Preregistration report

Study Information

1. **Title:** Replication Study of “Conformity to the descriptive norms of people with opposing political or social beliefs” (Pryor et al., 2019)

2. **Authors:** Eva-Marie Constanze von Butler, Nele Felicitas Werner, Juri Morisse, Felix Naujoks

3. **Description:** (optional)
   The field of social psychology investigates the question of how other people’s opinions affect our own decision making. In this experiment we examine the role of two theories that emerged from this line of research. One is the descriptive norm effect according to which people simply tend to make the decision that is most popular (APA Dictionary of Psychology, 2020). The other is the self-categorization theory (Turner & Reynolds, 2012) which states that an individual’s decisions are affected by identification with a social group (ingroup) and segregation from other social groups (outgroups) (Ingroup outgroup, n.d.). In particular, it states that an individual tends to make the same decision as the majority of its ingroup while avoiding making the decision the majority of the outgroup made.
   To investigate these two effects, we first identify each participant’s in- and outgroup before presenting the participants with two moral dilemma situations for which they have to make a decision. The dilemma is either a philosophical or a daily situation. Before making that decision, half of the participants are informed on how their ingroup decided while the other subjects are additionally informed on how the outgroup decided.
   According to the self-categorization theory the proportion of people who made the same decision as their ingroup should be higher in the group that was shown information on both in- and outgroup. Just following the descriptive norm effect would suggest the opposite, so that the proportion of people who make the same decision as their ingroup should be higher among the participants who were only shown ingroup information.
   This study is a replication of the paper “Conformity to the descriptive norms of people with opposing beliefs" by Pryor, Perfors and Howe (2019).

4. **Hypotheses:** (required)

We are here concerned with the self-categorization theory and the descriptive norm effect. In particular, we are going to compare two hypotheses that capture the different predictions about participant behaviour made by the two competing accounts. Those hypotheses are:

1. Main hypothesis (self-categorization theory): Participants will conform more to the ingroup descriptive norm when an opposing outgroup descriptive norm is shown (Self-categorization theory).

2. Alternative hypothesis (descriptive norm effect): People will conform to the overall descriptive norm, such that conformity with the ingroup descriptive norm will decrease when an opposing outgroup descriptive norm is presented.

The goal of this study is to investigate which hypothesis better predicts our data. As the self-categorization theory builds up on the descriptive norm effect by including additional effects (Pryor et al., 2019) and therefore can be seen as going beyond the descriptive norm effect, we regard the hypothesis of the self-categorization theory as our main hypothesis while the hypothesis of the descriptive norm effect is the alternative hypothesis.
We expect to find that our main hypothesis gives a better prediction of our data which would be the case if we find a Bayes Factor larger than 1 in favor of our main hypothesis over the alternative hypothesis. However, the strength of evidence a Bayes Factor poses depends on how large it is and therefore in order to have strong evidence in favor of our main hypothesis, we require a Bayes Factor of at least 10 in favor of our main hypothesis over the alternative hypothesis. More details about how we will compare the hypotheses and interpret our results can be found in the part ‘Analysis Plan’.

Design Plan

5. **Study type:** (required)
    5.1. Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

6. **Blinding:** (required)
    6.1. Blinding describes who is aware of the experimental manipulations within a study.
        6.1.1. For studies that involve human subjects, they will not know the treatment group to which they have been assigned.

6.1.2. Personnel who interact directly with the study subjects (either human or non-human subjects) will not be aware of the assigned treatments. (Commonly known as “double blind”)

7. **Study design:** (required)
Follow: https://github.com/jumorisse/XPL_Replication_Project/tree/main/Writing
to get to the complete design document.

Our experiment has a between subject design with 2 factors [2 (INGROUP DESCRIPTIVE NORM) x 2 (BOTH NORMS SHOWN) factorial design].
For this 2x2 factorial design, there are 2 * 2 = 4 different experimental conditions.
We decided to include two different dilemmata to make sure that the effect is not specific to the dilemma. The dilemma isn’t included as a factor, as this is only an exploratory investigation of a possible effect of “dilemma”.

Each participant sees boths dilemmata, resulting in a within-subjects design, turning the experiment including the exploratory investigation, in a mixed experimental design.

The variable INGROUP DESCRIPTIVE NORM refers to which of the two decisions was apparently favoured by the ingroup.

For the philosophical dilemma this would result in the INGROUP DESCRIPTIVE NORM to either be to pull the lever (INGROUP DESCRIPTIVE NORM = -1) or to do nothing (INGROUP DESCRIPTIVE NORM = 1).

For the daily dilemma this would result in the INGROUP DESCRIPTIVE NORM to either go to the police (INGROUP DESCRIPTIVE NORM = -1) or say nothing (INGROUP DESCRIPTIVE NORM = 1).

The independent variable BOTH NORMS SHOWN represents the distinction of whether only an ingroup descriptive norm was shown (BOTH NORMS SHOWN = 0) or both an ingroup descriptive norm and an outgroup descriptive norm were shown (BOTH NORMS SHOWN = 1).

8. **Randomization:** (optional)
For the different experimental conditions in our study we used the javascript function Math.random() and the decision value 0.5.

Math.random randomly selects a value between 0 and 1. In trials where the value will be below 0.5 the participant will be assigned group one, in trials where the value will be above or equal 0.5 the participant will be assigned group two, we have two variables where this group assignment applies. We used this kind of randomization to create a

completely random group assignment while the experiment is running and keep the group sizes as similar as possible.

The following variables will be randomly selected:
INGROUP DESCRIPTIVE NORM variable, encoding which decision the ingroup favoured and the BOTH NORMS SHOWN variable, encoding if both ingroup norm and outgroup norm are represented or only the ingroup norm.

To control for potential order effects we randomly ordered the ingroup and outgroup descriptive norms and the sequence of the dilemmata. The random order of descriptive norms only applies for the BOTH NORMS SHOWN = 1 condition. These two factors are therefore ignored in the data analysis.

Sampling Plan

The experiment will be conducted as a web-application. This web-application is hosted on netlify and can be accessed here: https://group-15-main.netlify.app/. This portable and interactive experiment will run inside any modern browser.

9. **Existing data:**
Data from the previous study which gets replicated was available but was not analysed or used for anything.
Data from the pilot study (N=4) was available and guided the specification of the statistical model. This data is not included in the final analysis. No data from the experiment to be preregistered here was available at this time.

10. **Explanation of existing data:**
Existing data does not contribute to further analysis.

11. **Data collection procedures:** (required)

    11.1. Participants will be recruited through a Link to the web-application. They won't be paid for participation. Participants must be at least 16 years old and must indicate German as their native language. We decided to only include participants above 16, as 16 is the youngest possible voting age in Germany and the topics in our study are partially political. We therefore only want to include participants who would be eligible to participate in the current democratic system. As the contents of our study requires the participant to be fluent in German, we are using the indicated native language as exclusion measurement.

12. **Sample size:** (required)
    12.1. In the original study they used a sample size of 301 participants. We do not think that we will be able to get such a high number of participants and do not set a specific number as ideal sample size. We rather collect as many data entries as possible. The link will be shared with other students, friends, family, and through personal social media profiles, we will start data collection on 2021-08-24 at 12:01 am and will stop data collection on 2021-08-27 at 11:59 pm.

13. **Stopping rule:** (optional)
    13.1. We start data collection at 2021-08-24 at 12:01 am and will stop data collection at 2021-08-27 at 11:59 pm.

Variables

14. **Manipulated variables:** (optional)
    In our experiment we used the two manipulated (independent) variables INGROUP DESCRIPTIVE NORM and BOTH NORMS SHOWN.

    We have four different conditions. The participants see either only the ingroup descriptive norm, or they see both the favoured ingroup and the outgroup descriptive norm, the ingroup and outgroup descriptive norms are always contradictory, meaning the choice of behaviour is exactly opposite to each other.
    Possible manipulation level/ experimental conditions :
    - Ingroup favours behaviour 1
    - Ingroup favours behaviour 2
    - Ingroup favours behaviour 1 and outgroup favours behaviour 2
    - Ingroup favours behaviour 2 and outgroup favours behaviour 1

15. **Measured variables:** (required)
    The dependent (measured) variable is the participants' responses on the Likert scale rating the certainty with which they would act in a certain way in the dilemma situations.
    We will measure this by asking the participant "How would you choose?" with respect to the two given ways to act.
    They have six different scaled options to choose from, going from 1 = definitely doing the first option to 6 = definitely doing the other option.
    This variable is measured as an ordinal variable.

Analysis Plan

16. **Statistical models:** (required)

In order to determine the extent to which self-categorization theory provides a better or worse explanation of our data than the alternative hypothesis, we implement two competing models that capture the two hypotheses and investigate which is better suited to explain our obtained data. Both models will be instances of Bayesian ordinal logistic regression implemented in R by using the stan and brms packages.
Both models try to predict the responses in the dependent variable dilemma_rating. Specifically, they try to predict the natural log odds of responding with a higher rating which in our case represents a more passive action to solve the dilemma (e.g. not report your friend to the police).

In the alternative model, the dependent variable dilemma_rating is regressed against the two independent variables ingroup_norm (I) and both_infos (B) as well as their interaction $(I{\cdot}B)$ as can be seen in Equation 1. The interaction term represents the effect of the outgroup_norm as it contains information on its existence (given by B) as well as direction (given by I and the positivity/negativity of b_{out}).

Equation 1:

$$log_e(odds\ of\ deciding\ for\ the\ passive\ option) = b_{in}I + b_{both}B + b_{out}I{\cdot}B$$

The model for the self-categorization theory includes two additional independent variables ingroup_agree and outgroup_disagree that interact with the ingroup_norm and the interaction term respectively (see Equation 2). These two variables indicate whether the subject actually identifies with its assumed ingroup and not identifies with its outgroup. They are included in the self-categorization theory model to account for the expected contrary effects of ingroup and outgroup opinions on the subject's decision.

Equation 2:

$$log_e(odds\ of\ deciding\ for\ the\ passive\ option) = b_{in}I{\cdot}Ingroup\_Agree + b_{both}B + b_{out}I{\cdot}B{\cdot}Outgroup\_Disagree$$

Both models make use of three priors that we sample from the same distributions as Pryor et al. (2019) did.

17. **Transformations:** (optional)

For the variables used in the regression, we used the same coding as Pryor et al. Namely, ingroup_norm is coded as -1 (the ingroup prefers for the active option) or 1 (the

ingroup prefers the passive option) and both_infos is either 0 (only the ingroup_norm was shown) or 1 (both ingroup and outgroup norm were shown). The two additional variables for the model of the self-categorization theory are both coded as either 0 or 1. The ingroup_agree variable is 0 if the subject does not identify with its assumed ingroup and 1 if it does. Outgroup_disagree is 0 if the subject does actually identify with its assumed outgroup and 1 if it does not (as expected). Therefore, the ingroup_agree and outgroup_disagree variables implement the assumptions of the self-categorization theory by acting as a control that can turn off both the effect of the ingroup_norm variable (if the subject does not identify with its supposed ingroup) or the effect of the interaction term (if the subject does actually identify with its supposed outgroup).
The values of the dependent variable dilemma_rating are the result of the rating response and lie between 1 (definitely deciding for the active option) and 6 (definitely deciding for the passive option).

18. **Inference criteria:** (optional)

To investigate which hypothesis is stronger supported by our data, we compare how well their respective models explain the obtained data. We do so by computing the Bayes Factor in favor of our self-categorization model as seen in Equation 3.

Equation 3:

$$BF = \frac{p(data|self-categorization\ model)}{p(data|alternative\ model)}$$

We interpret the resulting Bayes Factor based on Jeffreys' scale (Jeffreys, 1961).

19. **Data exclusion:** (optional)

First, we will exclude all participants that indicated that their native language is not German. Then we will exclude the data of participants who failed the understanding check. Since this study design assumes the participants' in- and outgroup determination is based on their opinion on a social issue statement, we want to ensure that all participants actually have a strong opinion about their chosen social issue. Therefore, we will also exclude all participants that rate the social issue statement between (including) -2 and 2.
Only the data of the remaining participants will be used in the analysis.

20. **Missing data:** (optional)

Since our means are limited and we want to gather as much data as possible, we decided to not force our participants to provide personal data like (e.g., age, education and native language). Consequently, we expect that for some participants this data will be missing. If that is the case, we will extend the demographics summary with information on how many participants provided information in the different categories (e.g., 'From the 100 participants 95 provided information on their gender. Of these 95, 40 were male and 55 were female.').

21. **Exploratory analysis:** (optional)

Contrary to the original study by Pryor et al., we did not confront the participants with just one but instead with two dilemmas. We chose these two dilemmas so that they differ in respect to their likelihood of occuring in a real life situation. One is the traditional trolley dilemma, which can be described as an instance of a purely philosophical dilemma that is unlikely to ever actually happen. The other dilemma poses a more realistic situation and therefore we refer to it as the realistic dilemma. We perform two exploratory analyses with regards to these two types of dilemmas.

For the first exploratory analysis, we are interested in how well the models explain data of only one type of dilemma. Therefore, we split the total dataset into two datasets each of which contain data of only one dilemma type (realistic or philosophical). Then we perform the same analysis we did for the whole dataset on these two subsets. Consequently, we obtain one Bayes Factor for each of the two subsets, which indicates what model better explains the data of this subset. The aim is to investigate whether there might be an indication of differing performance of the two competing models with regards to the type of dilemma used. While this analysis cannot make a definite statement about the differences in performance between the datasets, it might show that further research in that direction could be promising.

In the second exploratory analysis, we extend the Bayesian ordinal logistic regression models that were used earlier with a dilemma type term. The resulting new regression formulas can be seen in Equation 4 (for the exploratory alternative model) and Equation 5 (for the exploratory sct-model) with Trepresenting the dilemma_type variable.

Equation 4:

$$log_e(odds\ of\ deciding\ for\ the\ passive\ option) = b_{in}I + b_{both}B + b_{out}I{\cdot}B + b_{typ}T$$

Equation 5:

$$log_e(odds\ of\ deciding\ for\ the\ passive\ option) = b_{in}I{\cdot}Ingroup\_Agree + b_{both}B + b_{out}I{\cdot}B{\cdot}Outgroup\_Disagree + b_{ty}$$

These exploratory models can be understood as versions of the initial models for which we assume a main effect of the dilemma_type variable. We fit these exploratory models

the same way as the initial models and then investigate whether the exploratory models explain the data better than the initial models. We do so by computing the Bayes Factor in favor of the exploratory over the initial model and interpret the result based on Jeffreys' scale (Jeffreys, 1961). The reasoning behind this analysis is the following: If we indeed find that the Bayes Factor is high enough to suggest that the exploratory models are better at explaining our data, this is an indication that the dilemma type has a main effect on the participants decision and therefore the model's respective theory makes different predictions for the two types of dilemmas. Assuming we find such an interesting Bayes Factor in favor of the exploratory models, we will also investigate the direction of a possible dilemma_type main effect.

Other

22. **Other:** (optional)

22.1. If there is any additional information that you feel needs to be included in your preregistration, please enter it here. Literature cited, disclosures of any related work such as replications or work that uses the same data, or other context that will be helpful for future readers would be appropriate here.

22.2. **Statistical Analysis**
Under the following link we provide our R-Script, as open science procedures enable others to understand and possibly reproduce our results:
https://github.com/jumorisse/XPL_Replication_Project/tree/main/Analysis

22.3. **Way in which we chose the social factors:** (optional)
We selected a set of eight social issues which are currently relevant in Germany, to give the participants the option to choose a topic they are interested in. We define these by corresponding research and relevant articles. When selecting the topics, we made sure that they do not only appeal to one gender.

22.4. **Way in which we chose the dilemmata:** (optional)
As a follow-up on the original study, we use two moral dilemmata, meaning situations for which the following action is hard to decide. We purposely chose two different types of dilemmata, the first is a rather daily situation, while the second is a typical philosophical dilemma (trolley dilemma). In our exploratory analysis we will make use of having two different types of dilemmas to investigate possibly differing performances of the hypotheses with regards to the dilemma type and also a possible main effect of the dilemma type in predicting the participants' responses.

We also made sure that the dilemma don't include a social factor topic. Which was the case in the original study (gun control and robbery).

References

American Psychological Association. (2020). *APA Dictionary of Psychology.* Retrieved from https://dictionary.apa.org/descriptive-norm. (Retrieved 2021, August 22).

Jeffreys, H. (1961). *Theory of Probability* (3rd edition). Oxford: Clarendon Press.

Pryor et al. (2019). *Conformity to the descriptive norms of people with opposing political or social beliefs.* PLOS ONE. Retrieved from https://doi.org/10.1371/journal.pone.0219464. (Retrieved 2021, August 22).

Turner & Reynolds. (2012). *Self-categorization theory.* ResearchGate. Retrieved from https://www.researchgate.net/publication/293162479_Self-categorization_theory. (Retrieved 2021, August 22).

Wikia.org. (n.d.). *Ingroup outgroup*. Retrieved from https://psychology.wikia.org/wiki/Ingroup_outgroup. (Retrieved 2021, August 22).