Preprint

This is the submitted version of a paper published in *Organizational Behavior and Human Decision Processes*.

Citation for the original published paper (version of record):

Bidirectional associations between descriptive and injunctive norms.
*Organizational Behavior and Human Decision Processes*, 127: 59-69
http://dx.doi.org/10.1016/j.obhdp.2014.09.011

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:mdh:diva-28180
The naturalistic fallacy intuition: Injunctive norms inferred from descriptive norms

Running head: The naturalistic fallacy intuition

Kimmo Eriksson$^{1,2}$, Pontus Strimling$^1$ and Julie C. Coultas$^{1,3}$

$^1$Stockholm University, Sweden
$^2$Mälardalen University, Västerås, Sweden
$^3$University of Sussex, Brighton, UK

Correspondence should be addressed to Kimmo Eriksson, Centre for the Study of Cultural Evolution, Stockholm University, Wallenberglab., SE-10691 Stockholm, Sweden (e-mail: kimmo.eriksson@mdh.se).

Acknowledgements
This work was supported by the Swedish Research Council grants 2009-2390 and 2009-2678. We are grateful to Jonathan Haidt and Robert Cialdini for encouragement and helpful comments.
The naturalistic fallacy intuition: Injunctive norms inferred from descriptive norms

Abstract
According to social intuitionist research, moral (or “injunctive”) norms are often not rationally motivated. Where do these norms come from then? We propose that one way in which injunctive norms emerge is through inference from beliefs or information about how people typically behave (i.e., from “descriptive” norms). Specifically, we hypothesize that such inferences occur in accordance with Hume's naturalistic fallacy, such that if a behavior is known to be common it is interpreted as being prescribed and if the behavior is known to be uncommon it is interpreted as being condemned. This hypothesis was supported in a series of five experimental studies demonstrating how expectations of injunctive norms, recall of injunctive norms, and own moral judgments are all affected by manipulation of descriptive norms.

Keywords
descriptive norms; injunctive norms; naturalistic fallacy; priming; recall; moral judgments
Introduction

The existence of social norms that govern behavior seem to be a key feature of all human societies and constitute an important difference between humans and other animals. A number of different features of the human psychology seem to be involved in sustaining norms. For instance, people seem to be particularly good at recalling information about norms (O’Gorman, Wilson, & Miller, 2008) and they tend to intuitively understand the logic of normative rules (Cosmides, 1989). People are willing to use social control against norm breakers (Brauer & Chaurand, 2010). They are also sensitive to social control and tend to adapt their behavior to avoid it (Cialdini, 1993). In short, a good deal is understood about what makes humans capable of sustaining norms of behavior. Our interest instead lies in the question how morally loaded norms emerge and how they can change. A simple idea is that norms emerge and change for rational reasons, with the purpose of regulating behavior that needs to be regulated. However, this idea may have limited explanatory power. According to the social intuitionist research program, moral judgments are often not rationally motivated but intuitive (Haidt, 2001). Thus, to understand the emergence of a new morally loaded norm we may have to look at the mechanisms whereby moral intuitions are created. Here we propose one such mechanism: the naturalistic fallacy intuition.

The naturalistic fallacy intuition

David Hume famously pointed out that what is and what ought to be are not logically connected to each other. Deriving an ‘ought’ from an ‘is’ has become known as the naturalistic
fallacy.\textsuperscript{1} We believe the tendency to commit this fallacy is deeply ingrained in the human psychology. Specifically, we propose that a naturalistic fallacy intuition is at work in the cognition of norms, such that a behavior being very common is intuitively interpreted as the behavior being morally right. Similarly, if a behavior is very uncommon it is intuitively interpreted as being morally questionable. These intuitions, we propose, affect both people’s own moral judgments and their expectations of others’ judgments.

The naturalistic fallacy intuition can be discussed in terms of the distinction between injunctive and descriptive norms (Cialdini, Reno, & Kallgren, 1990; Reno, Cialdini, & Kallgren, 1993). Descriptive norms indicate what is the behavior of most people, i.e., they refer to the frequency with which given behaviors occur. Injunctive norms indicate how people ought to behave, i.e., they refer to moral values of right or wrong, what is socially obligated or socially unacceptable. In these terms, the naturalistic fallacy intuition can be described as a general tendency to infer injunctive norms from descriptive norms.

One famous study of descriptive and injunctive norms concerned alcohol use in college fraternities (Borsari & Carey, 2001). Alcohol use was common in these fraternities (i.e., a descriptive norm). There was also peer pressure (i.e., an injunctive norm) to use alcohol. Fraternity members used both types of norms to justify their personal alcohol use. In general, whether descriptive or injunctive norms exert the strongest influence on people's behavior seems to be determined by the which norm is salient (Cialdini, 2003; Cialdini et al.,1990; Reno et al., 1993; Kallgren, Reno, & Cialdini, 2000). But why is the descriptive norm influential at all? Naturalistic fallacy intuitions would provide one explanation, namely, that if you perceive that most people drink alcohol you intuitively infer that there is a social obligation to do so.

\textsuperscript{1} The term \textit{naturalistic fallacy} was coined by G.E. Moore and has been used in several ways, see Curry (2006).
Injunctive norms can be subdivided into *prescriptive* norms about what you are obliged to do and *proscriptive* norms about what you should not do (Janoff-Bulman, Sheikh, & Hepp, 2009). Again, consider norms about alcohol use. Whereas college students perceived a prescriptive norm about drinking alcohol in their fraternities, there are obviously proscriptive norms against the use of alcohol in many other contexts and communities. We expect the naturalistic fallacy intuition to apply both to proscriptive and prescriptive norms.

We have found surprisingly little prior research on how moral judgments of behaviors depend on how commonly the behaviors occur. The few studies we have found are consistent with our hypothesis. For instance, a large cross-cultural survey on uncivil behaviors showed a clear correlation such that behaviors were regarded as more deviant the rarer they were perceived to be (Brauer & Chaurand, 2010). Thus, although a connection between descriptive and injunctive norms is not a logical necessity, it seems there tend to be a connection anyway. Similarly, Thøgersen (2008) found a strongly positive correlation between beliefs about descriptive and injunctive norms for environmentally responsible behaviors. Conflict between injunctive and descriptive norm about energy conservation has been found to weaken behavioral intentions (Smith, Louis, Terry, Greenaway, Clarke, & Cheng, 2012). Finally, a study of tax evasion found that the more prevalent an individual perceives tax evasion to be within the community, the less likely the individual will be to judge the act harshly or fear informal sanctions directed against it (Welch, Xu, Bjarnason, Petee, O’Donnell, & Magro, 2005). These authors called this process “defining deviancy down”, conceiving of it as a breakdown of proscriptive norms. Our conception of the naturalistic fallacy intuition is more general and may serve to strengthen norms as well: perception of a proscribed behavior as uncommon should reinforce the proscriptive norm, and similarly for perception of a prescribed behavior as common.
Summary of studies

In this paper we report five studies designed to demonstrate that people miscode information about frequency as information about the moral standing of a behavior and, furthermore, that this coding affects the judgment of people who use, or do not use, the focal behavior.

Study 1 is designed to demonstrate that people connect descriptive and injunctive norms. If participants are told that a behavior is common in a strange country, we predict their intuition to be that the behavior is also supported by moral norms (i.e., prescribed if it is a generally prosocial behavior, or not proscribed if it is a generally antisocial behavior).

The next two studies use recall of information about norms to show that the intuitive connection between descriptive and injunctive norms is present already in how norms are encoded in people’s minds. Study 2 aims to demonstrate that information about descriptive norms is often recalled as information about injunctive norms. If participants are told about various descriptive and injunctive norms in a strange country, we predict that in a subsequent recall task descriptive norms will often be recalled as the injunctive norms to which they correspond according to the naturalistic fallacy intuition (e.g., a prosocial behavior should be likely to be recalled as obligated if it is presented as common, and likely to be recalled as not obligated if it is presented as uncommon). The third study gets at the same issue by a different method, through investigation of whether information about descriptive norms interferes with people’s recall of information about injunctive norms for the same behaviors. If participants are told about a strange country where a certain behavior is either supported or unsupported by moral norms and either common or uncommon, they are more likely to recall the moral norm incorrectly when it is inconsistent with the naturalistic fallacy intuition (e.g., because the combination that a behavior is both wrong and common is inconsistent with the
naturalistic fallacy, the injunctive norm of a behavior presented in this way should be likely to be incorrectly recalled as OK instead of wrong).

Finally we conduct two studies to show that naturalistic fallacy intuitions extend to people’s own moral judgments. This is very important from the point of view of the dynamics of norms. For an injunctive norm to be established, it is necessary both that people expect others to endorse the norm and that the norm is actually supported by moral judgments.

Study 4 aims to demonstrate that information about descriptive norms affects people's moral judgments in accordance with naturalistic fallacy intuitions. If participants are told about a novel situation and that a certain prosocial behavior is either common or uncommon in this situation, their judgment of the moral obligation should vary accordingly. The fifth and final study is designed to show that this effect is automatic. Specifically, we predict that moral judgments will be affected even when information about whether or not selfish behavior is common is not given explicitly but only through priming in an unrelated task.

All studies were carried out online with American participants recruited through the Amazon Mechanical Turk (https://www.mturk.com). As others have noted, Mturk provides a convenient and reliable data source for behavioral studies that can be conducted online (Buhrmester et al., 2011; Paolacci, Chandler & Ipeirotis, 2010). However, for research involving multiple studies it is a potential problem that the same individual may take part in several studies. Participants who had already taken part in a previous study in this project were identified by the user ID number and excluded. Thus, in the data presented in this paper, all studies have unique sets of participants.
Study 1

The first study was a survey to assess participants' conscious intuitions about the connection between injunctive and descriptive norms. As discussed in the introduction, we are interested in intuitions both about prescriptive norms and proscriptive norms.

Method

Participants. Two hundred participants (35% female) with age ranging from 18 to 74 years (M=30 years, SD=11 years) were recruited online among American users of Mturk at a compensation of half a US dollar.

Materials. Participants were asked to imagine that they traveled to a new place where customs may be different from what they are used to. The questionnaire introduced two behaviors called “phooshing” and “quining”. They were described as a generic antisocial behavior (“other people might appreciate if you don’t do it”) and a generic prosocial behavior (“other people might appreciate if you did it”), respectively. Participants were first asked what would be most important for them to know about the new place: either the actual frequencies of these behaviors or the general opinions about whether you ought or ought not to behave in these ways. This question served the purpose of emphasizing the conceptual difference between descriptive and injunctive norms.

To assess participants’ intuitions about the relationship between the strength of injunctive and descriptive norms, the following two questions were asked in two different ways depending on condition (N=50 participants in each of 2×2=4 conditions). In one condition the first question stated the assumption that "phooshing"—the antisocial behavior—is common, and asked whether the general opinion is most likely to be that one ought not to do it or that it is up to you whether you want to do it. Similarly, the second question stated the assumption that "quining"—the prosocial behavior—is uncommon, and asked whether the
general opinion most likely is that one ought to do it or that it is up to you whether you want to do it. The four conditions together presented all four combinations of assumptions about the frequencies (common or uncommon) of phooshing and quining.

**Results**

For the overwhelming majority of participants, intuitions followed the naturalistic fallacy. For a prosocial behavior, the injunctive norm was thought most likely to be prescriptive if the behavior was common (67%, $p=.001$, binomial test), but not prescriptive if the behavior was uncommon (80%, $p<.001$). For an antisocial behavior, the analogous result was obtained: The injunctive norm was thought most likely to be prescriptive if the behavior was uncommon (80%, $p<.001$), but not prescriptive if the behavior was common (85%, $p<.001$). See Figure 1.

**FIG. 1 ABOUT HERE**

**Discussion**

This simple study confirmed that people’s conscious intuitions about the relationship between descriptive and injunctive norms tend to be consistent with the naturalistic fallacy.

**Study 2**

The previous study showed that people expect injunctive and descriptive norms to be connected in a systematic way. Our next aim is to use recall tasks to establish that this connection is present already in the encoding of norms.

In this study prosocial behaviors were presented together with manipulated information about norms in a strange land. Some behaviors were presented as common, some were presented as uncommon, and yet others presented as obligated or not obligated. Participants were later asked to recall the items on a multiple-choice format. Based on our hypothesis about the naturalistic fallacy
intuition, we predicted items presented as ‘common’ to be recalled as ‘obligated’ more often than other incorrect responses. Similarly, we predicted items presented as ‘uncommon’ often to be recalled as ‘not obligated’. These two kinds of errors are henceforth referred to as naturalistic fallacy errors. We will refer to the other possible kinds of errors as opposite naturalistic fallacy (e.g., to recall ‘common’ as ‘not obligated’) and opposite descriptive (e.g., to recall ‘common’ as ‘uncommon’). The naturalistic fallacy intuition would favor the naturalistic fallacy error, whereas complete memory failure would not be expected to favor any particular kind of error. Thus, we expect the dominance of naturalistic fallacy errors to be particularly strong among items where complete memory failure can be ruled out. To test this prediction, participants were cued to recall the list of behaviors before the multiple-choice recall task. Among items where cued recall of the behavior was successful, we therefore expected subsequent norm recall errors to be particularly strongly dominated by naturalistic fallacy errors.

Method

Participants. Participants were recruited online among US users of the Amazon Mechanical Turk at a compensation of one US dollar. There were 40 participants (30% female) with age ranging from 18 to 58 years (M=32 years, SD=12 years).

Materials, Design and Procedure. The survey consisted of three parts. The first part ostensibly presented an excerpt from a guidebook to some unknown country. This excerpt listed 16 prosocial behaviors adapted from Janoff-Bulman et al. (2009). For example, one item read *It is common [uncommon/obligated/not obligated] to keep your garden looking proper and well-kept.* Among different versions of the survey, each item was presented equally often as ‘uncommon’, ‘common’, ‘obligated’, and ‘not obligated’. Within participants, each of these four norm descriptions was used for exactly one quarter of the items (i.e., four items out of sixteen).
Participants were told that the task was to guess which country the excerpted guidebook was describing and therefore encouraged to read the excerpt carefully. After guessing on a country, participants were asked to write a short essay about their general experiences with guidebooks. This served as a distracter task before the subsequent recall tasks.

In a second part of the survey, participants were cued to recall as many as possible of the behaviors presented in the previous part. This task was explicitly about recalling behaviors only; participants were told not to bother with recalling the norms. Finally, in the third part, participants were presented with a list of 27 behaviors. This list included all of the 16 original behaviors, reordered and interlaced with 11 new behaviors from the same domain. For each behavior, participants were asked a multiple-choice question about how the behavior had been presented in the first part: as common, uncommon, obligated, not obligated, or not mentioned at all.

**Coding.** For each participant we were interested only in the eight items that had been presented as ‘common’ or ‘uncommon’ (i.e., descriptive norms). In the cued recall task, each of these items was binary coded as either successfully recalled by the participant (1) or not (0). The coder was unaware of the hypothesis. As a reliability test, a second coder recoded 25% of the data; codings were 98% identical.

In the multiple-choice recall task, any errors were categorized as follows: *opposite descriptive* (‘common’ recalled as ‘uncommon’ or vice versa), *naturalistic fallacy* (‘common’ recalled as ‘obligated’ or ‘uncommon’ recalled as ‘not obligated’), *opposite naturalistic fallacy* (‘common’ recalled as ‘not obligated’ or ‘uncommon’ recalled as ‘obligated’), and *not mentioned before*.

**Results**

The total error rate in the subsequent norm recall task was 33% among all items that had been successfully recalled in the behavioral recall task, compared to a total error rate of 51% among the remaining items. Our interest lies in how the errors in the norm recall task were
distributed among the four error categories. As shown in Fig. 2, the dominating error category was the naturalistic fallacy error, and this pattern was particularly strong among items where behavior recall had been successful. We conducted an ANOVA of the number of norm recall errors with category (omitting the ‘not mentioned’ category) and behavior recall success as within-subject factors. This analysis confirmed that the difference between categories was unlikely to occur by chance, $F(2,38)=28.7, p<.001$, and similarly for the interaction between category and behavior recall success, $F(2,38)=8.6, p=.001$.

**Fig. 2 ABOUT HERE**

**Discussion**

This study provided evidence that recall of information about descriptive norms is affected by the naturalistic fallacy intuition. Behaviors presented as being common or uncommon in a foreign country were often recalled as having been presented as obligated or not obligated, respectively. This finding supports the notion that people may encode information about the frequency of a behavior in injunctive terms.

**Study 3**

The idea that descriptive norms may be encoded as injunctive norms also predicts recall of information about injunctive norms to be sensitive to interference from information about descriptive norms. Specifically, information about an injunctive norm should be recalled incorrectly more often when inconsistent with naturalistic fallacy intuitions derived from information about the descriptive norm for the same behavior.
Method

Participants. Participants were recruited online among US users of the Amazon Mechanical Turk at a compensation of one US dollar. There were 100 participants (55% female) with age ranging from 18 to 69 years (M=33 years, SD=12 years).

Materials, Design and Procedure. Participants completed an online survey in three parts. As described in detail below, the first two parts presented manipulated information about norms in a foreign country whereas the last part was a recall task. Each part dealt with the same set of sixteen behaviors taken from the Moralism scale (Janoff-Bulman et al., 2009). Of the 16 items, half were taken from the “prescriptive domain” of generally prosocial behaviors and half from the “proscriptive domain” of generally antisocial behaviors. The following are examples of an item from each domain.

Prescriptive: Cory is in the supermarket, where he sees an elderly woman having trouble carrying her groceries. He is in a hurry and knows he could ignore her, but considers instead helping the elderly woman carry her groceries.

Proscriptive: Melanie and Scott have just bought a house in a quiet, middle-class neighborhood. The homes are not fancy, but are modest and well-kept. Melanie and Scott are considering ignoring the community and painting their house bright orange with green trim.

First part: injunctive norms. The first part presented an injunctive norm for each behavior. Specifically, participants were told that respondents in a foreign country had made moral judgments of the behaviors. For each behavior they were told that a certain moral judgment had been typical in the foreign group (either OK or wrong for proscriptive items; either obligated or up to you for prescriptive items). In reality, judgments of each behavior were manipulated such that half of the participants were told one judgment was typical (say, ‘OK’), and the other half was told the opposite judgment was typical (say, ‘wrong’). Participants were then asked whether they agreed with the judgment of the foreign group. For instance, in this part of the survey the proscriptive item
above continued: “To paint one's house bright orange with green trim in this situation was generally considered OK. Do you agree with this judgment?” Half of the participants read that it was ‘wrong’ instead of ‘OK’.

Second part: descriptive norms. The second wave presented the same behaviors with information about descriptive norms. Specifically, participants were told that the foreign respondents had also estimated how frequently the behaviors occur. For each behavior they were told that a certain frequency estimate had been typical in the foreign group (either common or uncommon). Participants were asked whether they agreed with this estimate.² For instance, in this part of the survey the prescriptive item above continued: “To help the elderly woman with her groceries in this situation was generally considered to be a common behavior. Do you agree with this estimation?” Similar to the first part, the information was manipulated such that half of the participants were told the typical estimate was ‘common’ and the other half was told the typical estimate was ‘uncommon’. Importantly, the manipulations in the two parts were independent. That is, any given behavior was presented equally often with norms consistent with the naturalistic fallacy (e.g., wrong+uncommon or OK+common) and with norms inconsistent with the naturalistic fallacy (e.g., wrong+common or OK+uncommon). Within participants, exactly half of the items (i.e., eight items out of sixteen, four in each domain) were presented with information about norms consistent with the naturalistic fallacy.

Third part: recall. In the final part, participants were asked to recall for each item the typical moral judgment of the Swedish group as it was presented in the first part. The recall task used a binary choice. Depending on the domain of the item, the choice was either between ‘OK’ and ‘wrong’ or between ‘obligated’ and ‘up to you’.

² The questions about whether participants agreed with moral judgments and frequency estimates served double purposes: forcing participants to pay attention to the manipulated information, as well as yielding data on perceptions of injunctive and descriptive norms for a set of specific behaviors.
Results

The study was designed to test the prediction that recall of information about injunctive norms is worse, regardless of domain, if information about descriptive norms is inconsistent with the naturalistic fallacy. Thus, the dependent measure was the error rate, i.e., the proportion of items that were incorrectly recalled by a participant. We conducted an ANOVA with two within-subject factors: \textit{inconsistency} (2 levels: item presented either consistent or inconsistent with the naturalistic fallacy) and \textit{domain} (2 levels: prescriptive or proscriptive). As predicted, there was a main effect of inconsistency, $F(1,99)=42.94$, $p<.001$, such that the error rate was higher for inconsistent items than for consistent items (.35 vs. .20). There was a main effect of domain, $F(1,99)=4.92$, $p=.03$, such that the error rate was slightly higher for items in the prescriptive domain than for items in the proscriptive domain items (0.29 vs. 0.25). As expected, there was no interaction between inconsistency and domain, $F(1,99)=0.00$, $p=1.00$. See Fig. 3.

\textbf{FIG. 3 ABOUT HERE}

From participants’ responses on whether or not they agreed with the presented estimates and judgments we inferred their own perceptions of descriptive and injunctive norms (e.g., the moral judgment of a participant who was presented with the judgment ‘OK’ was coded as ‘OK’ if the participant agreed, and coded as ‘wrong’ if the participant did not agree). Fig. 4 shows how the proportion of participants who judged a behavior as OK/obligated tended to correlate with the proportion of participants who estimated it to be common. One single outlier (marked by a star in Fig. 4) does not follow this pattern, namely the proscriptive item about \textit{painting one’s house bright orange with green trim}, which tended to be judged as OK but uncommon. Among the other 15 data points, the correlation between injunctive and descriptive norms was as high as $r=.76$.

\textbf{FIG. 4 ABOUT HERE}
Discussion

In this study we found that when information about the descriptive norm for a behavior is presented, it interferes with recall of previously presented information about the injunctive norm for the same behavior. Error rates in recall were almost twice as high when the naturalistic fallacy intuition would give the incorrect answer. This finding provides additional support for the hypothesis that the naturalistic fallacy intuition works already at encoding of descriptive norms.

Study 4

As explained in the introduction, the naturalistic fallacy intuition can be an important mechanism in the emergence of norms only if people both tend to expect a moral norm to exist (because of biased encoding, as in the previous studies) and tend to support it by their own moral judgments. The fourth study investigated whether information about descriptive norms would affect people's moral judgment in a novel situation. Specifically, we tested moral judgments of behavior in the so called dictator game, an artificial situation only used in economic lab studies and therefore likely to be novel to participants. To control for individual baseline differences in strictness of moral judgments, we also included a related economic game as a control scenario.

Method

Participants. Two hundred participants (33% female) with age ranging from 18 to 75 years (M=31 years, SD=12 years) were recruited online among American users of Mturk at a compensation of 0.30 US dollar.

Materials. Participants were presented with two scenarios describing economic lab experiments that we had ostensibly ran in our lab. The first scenario presented a version of the investment game: “Two participants, anonymous to each other, are in different rooms in the
lab. The experimenter gives 10 dollars to the first participant with the following instructions: 

*You have two options. Either (a) you keep this money (10 dollars) or (b) you ‘invest the money’, in which case I'll take them and give the second participant 30 dollars. The second participant will then have to decide how much to give back to you: Either (b1) give back 10 dollars to you, and keep 20 dollars, or (b2) give back 20 dollars to you, and keep 10 dollars.*

When we ran this experiment, almost all first participants made the choice to ‘invest the money’. In the present survey we are interested in your thoughts about the moral norms in the situation of the second participants. How strong is the moral wrongness of keeping 20 dollars and give back only 10 dollars to the first participant?" Responses were given on a four-step Likert type scale anchored in 0=*not wrong at all* and 3=*very wrong.*

The first scenario served to provide an individual baseline in moral judgments. The manipulation of descriptive norms came in the second scenario. This scenario presented a version of the dictator game, along with the information that in an experiment either 10 percent or 90 percent (N=50 participants in each condition) chose to give up money: “Two participants, anonymous to each other, are in different rooms in the lab. The experimenter gives 20 dollars to one of the participants with the following instructions: *You have two options. Either (a) you keep this money (20 dollars), or (b) you give it back to me, in which case I'll give you 5 dollars instead and I'll give 15 dollars to the other participant (who will otherwise get nothing).* When we ran this experiment, about 10 [90] percent chose to give the money back (i.e., take 5 dollars instead of 20 dollars, so that the other participant would get 15 dollars). Thus, about 90 [10] percent chose to keep the 20 dollars (so that the other participant got nothing). In the present survey we are interested in your thoughts about the moral norms in this situation. How strong is the moral obligation to give the money back, as about 10 [90] percent did (and 90 [10] percent did not do)?” Responses were given on a four-step Likert type scale anchored in 0=*no obligation at all* and 3=*very strong obligation.*
**Results**

As predicted, moral judgments of the obligation to give money in the dictator game were stricter in the condition where giving was said to be common (M=1.21, SD=1.09) than in the condition where giving was said to be uncommon (M=0.86, SD=0.88), t(198)=2.51, p=.013. Inclusion of the baseline moral judgment (i.e., the judgment of the investment scenario) as a control variable only strengthened the effect, as reported in Table 1 (left column).

**Discussion**

When told that most people choose to give in the dictator game, our participants tended to judge the moral obligation to give as stronger than when told that most people choose not to give. Thus it seems that when people are presented with novel situations their moral judgment depends on their belief about the descriptive norms in these situations, as predicted by the naturalistic fallacy intuition.

**Study 5**

The previous study showed that people care about information about descriptive norms when forming moral judgments about prosocial behavior in a novel situation. The aim of the fifth study is to demonstrate that this is an automatic effect that can be elicited just by priming descriptive norms about unselfish behavior in a seemingly unrelated word scrambling task.

**Method**

**Participants.** Eighty participants (29% female) with age ranging from 18 to 57 years (M=26 years, SD=7 years) were recruited online among American users of Mturk at a compensation of half a US dollar.
**Materials.** Exactly as in Study 4, the survey asked for moral judgments of behavior in the investment game (to obtain a baseline) and the dictator game. Only the manipulation differed. Instead of presenting a descriptive norm in the dictator game, a seemingly unrelated “scrambled sentence test of verbal creativity” was inserted before the dictator game. The scrambled sentence test is a well-known priming technique (e.g., Bargh, Chen, & Burrows, 1996). It presents a number of word sets, for which the task is to form sentences by dropping one word and reordering the remaining words of the set. Our test consisted of ten sets of words. Whereas many of the words suggested a theme of animals and sports, four out of ten sentences to be formed stated descriptive norms about selfishness. In each of these four word sets, one word differed between a “selfishness is common” condition and a “selfishness is uncommon” condition (N=40 participants in each condition).³

**Results**

Results followed the same pattern as in Study 4. Moral judgments of the obligation to give money in the dictator game were stricter in the condition priming that unselfishness is common (M=1.30, SD=0.82) than in the condition priming that unselfishness is uncommon (M=0.88, SD=0.97), t(78)=2.12, p=.037. Again, the effect was even clearer when the baseline moral judgment was included as a control variable, see Table 1 (right column).

**Discussion**

Our participants tended to judge the moral obligation to give in the dictator game as stronger when they were primed that most people tend to be unselfish rather than selfish. This finding supports the notion that the naturalistic fallacy intuition serves as an automatic basis for moral judgments.

³ The four key sentences to be unscrambled were: “few [many] people are selfish”; “everybody [nobody] turned up to help”; “greed is really uncommon [common]”; “most [few] Americans make donations.”
General discussion

In the introduction we argued that whereas the psychology that sustains norms is pretty well understood, little is known about the mechanisms whereby norms arise or change. We proposed that morally charged norms can arise simply because behaviors happen to be common or uncommon. Specifically, we hypothesized a naturalistic fallacy intuition in the sense that people will make automatic inferences of moral judgments from information about the frequency of behaviors. We then presented five experiments which indeed suggest that information about frequency of behaviors tend to be encoded in injunctive terms and automatically used when moral judgments are formed. In other words, there is a tendency for common behaviors to be thought of as things you ought to do and that those who don’t are judged to violate a moral code. We therefore conclude that some injunctive norms could have originated from behaviors that where common in the population and then went through the prism of the naturalistic fallacy intuition, so that people stopped thinking of them just as common behaviors and started to think of them as morally correct behaviors and punish the rare people who behaved differently. One example could be norms about which hand to hold the fork in. It is plausible that when food was tough to cut it became common to use the right hand for the knife, as this is the stronger hand for most people. Once it was common, this behavior might have started to be thought of as normatively right, so that using the left hand became a break of etiquette. Of course, there may be several other psychological mechanisms, besides the naturalistic fallacy intuition, that contribute to the emergence of injunctive norms. A long term goal of our research is to map all such mechanisms.

It is noteworthy that the naturalistic fallacy intuition would also work to sustain norms once established. Societies where corruption is endemic might be one example. It is notoriously difficult to eradicate corruption in such societies (Persson, Rothstein, & Teorell, in press). One reason might be that when most people are corrupt, campaigns that plead to people’s moral judgment will be in vain because the intuition will be that corrupt behavior is not very wrong.
According to the social intuitionist model, moral judgment is the result of quick, automatic evaluations (Haidt, 2001). Our finding contributes to the social intuitionist literature by introducing perceptions of descriptive norms as one source of social intuitions. Future research may dig deeper in this issue. For instance, how is the naturalistic fallacy intuition related to moral foundations such as authority and purity (Haidt & Graham, 2007)?

As we mentioned briefly in the introduction, the naturalistic fallacy intuition also offers a novel reason why descriptive norms influence people's behavior (for a recent review of research on conformity and social influence, see Claidière and Whiten, 2012). The naturalistic fallacy intuition would lead people to expect, without any other cues, that majority behavior is indeed socially obligated.

Finally, we note that the naturalistic fallacy intuition can be seen as an intuitive justification of the status quo of behavioral patterns. As such, it might also be integrated into the theoretical framework on mechanisms behind such justification (Jost, Banaji & Nosek, 2004).
References


Table 1. Linear regressions predicting moral judgments of the obligation to give in a dictator game in Studies 4 and 5.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Study 4</th>
<th>Study 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>.18**</td>
<td>.27**</td>
</tr>
<tr>
<td>Baseline moral judgment</td>
<td>.25***</td>
<td>.41***</td>
</tr>
</tbody>
</table>

**: p<.01, ***: p<.001.

Note. Entries are standardized coefficients. Condition is dummy coded 1=‘unselfishness is common’, 0=‘unselfishness is uncommon’.
Fig. 1. Inferences of injunctive norms from descriptive norms in Study 1.
Fig. 2. Distribution of categories of norm recall errors in Study 2, for items for which behavior recall was successful (left diagram) or unsuccessful (right diagram), respectively.
Fig. 3. Error rates for items in the prescriptive and proscriptive domains depending on whether they were presented with norms consistent or inconsistent with the naturalistic fallacy (NF). Error bars: ±1 SE.
Fig. 4. Scatter plot of moral judgments and frequency estimates for the 16 behaviors in Study 3. An outlier is starred.