The Coarseness of Moral Language: Pre-Analysis Plan

Friederike Reichel*

January 10, 2023

1 Introduction

This project studies how the coarseness of moral categories affects moral behavior. The conceptual inflation hypothesis (Miles, 1989) and our social signaling model suggest that broadening a negative moral category can have unintended consequences. As a reaction to the new, coarsened moral category, some people may take a less moral action than previously, thereby diluting the category's meaning. We call this behavioral pattern the diluting effect, to be precisely defined below. We will test for the existence of the diluting effect in an online experiment with students based in Munich, Germany.

2 Experimental Design

We will schedule the experiment in sessions taking place on Zoom. Each subject will participate in one session. Throughout the Zoom meeting, subjects will have to turn on their camera.

After consenting, they will read the instructions and answer a few comprehension checks. Subjects will know that they will choose to solve any number of real effort tasks out of a maximum number of 20. They will have to make this choice three times under different circumstances. Denote the number of real effort tasks a subject chooses to solve as an action. There will be three experimental conditions: the No Threshold Condition, the Low Threshold Condition, and the High Threshold Condition. Our within subject design will require a subject to take an action in each condition. Actions will be incentive compatible via the strategy method. The conditions will be described to subjects as situations. One of the three conditions will be randomly selected. Later, subjects will have to solve as many

*LMU Munich. friederike.reichel@econ.lmu.de

real effort tasks as they chose to in the selected condition.

The real effort task is a variant of the counting zeros and ones task. For every real effort task subjects will solve, we will make a donation to Bund Naturschutz in Bayern e. V., a local environmental NGO. The donations will go towards moorland conservation, an effective natural climate solution. It will be made clear to subjects that the fewer tasks they choose to solve, the sooner will they be able to leave the experiment. With this information in mind, subjects will take their action in the No Threshold Condition, which we will call their *natural action*.

Only then will participants learn how the two threshold conditions, described to them as situation 2 and situation 3, differ from the No Threshold Condition. The Low (High) Threshold Condition will be characterized by a low (high) threshold action $r_l = 11$ ($r_h = 17$). Stratified by gender, subjects will be randomized in equal proportions to the order in which they will encounter the two thresholds. Subjects in the Low (High) Threshold First Condition, will face r_l (r_h) in situation 2 and r_h (r_l) in situation 3.

In case either of the two threshold conditions will be selected, subjects' actions will be partially made public. The experimenter will call up each subject one by one. If a subject's action is weakly greater than r, the subject will be announced as "protecting the climate". Else, they will be announced as "not protecting the climate". During the announcement, the subject's video will be spotlighted, i.e. displayed in large to everyone in the session.

After learning the threshold r in a given situation, but before taking their action, subjects will indicate how socially appropriate they think others will deem solving less than or weakly more than r real effort tasks, incentivized by the Krupka and Weber (2013) method. They will also state their belief on what share of subjects will solve fewer than r real effort tasks. Having taken an action in all of the three conditions, subjects will state their beliefs on the average action in each condition, and their belief on the average natural action of subjects choosing to solve less than r real effort tasks in each of the two threshold conditions. Then, should one of the two threshold conditions be selected, the public announcements will be made. Hereafter, subjects will leave the Zoom meeting and complete the experiment on their own. They will proceed with answering a few questions on social norms around climate change taken from Andre et al. (2021), filling out a brief demographics questionnaire, and solving the real effort tasks. Subjects will receive a show-up fee of €6, €12 for completion and up to €5.5 for the incentivized elicitations.

3 Outcome Variables and Hypotheses

3.1 Primary Outcome Variables

3.1.1 Purifying Effect and Diluting Effect

We will classify a subject's actions as consistent with the diluting effect, if their action is weakly greater than r_l in the Low Threshold Condition and smaller than r_l in the High Threshold Condition. Define $DE_i \in \{0, 1\}$ as an indicator variable to take on the value of 1, if and only if a subject's actions are consistent with the diluting effect.
We will classify a subject's actions as consistent with the purifying effect, if their action is smaller than r_h in the Low Threshold Condition and weakly greater than r_h in the High Threshold Condition. Define $PE_i \in \{0, 1\}$ as an indicator variable to take on the value of 1, if and only if a subject's actions are consistent with the purifying effect.
We hypothesize that the diluting effect and the purifying effect exist.

3.1.2 Actions

We will analyze the actions, i.e. the number of tasks a subject chooses to solve, in each of the three conditions. We will compare average actions across conditions. We will call the action in the No Threshold Condition the *natural action* and interpret it as a subject's type. We hypothesize that subjects react to our threshold incentives.

3.2 Secondary Outcome Variables

3.2.1 Beliefs

We will record subjects' beliefs on (i) average actions across conditions, on (ii) the share of subjects choosing an action smaller than r in each threshold condition, and on (iii) the average natural action conditional on choosing an action smaller than r in each threshold condition. We will compare these beliefs with the empirical values.

3.2.2 Perceived Shifts in Meaning

We will measure the perceptions of what it means to solve less than r_l real effort tasks in the Low Threshold Condition as opposed to solving less than r_h real effort tasks in the High Threshold Condition. To this end, we will analyze subjects' beliefs on (i) the share of subjects choosing an action smaller than r in each threshold condition, (ii) the average natural action conditional on choosing an action smaller than r in each threshold condition, and (iii) the modal response to the social appropriateness question. We hypothesize that

with an increase in r, (i) the perceived share of subjects choosing an action smaller than r increases, (ii) the perceived average natural action conditional on choosing an action smaller than r increases, and (iii) the perceived social appropriateness of taking an action smaller than r increases.

4 Sample and Empirical Strategy

4.1 Sample

Subjects will be recruited via ORSEE (Greiner, 2004) by the Munich Experimental Laboratory for Economic and Social Sciences (MELESSA). Participation will be restricted to university students based in Munich. The experiment will be scheduled in sessions with roughly 20 subjects per session. We aim for gender balance at the session level. In total, we aim to recruit approximately 240 subjects.

4.2 Power Analysis

A sample size of 240 allows us to detect a difference in average actions between the Low Threshold Condition and the High Threshold Condition of 0.8 at a 0.05 significance level with 80% power.

4.3 Empirical Strategy

4.3.1 Main Test

Diluting effect: As a first step, we will report the share of subjects whose actions are consistent with the diluting effect, as specified in Section 3.1.1. To account for decision noise, we will construct two benchmarks.

As a first naive benchmark, we will compare the observed relative frequency of the diluting effect with the relative frequency that would arise if actions were purely random. We will restrict the sample to those whose actions differ across the Low Threshold Condition and the High Threshold Condition. For each individual, we will simulate a low threshold action and a high threshold action with independent draws from the uniform distribution over the interval $[0, 20]$. We will repeat this simulation $10,000$ times and record the relative frequency of the diluting effect. The average relative frequency of the diluting effect across all repetitions will serve as the relative frequency of the diluting effect under the null hypothesis, p_0. Using the binomial test, we will test whether the observed relative frequency of the diluting effect is significantly different from p_0.

As a second benchmark, we will compare the observed relative frequency of the diluting effect with the relative frequency that would arise if threshold actions were randomly distributed around the natural action. For each individual, we will simulate a low threshold action and a high threshold action by independently adding uniformly distributed noise to their natural action. We will repeat this simulation $10,000$ times and record the relative frequency of the diluting effect. The average relative frequency of the diluting effect across all repetitions will serve as the relative frequency of the diluting effect under the null hypothesis, p_0. Using the binomial test, we will test whether the observed relative frequency of the diluting effect is significantly different from p_0. We will run this test once by uniformly drawing noise from the set $\{-1, 0, 1\}$ and once by uniformly drawing noise from the set $\{-2, -1, 0, 1, 2\}$.

Purifying effect: We will test for the existence of the purifying effect analogously.
In addition, we will test whether the subjects whose actions are consistent with the diluting effect have a higher type, i.e. natural action, than those whose low threshold action is below r_l.

4.3.2 Caliper Test

In order to test whether subjects react to our threshold incentives, we will employ a caliper test (Gerber and Malhotra, 2008). Our No Threshold Condition serves as the control condition. For each threshold condition, we will test with a one-sided t-test of proportions whether, within a specified range, the percentage of observations weakly greater than r is larger than in the control condition. The null hypothesis is that this frequency is higher in the control condition. We pre-specify the ranges $[10, r_l = 11, 12]$ and $[9, 10, r_l = 11, 12, 13]$ for the Low Threshold Condition, and the ranges $[16, r_h = 17, 18]$ and $[15, 16, r_h = 17, 18, 19]$ for the High Treshold Condition.

4.3.3 Average Treatment Effects

We will compare aggregate actions across all conditions. For a pure across subjects analysis, we will restrict our analysis to the first threshold action subjects take.
We will test for a differential dynamic of threshold actions by the order of threshold conditions. Let Y_i be the action chosen by subject i. Define $T_i \in \{0, 1\}$ as a treatment indicator taking on the value of 1 if subject i is in the High Threshold Condition and 0 if subject i is in the Low Threshold Condition. Define $HF_i \in \{0, 1\}$ as an indicator variable taking on the value of 1 if and only if subject i saw the high threshold first. Note that HF_i is fixed at the subject level. Let $\boldsymbol{X}_i$ be a vector of controls, specified in more detail below. We then

estimate

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 HF_i + \beta_3 (T_i \times HF_i) + \beta_4 \boldsymbol{X}_i + \epsilon_i \quad (1)$$

In Equation (1), β_3 measures whether the treatment effect differs by the order in which subjects encounter the thresholds. The controls $\boldsymbol{X}_i$ will include gender and age. We will cluster standard errors at the subject level.

4.3.4 Secondary Analysis

As a manipulation check, we will compare the average natural action conditional on an action smaller than r across the two threshold conditions. We expect the average natural action conditional on solving less than r_l real effort tasks in the Low Threshold Condition to be lower than the average natural action conditional on solving less than r_h real effort tasks in the High Threshold Condition. In addition, we will compare the share of subjects choosing an action smaller than r across the two threshold conditions. We expect this share to be larger in the High Threshold Condition than in the Low Threshold Condition. We will run Kolmogorov-Smirnov tests for equality of distribution across the three conditions. We will run simple mean comparisons on our secondary outcome variables specified in Section 3.2. We will apply a concave loss function on actions in order to mimic a concave externality function. We will then employ the tests specified in Section 4.3.3. We will analyze the magnitude of the diluting effect and the purifying effect by considering the absolute increases or decreases in actions associated with those effects.

References

Andre, Peter, Teodora Boneva, Felix Chopra, and Armin Falk. 2021. "Fighting climate change: The role of norms, preferences, and moral values."

Gerber, Alan S, and Neil Malhotra. 2008. "Publication bias in empirical sociological research: Do arbitrary significance levels distort published results?" *Sociological Methods & Research*, 37(1): 3–30.

Greiner, Ben. 2004. "An online recruitment system for economic experiments."

Krupka, Erin L, and Roberto A Weber. 2013. "Identifying social norms using coordination games: Why does dictator game sharing vary?" *Journal of the European Economic Association*, 11(3): 495–524.

Miles, Robert. 1989. "Conceptual inflation." In *Racism*. 41–68. Routledge.