

Collaborative registered replication of Griskevicius et al (2010): Can pro-environmental behavior be promoted by inducing status competition?

Jordan Wagge

Last Updated 2023-01-18

Corresponding author:

Ljiljana B. Lazarević, Faculty of Philosophy, University of Belgrade Čika Ljubina 18-20, 11000 Belgrade, Serbia Email: ljiljana.lazarevic@f.bg.ac.rs Phone: +381 64 1287 614

Author contribution: to be filled in

Conflict of interest statement: The authors report no conflict of interest.

Funding statement: This project was funded in part by NSF Award 2141930 to J. Wagge.

Ethics committee approval: This research was approved under an “umbrella” ethics proposal at the Institutional Review Board of the Department of Psychology, University of Belgrade (Protocol #2021-069). At each site where there was a local ethics board and where ethics approval was required, the local board approved the study (all ethics approvals can be found on the project pages from the individual sites: <https://osf.io/4bfwu/>). This study complied with the ethics code outlined in the Declaration of Helsinki.

Acknowledgment: The authors would like to thank all students, supervisors, and reviewers who contributed to this CREP project. The authors would like to specifically thank Drs. Jon Grahe, Mark Brandt, and Hans IJzerman for their vision and early dedication to the CREP.

Introduction

The Collaborative Replications and Education Project (Wagge et al., 2019b), or CREP, is one of several initiatives that seeks to improve undergraduate training in research methods. CREP’s role includes selecting studies to be closely replicated by students, structuring the methods for the replication, reviewing projects both before and after data collection, and facilitating the publication of a pooled analysis (such as the present paper). While CREP has a pedagogical mission, we hope to also advance science by providing additional evidence for the boundaries of published effects.

CREP participants are students and instructors at institutions around the world. We offer CREP as one alternative to the traditional undergraduate research project; while different models may work better for some instructors, institutions, and students, the CREP offers some benefits that other models may not. First, students learn methods by closely matching the work of scholars in the field. Second, students get to participate in authentic research (Grahe, CITE) that may eventually be published in a pooled analysis and will also be available on the Open Science Framework (osf.io) to meta-analytic researchers in the future. Third, students learn the importance of many key open science practices and issues such as preregistration, replication, open methods, and open data (Kidwell et al., 2016). Fourth, students engage with reviewers (CREP team members) external to their institution. Several CREP studies have been published (e.g., Ghelfi et al., 2020; Leighton et al., 2018; Wagge et al., 2019a) and others have been included in meta-analyses (Lehman et al., 2018).

The present report documents a pooled analysis of data collected by teams who signed up to replicate one of CREP's earliest selections for replication: Experiment 1 from Griskevicius et al. (2010).

The original study: Status competition and pro-environmental behavior

Can pro-environmental behavior be promoted by inducing status competition? Griskevicius and colleagues (2010) tried to give an answer to this question and reported the results of three experiments suggesting that status competition can be used to promote pro-environmental behavior. Namely, the authors of the original study showed that activating status motives incites people to choose green products over more luxurious non-green products. In other words, Griskevicius and his associates claimed that green purchases are motivated by competitive altruism, that is the notion that people are trying to appear more altruistic when competing for status. Showing publicly pro-environmental behavior suggests people care, these altruistic tendencies are positively valued by others and it gives people prestige and status. These authors claimed that people are even more motivated to shop for green products when they are costly, and not when these products are cheaper than luxurious products. To conclude, the original study suggested that status competition can promote pro-environmental behavior.

In CREP, the first experiment from the Griskevicius et al (2010) study was selected. Experiment 1 investigated how activating a motive for status influenced respondents to choose between relatively luxurious non-green and less-luxurious pro-environmental products. Importantly, the prices of both groups of products were equal. The experiment included three conditions, i.e., condition activating status motives and two control motive conditions. Non-green products were selected so they are more desirable than their green counterparts. In the experimental condition participants read a cover "status" story that elicited "a desire for social status", and it was predicted that it will increase the likelihood of choosing prosocial, green products. In the first control condition, participants read a story that elicited similar levels of affect but did not elicit status motives. It was predicted that participants will select non-green products more frequently in the control motive condition. The second control condition did not include a cover story, and participants simply selected among products. In the original paper, using a one-way analysis of variance (ANOVA) on the product composite found a large effect of status prime ($d=0.47$). To summarize, the first experiment of Griskevicius et al (2010) suggested that activating status motives inclines people towards choosing pro-environmental products over more luxurious nongreen products that might signal prosocial, self-sacrificing behavior.

Method

CREP Procedures

This study was selected for replication using the process outlined in Wagge and colleagues (2019). Briefly, in 2014, CREP volunteers located citations for the most-cited 2010 papers from the flagship empirical journals in psychology and its subdisciplines. In this way, CREP was able to narrow down the pool for selection while also guaranteeing that the paper we selected was both recent and high-interest. The goal was not to find the **best** paper for replication in terms of the field's specific needs, but rather to find a set of recent, high-interest papers, that is, **good** candidates for replication. Once the pool of citations had been gathered, CREP coders read the Method section for each paper and indicated whether it was feasible for student teams to replicate in a semester on a scale of 1 (low) -5 (high) and also whether they thought it would be interesting to students, again, on a scale of 1 (low) -5(high). The feasibility ratings were used to narrow down a final pool, while the interest ratings were used in the final decision-making process if all other factors weighted equally. Documentation of the coding process can be found here: <https://osf.io/9kzje/wiki/home/>. Once a final pool of papers was narrowed down, CREP recruited psychology professors to look again at the feasibility and interest to narrow it down more, and then the CREP leadership team engaged in internal conversation to select the study or studies that would be selected that year. The process in 2014 resulted in several selected studies, including Experiment 1 from Griskevicius and colleagues.

Once this study was selected, the CREP leadership team sent an email to Dr. Griskevicius as the corresponding author of the original work. The CREP team communicated our process and goals and asked for input on

possible extension hypotheses and guidance for replication teams. The key parts of the correspondence from Dr. Griskevicius are available here: <https://osf.io/vdo0i/wiki/home/>. Two key points from this correspondence will be discussed in the context of our results in this paper. First, Dr. Griskevicius noted that the effect may not replicate if participants didn't equate "green" choices with prosocial behavior. Second, Dr. Griskevicius noted that the connection between "green" choices and status was unique to politically liberal groups. The first note will provide us with a lens through which we will interpret our results, while the second provides some context for why many student teams included political ideology in extension hypotheses.

CREP Project Process

Groups signed up for the project and prepared OSF project pages for pre-data collection review. Prepared pre-data collection project pages had to include materials, analytical strategy, video of procedure, and IRB approval. Project pages were reviewed by two reviewers and the executive reviewer. After revisions, project pages were reviewed again and after obtaining positive feedback, groups were cleared for data collection. Groups had to pre-register project pages before data-collection. After data-collection was completed, project pages were reviewed again by two reviewers and the executive reviewer. At this stage, project pages were revised to include the dataset, a short report about the obtained results, and a completion pledge. Following a positive review, replication was considered successful and the project was completed. Since 2013, 49 groups from nine different countries expressed interest in conducting CREP Griskevicius et al (2010). Eleven groups did not provide their data and one group did not provide the codebook with us, five groups did not complete data collection, three groups did not create the OSF page, four groups dropped out before data collection, and one group did not follow the CREP procedure. The final sample of completed projects included data collected by 24 groups from 6 different countries: USA, UK, Germany, Canada, Netherlands, and Iceland. The overview of groups participating in this project is provided in Table XX in Supplementary materials (<https://osf.io/vc5rh/>) (XXX THIS LINK IS BAD). Over the years, this particular project included about 30 reviewers, 3 CREP assistants, and 3 executive reviewers.

Target sample size The sample size for each group was set to 82 respondents. In CREP, for each project, the sample size was determined to be at least half of the original N. For this study, the targeted sample size was set to 84, but unfortunately, in the documentation file, there was a typo (N=82). However, with multiple groups collecting data, and in this particular case with 3774 participants, we have adequate statistical power to detect a very small focal effect.

Differences from the original study The original study was conducted in a lab, where participants were tested in small groups. Replications were conducted both in the lab and online. Some groups conducted direct replication and some groups included extension variables. Characteristics of the replications are given in Table XX in Supplementary materials (<https://osf.io/7zybp/>).

Disclosures

Preregistrations Each lab preregistered its materials, protocol, and analytical strategy on Open Science Framework (OSF) before data collection. Additionally, this meta-analysis was pre-registered: <https://osf.io/ach3n>.

Data, Materials and Resources The authors of the original study provided materials for replications. All groups who completed data collection, uploaded data, analysis, and a short description of results to their OSF project page. All datasets, materials, analytical scripts, and other materials can be found on the central project OSF page (<https://osf.io/rh2nw/>). The complete codebook for the dataset is available at <https://osf.io/7nkhq/>.

Reporting We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

Ethical approval Data were collected in accordance with the Declaration of Helsinki. This research was approved under an “umbrella” ethics proposal at the University of Belgrade, Serbia. Overarching IRB approval was obtained from the IRB of the Department of Psychology, University of Belgrade, Serbia (<https://osf.io/pcwhg/>). In addition, all groups obtained local IRB approvals which are available on our OSF project page: <https://osf.io/4bfwu/>

Sample

Data were collected by student groups in USA, UK, Germany, Canada, Netherlands, and Iceland who participated in the CREP project. In total, the merged data set contained data collected from 3774 participants. We use datasets collected by 24 groups. All individual data sets are available here: <https://osf.io/vc5rh/>. All groups collected at least 82 respondents, except for one that collected 70 participants. Nine groups did not collect data on at least one of the following: age, gender, or race/ethnicity. Details about the missing data are available here (XXX- do we need a link here?). The complete dataset is available here: <https://osf.io/azq76/>.

We removed 21 participants prior to posting the public compiled dataset because they reported ages of 16 ($n = 3$) or 17 ($n = 18$), for a total final dataset with 3774 participants. The final sample included 2302 women and 973 men with 18 participants reporting a gender of gender variant/nonconforming, “other,” “prefer not to say,” or transgender/non-binary. An additional 481 participants did not indicate their gender, either by choice or because the survey they completed did not request this information. The mean reported age for participants was 23.98 years ($SD = 10.21$). 956 participants did not report their age.

Information about participant race/ethnicity was reported by some sites and not others, and within each site the question could vary in how it was asked. We therefore recoded the original race information to include the following categories: White ($n = 1339$, 71.34%), Black ($n = 129$, 6.93%), Asian ($n = 131$, 6.98%), Hispanic/Latino ($n = 151$, 8.04%), Middle Eastern ($n = 28$, 1.49%), Multiracial ($n = 8$, 0.43%), Native Hawaiian or Pacific Islander ($n = 5$, 0.27%), and “other” ($n = 71$, 3.78%). These reported percentages do not include the participants who either did not report their race/ethnicity, or were not asked this question ($n = 1897$, 50.26% of total).

Procedure

Data were collected in the lab, or online. Each group documented the type of data collection. Respondents were recruited from the student or general population. Each group underwent a pre-data-collection review procedure by CREP reviewers and executive reviewers and made preregistration of their project. After data collection was completed, project pages were again reviewed (post-data-collection review) by the executive reviewer and CREP reviewers. Descriptions of individual datasets, along with all deviations are available in Table XX in Supplementary materials (<https://osf.io/vc5rh/>).

Treatment of missing data, data preparation, and data analysis

JORDAN TO-DO: work on this section, address comments from Nate: “What does this mean? What metrics were used to determine if the data from a given study was of a sufficient quality to be included in the analysis? Were there any datasets that did not surpass these metrics? Also, how does the online versus in lab data collection impact the data quality? Were there any methods to ensure online data collection consisted of high quality respondents? Were online participants students from universities, or recruited from crowdsourcing platforms?”

We excluded participants without responses for the dependent variables. To create a merged dataset, we accessed all individual datasets through each team’s OSF page when this was available, and contacted teams directly for their data when it was not. When necessary, we also contacted groups to obtain information on whether the dataset was raw or transformed and to obtain codebooks. We left missing data as missing. The merged dataset is available on the OSF project page (**ADD LINK HERE**). A codebook for our merged data set with all variables collected as part of the project is available on <https://osf.io/7nkhq/>. This link includes number of participants, country, institution, race, gender, conditions, and any extension variables added by each team.

To note, the original study conducted data collection in a computer lab and tested participants in small groups; therefore, we use ANOVA to explore differences between settings (online/lab, testing in-group/testing individually). We also conduct the same analyses as the original paper: for each product (car, soap, and dishwasher) we report percentages by condition, Chi-square tests, and Phi coefficients. We also compute a composite green score and run ANOVA across conditions (story, no story, and control story) and report effect size and posthoc comparisons using the LSD test.

In our exploratory analyses, we test the moderating role of political orientation. We also run 2 x 2 factorial ANOVAs to determine whether political orientation (liberal/conservative) or political party (democrat/republican) interacts with the condition in its association with the composite green score.

Results

Confirmatory Analyses

Composite Green Score Despite low internal consistency of the three items (Cronbach’s $\alpha = 0.38$), we followed the procedure used in the original study and computed a composite green score by assigning the value “1” to all “green” selections and a score of “0” to all non-“green” selections. Because participants had three dichotomous choices for three products, scores ranged from 0 (no “green” products) to 3 (all “green” products). The mean composite score was 1.64 ($SD = 0.996$). A distribution of composite scores can be seen in Figure 1.

The original paper by Griskevicius and colleagues reported a significant effect of status on the composite score when the status prime condition was compared to both control conditions (control story and no story) together, $F(1, 166) = 8.53, p = .004, d = 0.47$. The same test using our data did not reveal a significant effect, $F(1, 3, 718) = 1.16, p = .282, \eta^2 = .000$. However, when we examined our three conditions (control story, status story, and no story), a small effect was found, $F(2, 3, 717) = 3.21, p = .040, \eta^2 = .002$. Post-hoc testing using Tukey’s HSD correction for multiple pairwise comparisons resulted in no pairwise comparisons with a $p < .05$ (no story/control = 0.06, status story/control = 0.97, status story/no story = 0.09).

The rest of our analyses combine the “control story” and “no story” conditions into one control group.

Green car The original paper found that the green car selection was higher in the status condition (54.5%) than in the control condition (37.2%), $\chi^2(1, N = 168) = 4.56, p = .033, \phi = .165$. In our sample, 55.5% of control participants selected the green car, compared to 54.83% of the status participants. We were unable to detect an effect of condition on green selection, $\chi^2(1, N = 3726) = 0.13, p = 0.72, \phi = -0.006$.

Green cleaner Similarly, the original paper found that participants in the status condition selected the green cleaner more frequently than participants in the control condition (41.8% and 25.7%, respectively), $\chi^2(1, N = 168) = 4.52, p = .034, \phi = .164$. In our sample, 56.75% of the control and 54.95% of the status participants selected the green cleaner. As with the “car” selection, we were unable to detect an effect of condition on choice of cleaner, $\chi^2(1, N = 3725) = 1.06, p = 0.3, \phi = -0.018$.

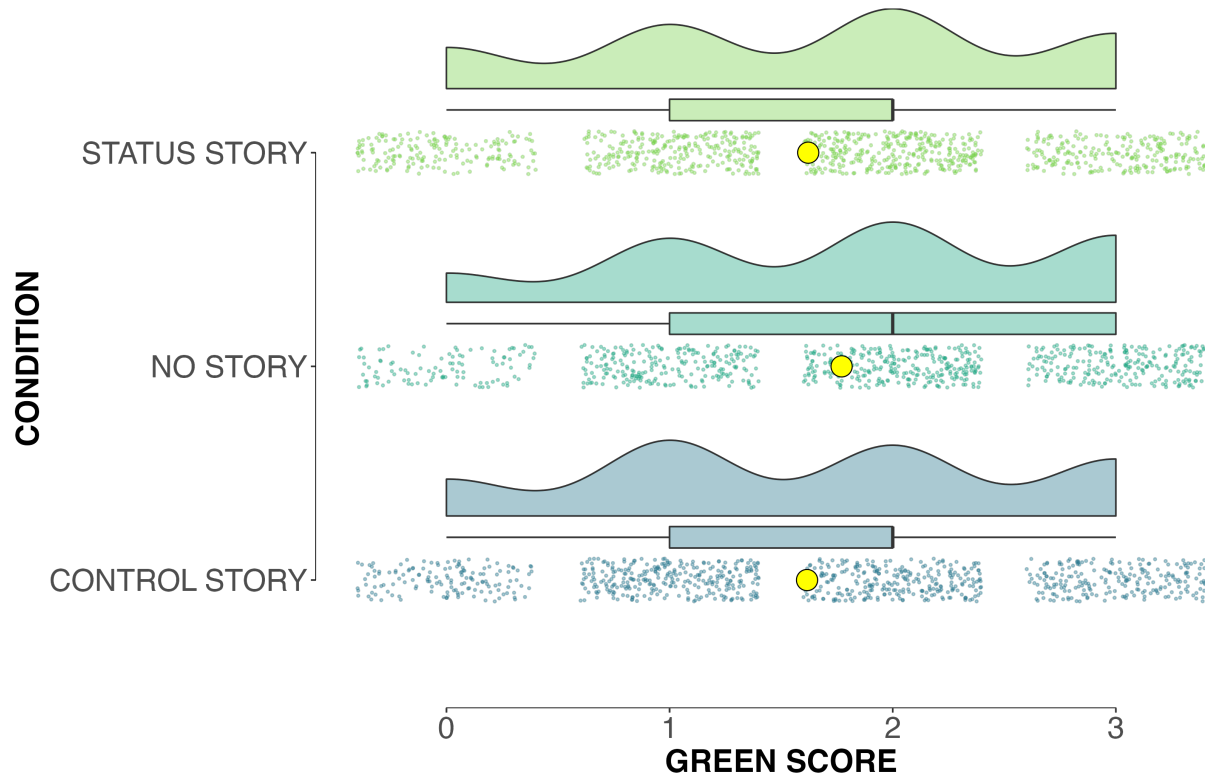


Figure 1: **Figure 1.** The rain cloud diagram shows the distribution of responses under different conditions. Yellow dots represent mean values surrounded by individual data points with jitter. Vertical lines in the centers of the boxplots indicate the medians. Shaded areas around them correspond to interquartile ranges, and whiskers encompass 1.5 times the interquartile ranges. The shaded areas above the boxplots indicate the density of the data.

Green dishwasher Finally, the original paper also reported greater that selection of the green dishwasher was more likely in the status condition than the control condition (49.1% and 34.5%, respectively), $\chi^2(1, N = 168) = 3.30, p = .069, \phi = .140$, while we found no such effect when comparing the 52.9% of control participants who selected the green dishwasher to the 51.65% of status prime participants who did the same, $\chi^2(1, N = 3729) = 0.49, p = 0.48, \phi = -0.012$.

Exploratory Analyses

Because the original study was conducted only with individuals in the United States, we conducted an exploratory test of the data to determine whether there was an association between condition and green choice for only US-based participants. For US-based participants, we found no association between condition (two levels: control and status) and green car selection ($\chi^2(1, N = 1940) = 2.58, p = 0.11, \phi = 0.04$), green cleaner selection ($\chi^2(1, N = 1939) = 1.02, p = 0.31, \phi = 0.02$), or green dishwasher ($\chi^2(1, N = 1943) = 0.01, p = 0.92, \phi = 0$) selection.

To see if there was an effect that differed across time, we examined effects across projects from different years, as determined by their project codes assigned at the time they signed up for the study (e.g. project #14-3 was initiated in 2014). For each project year, we examined whether we could detect any differences in composite green scores across conditions (control and status). As demonstrated by Figure 2, there doesn't appear to be any meaningful change in effect size over time, such that an effect could be detected at an early point in time but not later or vice versa.

To determine if there were differences in setting type, we first collapsed all reported lab settings (group, individual, and just "lab") into one variable level, and then conducted chi-square tests of independence using a dichotomous setting variable (lab versus online). We found no relationship between green choice and setting type for all three types of product (ps of .24, .11, and .63 for the green car, soap, and dishwasher).

When the CREP team first contacted the original authors about the replication, they recommended examining political orientation as a possible extension variable. Therefore, many groups decided to add a question about political orientation (liberal/conservative) and others added a question about political party (Republican/Democrat/Independent); some groups added both questions, and some groups added questions with responses that were more fitting to their particular country (e.g., Canada). We decided to use this data to compare liberal/conservative on green selections. Some institutions measured this with a scale (from very liberal to very conservative, for example), while others measured it with categorical response types (e.g., liberal, conservative, neutral). We collapsed over these responses to create a variable with one category for "liberal" that included all responses indicating any degree of being liberal, another category for "conservative" that included any response indicating being conservative, and excluded any "neutral" or "other" responses from this variable.

Using this method of collapsing data, we found that $n = 551$ participants identified as conservative, while $n = 744$ participants identified as liberal.

Liberal/conservative We conducted a 2×2 factorial ANOVA (political orientation: liberal vs. conservative; condition: control vs. experimental) to determine whether political orientation interacted with condition in its association with the composite green score. As predicted, we did find a main effect of political orientation such that participants who identified as liberal selected significantly more green products on average ($M = 1.87, SD = 0.98$) than participants who identified as conservative ($M = 1.25, SD = 0.97$), $F(1, 1, 287) = 130.73, p < .001, \eta^2 = .092$. As we found earlier, there was also no main effect of condition; the mean scores for participants in the grouped control condition ($M = 1.65, SD = 0.99$) did not differ from those in the status condition ($M = 1.61, SD = 1$), $F(1, 1, 287) = 0.41, p = .522, \eta^2 = .000$. There was also no interaction between the two variables, $F(1, 1, 287) = 0.28, p = .597, \eta^2 = .000$.

NOTE from Michał re: grouping control conditions: but still I would report whether there was a difference between no story and control; I suggest reporting at least p-values and effect sizes

Effect of Condition (Control/Status) on Composite Green Score by Project Initiation Year

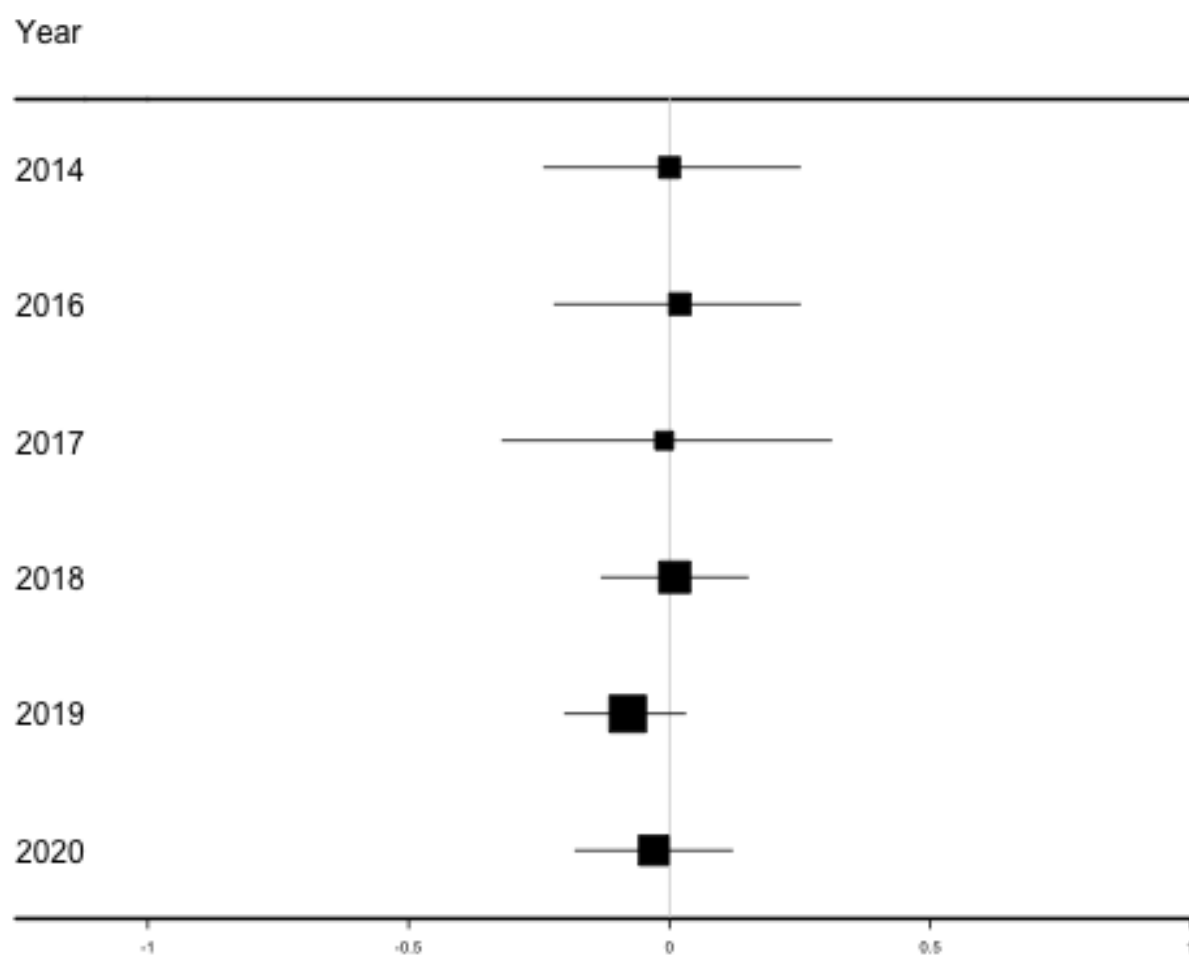


Figure 2: **Figure 2.** Forest plot caption stuff

Democrat/republican Similarly, for political party (“Democrat” or “Republican”), we conducted a 2 x 2 factorial ANOVA (political party: Democrat vs. Republican; condition: control vs. status). When we grouped participants into “Republican” ($n = 182$) and “Democrat” ($n = 332$) categories, we did find a main effect of political party such that participants who identified as Democrat selected significantly more green products ($M = 1.77$, $SD = 1.03$) than participants who identified as Republican ($M = 1.12$, $SD = 0.86$), $F(1, 505) = 51.95$, $p < .001$, $\eta^2 = 0.09$. We were unable to detect an interaction between political party and condition, $F(1, 505) = 0.14$, $p = .709$, $\eta^2 = 0$, and as reported in the “Liberal/Conservative” subsection above, we found no main effect of condition (control: $M = 1.65$, $SD = 0.99$); status: $M = 1.61$, $SD = 1$), $F(1, 505) = 0.80$, $p = .370$, $\eta^2 = 0$.

Importantly, we did not have adequate power to detect any interaction effects. Using the Superpower package in R, we determined that we would need at least 418 participants *per cell* to detect an interaction effect between political party and condition. In our four cells created by the two variables political party (Republican or Democrat) and condition (status or control), our cell sizes were 209, 113, 123, and 69.

Discussion

Original study conducted by Griskevicius et al (2010) concluded that proenvironmental behavior can be promoted using status competition. That is, earning a good reputation via prosocial and proenvironmental behavior and action can increase an individual’s status in a group. Original study also showed that all selection of all three green-products (green car, green cleaner, green dishwasher) was higher in the status condition.

In this multi-lab replication study we followed the analytical strategy as in the original study. As part of pre-registered confirmatory analysis we tested whether effect exists when green score is calculated as composite, and our results resulted in lack of statistically significant effect. We also analyzed separate if effect exists for each of the green vs non-green products. For all three products, we did not detect significant effect for none of the products. Overall, our findings do not support main conclusion of Griskevicius et al (2010).

Per recommendation of original authors, exploratory analyses included testing if political orientation moderates significantly main effect. In this study, data on political orientation and/or political party were collected as extension variables. When political orientation was defined as liberal-conservative we detected a main effect of political orientation on selection of green products, but the main effect of condition was absent. When political party was tested as moderator, a main effect of political party was detected (i.e., people who identified as democrat more often select green products), but moderation effect of political party on product select was absent again.

Add: discuss reliability of composite score

Discuss that only some moderators had sufficient number of data collected to be analyzed

Discuss differences -online and lab

2 vs 3 conditions

What else? - effect over time, US versus rest

References

- Frank, M. C., & Saxe, R. (2012). Teaching Replication. *Perspectives on Psychological Science*, 7(6), 600–604. <https://doi.org/10.1177/1745691612460686>
- Ghelfi, E., Christopherson, C. D., Urry, H. L., Lenne, R. L., Legate, N., Ann Fischer, M., Wagemans, F. M. A., Wiggins, B., Barrett, T., Bornstein, M., de Haan, B., Guberman, J., Issa, N., Kim, J., Na, E., O’Brien, J., Paulk, A., Peck, T., Sashihara, M., . . . Sullivan, D. (2020). Reexamining the effect of gustatory disgust on moral judgment: A multilab direct replication of Eskine, Kaciniak, and Prinz (2011). *Advances in Methods and Practices in Psychological Science*, 3(1), 3–23. <https://doi.org/10.1177/2515245919881152>

- Grahe, J. E., Reifman, A., Hermann, A. D., Walker, M., Oleson, K. C., Nario-Redmond, M., & Wiebe, R. P. (2012). Harnessing the undiscovered resource of student research projects. *Perspectives on Psychological Science*, 7(6), 605–607. <https://doi.org/10.1177/1745691612459057>
- Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: Status, reputation, and conspicuous conservation. *Journal of Personality and Social Psychology*, 98(3), 392–404. <https://doi.org/10.1037/a0017346>
- Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., Kennett, C., Slowik, A., Sonnleitner, C., Hess-Holden, C., Errington, T. M., Fiedler, S., & Nosek, B. A. (2016). Badges to acknowledge open practices: A simple, low-Cost, effective method for increasing transparency. *PLOS Biology*, 14(5), e1002456. <https://doi.org/10.1371/journal.pbio.1002456>
- Lehmann, G. K., Elliot, A. J., & Calin-Jageman, R. J. (2018). Meta-analysis of the effect of red on perceived attractiveness. *Evolutionary Psychology*, 16(4), 1474704918802412. <https://doi.org/10.1177/1474704918802412>
- Leighton, D. C., Legate, N., LePine, S., Anderson, S. F., & Grahe, J. (2018). Self-esteem, self-disclosure, self-expression, and connection on Facebook: A collaborative replication meta-analysis. *Psi Chi Journal of Psychological Research*, 23(2), 98–109. <https://doi.org/10.24839/2325-7342.JN23.2.98>
- Perlman, B., & McCann, L. I. (1999). The most frequently listed courses in the undergraduate psychology curriculum. *Teaching of Psychology*, 26(3), 177–182. <https://doi.org/10.1207/S15328023TOP260303>
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. <http://www.rstudio.com/>
- Standing, L. G., Grenier, M., Lane, E. A., Roberts, M. S., & Sykes, S. J. (2014). Using replication projects in teaching research methods. *Psychology Teaching Review*, 20(1), 96–104.
- Wagge, J. R., Baciú, C., Banas, K., Nadler, J. T., Schwarz, S., Weisberg, Y., IJzerman, H., Legate, N., & Grahe, J. (2019). A demonstration of the Collaborative Replication and Education Project: Replication attempts of the red-romance effect. *Collabra: Psychology*, 5(1), 5. <https://doi.org/10.1525/collabra.177>
- Wagge, J. R., Brandt, M. J., Lazarevic, L. B., Legate, N., Christopherson, C., Wiggins, B., & Grahe, J. E. (2019). Publishing research with undergraduate students via replication work: The Collaborative Replications and Education Project. *Frontiers in Psychology*, 10, 247. <https://doi.org/10.3389/fpsyg.2019.00247>
- Wagge, J. R., Hurst, M. A., Brandt, M. J., Lazarevic, L. B., Legate, N., & Grahe, J. E. (2022). Teaching research in principle and in practice: What do psychology instructors think of research projects in their courses? *Psychology Learning & Teaching*, 147572572211019. <https://doi.org/10.1177/14757257221101942>