Testing conformity to descriptive norms in people with opposing beliefs: A replication study

Marie Bensien, Iolanta Martirosov, Clara Matheis and Alina Ohnesorge

Termpaper for "Experimental Psychology Lab"

a course by Prof. Michael Franke, University of Osnabrück

Abstract

Our social group can have an influence on our choice of behavior. Although this is a well known fact, the proper mechanisms behind it are not yet studied to their full extent. In their study from 2019, Pryor, C., Perfors, A., & Howe, P. D. (2019) investigated this influence, which we replicated in order to examine their findings. Pryor et al (2019) tested the implications which the self-categorization theory makes about the effect of ingroup and outgroup norms on choice of behavior. This study is a direct replication of their study. In the experiment, participants are divided into two groups. In group one, participants are presented with an ingroup descriptive norm only, while in group two, participants are presented with both an ingroup descriptive norm favoring some behavior and an outgroup descriptive norm favoring the other behavior.This replication resulted in the finding that participants would prefer an action when being presented with the claim that within the majority of people the alternate option was unpopular, regardless of not agreeing with that majority concerning a chosen social issue.This finding speaks against the self-categorization theory, proposing that the wish to harmonize with the opinion of others (the majority) might defeat the influence of ingroup against outgroup mentality.

Keywords: self-categorization theory, descriptive norm effect

Testing conformity to descriptive norms in people with opposing beliefs:

A replication study

Introduction

Human behavior is one of the least predictable variables in human life. It is influenced by a magnitude of factors in a multitude of ways, which makes it nearly impossible to securely predict any behavior or reaction reliably. That is one of the reasons researchers study human behavior in relation to different factors and develop theories to achieve a more tangible understanding. Pre-established findings state that individuals tend to prefer behavior that mirrors their knowledge of how other people would behave. This is called the descriptive norm effect (Rimal, 2008).

Self-categorization theory combines the descriptive norm effect with other social phenomena and intuitively predicts that people will choose to conform to the behavior they expect from people they identify with (ingroup), e.g. judging by political views, thereby facilitating ingroup norm conformity. Moreover, self-categorization theory predicts a correlation between the degree of behavioral conformity with the group and the extent to which the individuals perceive themselves as members of that ingroup. In other words, the more an individual identifies with a group, the more they will conform with its behavior.

As another key prediction, self-categorization theory conversely predicts individuals to desist from behavior popular by groups other than, or in opposition to, their ingroup (outgroups). This enhances a distinguishable ingroup identity and hereby the individuals' sense of identity.

These theories do not exclude each other, but can contradict each other in certain cases. For example, someone could identify with feminism but be situated in a social group that does not identify similarly, such that the majority of people and therefore the behavior the individual is exposed to is not compatible with their personal beliefs.

Some explanations of the descriptive norm effect disregard self-categorization theory and focus more on the general assumption that people will conform to behavior shown by the majority of people, independent of the ingroup and outgroup

identification. Research distinguishing the effects of ingroup and outgroup descriptive norm effect on people's behavior is, to our current knowledge, scarce and previous research regarding descriptive norms and their effect on other social phenomena is inconsistent so far.

Therefore it was unclear if an individual would generally prefer behavior that fits the preference of people they identify with rather than conform to behavior prefered by the majority of people, until Pryor, Perfors, and Howe (2019a) aimed at answering that question in their paper "Conformity to the descriptive norms of people with opposing political or social beliefs" published in 2019 (Pryor et al., 2019a). Their data was not consistent with the self-categorization, with strong evidence in favor of the descriptive norm effect model. While the chosen behavior of a participant's ingroup had a significant effect on the participant's behavior choice, showing both ingroup and outgroup norms did not have a significant effect. Notably, even though the interaction effect between ingroup norm and outgroup norm, reflecting the effect of the behavior favored in a participant's outgroup, was consistent with the self categorization theory, it was not significant. As critical research practice requires reproducibility, we decided to replicate the aforementioned study in the context of the seminar "Experimental Psychology Lab" at the University of Osnabrück.

Aim and hypotheses

The aim of this study was to replicate a large-scale experimental study on a smaller scale to test the findings of Pryor et al. and thereby either confirm or refute the hypotheses presented in the following. As stated above, self-categorization theory predicts that, across all groups, participants will favor the behavior of their ingroup. The alternate hypothesis assumes participants will conform to the overall descriptive norm favored by the majority of people, even if these are from the outgroup.

To test this hypothesis, half of the participants were presented with an ingroup descriptive norm in favor of a certain option (group one), while the other half were presented with both the ingroup descriptive norm and an outgroup descriptive norm in

favor of the opposing option (group two). Additionally, self-categorization theory predicts a stronger descriptive norm effect in the group two, with the intention to distinguish themselves from the outgroup. Alternatively, the ingroup descriptive norm effect is weaker in group two, such that they conform more with the overall descriptive norm.

Method

Participants

66 participants (33 women, 29 men, 1 non-binary, 3 other, median age = 23 years, age range = 18 - 34 years) were recruited via the Cognitive Science student mailing list and via personal contacts of the four experimenters. They were not compensated for their participation. The experiment was completed in a web browser and took around 3 minutes to execute. All participants gave informed consent. Experimenters were not blinded to the conditions of the experiments, therefore data collection and analysis were not performed blind as well. There were six additional participants recruited in the same way which tested a preliminary version of the experiment. Their data was excluded from the final analysis, as well as example trails from the researchers.

Materials

The experiment was executed as a browser based implementation using the _magpie architecture (Ilieva, Ji, Rautenstrauch, & Franke, 2021) and hosted on Netlify (Biilmann & Bach, 2021). The experiment was adapted from Pryor et al. (Pryor et al., 2019a). Therefore, their pre-registration report (Pryor, Perfors, & Howe, 2019b) and final paper (Pryor et al., 2019a) were used as a guideline for the implementation and execution of the experiment.

Procedure

Participants were first asked to provide basic demographics about age and gender and then had to choose one out of nine social issues which they cared about most. They

then were asked how much they agreed or disagreed with a statement about their chosen issue and had to give an answer on a 11-point Likert- scale ranging from -5 („strongly disagree) to +5 („strongly agree“). If a participant for example chose „Feminism“ as the social issue they cared about most, the statement was „Feminism is important and beneficial to modern society“. The rating of the social issue was then used to determine the ingroup and outgroup.

Afterwards, general instructions were shown to the participants as well as a fictitious background story claiming that this study is following up on a previous study which looked at how people feel about moral dilemmas. Participants were not informed about the actual research issue to ensure unbiased reactions, so this approach was used to explain the source of the later presented descriptive norms. As the general instructions stated, participants were presented with a scenario describing a moral dilemma. They then chose the preferred action and provided a rating of how good or bad they imagined they would feel after taking that action.

The moral dilemma read like this:

„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.“

Regarding this dilemma description, an ingroup descriptive norm was also shown to the participants telling them that “Approximately 60% of participants who agreed with you about [social issue] chose to [action]”. The group of participants was randomized such that each action was shown to half of the participants to reduce bias towards any action. In other words, half of the participants of each group were told their ingroup chose to call the police and report the robber while the other half was told their ingroup chose to do nothing and leave the robber alone.

Half of the participants were additionally informed that 85% of participants of the previous study that did not agree with them about their chosen social issue preferred the alternative option of action as a response to the moral dilemma.

An understanding check was included at the end, to ensure that participants paid attention. Participants answering at least one of the following questions wrong were excluded from the study.

Which of these options was true about the previous study described in the instructions?

1. *Participants chose which action they preferred (correct)*
2. *Due to a computer error, participants were not allocated equally to imagine performing the different actions (incorrect)*
3. *No data was saved during the experiment (incorrect)*
4. *The participants completed the experiment with their eyes closed (incorrect)*

Lastly, a single-item self-identification measure was used to control for a sense of identification with the determined ingroup. Participants had to rate their agreement on a 7-point Likert scale ranging from 1 (fully disagree) to 7 (fully agree) with the statement "I identify with [ingroup description]" and "I identify with [outgroup description]" . The brackets were replaced by the earlier determined ingroup and outgroup descriptions.

Design

This study is structured as an experimental two factor (2 X 2 factor design) between-subject study with the following components:

The first independent variable is a two-level factor with the levels directly related to the trial condition, distinguishing between the trials in which only ingroup information was shown (both_norms_shown = 0) and both ingroup and outgroup information were shown (both_norms_shown = 1).

The second independent variable is also a two-level factor indicating which behavior was allocated to the ingroup norm, with the two levels "calling the police" (ingroup_norm_new = -1) and "do nothing" (ingroup_norm_new = 1).

The dependent variable represents the choice of behavior of each participant. In order to control for the order of the two given norms having an effect on the dependent variable, the order of appearance is randomized. All randomization will be simple and computer-generated as a coin flip.

Data preparation

Before analysing the data, suitable trials were extracted by excluding some of the participants. As mentioned earlier, any trial with incorrect answers in the understanding check was excluded as they were possibly not paying enough attention during the experiment. Another reason to exclude trials was if participants were neutral about their chosen social issue, as this prevented a clear distinction of their ingroup and outgroup respectively.

Results

Before running the analysis, 2 participants were excluded as they rated their opinion towards the chosen topic as neutral. No trials were excluded for failing the understanding check. The way in which the remaining participants chose to respond to the moral dilemma is displayed in Figure 1.

Models

To determine if the self-categorization theory is applicable to explain our data in a sensible way, we directly compared two Bayesian models based on the different assumptions of self-categorization theory and descriptive norm effect respectively. Like in the original paper, we used a Bayesian version of ordinal logistic regression. By modelling the two hypotheses, we can predict answers on the ordinal answer scale ranging from “Definitely call the police” (negative direction) to “Definitely do nothing” (positive direction) and thus, we can compute the posterior probability of the two models, given our collected data. We then used this analysis to compute a Bayes Factor from the two hypotheses, indicating the relative evidence in favor of each hypothesis. All analysis was implemented in R, using the “Stan” package to model the hypotheses

and the "Bridge Sampling package" to calculate the Bayes Factor, as the original authors did.

Each of the two hypotheses was modeled with the help of Bayesian ordinal logistic regression and includes parameters representing the influence of behavior allocated to ingroup norm, whether both norms were shown and presenting an outgroup norm which favors the opposite behavior from the ingroup norm on the choice of behavior, respectively. The priors for these parameters are directly taken from the original paper and differ between the models mainly in how the presentation of an outgroup norm influences the answer choice. Additionally, the self categorization model also includes an additional influence of the identification variable, which effectively can cancel out the influence of ingroup and outgroup norm. This identification is not considered in the alternative hypothesis and thus, its influence is not included in the alternative model.

The parameters for the two models representing the hypotheses are the following. An explanation of their origin can be found in Pryor et al. (2019), while Table 1 describes the meaning of used variables.

Self-categorization model: $log_e(odds) = b_{in}I + b_{both}B + b_{out}I \times B$

Alternative model: $log_e(odds) = b_{in}I + b_{both}B + b_{out}I \times B$

Our analysis resulted in a Bayes factor of 39.50, which indicates that our collected data is 39.50 times more likely when the alternative model is assumed. Interpreted according to the scale proposed in a paper by (Andraszewicz et al., 2015), this means extreme evidence for the alternative model. Compared to the Bayes Factor of (Pryor et al., 2019a) which was 34.97, both can be seen as extreme evidence for the alternative model.

Effect sizes

In addition to the Bayesian models, we ran a frequentist ordinal logistic regression to analyse effect sizes of the parameters in the self-categorization model. For the ingroup descriptive norm we found a significant effect, which described that participants decided to shift away from preferring the decision popular within the ingroup

($N = 64, OR = 0.814, 95\%CI[-0.81, 0.39]$). For participants presented with both norms ($N = 64, OR = 1.734, 95\%CI[-0.33, 1.45]$) there was no significant shift in preferring decisions based on this exposure. Considering the interaction of the ingroup descriptive norm and the instance of both norms shown which allocate opposite behaviors ($N = 64, OR = 1.433, 95\%CI[-0.52, 1.25]$), the odds ratio suggested a higher probability of a link between exposure to this interaction and the results. Nonetheless, this effect was not significant.

Discussion

Self-categorization theory would expect participants to shift their choice of action towards the ingroup norm (Hogg, Turner, & Davidson, 1990). This effect could not be observed. On the contrary: Participants shifted significantly away from the ingroup preference. A possible explanation for this is that participants might make decisions contrasting with those of others to stand out of a group. Another explanation could lie in the study design: The percentages used to describe the ingroup and outgroup norms were rather low (60% and 85% respectively). Further studies could explore the effects of portraying ingroup norms as stronger, which might change the level of felt belonging of the participants towards their respective ingroup norms.

The proposed actions to moral dilemmas were randomized in this study, however it might be necessary to present norms which reflect stereotypical behavior of the specific ingroups. A possible bias of people generally preferring social support of people over pro-actively helping to enforce the law could be observed: Reviewing the overall choice of participants when presented with the moral dilemma, around 60% of participants chose to not report the robber, independent of the shown ingroup norm. This percentage rose even higher when participants were presented with opposing opinions of ingroup and outgroup (73% and 85% chose to leave the robber alone), suggesting that being presented generally with opposing opinions could have an impact on the outcome as well.

In our experiment and analysis we discovered that people's opinion shifted away

from the option which was popular within their ingroup. However, we also found that a more general perception of a descriptive norm effect explained our data in a better way. This shift away from the ingroup norm happened even though ingroup and outgroup were determined by social issues which were important to participants. This opposes previous findings suggesting that people strongly act against norms of their outgroups. Our replication seconds Pryor et al's (Pryor et al., 2019a) findings, self-categorization theory is unable to explain our data. Pryor et al. (2019a) suggested that more general aims to conform with people's opinions come into play, however our findings suggest that people did not change their opinion when being only presented with the ingroup norm. Here other components referring to the percentage of people supporting an opinion or different opinions being displayed at all could have an effect.

References

Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., & Wagenmakers, E.-J. (2015). An introduction to bayesian hypothesis testing for management research. *Journal of Management*, *41*(2), 521–543.

Biilmann, M., & Bach, C. (2021, August). *netlify.* Retrieved 2021-08-31, from `https://www.netlify.com/`

Hogg, M. A., Turner, J. C., & Davidson, B. (1990). Polarized norms and social frames of reference: A test of the self-categorization theory of group polarization. *Basic and Applied Social Psychology*, *11*(1), 77–100.

Ilieva, S., Ji, X., Rautenstrauch, J., & Franke, M. (2021, August). *_magpie architecture.* Retrieved 2021-08-31, from `https://magpie-ea.github.io/magpie-site/`

Pryor, C., Perfors, A., & Howe, P. D. (2019a, July). Conformity to the descriptive norms of people with opposing political or social beliefs. *PloS one*, *14*(7), e0219464.

Pryor, C., Perfors, A., & Howe, P. D. (2019b, April). *Outgroup norms.* Retrieved 2021-08-31, from `https://doi.org/10.17605/OSF.IO/J9UGV`

Rimal, R. N. (2008). Modeling the relationship between descriptive norms and behaviors: A test and extension of the theory of normative social behavior (tnsb). *Health Communication*, *23*(2), 103–116.

Table 1

List of equation variables and what they represent.

Equation variable	Represents
b_{in}	Influence of ingroup norm on behavior choice. Prior is a folded normal distribution with $mean = 0.816$ and $SD = 0.5$
b_{both}	Influence of showing both norms on behavior choice. Prior is a normal distribution with $mean = 0$ and $SD = 0.5$
b_{out}	Influence of showing outgroup norm that is opposite of ingroup norm. Prior for self categorization model is a half normal distribution with $mean = 0$ and $SD = 0.5$. Prior for alternative model is $-0.850.6b_{in}$
I	Ingroup norm variable from data. The levels are -1 ("call the police") and 1 ("do nothing")
B	Both norms shown variable from data. The levels are 0 (only ingroup norm shown) and 1 (both norms shown)
ingroup agree	Identification variable from data. The levels are 1 (participant identifies with ingroup) or 0 (participant does not identify with ingroup)
outgroup disagree	Identification variable from data. The levels are 1 (participant does not identify with outgroup) or 0 (participant identifies with outgroup)

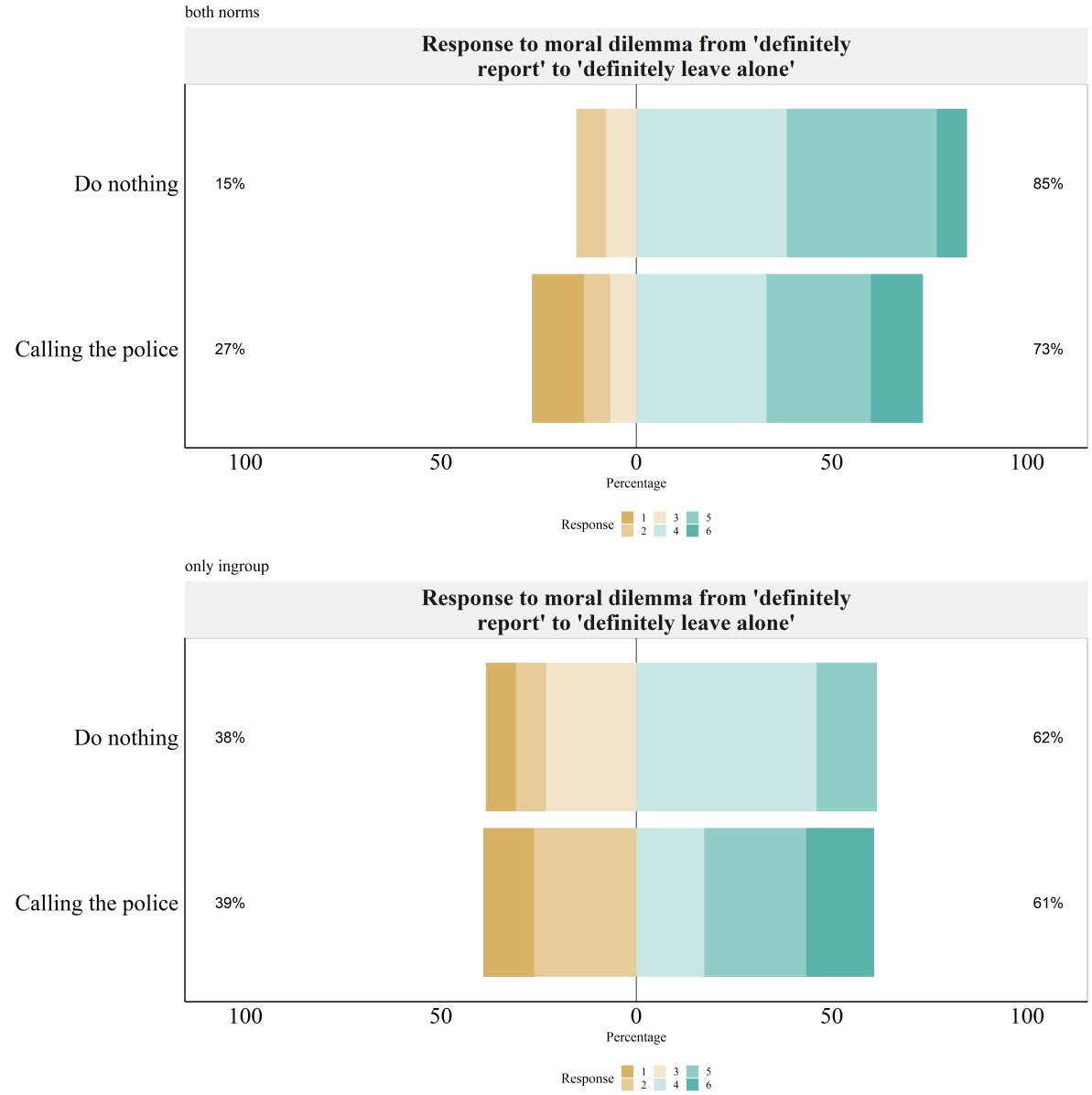

Figure 1. Likert scale answers by condition given responses to moral dilemma. The coloured areas of the bars represent the chosen actions for the dilemma according to the descriptive norms participants were shown. The responses range from 1 (definitely report) to 6 (definitely leave alone).The two plots represent the two conditions of only the ingroup norm being shown and both norms being shown.