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

Jordan Wagge

Last Updated 2023-12-20

**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](mailto: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. 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. Mark Brandt and Hans IJzerman for their vision and early dedication to the CREP.

**Author contribution:**

Conceptualization: LBL, JRW, JG, and SL

Data curation: LBL, JRW, BB, VS, MF, and JG

Formal analysis: JRW, MF, and AL

Funding acquisition: JRW

Investigation: SCW, PD, AJK, KT, JT, JHF, SCW, NS, EH, LB, CDC, AG, JR, KS, and ADP

Methodology: LBL, JRW, and MF

Project Administration: LBL, BB, SCW, AJK, JHF, SCW, MBP, EH, JG, and CDC

Resources: SCW, PD, AJK, KT, JT, JHF, SCW, EH, LB, JG, AG and ADP

Supervision: LBL, VS, SCW, AJK, SCW, EH, JG, CDC, and MCR

Validation: JRW, BB, and VS

Visualization: JWR, MF

Writing - Original draft: LBL and JWR

Writing - Review and Editing: LBL, JWR, MF, AL, SCW, PD, AJK, KT, JT, SL, JHF, SCW, NS, MBP, EH, LB, JG, CDC, BB, MR, NRG, and AG

**Abstract** The present study presents results of a collaborative registered replication of Griskevicius et al.’s (2010) findings from Experiment 1 conducted by 24 student groups from six countries (N=3774) investigating if pro-environmental behavior can be promoted by inducing status motives. Student projects were part of the Collaborative Replication and Education Project. Results of this large multi-site replication showed no evidence to support the hypothesis that pro-environmental behavior can be stimulated by inducing status motives. We performed several exploratory analyses to investigate whether extension variables (i.e., equating “green” choices with prosocial behavior, political beliefs, sampling methods, US vs. other countries, duration of data collection and gender) moderated the effect of status motives on pro-environmental choices, but our findings reproduced null results.

## Introduction

The Collaborative Replications and Education Project (Wagge, Brandt et al., 2019), or CREP, is an initiative seeking 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). Replications are regarded as a gold standard warranting the reliability of published findings (Open Science Collaboration, 2015). In addition, it was found that conducting a replication of highly-cited studies is perceived as exciting by students, and it also gives them a real opportunity to make substantial contributions to the field (Grahe et al., 2012; Frank et al., 2012).

CREP participants include students and instructors at institutions worldwide[[1]](#footnote-21). CREP is an alternative to the traditional undergraduate research experience offering several advantages. First, the CREP model allows students to learn about best practices by directly replicating published work by scholars in their field. This active learning experience—also known in teaching literature as mastery experience or high-impact learning practice—can be an important predictor of academic achievement (Loo & Choy, 2013). Second, students get to participate in authentic scientific research (Grahe, 2017) that may eventually be published in a pooled analysis and will also be openly available on the Open Science Framework (which is an open-source management tool that facilitates open and transparent collaboration in scientific research) to meta-analytic researchers in the future. Third, students have the option of performing a “replication+” study to add their own additional research question, condition, or measure to the replication study (assuming it does not interfere with the direct replication). Furthermore, students acquire an understanding of the significance of open science practices such as preregistration, replication, open methods, and open data, all foundational to research integrity (Kidwell et al., 2016). They actively engage with external reviewers (i.e., CREP team members), collaborating and learning beyond their institution through engagement with CREP members. Several CREP studies have been published (e.g., Ghelfi et al., 2020; Leighton et al., 2018; Wagge, Baciu, et al., 2019), with some included in meta-analyses (Lehmann et al., 2018). This underscores the project’s significant contributions to psychological science.

The present report documents a pooled analysis of data collected by 24 teams who signed up to replicate one of CREP’s earliest selections for replication: Experiment 1 from Griskevicius et al. (2010). This collaborative approach not only enhances the robustness of our findings but also exemplifies the strength of collective efforts in advancing transparent and replicable research practices within the scientific community. In addition, the methodology employed in this pooled analysis underscores the commitment to fostering a culture of openness and team building in psychological research.

### The original study: Promoting pro-environmental behavior through status competition

Griskevicius and colleagues (2010) reported the results of three experiments suggesting that status competition can promote pro-environmental behavior. Per CREP procedures (see details in Method - CREP Procedures), only one study in a multi-study paper is typically selected for replication. In this case, the CREP team selected the first experiment from Griskevicius et al. (2010) as it was the focal study that was conceptually replicated and extended in their subsequent experiments (see Method for justification). Specifically, the original authors showed that activating status motives (desires for social status and prestige) leads people to choose “green” (environmentally-friendly) products over more luxurious non-green products (Study 1). The authors predicted that reading a story that activated status motives (e.g., a short story about obtaining a successful job) would increase the likelihood of choosing pro-environmental (green) products. In control conditions that did not activate status motives (e.g., reading a story of a similar length about losing tickets to an upcoming concert, or reading no story at all), non-green products were predicted to be more desirable than their green counterparts. Participants were assigned to either the status story or one of the two control conditions (non-status story, or no story). After reading their respective stories, participants were asked to select between green products (e.g., a low-emission car) and non-green products (e.g., a more luxurious sports car). Participants in the no-story control condition proceeded to the product selection phase of the study immediately. A composite “green” score was calculated by adding up the number of “green” choices the participant made (no green choices = 0; all green choices = 3). Participants whose status motives were induced were substantially more likely to choose green products (*d* = 0.47) compared to participants in the control conditions. In summary, Griskevicius and colleagues (2010) found that activating status motives inclines people to choose pro-environmental products over more luxurious non-green products. Theoretically, this effect may reflect efforts to signal prosocial, self-sacrificing behavior. The authors interpreted their findings with respect to the “competitive altruism” hypothesis (see Van Vugt et al., 2007), which suggests that individuals attempt to appear more altruistic when competing for status. Publicly demonstrating pro-environmental behavior signals to others that you prioritize environmental concerns, indicating a willingness to incur costs (such as opting for a less luxurious product at a higher price), for the benefit of others. These altruistic tendencies are highly regarded by society, conferring prestige and status, as outlined by Griskevicius and colleagues (2010).

The present multi-lab study aims to assess the robustness of the effect observed by Griskevicius et al. (2010) in their Experiment 1 and to provide a more precise estimation of the effect size. Successful replication would entail that respondents in the social status condition would choose the less-luxurious but green product over the more luxurious but non-green alternative. Since its publication 15 years ago, the original study continues to garner significant attention, with an impressive citation count of almost 2,500 and an average of nearly 300 citations per year since 2020, according to Google Scholar data. This sustained impact on the field further warrants the need to replicate these findings in a larger, multi-site study.

## Method

### CREP Procedures

Experiment 1 of Griskevicius et al. (2010) was selected for replication by following the standard CREP selection process, outlined by Wagge and colleagues (2019). This process, which took place in 2013, resulted in the selection of several studies, including this experiment.

Once the focal study was selected, the CREP leadership team emailed the corresponding author of the original work. The CREP team communicated their process and goals and asked for input on possible extension hypotheses and guidance for replication teams. The key components of this correspondence are available here: <https://osf.io/vdo0i/wiki/home/>. There are two key points from the correspondence. First, the main author noted that the effect may not replicate if participants do not equate “green” choices with prosocial behavior. Second, the original first author noted that the connection between “green” choices and status was unique to politically liberal groups. The former will provide us with a lens through which we will interpret our results, while the latter explains why many student teams included political ideology in their extension hypotheses.

Between 2014 and 2020, groups signed up for the project and prepared OSF project pages for pre-data collection review. These pre-data collection project pages had to include materials, a planned analytical strategy, a video documenting the procedure, and local Institutional Review Board (IRB) approval. The project pages were examined by two reviewers and the executive reviewer, at which point projects were either approved or sent back for revisions until approval could be granted. Groups were asked to pre-register project pages before data collection. After data collection was completed, project pages were again reviewed by two reviewers and the executive reviewer. At this stage, project pages were revised to include the dataset, a short report describing the obtained results, and a signed completion pledge confirming that the research followed ethical guidelines and the predefined study protocol. Following a positive review, individual projects were deemed complete.

#### Target sample size

Per early CREP policies, the required sample size was set as a rule of thumb to be at least half of the original *N*. For this study, the targeted sample size should have been 84 (half of the original *N* of 168), but due to a typo in the documentation file the target sample size for each group was set to 82 respondents.

#### 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 replications, and others included extension variables (e.g., pro-environmental literacy, political ideology), but always after the original protocol was completed. Characteristics of the replications can be found at <https://osf.io/gt39p>.

### Disclosures

#### Preregistrations

Each lab preregistered its materials, protocol, and analytical strategy on the OSF before data collection. Additionally, this analysis was pre-registered: <https://osf.io/ach3n>.

#### Data, Materials, and Resources

The authors of the original study provided materials for replication. All groups who completed projects uploaded their data, analyses, and a short description of their 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, anonymized dataset is available here: <https://osf.io/3xrdf>. A codebook for our merged dataset with all variables collected as part of the project is available at <https://osf.io/7nkhq/>.

#### Reporting

A description of the sample size, data exclusions, manipulations, and measures are reported in the study.

#### Ethical approval

Data were collected in accordance with the 1964 Declaration of Helsinki. Overarching IRB approval for this study was obtained from the IRB of the Department of Psychology, University of Belgrade, Serbia (<https://osf.io/pcwhg/>). In addition, as part of the CREP procedure, all groups obtained local IRB approval when necessary.

### Sample

From 2014 to 2020, 49 groups from nine different countries expressed interest in conducting a CREP replication of Griskevicius et al. (2010) Experiment 1. Of these 49 groups, 25 did not contribute data to the final analysis: specifically, three did not create their OSF page, four did not begin data collection, five did not complete data collection, 11 did not provide their data, one did not provide their codebook, and one did not follow the CREP procedure. The final sample of completed projects included data collected by 24 groups from six countries: USA (*N* = 18), UK (*N* = 1), Germany (*N* = 1), Canada (*N* = 2), Netherlands (*N* = 1), and Iceland (*N* = 1). The overview of groups participating in this project can be found at <https://osf.io/vc5rh>. Overall, the present project included approximately 30 reviewers, three CREP assistants, and three executive reviewers.

A total of 21 participants were excluded prior to publicly posting the compiled dataset because they reported ages of 16 (*N* = 3) or 17 (*N* = 18). The final data set consisted of data collected from 3774 participants, by 24 different student groups. All individual data sets are available here: <https://osf.io/dp3rc>. Each student group recruited at least 82 respondents, except for one group that recruited 70 (due to the COVID-19 outbreak that interfered with the project). Nine groups failed to collect data on at least one of the following demographic characteristics: age, gender, or race/ethnicity, and within each site, the question varied in how it was asked[[2]](#footnote-41). Details about the missing data are available here: <https://osf.io/vc5rh>.

The final sample consisted of 2302 individuals who identified as women and 973 who identified as men; 18 participants identified as gender variant/non-conforming, “other”, “prefer not to say,” or transgender/non-binary. An additional 481 participants did not indicate their gender. The mean reported age for participants was 23.98 years (*SD* = 10.21), though age was not collected for 956 of the participants.

We recoded the original race/ethnicity 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 the total).

### Procedure

The experiment employed a between-subjects design and included one status and two control conditions, using the exact conditions from Experiment 1 of Griskevicius et al. (2010). In the status condition, participants read a “cover” story designed to elicit “a desire for social status” (Griskevicius et al., 2010, p. 395), in which participants imagined graduating from college and obtaining a successful job with the opportunity for promotion. In one of the control conditions, participants read a control story (involving losing and eventually finding tickets to a forthcoming concert and then attending the concert with a same-sex peer) that elicited similar affect but did not activate status motives. See Griskevicius et al. (2010) for more complete details of the cover stories.

To eliminate potential suspicions, participants were instructed that they were participating in a study investigating memory and that they would be asked to recall information from the story later in the session. To ensure that some particular aspect of the story did not influence respondents’ answers, there was a second control condition in which participants were not presented with a story to read. After reading their respective story (or in the no-story control condition, participants proceeded immediately), participants were asked to choose between three pairs of products (a car, soap, and a dishwasher). For each product, participants could choose between a pro-environmental (green) and a conventional (the more luxurious nongreen) version of the product, each of which was equally priced and manufactured by the same company.

Data were collected in the lab and online, using various survey platforms; the majority (*N* = 15) of groups used Qualtrics, while three used paper surveys and only one team used each of the following: Medialab, SuperLab, SurveyGizmo, SoSciSurvey, and PsyToolkit. One team used an unidentified online platform for data collection. Each group documented the mode of data collection that they used (online vs. in-person). Respondents were recruited from both student and general populations. Descriptions of individual datasets, along with all deviations are available at <https://osf.io/t8h3k>.

Any additional variables were included after the main part of the study, that is, at the end of the survey.

### Treatment of missing data, data preparation, and data analysis

We excluded participants who did not respond to the dependent variables. To create a merged dataset, we accessed all individual datasets from each team’s OSF page or reached out directly to teams for data if it was not available on OSF. In cases where clarification was needed, we contacted groups to ascertain whether the dataset was in its raw form or transformed, and to request codebooks. Any instances of missing data were retained in the analysis.

We conducted the same analyses as the original paper: for each product (car, soap, and dishwasher) we reported percentages of “green” versus “non-green” selections by condition, Chi-square tests, and Phi coefficients. We also computed a composite green score, which we analyzed with a between-subject ANOVA across conditions (story, no story, and control story), and reported effect size and post-hoc comparisons using the LSD test.

One advantage of the multi-lab approach is that natural variation in the data collection process across different labs can be leveraged to test for moderators or boundary conditions (Kline et al., 2018). Accordingly, we also performed several exploratory analyses. First, as per the recommendation of the original authors, we tested if participants equated “green” choices with prosocial behavior which is considered fundamental for replicating results. Second, we ran 2 x 2 factorial ANOVAs to determine whether political orientation (liberal/conservative) or political party (democrat/republican) interacted with condition (control/status). Third, we ran an ANOVA to explore whether different settings (online/lab) yielded different results. Fourth, we explored whether the main effects replicated in other countries in which we collected data, and tested potential differences between the US and other countries. Fifth, the long-lasting nature of our project (data collection was done between 2014 and 2020) enabled testing whether the effects differed across time. Lastly, the original authors stated that it would be worth testing for potential gender differences; thus, as part of our exploratory analyses, we explored gender differences in status motives using a 2 x 2 factorial ANOVA (gender: woman or man; condition: control vs. status). All analyses were done using R markdown and R Studio (RStudio Team, 2020, R Core Team, 2021). The analysis code is available here: <https://osf.io/kb3ct>.

## Results

### Confirmatory Analyses

#### Green car

The original paper found that participants selected the green car more often in the status condition (54.5%) than in the control condition (37.2%), (1, N = 168) = 4.56, *p* = .033, = .165. In our sample, 55.50% of control participants selected the green car, compared to 54.83% of the status participants. We did not detect an effect of condition on green car selection, (1, *N* = 3,726) = 0.13, *p* = .720, = -0.006.

#### Green cleaner

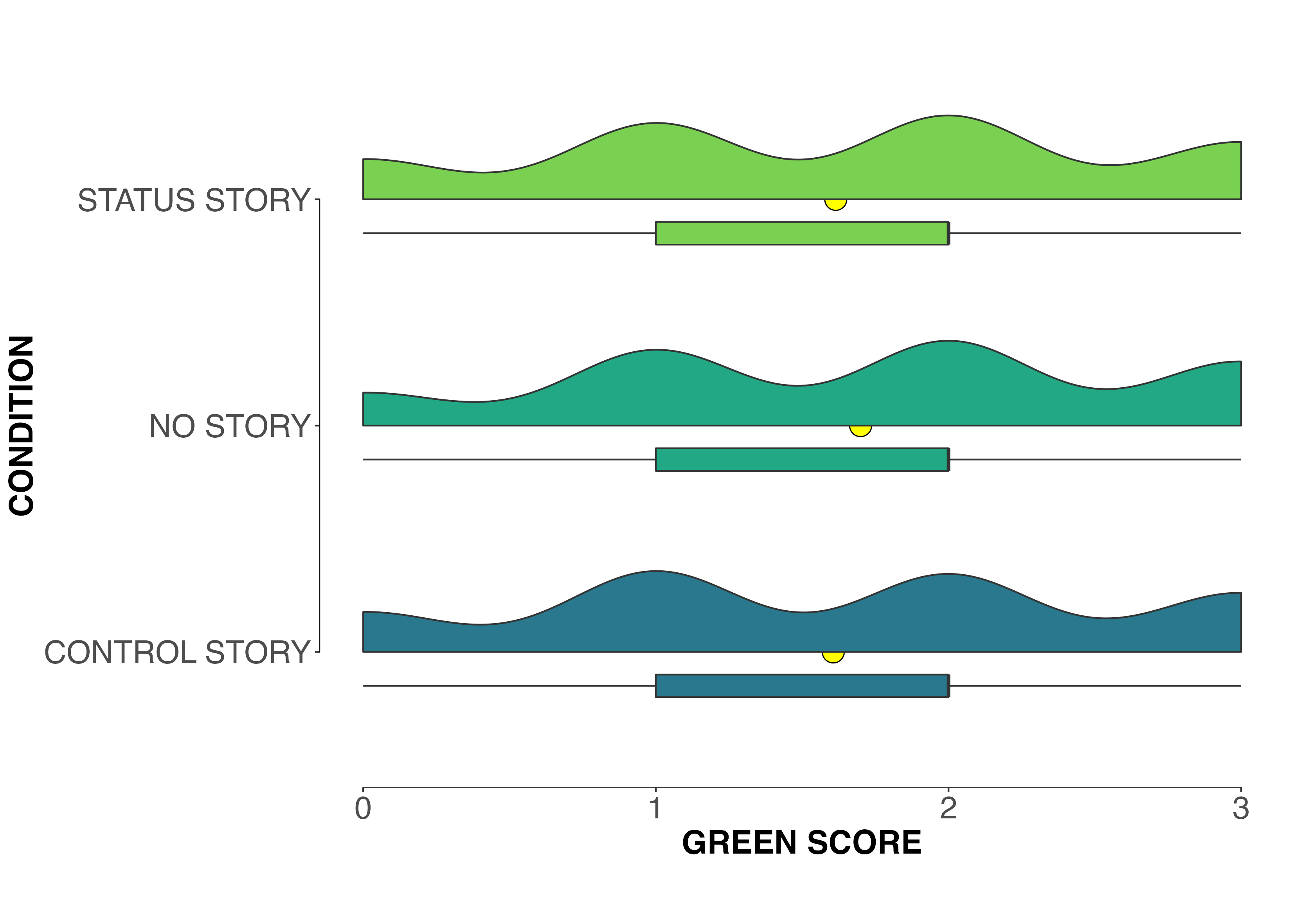
Similarly, the original paper found that participants selected the green car more often in the status condition (54.5%) than in the control condition (37.2%), (1, *N* = 168) = 4.52, *p* = .034, = .164. In our sample, 56.75% of the control and 54.95% of the status participants selected the green soap. As with the choice of car, we did not detect an effect of condition on the choice of cleaner, (1, *N* = 3,725) = 1.06, *p* = .303, = -0.018.

#### Green dishwasher

Finally, in the original paper, participants selected the green dishwasher more in the status condition than the control condition (49.1% and 34.5%, respectively), (1, *N* = 168) = 3.30, *p* = .069, = .140. We found no such effect when comparing the control participants who selected the green dishwasher (52.90%) to the status condition participants who did the same (51.65%), (1, *N* = 3,729) = 0.49, *p* = .482, = -0.012.

#### Composite Green Score

Internal consistency of the three items (i.e., products) was low (Cronbach’s alpha = 0.38) which was expected (Rammstedt & Beierlein, 2014; Malkewitz et al., 2023). Despite this, we followed the procedure used in the original study and computed a composite green score by assigning a score of “1” to all “green” selections and a score of “0” to all non-“green” selections. Because participants made a dichotomous choice for each of the three products, scores ranged from 0 (no “green” products) to 3 (all “green” products). The mean composite score was 1.64 (*SD* = 1.00). The distribution of composite scores can be seen in Figure 1.



\*\*Figure 1. The rain cloud diagram shows the distribution of responses under different conditions. Yellow dots represent mean values. Shaded areas below them correspond to 1.5 times the interquartile ranges.

The original paper by Griskevicius and colleagues (2010) 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) combined, *F*(1, 166) = 8.53, *p* = .004, *d* = 0.47. The same test on our data did not reveal a significant effect, *F* = (1, 3,718) = 1.16, = 0.000, *d* = 0.04. Notably, this composite score has low reliability with a Cronbach’s alpha of 0.38, so it is unsurprising that we did not find an effect. The original paper did not report a measure of reliability for the composite score. However, when we performed this analysis without combining the control conditions, we found a small effect, *F* = (2, 3,717) = 3.21, *p* = .040, = 0.002. Post-hoc testing using Tukey’s LSD corrections for multiple pairwise comparisons resulted in no pairwise comparisons with a *p* < .05 (no story/control story = .057, status story/control story = .974, status story/no story = .085).

### Exploratory Analyses

The original authors made two important notes regarding conditions that might influence replicability of the key effect: prosociality of “green” choices and politics. Regarding prosociality, they stressed that equating “green” choices with prosocial behavior is fundamental for replicating results. That is, if participants do not equate “green” choices with prosocial behavior, effects might be weakened and consequently unreplicated. Regarding this, our multi-site study offers limited response as only two groups added an extension variable to determine whether participants found owners of the green products in the study to be more nice, caring, or altruistic than owners of the non-green products; however, they added different questions that cannot be analyzed together, so we examined each separately.

The first site (project #15, *N* = 232) asked participants to respond to the following question (translated from German: “Please indicate on a scale from 1 (disagree completely) to 9 (agree completely) how much you agree with the following statements: ‘Consumers of pro-environmental products are [kinder/more empathetic/more altruistic] than consumers of conventional products.’” Here, the terms “kind” and “more empathetic” are used as measures of “nice” and “caring,” respectively. This rating does not allow us to make any inferences about how participants would rate green product choosers versus non-green product choosers, but the means for each characteristic rating are below 5 (kind/nice: *M* = 3.58, *SD* = 2.19; empathetic/caring: *M* = 4.82, *SD* = 2.32; altruistic: *M* = 4.59, *SD* = 2.17). Importantly, the correlations between the total composite score and these characteristic ratings were small and nonsignificant (kind/nice: *r*(125) = .11, *p* = .235; empathetic/caring: *r*(125) = .00, *p* = .975; altruistic: *r*(125) = .06, *p* = .496), indicating that for these participants, green product choices may not have been associated with how nice, caring, or altruistic the choices appeared to be.

The researchers at this site also included a general question assessing prosociality of green choices. Using the same rating scale as above, the asked to what extent participants agreed with this statement (again, translated from German): “Choosing to buy pro-environmental products is a behavior that is beneficial for the general public.” Participant ratings for this characteristic were much higher (*M* = 7.38, *SD* = 1.98) than the next highest rating (empathic/caring, reported above), *t*(126) = 11.03, *p* < .001, *d* = 2.56. This site did not replicate the main findings of the original study (results of one-way ANOVA comparing composite scores across condition[status/control]: *F*(1, 230) = 0.22, *p* = .639, = .001.

The next question we asked was whether participants’ ratings of prosociality moderated the relationship between condition and composite scores. We found that, while ratings of prosociality did predict composite score such that participants who rated green choices as more prosocial tended to make more of those choices ( = 0.37, *p* < .001) and condition did not ( = -0.05, *p* = 0.776), there was no interaction between condition and ratings of prosociality such that those who found green choices more prosocial weren’t more or less likely to choose those products across the different conditions (control and status) ( = -0.32), *p* = 0.062).

The second site (project #13, *N* = 227) had participants rate these three characteristics from 1 (totally agree that the owner has this quality) to 9 (totally disagree that the owner has this quality). This group used the same items as the original study. We added these ratings together for each participant, and found that participants rated the owners of green products as nicer (*M* = 9.19, *SD* = 4.21) than owners of non-green products (*M* = 12.19, *SD* = 4.19), *t*(226) = -10.85, *p* < .001, *d* = 0.72. Likewise, participants rated green product owners as more caring (*M* = 7.95, *SD* = 3.91) than non-green product owners (*M* = 13.47, *SD* = 4.32), *t*(225) = -16.35, *p* < .001, *d* = 1.09, and more altruistic (*M* = 9.83, *SD* = 4.43) than non-green product owners (*M* = 13.87, *SD* = 4.19) as well, *t*(224) = -10.16, *p* < .001, *d* = 0.68. This site did not replicate the main findings of the original study (results of one-way ANOVA comparing composite scores across condition[status/control]: *F*(1, 225) = 0.78, *p* = .378, ) = .003, thereby weakening the argument that the effects might not replicate because participants might not find these products to be associated with prosocial behavior. However, we have to be cautious regarding the conclusions because of the low power we had for this analysis.

Next, the original authors regarded political ideology as a relevant factor in explaining prosocial behavior. Thus, many student teams included political ideology as extension variables.

#### Political orientation and party affiliation

When the CREP team first contacted the original authors about the replication, the original authors recommended examining political orientation as a possible extension variable. A recent meta-analysis by Cruz (2017) suggested that both political ideology and party affiliation have associations with environmental concerns. Political orientation describes where someone falls on the spectrum of political beliefs ranging from strongly liberal to strongly conservative, and political party affiliation specifies the major political party with which someone identifies (Cruz, 2017). However, defining political orientation is complex, as specific attitudes and beliefs associated with it vary across time and place (Jost et al., 2003). For example, the liberal/conservative dimension conveys different political attitudes in the US and European countries (Greenberg & Jonas, 2003).

In our study, many groups (*N* = 10) added a question about political orientation (liberal/conservative), and some (*N* = 4) added a question about political party affiliation (Republican/Democrat/Independent); several groups added both questions (*N* = 4), and some groups added questions with responses that were more fitting to their particular country of origin (e.g., Canada) (more information on characteristics of datasets and extension variables is available here: <https://osf.io/yqf8a>). We used these data to test whether green product selections differed for liberals and conservatives. Some institutions measured this with a scale (e.g., from very liberal to very conservative), while others measured it with categorical response types (e.g., liberal, conservative, neutral). We collapsed over these responses to create a novel variable with two levels: a “liberal” level that included all responses indicating any degree of being liberal, and a “conservative” level that included all responses indicating any degree of being conservative. We excluded any “neutral” or “other” responses. Using this method of collapsing data, we found that *N* = 551 participants were classified as conservative, while *N* = 744 participants were classified as liberal. Since this constitutes just 34.3% of the overall sample of 3774 participants, it is unknown whether the effects of political beliefs on status motives, if detected, would hold for the remaining 65.7% of participants.

##### Liberal/conservative

We conducted a 2 x 2 factorial ANOVA (political orientation: liberal vs. conservative; condition: control vs. status) to determine whether political orientation interacted with the condition to predict the composite green score.

Importantly, we did not have adequate power to detect the interactions tested here. Using the *Superpower* package (Lakens & Caldwell, 2021) in R, we determined that we would have needed at least 418 participants *per cell* to detect an interaction effect between political orientation and condition. In the four cells created between these two variables, only one cell size reached this threshold (Conservative/Control, *N* = 365, Conservative/Status, *N* = 186; Liberal/Control, *N* = 492; Liberal/Status: *N* = 252). Thus, we did not further reduce power by conducting a test with ungrouped control conditions.

We found 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, η^2 = .092. As we found earlier, there was 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* = 1.00), *F*(1, 1,287) = 0.41, *p* = .522, η^2 = .000. There was also no interaction between the two variables, *F*(1, 1,287) = 0.28, *p* = .597, η^2 = .000.

We found a similar pattern of results when the control conditions (no story and control story) were ungrouped: a main effect of liberal/conservative (*F*(1, 1,285) = 130.85, *p* < .001, η^2 = .092), no main effect of status condition (*F*(2, 1,285) = 1.44, *p* = .238, η^2 = .002), and no interaction between the two factors (*F*(2,1285) = 0.52, *p* = .598, η^2 = .001).

##### Democrat/republican

Similarly, for political party affiliation (“Democrat” or “Republican”), we conducted a 2 x 2 factorial ANOVA (political party: Democrat *N* = 332 vs. Republican *N* = 182; condition: control vs. status) including only participants from the US.

In the four cells created by political party and condition, Republican/Status (*N* = 69) or Democrat/Status (*N* = 123), Republican/Control (*N* = 113) or Democrat/Control (*N* = 209), our cell sizes did not reach the threshold of 418 participants per cell that the previous, similar analysis also required.

We found 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, η^2 = 0.09. We did not detect an interaction between political party and condition, *F*(1, 505) = 0.14, *p* = .709, η^2 = .000, and we detected no main effect of condition (control: *M* = 1.65, *SD* = 0.99); status: *M* = 1.61, *SD* = 1.00), *F*(1, 505) = 0.80, *p* = .370, η^2 = .000.

#### Lab vs. Online

To determine whether there were differences in results across different testing settings, we first collapsed all reported lab settings (group, individual, and just “lab” without specifying whether data were collected in a group or individually) into one variable level. We then conducted chi-square tests of independence using a dichotomous setting variable (lab [*N* = 2024] versus online [*N* = 1581]). We found no relationship between green choice and setting for any of the three types of products (*p*s of 0.26, 0.13, and 0.67 for the green car, soap, and dishwasher, respectively).

To see if there was an interaction between condition and setting, we conducted a 2 (condition: status vs. control) x 2 (lab vs. online) ANOVA using the composite score as a dependent variable. We were unable to detect a mean difference in composite score between those who completed the study in the lab (*M* = 1.66, *SD* = 0.98) and those who completed the study online (*M* = 1.64, *SD* = 1.01), *F*(1, 3,547) = 0.09, *p* = .766, η2 = .000. 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* = 1.00), *F*(1, 3547) = 1.27, *p* = .260, η2 = .000. There was also no interaction between the two variables, *F*(1, 3,547) = 1.19, *p* = .275, η2 = .000.

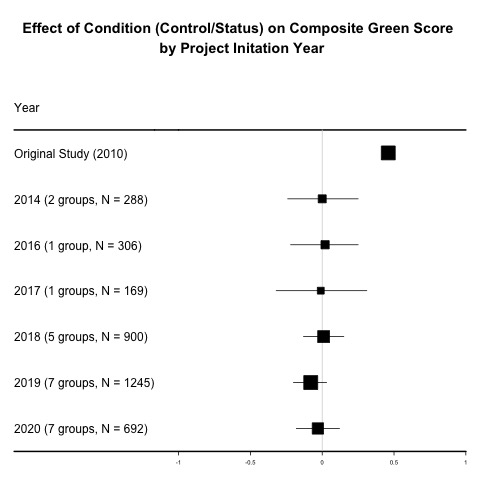
#### US vs. others

The original study was conducted only with individuals in the United States. Here, we conducted an exploratory analysis to determine whether the tested effects differed for US-based participants versus participants from other countries. Using data from US-based participants only (*N* = 2,151), we found no association between condition (control versus status) and green car selection ((1, *N* = 2,104)) = 0.12, *p* = .728, = 0.01), green cleaner selection ((1, N = 2,103)) = 0.02, *p* = .884, = 0.00), or green dishwasher ((1, N = 2,107)) = 0.00, *p* = .952, = 0.00) selection.

We then conducted a 2 x 2 ANOVA to test whether there was a main effect of condition (control or status) or interaction with geographic setting (US or non-US). We detected no significant main effect of condition (*p* = .558) or interaction (*p* = .409). We found a small main effect of location, such that individuals from non-US countries selected significantly more green products (*M* = 1.83, *SD* = 0.95) than individuals from the US (*M* = 1.48, *SD* = 1.01), *F*(1, 3,714) = 115.62, *p* < .001, η^2 = .030.

#### Stability of effect over time.

To determine if any effects differed across years, we examined effects across projects from different years, as determined by their project codes assigned at the time the teams signed up for the study (e.g., project #14-3 was initiated in 2014). With year treated as a factor, we examined whether we could detect any differences in composite green scores across conditions (control and status). As demonstrated in Figure 2, there does not appear to be any meaningful change in effect size over time.



**Figure 2.** Effect sizes (Cohen’s d) by year with 95% confidence intervals. The effect size for the original study, 0.46, is shown without a confidence interval on the top line of the forest plot for reference.

#### Gender differences

In their original paper, Griskevicius and colleagues (2010) stated that exploring gender differences would be interesting in future studies. Although they did not find any gender differences, Griskevicius et al. (2010) hypothesized that men could be more likely to engage in pro-environmental “showing-off” behavior than women because men are generally more concerned about status motives. Because of our additional statistical power, we sought to test these potential gender differences as part of our exploratory analyses.

To test this question, we conducted a 2 x 2 factorial ANOVA (gender: woman or man; condition: control vs. status). We found a main effect of gender, such that participants who identified as women selected significantly more green products (*M* = 1.69, *SD* = 0.99) than participants who identified as men (*M* = 1.46, *SD* = 1.01), *F*(1, 3,263) = 36.22, *p* < .001, η^2 = 0.01. We did not detect an interaction between gender and condition, *F*(1, 3,263) = 0.00, *p* = .964, η^2 = .000.

#formatting p values for next paragraph  
aov2x3.nscs.p <- round(tukeyUG$condition[1,4], digits = 3)  
aov2x3.nscs.p.formatted <- sub("^0\\.", ".", as.character(aov2x3.nscs.p))  
  
aov2x3.nsss.p <- round(tukeyUG$condition[3,4], digits = 3)  
aov2x3.nsss.p.formatted <- sub("^0\\.", ".", as.character(aov2x3.nsss.p))  
  
aov2x3.csss.p <- round(tukeyUG$condition[2,4], digits = 3)  
aov2x3.csss.p.formatted <- sub("^0\\.", ".", as.character(aov2x3.csss.p))

When we conducted a 2 x 3 factorial ANOVA with the independent conditions (control, status, and no story), we found an effect of condition, *F*(2, 3,261) = 4.23, *p* = .015, η^2 = .000. Post-hoc testing with Tukey’s LSD correction revealed significant differences between the “no story” condition and the other two conditions (95% CI of the difference for no story vs. control story [0.01, 0.21], *p* = .028); 95% CI of the difference for no story vs. status: [-0.21, -0.01], *p* = .032). There was no difference between the control story and the status story conditions, *p* = .995). Participants in the “no story” condition made more “green” choices (*M* = 1.7, *SD* = 0.98) than participants in the control condition (*M* = 1.58, *SD* = 1.01) or status condition (*M* = 1.59, *SD* = 1.00). This effect was not detected when the ungrouped conditions were examined in the entire data set; it was detected only when we excluded participants who did not respond to the question about gender or listed a gender other than “Male” or “Female.”

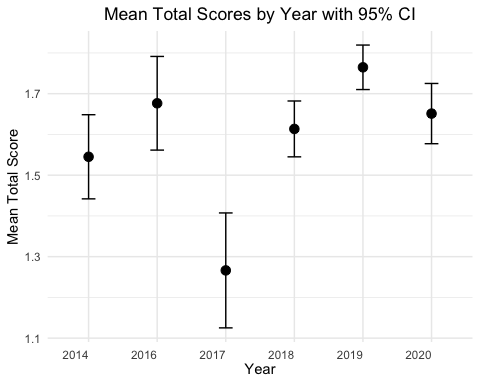
## Discussion

In their original study, Griskevicius et al. (2010) suggested that pro-environmental behavior can be promoted using status competition. They linked this finding to the “competitive altruism” hypothesis, according to which individuals’ altruistic behaviors arise from status competition (i.e., attempts to be perceived as good; Hardy & Van Vugt, 2006). In essence, cultivating a positive reputation through prosocial and pro-environmental behavior and action, coupled with a demonstrated willingness to incur costs for the public good, can elevate an individual’s standing within a group. The original study revealed a higher preference for all three green products (green car, green cleaner, green dishwasher) in the status condition compared to the control conditions.

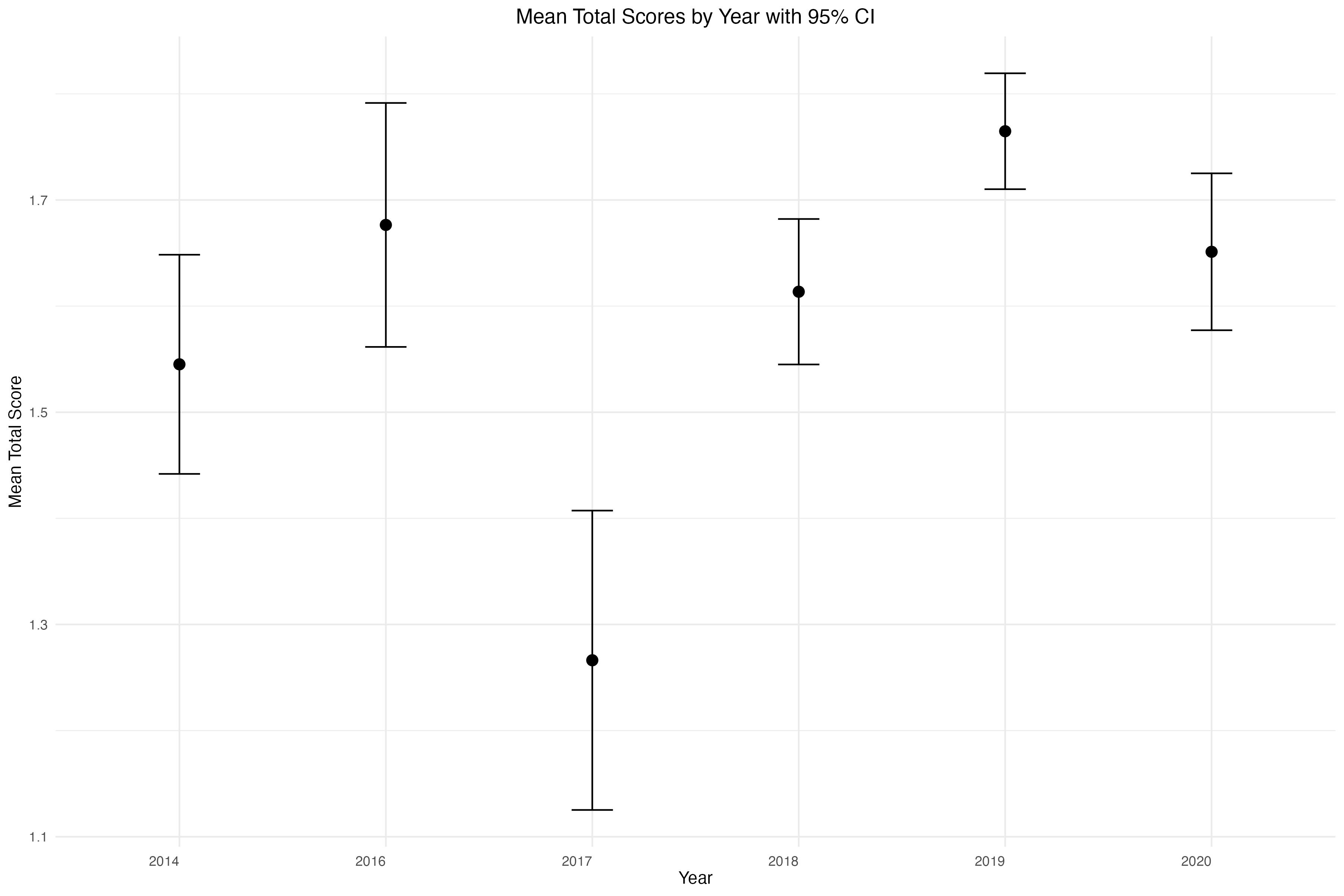
In this multi-lab replication study, we recruited a large sample of participants across six different countries. We employed the same methods and analytical strategy as the original study. As part of our pre-registered confirmatory analysis, we tested whether an effect of status on pro-environmental choice behavior existed when the green score was calculated as a composite. Critically, this analysis did not detect a statistically significant effect of status, failing to replicate the original findings from Experiment 1 of Griskevicius et al. (2010). We also analyzed if an effect of status existed for each of the green vs. non-green products separately. For all three products, we failed to detect a significant effect of status. Participants in the status condition and the control conditions did not differ in their tendency to select green products. As in the original study, we found no differences in green choice behavior between the two control conditions (non-status story vs. no story). However, there appeared to be potential (but not statistically significant) differences between both “story” conditions (status and control) and the “no story” condition such that exposure to any story resulted in fewer “green” choices. Thus, overall, we did not replicate the main conclusion of Experiment 1 from Griskevicius et al. (2010) with a much more highly powered analysis (original *N* = 168, present *N* = 3774).

One possible explanation for this difference in results is that our control participants selected the green products at much higher rates than in the original study. Our control participants’ rates of selecting the green car, soap, and dishwasher were 55.5%, 56.8%, and 52.9%, respectively, while the original control participants’ rates were much lower at 37.2%, 25.7%, and 34.5%, respectively. There could be a ceiling effect with green product choices, such that a certain proportion of the population will never choose to buy green products. Accordingly, in our sample, the proportions of “green” choice selections among participants in both the control and status conditions were potentially near this ceiling. Meanwhile, green products may have been perceived as more novel when Griskevicius and colleagues ran their study (~2008), such that a smaller group of the population would buy them when given the choice. In this case, perhaps status primes moved some people over the fence, but by the time our data collection occurred, the people who could have been persuaded by a status prime were already making these choices.

Given this possible alternative explanation, we examined whether there appeared to be any overall pattern in the total composite score over time (collapsing across conditions). As shown in Figure 3, some years do differ from others, but the early years (2014 and 2016) are not significantly different from the last two years (2021 and 2022). Nevertheless, it is still possible that attitudes toward pro-environmental behaviors had sufficiently changed between the time when Griskevicius et al. (2010) conducted their study (which appeared to have occurred in 2008 or earlier, based on the date of original submission to the *Journal of Personality and Social Psychology*) and 2014, the earliest starting date of data collection in our study.



Forest plot example



**Figure 3.** Mean composite scores by year. There were no noticeable differences in mean composite scores between the earliest projects (2014 and 2015) and the latest projects (2019 and 2020).

The original authors stressed that equating pro-environmental choices with prosocial behavior is fundamental for the effect to appear. Despite being recommended by the original authors, only one group decided to explore this extension hypothesis. Although limited by low power, our findings do not give support for the original authors’ claim.

Per the recommendation of the original authors, we performed exploratory analyses to test whether political beliefs significantly moderated the effect of status motives on pro-environmental choices. In this study, some testing sites collected data on political orientation (e.g., liberal/conservative) and/or political party (e.g., Democrat/Republican) as extension variables. We found that participants who identified as liberal or as Democrats selected more green products than participants who identified as conservatives or as Republicans. These independent effects of political belief (either orientation or party affiliation) did not, however, interact with any effect of activating status motives on green choice behavior.

One difference between the original study and our multi-site replication was that in the original study, participants were tested in small groups in a lab setting. Meanwhile, in our study, many sites tested participants in the lab, sometimes individually and sometimes in small groups, yet other sites tested participants online. It is conceivable that being tested in a laboratory context, which is presumably more social (in that, at minimum, the participant has some interaction with a research assistant, and possibly with other participants), may more strongly influence status motives on pro-environmental behaviors, compared to being tested online. If this is the case, then we would expect to have found differences in the proportion of green choice selections between in-person and online participants in our study (e.g., Belletier et al., 2015; Belletier & Camos, 2018). However, when we explored whether the setting in which participants were tested (lab—presumably at least somewhat social—versus online) influenced the results, we found no significant difference in composite green scores across these settings. Thus, we cannot conclude that the difference between the present study and that of the original study could be explained by different sampling methods. Online data collection has become increasingly popular over the last decade and offers several advantages over traditional lab settings, such as reduced costs, more effective testing, increased generalizability of results, and more efficient (i.e., automated) data collection (e.g., Buhrmester et al., 2011; Dandurand et al., 2008; Gosling et al., 2004; Nayak & Narayan, 2019). Our findings of no difference in results between the lab and online contexts are consistent with those of many previous studies comparing the quality of online and lab-based samples of participants (e.g., Peer et al., 2017; Riva et al., 2003).

It is important to note that the original study was conducted in the US, while this multi-site replication project collected data in five other countries: Canada, the Netherlands, the UK, Germany, and Iceland. We found no significant differences in results between the samples from different countries. We believe this finding is expected as all countries are WEIRD (Western, Educated, Industrialized, Rich, and Democratic; Henrich et al., 2010) and highly sustainable, with sound environmental policies according to the Fragile State Index (Fund for Peace, 2022).

An additional aspect of this multi-site replication study is the extended duration of data collection spanning several years. This allowed us to investigate potential variations in the tested effects over time. However, we did not find substantial differences in effect sizes obtained from replication studies completed between 2014 and 2020.

Finally, in their paper, Griskevicius and associates stated that gender differences in status motives would be worth exploring. Our findings indicate a greater inclination among women toward green products compared to men, but a moderating effect of gender on the relationship between experimental conditions and product preference was not obtained.

### Limitations of the study

The present study has some limitations. First, as stated previously, all of the countries in which the study was conducted are WEIRD and highly sustainable countries. Our findings indicate that status motives do not drive individuals to choose pro-environmental products. However, one should exercise caution in generalizing these findings to broader populations. One possibility is that participants from non-WEIRD and fragile countries, many of which have recently increased efforts to promote pro-environmental behaviors (e.g., Diaz et al., 2020), could be more sensitive to any potential relation between social status and pro-environmental action. Therefore, it will be worth exploring in future work whether the results of the original study would replicate in non-WEIRD and countries that are considered fragile. Additionally, it is crucial to acknowledge the potential impact of socio-cultural factors in these diverse contexts. Evidence suggests that pro-environmental behavior is influenced by social norms (e.g., Saracevic et al., 2022) and individual values (e.g., Nordlund & Garvill, 2002). Thus, the specific recommendation for future research is that researchers can explore how local values and socio-economic conditions might influence the relationship between social status and pro-environmental actions in different regions.

Second, we tested moderation effects only for those variables for which we had sufficient statistical power. Therefore, most of the extension variables that were measured could not be used in the analyses (e.g., personality traits, altruism, empathy). To more thoroughly explore the role of extension variables in the effects being replicated, future CREP studies could pose stricter guidelines on the usage of extension variables in student groups. Furthermore, incorporating a comprehensive set of guidelines for the inclusion of extension variables would enhance the robustness and applicability of findings across diverse research settings, fostering a more nuanced understanding of moderation effects.

### Conclusion

One goal for this project was to provide a medium for students to engage in high-quality replication research while contributing to the field. Student-led projects can and do offer adequate rigor; the quality of the studies was ensured by continuous and careful supervision of senior researchers (faculty members, two reviewers, one CREP board member), recommendations from the original first author, and documentation of the research process (e.g., Grahe et al., 2020; Wagge et al., 2019; Wagge et al., 2023).

The primary goal of this project, however, was to investigate the possibility of promotion of pro-environmental behavior through status competition. This study presents a large sample multi-site replication of the Griskevicius et al. (2010) study, conducted by student teams at various institutions. Overall, our study did not replicate the original work, as we observed no evidence supporting the notion that pro-environmental behavior can be stimulated by inducing status motives. While these results do not dismiss the theoretical framework proposed by Griskevicius et al., at the least they do indicate substantial boundary conditions. Given our inability to demonstrate a reliable effect, this theory warrants reconsideration to identify populations that may still exhibit the effect or to understand why the effect has diminished. The time lapse between the original study and our replications, spanning from 4 to 10 years, has seen increased attention directed toward climate change and environmental choices. It is conceivable that the heightened awareness on this topic has moderated the manifestation of the effect. Regardless of cause, it seems the original effect is either very sensitive to the conditions in which it is collected, or is not reliably detected in modern times.

## References

Belletier, C., Davranche, K., Tellier, I. S., Dumas, F., Vidal, F., Hasbroucq, T., & Huguet, P. (2015). Choking under monitoring pressure: Being watched by the experimenter reduces executive attention. *Psychonomic Bulletin & Review, 22,* 1410–1416. <https://doi.org/10.3758/s13423-015-0804-9>

Belletier, C., & Camos, V. (2018). Does the experimenter presence affect working memory? *Annals of the New York Academy of Sciences, 1424,* 212–220. <https://doi.org/10.1111/nyas.13627>

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3–5. <https://doi.org/10.1177/1745691610393980>

Cruz, S. M. (2017). The relationships of political ideology and party affiliation with environmental concern: A meta-analysis. *Journal of Environmental Psychology, 53,* 81-91. <https://doi.org/10.1016/j.jenvp.2017.06.010>

Dandurand, F., Shultz, T. R., & Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods, 40*(2), 428–434. <https://doi.org/10.3758/BRM.40.2.428>

Díaz, M. F., Charry, A., Sellitti, S., Ruzzante, M., Enciso, K., & Burkart, S. (2020). Psychological factors influencing pro-environmental behavior in developing countries: Evidence from Colombian and Nicaraguan students. *Frontiers in Psychology, 11,* 580730. <https://doi.org/10.3389/fpsyg.2020.580730>

Frank, M. C., & Saxe, R. (2012). Teaching Replication. *Perspectives on Psychological Science, 7*(6), 600–604. <https://doi.org/10.1177/1745691612460686>

Fund for Peace (2022). Fragile State Index 2022 - Annual Report. Retrieved from <https://fragilestatesindex.org/2022/07/13/fragile-states-index-2022-annual-report/>

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, Kacinik, and Prinz (2011). *Advances in Methods and Practices in Psychological Science, 3*(1), 3–23. <https://doi.org/10.1177/2515245919881152>

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should We Trust Web-Based Studies? A Comparative Analysis of Six Preconceptions About Internet Questionnaires. *American Psychologist, 59*(2), 93–104. <https://doi.org/10.1037/0003-066X.59.2.93>

Grahe, J. E. (2017). Authentic research projects benefit students, their instructors, and science. In How we teach now: The GSTA guide to student-centered teaching. (pp. 352–368). Society for the Teaching of Psychology. Retrieved from <https://teachpsych.org/ebooks/howweteachnow>

Grahe, J. E., Cuccolo, K., Leighton, D. C., & Cramblet Alvarez, L. D. (2020). Open science promotes diverse, just, and sustainable research and educational outcomes. *Psychology Learning & Teaching, 19*(1), 5–20. <https://doi.org/10.1177/1475725719869164>

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>

Greenberg, J., & Jonas, E. (2003). Psychological motives and political orientation–the left, the right, and the rigid: Comment on Jost et al.(2003). Psychological Bulletin, 129(3), 376-382. DOI: 10.1037/0033-2909.129.3.376

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>

Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin, 32*(10), 1402–1413. <https://doi.org/10.1177/0146167206291006>

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and Brain Sciences, 33*(2-3), 61-83. <https://doi.org/10.1017/s0140525x0999152x>

Jost, J. T., Glaser, J., Sulloway, F. J., & Kruglanski, A. W. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129*(3),339–375. <https://doi.org/10.1037/0033-2909.129.3.339>

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>

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams Jr, R. B., Alper, S., … & Sowden, W. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443-490. <https://doi.org/10.1177/2515245918810225>

Lakens, D., & Caldwell, A. R. (2021). Simulation-Based Power Analysis for Factorial Analysis of Variance Designs. *Advances in Methods and Practices in Psychological Science, 4*(1), 251524592095150. <https://doi.org/10.1177/2515245920951503>

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>

Loo, C. W., & Choy, J. L. F. (2013). Sources of self-efficacy influencing academic performance of engineering students. *American Journal of Educational Research, 1*(3), 86-92. <https://doi.org/10.12691/education-1-3-4>

Malkewitz, C. P., Schwall, P., Meesters, C., & Hardt, J. (2023). Estimating reliability: A comparison of Cronbach’s α, McDonald’s ωt and the greatest lower bound. *Social Sciences & Humanities Open, 7*(1). <https://doi.org/10.1016/j.ssaho.2022.100368>

Nayak, M. S. D. P., & Narayan, K. A. (2019). Strengths and weaknesses of online surveys. *Technology, 6*(7). <https://www.iosrjournals.org/iosr-jhss/papers/Vol.%2024%20Issue5/Series-5/E2405053138.pdf>

Nordlund, A. M., & Garvill, J. (2002). Value structures behind proenvironmental behavior. *Environment and Behavior, 34*(6), 740-756. <http://dx.doi.org/10.1177/001391602237244>

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), <https://doi.org/10.1126/science.aac4716>

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70,* 153–163. <https://doi.org/10.1016/j.jesp.2017.01.006>

Rammstedt, B. & Beierlein, C. (2014). Can’t We Make It Any Shorter?. *Journal of Individual Differences, 35*(4), 212-220. <https://doi.org/10.1027/1614-0001/a000141>

R Core Team. (2021). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <https://www.R-project.org/>

Riva, G., Teruzzi, T., & Anolli, L. (2003). The use of the Internet in psychology research: Comparison of online and offline questionnaires. *CyberPsychology & Behavior, 6*(1), 73–80. <https://doi.org/10.1089/109493103321167983>

RStudio Team. (2020). RStudio: Integrated Development Environment for R. RStudio, PBC. <http://www.rstudio.com/>

Saracevic, S., Schlegelmilch, B. B., & Wu, T. (2022). How normative appeals influence pro-environmental behavior: The role of individualism and collectivism. *Journal of Cleaner Production, 344,* 131086. <https://doi.org/10.1016/j.jclepro.2022.131086>

Van Vugt, M., Roberts, G., & Hardy, C. (2007). Competitive altruism: Development of reputation-based cooperation in groups. In R. Dunbar & L. Barrett (Eds.), *Handbook of evolutionary psychology* (pp. 531–540). Oxford, England: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198568308.013.0036>

Wagge, J. R., Baciu, 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. (2023). Teaching research in principle and in practice: What do psychology instructors think of research projects in their courses? *Psychology Learning & Teaching, 22*(1), 4–19. <https://doi.org/10.1177/14757257221101942>

1. As of October 2023, 480 projects at 174 institutions in 29 countries have been initiated. [↑](#footnote-ref-21)
2. Regarding demographics, the original study reported only information about gender of the participants (male and female). [↑](#footnote-ref-41)