Replication study

Conformity to the descriptive norms of people with opposing social or political beliefs

Niko Britt, Jens Huth, Thomas Nortmann, Simon Roske

University of Osnabrück

Abstract

The effect of descriptive norms influences the decision-making across subjects and situations. According to self-categorization theory, the descriptive norm is impacted by the group the subject feels most included in. Because ingroup-norms and outgroup-norms may have conflicting directions, the descriptive norm effect can take on both positions ending in an inconsistent theory. Researchers who investigated this question—which they believed is complex and not fully understood yet—obtained evidence showing that self-categorization cannot account for the effect of descriptive norms. The purpose of this paper is to re-examine this finding. By replicating the experiment, we obtained similar results, however not as strong as reported in the first place.

Introduction

This study is designed to replicate the paper *Conformity to the descriptive norms of people with opposing political or social beliefs* by Campbell Pryor, Amy Perfors, and Piers D. L. Howe published in 2019. Although the authors conducted two single experiments in their paper, we are only replicating the first one here. In their study, they investigated the phenomenon of descriptive norms with respect to the self-categorization theory, which hypothesizes that people behave opposite to outgroup norms, against the alternative hypothesis: People conform to the overall descriptive norm. Descriptive norms are understood as an individual's tendency to conform to the behavior popular in a reference group under similar circumstances. People choose or act a certain way when they observe other people doing the same. This effect has been observed in several studies using a variety of different selection scenarios—e.g., anti-social behavior (Abbink, Freidin, Gangadharan, Moro, 2018; Köbis, Van Prooijen, Righetti, Van Lange, 2015) or ecological questions (Schultz, Nolan, Cialdini, Goldstein, Griskevicius, 2007).

The theory of self-categorization claims that people are not only influenced by other people's actions or choices, but actively position themselves to conform with the norms of groups they identify with (ingroup norm) while opposing norms of groups they do not identify with (outgroup norm). Thus, this theory proposes that a person's conviction is linked to the norms of social groups that the person identifies with and will adjust to them appropriately (Turner, Hogg, Oakes, Reicher, Wetherell, 1987). The amount of coherence between the ingroup norm and the individual opinion is believed to depend on the degree to which a person identifies with the respective group (see Smith & Terry, 2003; Rimal, 2008).

The findings we aim to replicate and that were reported by Pryor, Perfors and Howe (2019) suggesta stronger tendency to conform to the behavior being more popular overall (alternative hypothesis) independent of the ingroup or outgroup-norm as opposed to always pursuing the ingroup norm. Since the issues in question were controversial political or social dilemmas (e.g. gun control, feminism, climate change etc.) chosen by the participants themselves, the authors are convinced that the results of their study seem to tend to oppose self-categorization theory rather than support it. Thus, they argue that self-categorization theory is not false but incomplete[1]. According to the authors,the theory itself may be less general than earlier thought since it cannot fully account for the results. The participants' choices had shifted away from what they ought to select if that theory was true. This is to a large degree consistent with the findings of Rimal & Real (2003,2005) proposing a richer explanation to justify the effects of norm conformity. It could also be that self-categorization theory applies only under harsher restrictions like for an absolute ingroup-outgroup norm separation. More explanations and ideas are provided in the original paper but will not be

revised here in detail. As standard for a replication study, we attempted to keep the study design as close as possible to the original one, using the same stimuli with as few adjustments as possible.

Method

Participants

Participants were recruited through social media platforms (Facebook, Instagram) as well as a mailing list for cognitive students at the University of Osnabrück and university students at Virginia Tech in the United States. The total number of participants we were able to recruit was 99 (47 female, 39 male, 5 non-binary, μ_{age}= 27.9 years, age range: 13 - 84). No financial compensation or other reward was given to participants. The experiment was carried out in a web browser using _magpie[2] and took between two to three minutes to complete when done properly. The subjects were free to break off the experiment at any time. Besides age and gender, no personal data was collected.

Materials and Procedure

The materials for this study were taken from Pryor, Perfors & Howe (2019) and adjusted such that the social and political issues were adequate and up-to-date. For a full description of the issued themes and statements, see A1.

First, participants were asked to provide basic demographic information including age and gender. After that, participants had to choose the one (out of nine) social-political issue they cared about the most. A controversial statement about that issue was then displayed (e.g. 'Feminism is important and beneficial to modern society' when feminism was chosen) to which the participants had to answer how much they agree or disagree on a scale from minus five (strongly disagree) to plus five (strongly agree). That rating was used to determine the ingroup and outgroup membership. Then, the instructions for the critical trial were presented. The participants were told that this experiment was following up on a study investigating how people feel during moral dilemmata. This was to present participants with a believable story when presenting the descriptive norms. After providing these instructions, the following moral dilemma was presented:

'Imagine you have witnessed a man rob a bank. However, you then saw him do something unexpected with the money. He donated it all to a run-down orphanage

that would benefit greatly from the money. You must decide whether to call the police and report the robber or do nothing and leave the robber alone.'

The choice had to be made on a scale from minus 3 (Definitely call the police and report the robber) to plus 3 (Definitely do nothing and leave the robber alone). Furthermore, all participants were presented with an ingroup-norm favoring one of these two options at an equal likelihood. For example, if feminism was the chosen social issue, the ingroup norm could look like the following: '*approximately 60% of participants who agreed with you about Feminism chose to call the police and report the robber*' or the exact opposite choice. In addition, half of the participants were also presented with an outgroup-norm contrasting the choice presented in the ingroup norm (both norms in a random order). Sticking to the example above, the outgroup norm presented would be: '*approximately 85% of participants who disagreed with you about Feminism chose to do nothing and leave the robber alone.*'

   Our study also stayed true to the original one in the sense that we have a 2 (*INGROUP DESCRIPTIVE NORM*) x 2 (*BOTH NORMS SHOWN*) between-subjects design. The variable *INGROUP DESCRIPTIVE NORM* indicated the ingroup decision displayed during the trial (reporting the robber = -1, leaving the robber alone = 1) and the variable *BOTH NORMS SHOWN* represents whether only the ingroup norm or ingroup and outgroup norms were shown to the participants (ingroup norm = 0, both norms shown = 1). After answering the moral dilemma, the subjects were asked how they felt about their decision. This was done to cement the idea of conducting a study focussing on moral dilemmata and their effects on a person's feelings. That question had to be answered on a scale ranking from minus 3 (I have a very bad feeling) to plus 3 (I have a very good feeling). Afterwards, an attentional trial was shown which checked whether the participants actually understood the preceding trials (see A1). Lastly, participants were questioned about their group identification to test whether they identified with their respective ingroup and did not do so with their respective outgroup. This was determined in two identification trials with the help of Postmes, Haslam and Jans' single-item social identification measure asking how strongly they agree or disagree with a statement in the form "I identify with [ingroup]" or "I identify with [outgroup]" (rating from minus 3 =strongly disagree to plus 3 =strongly agree) where ingroup and outgroup was replaced accordingly (e.g. feminists vs. Anti-Feminism Advocates).

Data preparation

In order to retrieve a usable data format from the _magpie csv file, we applied a python script[3] on the raw data, yielding a cleaned data file.

Only data of subjects who selected the correct option in the attention trials and were not neutral (0 between -5 and 5) towards their chosen issue were used in the analysis. Because of these rules 16 of the 99 participants were excluded. Moreover, we did not exclude participants that spent a long time (ten minutes or more) completing the experiment, as we do not assume this alters the results drastically.

Results

Models

Following the analysis plan of the study we are replicating, we instantiated two models reflecting the hypothesis of self-categorization theory as they are better suited in explaining the observed data. Thus, one of these models incorporates this assumption (self-categorization), while the other model employs the alternative descriptive norm hypothesis.

Both models are instances of a bayesian ordinal logistic model with priors which are compared using the Bayes Factor, yielding the probability of observing the data under the respective models. Both versions assume that the descriptive norm changes the odds of making higher (preferring to not report the robber) or lower responses (preferring to report the robber) on an ordinal scale. The ordinal logistic regression model predicts the distribution of these responses. The variables are parameterized with respect to the natural log odds in favor of a higher response as presented here:

log_e *(odds of responding higher)* $= b_{in}I + b_{both} B + b_{out} I \times B$ Equation 1

The letter *I* represens the parameter *INGROUP DESCRIPTIVE NORM* and *B* the parameter *BOTH NORMS SHOWN*. To account for the *OUTGROUP DESCRIPTIVE NORM* parameter, we represent it by portraying the interaction of the two independent variables (*I x B*). Lowercase letters are priors that serve as effects caused by changing those parameters that differ in both models, as stated below.

In addition, the self-categorization model takes ingroup-agree and outgroup-disagree terms into account as shown here in Equation 2 :

$$log_e \ (odds) = b_{in} \ I \ x \ INGROUP \ AGREE \ + \ b_{both} \ B \ + \ b_{out} \ I \ x \ B \ x \ OUTGROUP \ DISAGREE$$

The first one is defined as 1 when identifying with the ingroup and 0 otherwise. The same principle applies tothe latter. These terms are to be understood as an identity term revealing whether the subjects' position was in accordance with the self-categorization theory. The main difference to the alternative model is that the variables which heavily influence the outcome are ignored, since zero and one disregard or fully regard parts of the equation.

Figure 1 shows the proportions of the responses for each condition under both models. The blue bars represent the mean response, and the gray bars represent the relative proportions of each response. The means are fairly similar. However, the relative proportions express an unevenly distributed number of responses which may be the result of a small number of participants.
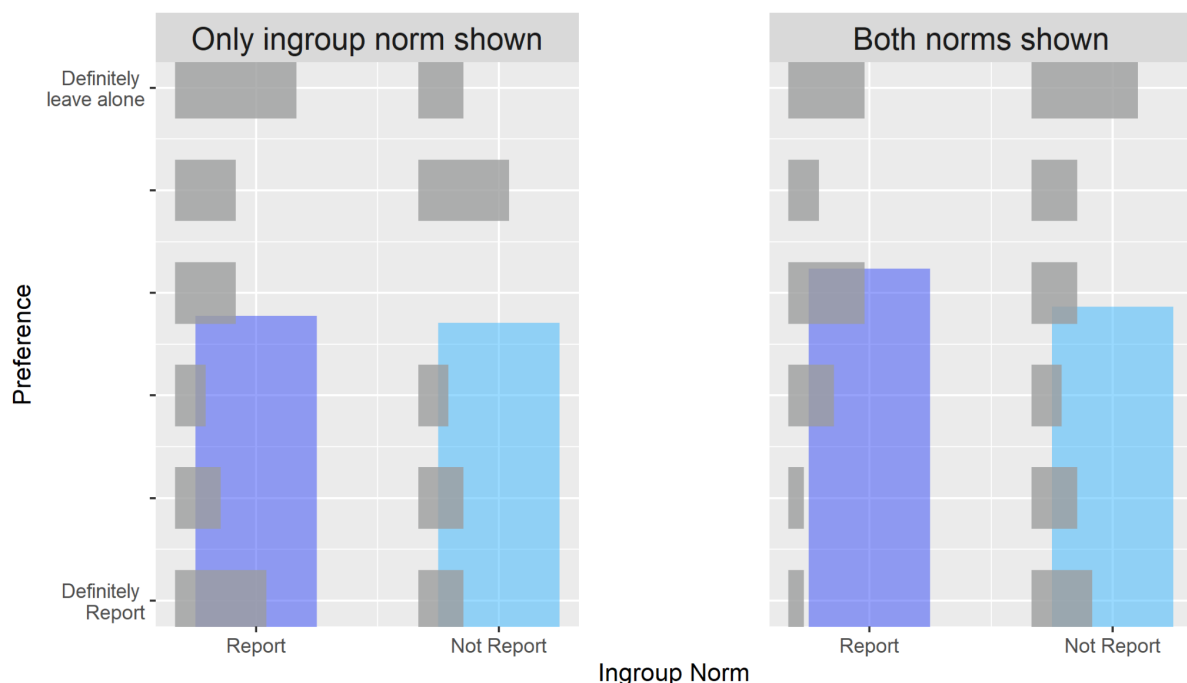
Figure 1. Moral dilemma responses shown for both conditions. Blue bars represent the mean response and gray bars the relative proportion of responses. No significant effect is visible.

Prior assumptions

   We took all priors from the paper and will therefore briefly mention every detail. We set the prior b_{in} to be a folded normal distribution with a mean of 0.816 and a standard deviation of 0.5 for both models. It is strictly greater than 0, since both models assume that the effect of the ingroup descriptive norm will have a positive direction.

The prior b_{both} depicts the effect of presenting both group norms allowing an outgroup norm to be more effective in one direction than in another. Both models share the normal distribution with a mean of 0 and a standard deviation of 0.5.

Lastly, the prior b_{out} shows the amount to which presenting an outgroup descriptive norm shifted preferences towards or away from the option favoured by the ingroup descriptive norm. The self-categorization model resembles this by having a half-normally distributed prior with a mean of 0 and a standard deviation of 0.5 restricted to be greater than 0 (increasing conformity with the ingroup norm). The alternative model was given a prior distribution in the form of the transformed b_{in}. Regarding the provided percentages during the trials, we applied the multiplication like so: - *0.85/0.6 b_{in}* (Ingroup and outgroup norms are expected to influence the participants' preferences equally with the exception of their strengths).

Model comparison

   To determine the model that explains the observed data best, we made use of the Bayes Factor calculated with the 'Bridge Sampling' package in R.The Bayes Factor (the likelihood of observing the data under the alternative model divided by the likelihood of observing the data under the self-categorization model) is 6.06522. Compared to the BF of 34.97 in the original paper, our results show a less significant discrepancy between the two models which weakens the evidence for the alternative model previously observed. To test how the variations of the prior parameter b_{out} affect the outcome, we ran the same analysis with different values for the mean (0, 0.5, 1) and the standard deviation (0.25, 0.5, 1, 2). This yielded values in the range of 4 to 21 favoring the alternative model. (This interval is a lot smaller than reported in the original study which is coherent with our primary finding of the models being less divers). We also plotted the prior and posterior distribution for each parameter (b_{in}, b_{both}, b_{out}) to look for major deviations. The light blue dashed lines represent the prior and the solid black lines the posterior distribution, see Figure 2 below:
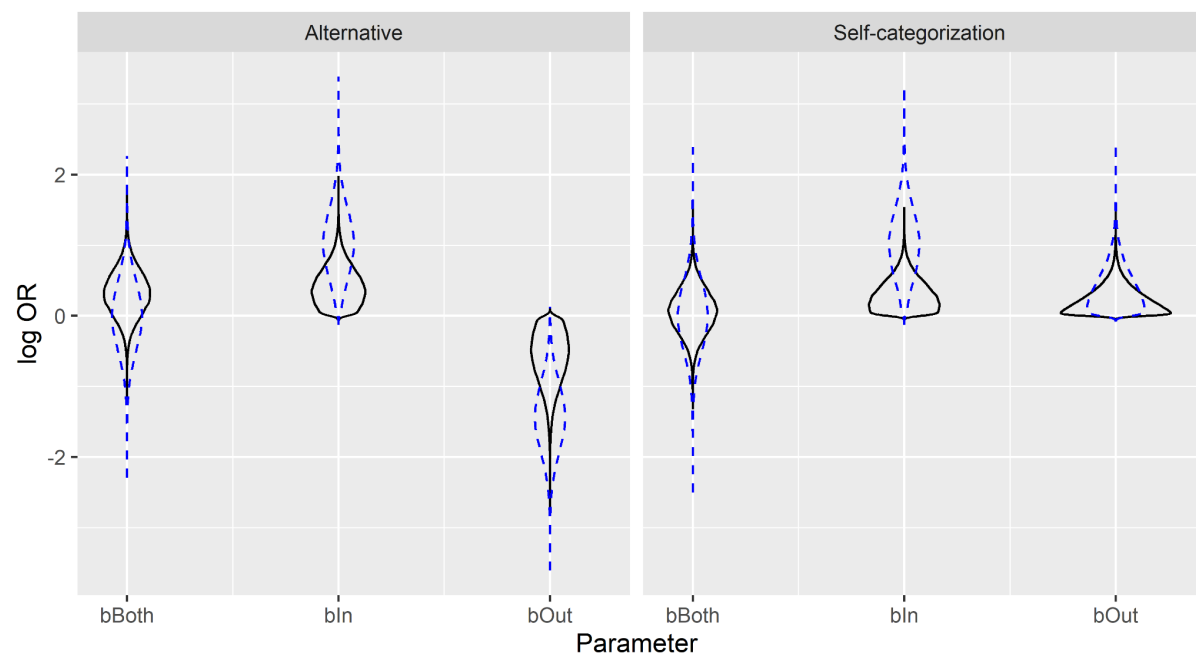
Figure 2. Prior and posterior density for all three parameters of the equation (both models). The y-axis gives the values on a logarithmic scale.

The outcome here favors the alternative model instead of the self-categorization model. That is, we could not find clear evidence supporting the hypothesis that self-categorization plays a major role in the decisions made by our participants. Overall, our analysis weakly supports the findings of the Pryor, Perfors and Howe (2019) that self-categorization cannot account for the results as well as the descriptive norm effect does, since our results point in the same direction but to a much lesser extent.

Effect sizes

 Despite having examined the Bayes Factor, we did a frequentist ordinal regression using the cumulative link model which led to no significant outcomes (see fig.3).

```
link  threshold nobs logLik  AIC    niter max.grad cond.H
logit flexible  83   -144.16 304.31 5(0)  2.70e-13 1.6e+02

Coefficients:
                       Estimate Std. Error z value Pr(>|z|)
ingroupNorm            -0.09685  0.54809   -0.177   0.860
bothShown               0.33046  0.53783    0.614   0.539
ingroupNorm:bothShown  -0.14059  0.78457   -0.179   0.858

Threshold coefficients:
    Estimate Std. Error z value
1|2  -1.5248  0.4108   -3.712
2|3  -0.8216  0.3796   -2.165
3|4  -0.3322  0.3693   -0.900
4|5   0.2508  0.3672    0.683
5|6   1.0369  0.3823    2.712
```

Figure 3. Summary of frequentist analysis using cumulative link model for ordinal regression

Conclusion

The authors of the original article reported findings across two experiments supporting the alternative hypothesis that descriptive norms having an effect on people's opinions in moral judgments cannot be purely explained by self-categorization theory. Surprisingly, this was found in social and politically controversial situations which the participants had reported to seriously care about. That a person's decision-making is only influenced by the group norm this person identifies with could not be confirmed, since compared with ingroup *and* outgroup norms presented, the bigger number of participants shifted their opinion towards following the overall consensus. To question this finding, we replicated the first of these experiments. By doing so, we observed smaller differences between the conditions *INGROUP DESCRIPTIVE NORM* and *BOTH NORMS SHOWN* which has us believe that the self-categorization effect (following the more popular decision in the ingroup) did not affect the subject's decisions as weakly as previously reported when competing with an outgroup norm.

Nevertheless, it is reasonable to support the finding that self-categorization is not sufficient for explaining the observed outcome, since our results provide an objection to self-categorization theory and are similar to the results of the original paper; that is, the alternative model is more likely to have generated the outcome. The degree to which the descriptive norm effects are controlled by ingroup norms was not specified in this experiment and remains uncertain. It could be plausible that a kind of association between the groups and the norms needs to be established which would result in us having one or more latent variables in our equations. Such a relation would be found easily assuming that a specific descriptive norm behavior was stereotypical for a group. Not having a separate ideology can harm group identities and lead to a worse suited prior configuration of groups. It is up to future researchers to investigate to which extent self-categorization theory is consistent with descriptive norms and what is missing. What we do claim is that the hypothesis we tested is very unlikely to be true, as there are better quantitative reasons to reject than to maintain it. Therefore, we side with Pryor, Perfors and Howe and suggest that a richer and more complex version of the self-categorization theory is needed when declaring it the source of descriptive norm effects.

We should emphasize once more that the fact that we had fewer participants and that they were predominantly young may very well have distorted our results. Since most of the subjects were recruited either over social media or universities and these institutions are made up of mostly young people, it is unlikely that our participants represent the general population. Keeping this discrepancy in mind, one can argue that the ingroup norm gets to take on a more influencing role compared to the general descriptive norm, because we propose young adults (μ_{age}= 27 years in our study) have a stronger tendency to conform with their respective ingroup than older adults (μ_{age} = 60 years in original study) do.

References

Abbink K, Freidin E, Gangadharan L, Moro R. The effect of social norms on bribe offers. The Journal of Law, Economics, and Organization. 2018; 34(3): 457–74

Gronau QF, Sarafoglou A, Matzke D, Ly A, Boehm U, Marsman M, et al. A tutorial on bridge sampling. Journal of mathematical psychology. 2017; 81: 80–97. https://doi.org/10.1016/j.jmp.2017.09.005 PMID: 29200501

Köbis NC, Van Prooijen J-W, Righetti F, Van Lange PA. "Who doesn't?"—The impact of descriptive norms on corruption. PloS one. 2015; 10(6): e0131830. https://doi.org/10.1371/journal.pone.0131830 PMID: 26121127

Postmes, T., Haslam, S. A., & Jans, L. (2012). A single-item measure of social identification: Reliability, validity, and utility. *British Journal of Social Psychology*, *52*(4), 597–617. https://doi.org/10.1111/bjso.12006

Pryor C, Perfors A, Howe PDL (2019) Conformity to the descriptive norms of people with opposing political or social beliefs. PLoS ONE 14(7): e0219464. https://doi.org/10.1371/journal.pone.0219464

Rimal RN. Modeling the relationship between descriptive norms and behaviors: A test and extension of the theory of normative social behavior (TNSB). Health Communication. 2008; 23(2): 103–16. https://doi.org/10.1080/10410230801967791 PMID: 18443998

Rimal RN, Real K. How behaviors are influenced by perceived norms: A test of the theory of normative social behavior. Communication Research. 2005; 32(3): 389–414

Rimal RN, Real K. Understanding the influence of perceived norms on behaviors. Communication Theory. 2003; 13(2): 184–203

Schultz PW, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V. The constructive, destructive, and reconstructive power of social norms. Psychological science. 2007; 18(5): 429–34. https://doi.org/10. 1111/j.1467-9280.2007.01917.x PMID: 17576283

Smith JR, Terry DJ. Attitude-behaviour consistency: The role of group norms, attitude accessibility, and mode of behavioural decision-making. European Journal of Social Psychology. 2003; 33(5): 591–608

Turner JC, Hogg MA, Oakes PJ, Reicher SD, Wetherell MS. Rediscovering the social group: A self-categorization theory. Oxford, UK: Basil Blackwell; 19871

Endnotes

1. See https://doi.org/10.1371/journal.pone.0219464, p. 11f.

2. https://magpie-ea.github.io/magpie-site/

3. The python script cleaner-script.ipynb can be found in our github repository in the

   analysis section