underestimators" and panel B for "overestimators." Underestimators are defined such that they hold beliefs below and overestimators above reported payment claims in their group. Estimated values are based on 480 decisions. Main and interaction effect are significant with $p < 0.001$.
Data source: Rauhut (2013); the figure can be computed from the linear regression model in Table 3 therein.

auxiliary assumption that most people in the population in question underestimate the extent of norm violations.[8]

Discussion

Our key point in this article is the demonstration of the contagiousness of norm violations by the exclusive mechanism of conditional norm compliance. This mechanism is the "purest" kind of contagiousness, where actors simply imitate norm violations of others. Our experimental design excluded any punishment threats by generating absolute anonymity of norm violations. This excluded any explanation based on rational belief updates or cost–benefit analyses of getting caught and punished. Our findings therefore demonstrate that normative behavior implies "automatic," non-strategic decision-making when it comes to compliance with or violation of social norms. Our results suggest that merely information of norm violations of others is sufficient to trigger its spread.

Our experiment investigates whether lifting the veil of ignorance of the extent of others' norm violations triggers subsequently higher rates of norm

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

violations (*ignorance hypothesis*). This can be confirmed: giving information about the extent of lying increases subsequent lying. This effect cannot be due to learning or habituation since it does not occur in the control condition, where subjects receive no information feedback. Furthermore, this effect was robust to the specific kind of information feedback; either knowledge about norm violations of the own group or of a group of similar others.

Our additional analyses show that a greater extent of revealed cheating does not seem to trigger even more cheating. This could mean that mere information feedback gives subjects the idea that others are cheating and this awareness is sufficient to trigger the erosion of the honesty norm. The actual extent of cheating seems to be less relevant. Furthermore, it could be that subjects have difficulties to calculate the extent of cheating from the distribution of reported payoffs, and therefore only take the fact into account that cheating does occur. Another alternative is that the effects are too small to be detected. With 466 cases, our sample size appears to be large, but our design trades off high anonymity against statistical efficiency. Subsequent studies should explore designs that yield more efficient estimates of the extent of norm violations. Finally, there is the possibility that a certain proportion of subjects overestimated the extent of norm violations, thereby mitigating the testability of the ignorance hypothesis.

In our study, we implicitly assumed that the rate of undetected norm violations is underestimated. This implies that underestimation stabilizes the norm. The larger this bias, the stronger the effect of ignorance on norm compliance. In contrast, if the extent of norm violations was overestimated, the reverse effect could occur; a preventive effect of knowledge. This interaction seems to hold, the implications of which were explored in a follow-up study (Rauhut, 2013). Because the presence of overestimators diminishes the ignorance effect, further replications should control for ex ante beliefs on the frequency of norm violations.

The erosion of social norms is typically a gradual, subtle, and slow process. Social norms do not disappear overnight. Possibly, the honesty norm would further deteriorate if we allowed for a substantial continuation of norm violations and respective information feedback; a conjecture which also holds in the experiment by Rauhut (2013).

The conclusion that contagiousness of norm violations can exclusively operate by conditional norm compliance relates our findings not only to social psychology but also to experimental economics. Here, the recent notion of *conditional cooperation* describes the experimental finding that people condition their contributions to public goods on their beliefs of what they think others would contribute (Fischbacher et al., 2001; Neugebauer et al., 2009). Individuals with cooperative intentions cease to cooperate if they become aware of sufficiently many free riders (Gächter, 2007). The

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

finding of conditional cooperation in experimental economics overlaps with sociological and social-psychological evidence showing that individuals align their behavior with the behavior of others (Goldstein et al., 2008; Schultz et al., 2007).

Our findings underline the existence of non-strategic motives of normative behavior. In our experiment, there is no "rational" explanation of why actors condition their behavior on others, because liars in the group did not affect payoffs of other members. This distinguishes our findings from strategic explanations of conditional cooperation in public goods experiments, where free riders in the group lower the payoffs of all members. This gives sound evidence of the "non-rational," imitative basis of normative behavior.

Acknowledgements

The alphabetical order of the authors is intentional and reflects equal contributions. Specific author contributions: all three authors designed the experiment together, W.P. programmed the experiment, W.P. and H.R. performed the experiment and analyzed the data, and A.D. and H.R. wrote the article. Please address correspondence to A.D., W.P., or H.R. The authors thank James A. Kitts, Ryan O. Murphy, Karl-Dieter Opp, Fabian Winter, and the anonymous reviewers for their valuable comments. The data source of this article is Diekmann et al. (2011), which is a German research note. This research note focuses on the work of the classical German sociologist Heinrich Popitz and illustrates how classical theories, such as Popitz' theory of social norms, can be tested by experimental data. The English article is a revised and extended version with novel data analyses, including the analyses of session variability and the computation of types in the Appendix.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. The concept of conditional norm compliance relates to the recent debate about conditional cooperation in experimental economics (Fischbacher et al., 2001; Gächter, 2007; Neugebauer et al., 2009). Our concept refers to pure imitation without requiring strategic motives of reciprocity. A non-strategic explanation of conditional cooperation was put forward by Fischbacher et al. (2001). They measured conditional cooperation in public goods experiments by the strategy method. They found that subjects contribute conditionally on the contribution levels of other group members. However, most subjects are only willing to contribute less than the average contribution level in their group. Fischbacher et al. (2001) call this behavior "conditional cooperation with a self-serving bias." If these behavioral types are put in a dynamic context, they can generate

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

the well-known observation of decreasing cooperation levels in repeated public good situations. Neugebauer et al. (2009) provide further evidence for this mechanism. They examined alternative hypotheses with the repeated public good game. The empirical evidence confirmed the hypothesis of "selfish-biased conditional cooperation." No evidence was found for the approach that strategic motives drive the process of the decay of cooperation (Kreps et al., 1982). Thöni and Gächter (2015) report findings of an experiment with the "gift-exchange" game. Their results on conformism with peer behavior are very much in accordance with the Popitz hypothesis. The authors conclude, "… some agents initially choose their effort according to a norm and turn to a selfish strategy once they observe others breaking the norm."

2. These observations built the basis of Seneca's political recommendation to uphold ignorance of the number of slaves:

   > A proposal was once made in the senate to distinguish slaves from free men by their dress; it then became apparent how great would be the impending danger if our slaves should begin to count our number. Be sure that we have a like danger to fear if no man's guilt is pardoned; it will soon become apparent how greatly the worse element of the state preponderates. (Seneca, 1928 [first 63], xxii. 3–xxiv. 1)

   More recently, in *The New York Times*, Friedman (2011) pointed out that, among other things, the information about the unequal distribution of land accessible through Google Earth lead to popular uprising against the ruling family in Bahrain. Note, however, that the question of how perceived inequality may deteriorate the legitimacy of an autocratic system is a related but different question than the one implied in our ignorance hypothesis. Here, we suggest that ignorance about norm violations prevents the diffusion of such behavior and thus stabilizes a normative system.
3. For a recent discussion of the theoretical implications of Popitz' ignorance hypothesis, see also Opp (2012).
4. Of course, the experimental instructions did not point out the possibility of false reporting. Therefore, some subjects may not have considered the option of cheating. In case subjects underestimated the rate of norm violations and were not aware of the cheating option, the ignorance effect will be biased downward. A downward bias may also result if subjects in the information condition become aware that dishonest behavior will be discovered at the group level.
5. In this sense, the method is comparable to the so-called randomized response method (Warner, 1965).
6. A replication study (Kroher and Wolbring, 2015) did not find decreasing payoffs in the control group and, therefore, did not provide evidence for the existence of a robust trend.
7. The session data are structured as follows. The experimental sessions consisted of 14–16 subjects. In each session, all subjects were partitioned into the three treatment conditions so that each treatment consisted of four to six subjects within each session (except session 19, which only consisted of treatments 1 and 2).

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

8. The replication by Kroher and Wolbring (2015) without belief measures showed a positive but not significant increase in norm violations after information feedback. Here, too, it may be assumed that the effect in the group of underestimators was (partially) compensated by the opposite effect in the group of overestimators.

References

Allingham MG and Sandmo A (1972) Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1(3–4): 323–338.

Becker GS (1968) Crime and punishment: An economic approach. *Journal of Political Economy* 76(2): 169–217.

Bikhchandani S, Hirshleifer D and Welch I (1998) Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives* 12: 151–170.

Bullough VL (1998) Alfred Kinsey and the Kinsey Report. Historical overview and lasting contributions. *Journal of Sex Research* 35: 127–131.

Cialdini RB, Reno RR and Kallgren CA (1990) A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology* 58: 1015–1026.

Diekmann A, Przepiorka W and Rauhut H (2011) Die Präventivwirkung des Nichtwissens im Experiment. *Zeitschrift für Soziologie* 40(1): 74–84.

Fischbacher U (2007) Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2): 171–178.

Fischbacher U and Föllmi-Heusi F (2013) Lies in disguise. An experimental study on cheating. *Journal of the European Economic Association* 11(3): 525–547.

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

Friedman TL (2011) This is just the start. *The New York Times*, 1 March. Available at: http://www.nytimes.com/2011/03/02/opinion/02friedman.html

Gächter S (2007) Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications. In: Frey BS and Stutzer A (eds) *Economics and Psychology: A Promising New Cross-Disciplinary Field*. Cambridge, MA: MIT Press, pp. 19–50.

Gino F, Ayal S and Ariely D (2009) Contagion and differentiation in unethical behavior. *Psychological Science* 20(3): 393–398.

Goldstein NJ, Robert B, Cialdini, et al. (2008) A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research* 35(3): 472–482.

Groeber P and Rauhut H (2010) Does ignorance promote norm compliance? *Computational Mathematical Organization Theory* 16: 1–28.

Keizer K, Lindenberg S and Steg L (2008) The spreading of disorder. *Science* 322: 1681–1685.

Kinsey AC, Pomeroy WB and Martin CE (1948) *Sexual Behavior in the Human Male*. Philadelphia, PA: Saunders.

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

Kinsey AC, Pomeroy WB, Martin CE, et al. (1953) *Sexual Behavior in the Human Female*. Philadelphia, PA: Saunders.

Kreps DM, Milgrom P, Roberts J, et al. (1982) Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory* 27: 245–252.

Kroher M and Wolbring T (2015) Social control, social learning, and cheating. Evidence from lab and online experiments on dishonesty. *Social Science Research*. doi: http://dx.doi.org/10.1016/j.ssresearch.2015.06.003

Marks D (1983) Broadcasting across the wall: the free flow of information between East and West Germany. *Journal of Communication* 33(1): 46–55.

Neugebauer T, Perote J, Schmidt U, et al. (2009) Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of Economic Psychology* 30: 52–60.

Opp K-D (2012) The beneficial and unintended consequences of false beliefs about norm violation. When is there a "preventive effect of ignorance"? In: Mica A, Peisert A and Winczorek J (eds) *Sociology and the Unintended: Robert Merton Revisited*. New York: Peter Lang, pp. 257–283.

Popitz H (1968) *Uber die Präventivwirkung des Nichtwissens: Dunkelziffer, Norm und Strafe*. Tübingen: Mohr Siebeck.

Rauhut H (2013) Beliefs about lying and spreading of dishonesty: Undetected lies and their constructive and destructive social dynamics in dice experiments. *PLoS ONE* 8(11): e77878.

Sampson RJ and Raudenbush SW (2004) Seeing disorder: Neighborhood stigma and the social construction of "Broken Windows." *Social Psychology Quarterly* 67: 319–342.

Schultz P, Wesley, Jessica M, et al. (2007) The constructive, destructive, and reconstructive power of social norms. *Psychological Science* 18(5): 429–434.

Seneca LA (1928) [first 63] *Moral Essays* (The Loeb Classical Library), vol. 1 (Trans. JW Basore). London: W. Heinemann (1928–1935, 3 vols). Available at: http://www.stoics.com/seneca_essays_book_1.html#MERCY1

Shalvi S, Dana J, Handgraaf M, et al. (2011) Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes* 115(2): 181–190.

Spiegel (1980) Vielköpfige Hydra. *Der Spiegel*, 3 November, pp. 42–46.

Thackeray WM (1869) On being found out. London, pp. 125–132. Available at: http://www.online-literature.com/thackeray/1327/

Thöni C and Gächter S (2015) Peer effects and social preferences in voluntary cooperation: A theoretical and experimental analysis. *Journal of Economic Psychology* 48(0): 72–88.

Warner SL (1965) Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association* 60: 63–69.

Wilson JQ and Kelling GL (1982) The police and neighborhood safety: Broken windows. *Atlantic Monthly* 127: 29–38.

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

Appendix 1

Regression models for ignorance hypothesis

Table 3. Regression models on the preventive effects of ignorance.

	OLS 1		OLS 2		Logit	
	(Payoff 2 − payoff 1)		(Payoff 2 − payoff 1)		(Payoff = 5)	
	coef.	s.e.	coef.	s.e.	coef.	s.e.
First throw					(ref.)	
Control group (info 0)	(ref.)		−0.234	0.138	−0.054	0.167
Experimental group (info 1)	0.487*	0.195	0.253	0.135	0.243	0.206
Experimental group (info 2)	0.501*	0.189	0.267	0.177	0.360*	0.182
Intercept	−0.234	0.138			−0.601*	0.105
N_1	466		466		932	
N_2	30		30		30	
Adj. R^2	0.01		0.01			
Pseudo R^2					0.004	
χ^2					5.827	

OLS: ordinary least square; s.e.: standard error; DV: dependent variable.
The table lists coefficient estimates with cluster-robust standard errors from OLS and logistic regression models (*$p < 0.05$, for two-sided tests). The DV in model OLS 1 is the difference between the first and the second reported payoffs. The control condition is the reference category so that the coefficients estimate the differences between the treatment conditions and the control. OLS 2 is estimated without intercept. Coefficients test whether reported payoff differences are different from 0. In model logit, the binary DV is 1 if the subject reported the maximum payoff of 5, and 0 otherwise. The reference category is the first throw. This means that the coefficients estimate the difference between first and second throws for each experimental condition. A positive coefficient indicates an increased likelihood to report a five in the second throw and a negative coefficient indicates the opposite.

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

Appendix 2

Distributions of payoff combinations by experimental groups

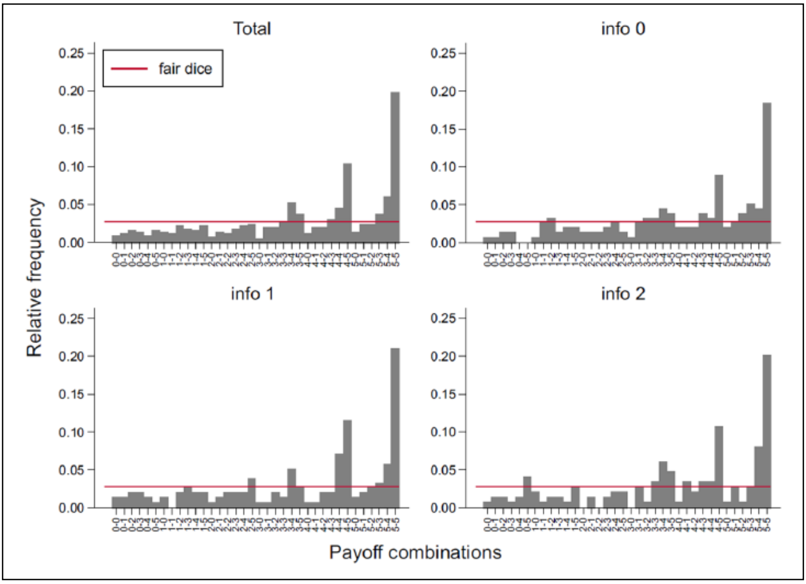

Figure 7. Payoff combinations.

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

Appendix 3

Screenshots of the experiment

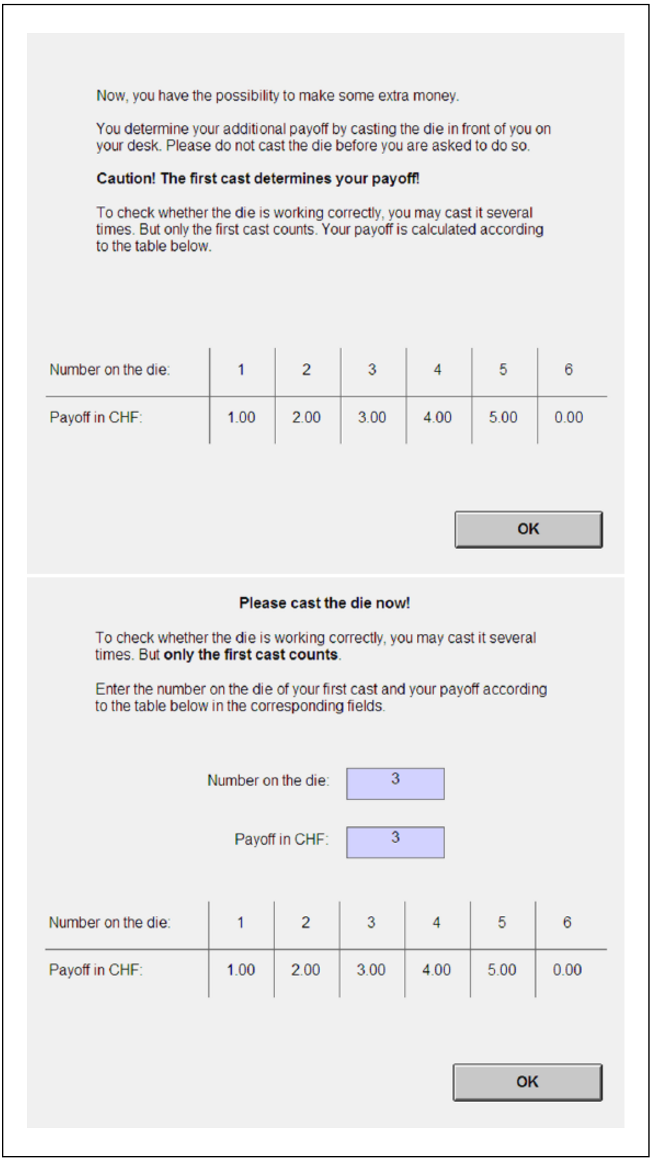

Figure 8. Instructions of the experiment (top) and the report of the first throw (bottom).

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

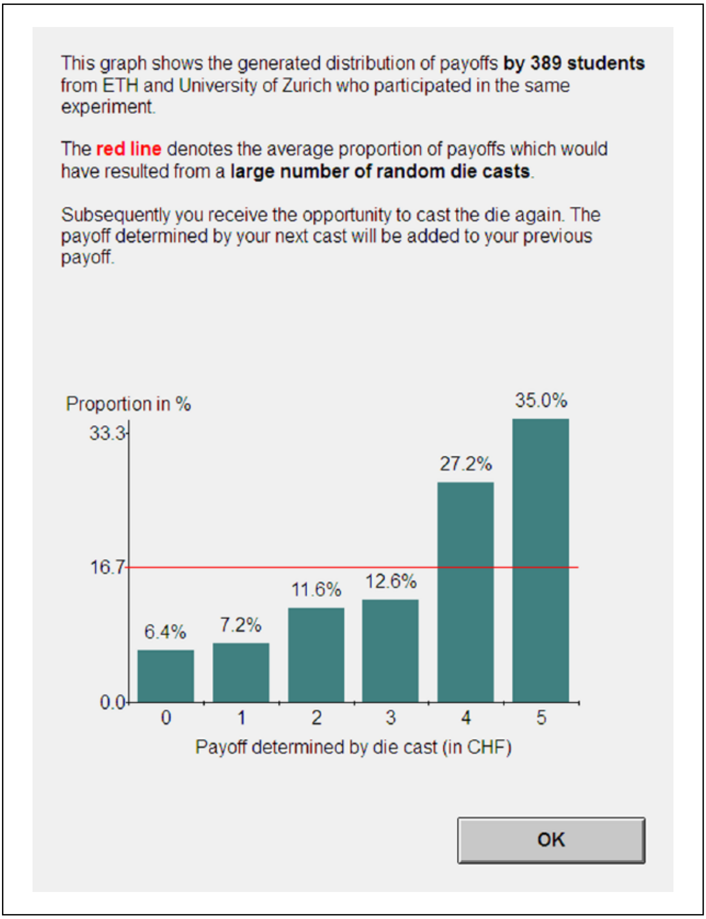

Figure 9. Screenshot of information treatment 1 (info 1) and the shown distribution of the experiment from Fischbacher and Föllmi-Heusi (2013).

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015

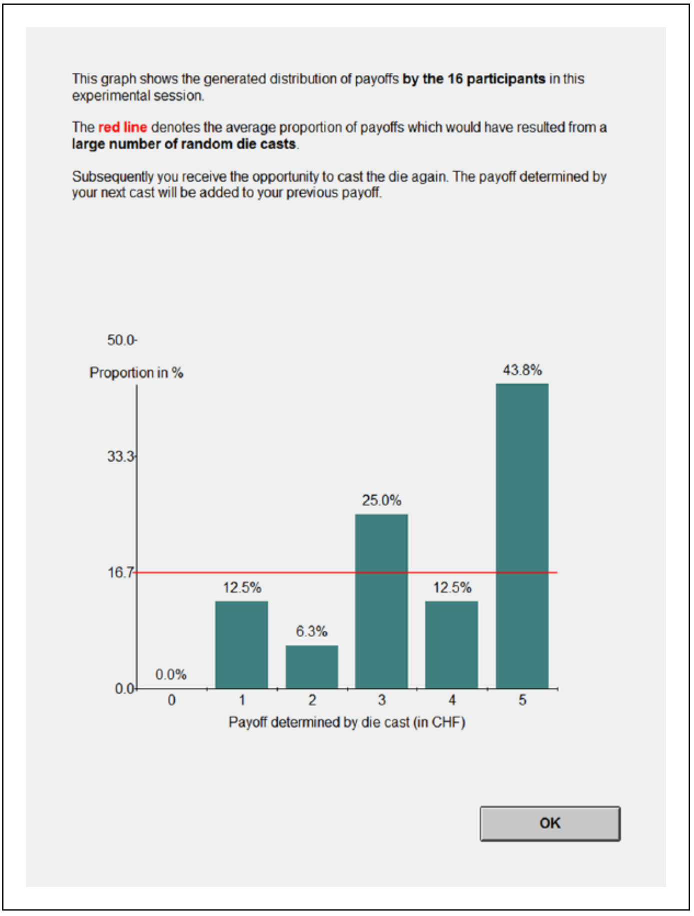

Figure 10. Screenshot of information treatment 2 (info 2) based on the exemplary distribution in session 17.

Downloaded from rss.sagepub.com at University Library Utrecht on December 16, 2015