Irregularities in Margaret Perez Brower (2024), "Reframing Gendered Issues: Intersectional Identity Frames and Policy Agendas," *Political Behavior*

Christopher Adolph[★], Caitlin Ainsley[†], and Aslı Cansunar[†]

University of Washington, Seattle

12 August 2025

Submitted for publication in *Political Behavior*

Note. The following is a report authored on 8 May 2025 and first provided to the editors of *Political Behavior* on 18 June 2025. A version of this report has been submitted for publication in *Political Behavior*.

[★] Professor of Political Science, Adjunct Professor of Statistics, and Associate Director of the Center for Statistics and the Social Sciences, University of Washington, Seattle. *Corresponding author:* cadolph@uw.edu.

[†] Associate Professor of Political Science, University of Washington, Seattle.

Irregularities in Margaret Perez Brower (2024), "Reframing Gendered Issues: Intersectional Identity Frames and Policy Agendas," *Political Behavior*

8 May 2025

Background on Perez Brower (2024)

Perez Brower's (2024) *Political Behavior* article investigates how the identity framing of policy agendas influences political mobilization, particularly among highly educated white women. In the published paper, Perez Brower claims to use original data from two survey experiments conducted with a national sample of 926 highly educated white women to examine how varying the race and class of female policy beneficiaries shapes engagement with two gender-related issues: (1) sexual harassment and (2) the gender pay gap. According to the published findings, highly educated white women tend to disengage more when the policy frame misaligns with their racial identity, and this effect is strongest when both race and class diverge from their own. Specifically, Perez Brower reports respondents were least likely to take political action on the gender pay gap when the framing highlighted low-status women of color as beneficiaries. The paper ultimately argues that intersectional identity framing – messaging that links policy relevance to multiple identity traits – can have strong, and at times negative, effects on political engagement. This is a timely study that offers potentially valuable insights for policy organizations seeking to design more effective campaigns and enhance the impact of their advocacy efforts.

Summary of Irregularities

We take as stipulated the methods of inference used in Perez Brower (2024): specifically, the comparison of mean outcomes across experimental treatment conditions using *t*-tests and the adoption of 0.05 significance levels to accept or reject hypotheses. In other words, in this report, we set aside any potential debates about the choice of appropriate research methods to focus on fundamental issues regarding proper conduct and data reporting in scientific inquiry.

A review of the publicly available replication materials and survey documentation described below revealed several notable irregularities. These issues raise questions about the consistency

between the author's reported analyses, pre-analysis plan, and the data collected in the survey experiment, as well as questions about the validity of the published findings.

Problem 1: Selective omission of survey respondents. For the second experiment, the paper suggests in several ways that no "ambiguous race and class" control group was fielded, such that results comparing such a control condition to the treatment condition can only be shown for the first experiment. However, this control condition was fielded for both experiments, and numerous pieces of evidence confirm the author's awareness of this. Results for the second experiment using this control, which are not reported, fail to confirm the author's Hypothesis 1.

Problem 2: Selective reporting of outcomes. The author proposes two dependent variables: a count of political activities endorsed by each respondent, and a binary indicator for respondents who answered they "would NOT participate in any activity related" to the issue for that experiment. Across both experiments, outcomes are selectively reported on both dependent variables, especially for the first dependent variable. Many results that fail to confirm the author's hypotheses are not reported.

Problem 3: Miscoding of the second dependent variable. The second dependent variable, which is collected using a non-standard and potentially confusing question format, is the basis of most of Perez Brower's reported result. However, the actual data for this dependent variable appears to have been omitted from the analysis altogether, and a different variable has been analyzed in its place. Replicating the analysis using the author's intended second dependent variable yields null results for all reported tests of Hypotheses 2 and 3.

As a result of the problems noted above, the findings of the paper are comprehensively invalidated: the experiments reported are *inconsistent* with respect to Hypothesis 1 and *do not support* the intersectionally-mediated disengagement conjectured in Hypothesis 2 or Hypothesis 3.

Potential Sources of Confusion in Perez Brower (2024)

Before proceeding further, it will help to clarify a number of potential sources of confusion in the way Perez Brower (2024) presents her experimental conditions, hypotheses, and results.

Inconsistent numbering of experimental conditions. In Perez Brower (2024) Tables 2 and 3, there are five conditions listed: a control group and four treatment groups. The first treatment group is also described as a reference group. Following the order of conditions listed in these two tables, we will consistently refer in this report to the following treatment frames by number:

Conditions	Description
Control	Ambiguous race and class frame
Treatment 1	White highly-educated women frame
Treatment 2	Black & Latina highly-educated women frame
Treatment 3	White low-wage women frame
Treatment 4	Black & Latina low-wage women frame

Unfortunately, in other places in Perez Brower (2024), including Tables 1 and 4, the author re-orders these treatments, referring to the "white low-wage" frame as "Treatment 1", the "Black & Latina highly-educated" frame as "Treatment 2", and the "Black & Latina low-wage" frame as "Treatment 3". As a result, great care must be taken in reading both the original results in the paper and the replication of results in this report to avoid confusing different treatments with each other.

Grouping of hypotheses by experiment. Perez Brower (2024) offers several hypotheses to be tested by her experiments. First, she claims that "[h]ighly educated white women will report they will take more political actions to support a women's organization when the framing" of a policy agenda "aligns with their race and class relative to the control group." Notably, although there is no reason to suppose this hypothesis is specific to the first experiment on sexual harassment, it is worded to apply only to that experiment in Perez Brower (2024), where it forms her Hypothesis 1.

Next, Perez Brower offers three linked hypotheses: that highly educated white women will disengage more when framing is misaligned with their class, misaligned with their race, and finally misaligned with both their class and race at the same time. Her Hypothesis 2 suggests these three claims apply to the sexual harassment experiment, and her Hypothesis 3 suggests these three claims apply to the gender pay gap experiment. Notably, the theoretical claims in each of Hypothesis 2 and 3 are identical, except for the policy agenda to which they are applied.

Presentation of results. Perez Brower (2024) shows results from her experiment in visual displays of point estimates with error bars (her Figures 1, 2, and 3), a tabular summary of which results are significant at the 0.05 level for Hypotheses 2 and 3 (her Table 4), and reporting of significance tests within the main text itself. Because her Figures 1, 2, and 3 show error bars rather than the more conventional 95% confidence intervals, and because these Figures omit several results discussed elsewhere in the paper, we focus our discussion on the reports of significance tests in the main paper and in her Table 4. However, our conclusions would not change if we framed our discussion around the figures.

Data and Documentation

We downloaded Perez Brower's replication files (an `R` script that reproduces the results and a `.csv` file containing the cleaned survey data) from her personal webpage[1]. The public availability

---

[1]Link to Prof. Brower's personal webpage containing the mentioned files, archived on April 19, 2025

of replication code and data is a prerequisite for publication in the journal. *Political Behavior*'s webpage specifies that:

> Authors of accepted manuscripts will be required to deposit all of the data and script files needed to replicate the published results in a trusted data repository such as ICPSR or Dataverse. Manuscripts will not be published, even online, until those data are made available.[2]

Political Behavior's replication requirements align with broader changes in both quantitative and qualitative research aimed at increasing transparency and reproducibility in the social sciences. Many political science journals now employ in-house replication assistants who verify quantitative analyses prior to publication, and there is growing momentum toward similar practices for qualitative research.[3] The broader goal of these efforts is to foster a research environment where scholars can learn from and build upon one another's empirical work, ultimately advancing the field through greater openness and accountability.

In addition to the replication files, we also retrieved the original survey materials uploaded to the Open Science Framework (OSF) by Time-sharing Experiments in the Social Sciences (TESS), the organization which funded the survey that underlies Perez Brower (2024).[4] These materials include the field report of the experiments, the full project description (with survey items, experimental frames, and pre-registered hypotheses), the IRB application, a codebook, and the dataset itself.[5] Notably, TESS has long embraced the value of preregistration and embeds many of its core principles into their data release practices.[6] As they explain:

> Since July 2014, TESS has posted accepted proposals for fielded projects online at the same time it posts the raw data. That is, one year after the data are delivered to investigators. We do this partly for transparency and partly because of the growing interest in pre-registration of studies. While TESS studies do not contain all the elements of pre-registration, our investigators have from the beginning of TESS been articulating their hypotheses before fielding their study, and the policy change credits their doing so.

While preregistration is often seen as a tool to prevent p-hacking in quantitative research, its deeper value is epistemological. Preregistration encourages better science by forcing researchers to clarify what they expect, why it matters, and how exactly they will test it (Nosek et al, 2018).

Although testing hypotheses is essential for advancing social scientific theories, preregistration does not constrain exploration. Researchers are still free – and encouraged – to uncover new patterns in their data. The key is transparency: clearly distinguishing between planned analyses from unanticipated exploratory findings helps preserve the diagnostic value of statistical tests and guards against selective reporting and confirmation bias.

---

[2]Link to the replication guidelines of Political Behavior, archived on November 23, 2020.

[3]See the American Journal of Political Science's Qualitative Data Verification Checklist, for example.

[4]Link to the related materials and data uploaded by TESS to OSF.

[5]Link to the project description uploaded by TESS to OSF.

[6]Link to project description hosted on TESS website.

Selective Omission of Survey Respondents

Perez Brower (2024, 339) describes the treatment conditions in the first experiment on sexual harassment as follows:

> Each experiment included a policy agenda from a fictional women's organization that presented the same policy agenda in each condition but framed the agenda as affecting different subgroups of women by class and race. In the first experiment, a control condition is included that frames the policy agenda using ambiguous racial and class characteristics. It is not a "pure" control, but instead a comparison frame. Framing experiments often use a control in this way to compare between different frames (Druckman, 2001). This condition is included to test the frame alignment hypothesis by comparing the control to a condition that explicitly frames the policy to benefit professional white women in order to prime both the race and class of survey participants.

> The remaining conditions in the first experiment are then compared to the professional white women condition (the reference group) to test for the effects of misalignment by class, race, and then class and race together in a policy frame. The conditions that are compared include a policy frame that benefits (1) working-class white women in pink- and blue-collar industries; (2) Latina and Black professional women in a white-collar industry; and finally (3) working-class Latina and Black women in pink- and blue-collar industries.

From the above, it is evident that there are five total conditions in the first experiment on sexual harassment: an ambiguously framed control group, a professional white women condition, a working class white women condition, a Latina and Black professional women condition, and a Latina and Black working class condition.

Regarding the second experiment, on the gender pay gap, the very next sentence of Perez Brower (2024, 339) states: "In the second experiment, participants similarly were randomly assigned to these three conditions or the reference group condition (i.e., professional white women)." That is, Perez Brower indicates there are four possible conditions in this experiment. The distribution of assignments for each experiment is further detailed in Table 3 of Perez Brower (2024, 341), which reports a total of 514 observations across *five* conditions in Experiment 1, and a total of 412 observations across *four* conditions in Experiment 2, with "–" indicating there were no respondents in the second experiment that received the control (See our Figure 1 for a picture of Perez Brower's published Table 3). In the text of the manuscript, Perez Brower (2024, 341) states that "[t]he survey sample includes a nationally representative sample of 926 U.S. white identifying women, aged 18 and older, with a bachelor's degree or higher," a total which corresponds to the sum of the "Total Observations" from the 5 conditions in first experiment (514) and 4 conditions in the second experiment (412).

In the final two pages of the Online Research Appendix, Perez Brower (2024) reputedly shows

Table 3 Sample and random assignment by experiment

Randomly assigned condition	Experiment 1 Sexual harassment	Experiment 2 Gender pay gap
Control	102	–
High-income white women condition (reference group)	103	104
High-income women of color condition	103	101
Low-income white women condition	103	104
Low-income women of color condition	103	103
Total observations	514	412

Figure 1. *Table 3 from Perez Brower (2024), listing Experiment 2 as having no control group.* In reality, there are 104 subjects in the control group for Experiment 2.

"All Experimental Conditions for Sexual Harassment & Gender Pay Gap."[7] This reaffirms Perez Brower's claim that there are a total of nine conditions: five for the first experiment, and four for the second experiment (See our Figure 2 for a reproduction of these pages).

However, Perez Brower's experiment, as documented in survey files posted by TESS to OSF and in Perez Brower's own replication files and online appendix, clearly contained a control group of 104 respondents for the second experiment. Each of these data files contains a total of 1030 respondents. The discrepancy of 104 between the actual size of the survey and the 926 respondents Perez Brower (2024) reports in her Table 3 and in the quote above is the result of the omission of these 104 surveyed individuals who received the "Gender Wage Gap Control." In addition to this being readily apparent in the data, this is clearly reported in the Final Programming Questionnaire uploaded by TESS to OSF. In the table spanning pages 5 and 6, the questionnaire shows a quota target of 105 for respondents to receive `P_Brower_GWGC`, which is the label assigned to the control condition omitted in the Perez Brower (2024) online appendix (our Figure 2), but reproduced from the TESS questionnaire (our Figure 3).

Contrary to the language in Perez Brower (2024) suggesting Experiment 2 contains no control group, we have found no less than four documents authored and published by Perez Brower over the course of the project which show a clear awareness that Experiment 2 contains a control group with an ambiguous race and class frame. We now describe these documents in rough chronological order:

(1) In the brief project description uploaded by TESS to OSF on August 23, 2020, Perez Brower describes the design of the second experiment:[8]

---

[7]Link to Online Research Appendix for Perez Brower (2024) hosted by *Political Behavior*.

[8]This project description is contained in the file `Project Description - Brower.docx` published by TESS to OSF.
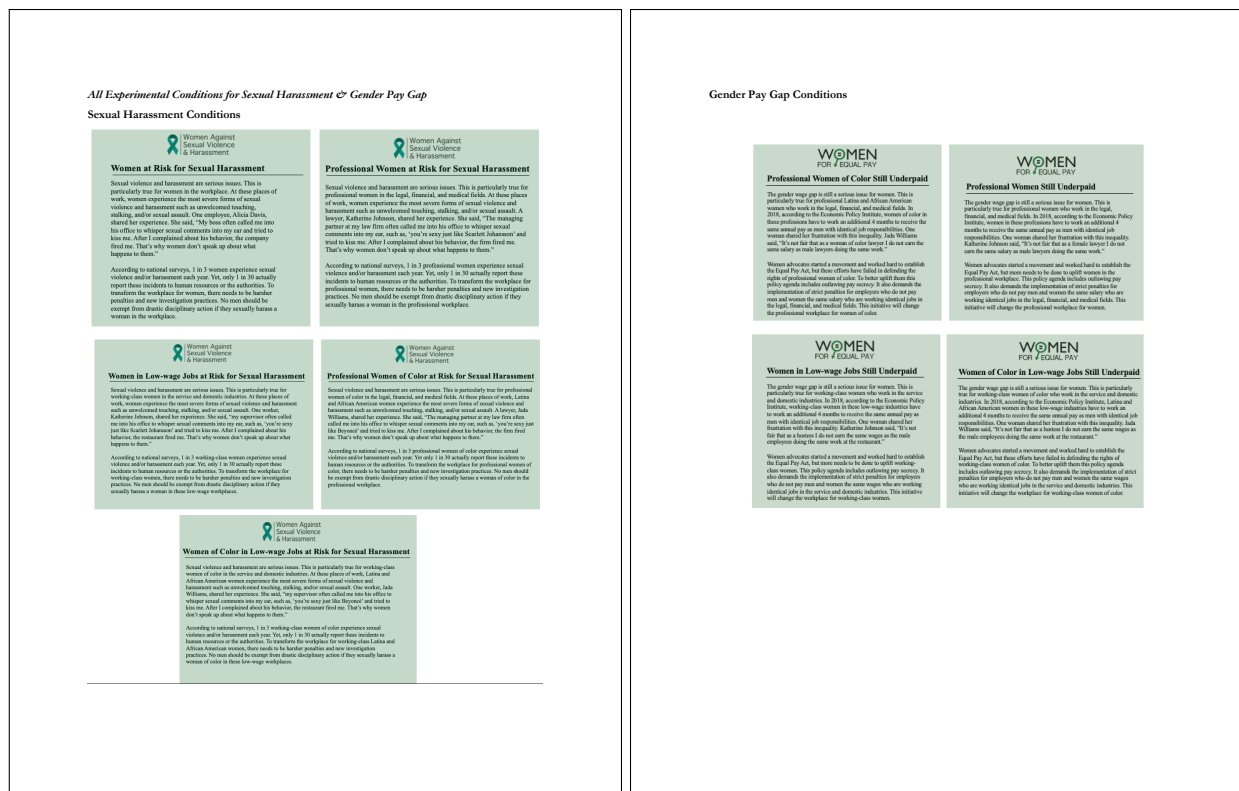
Figure 2. *Pages 7 and 8 from the Online Research Appendix to Perez Brower (2024).* Page 7, at left, states these material include "All Experimental Conditions for Sexual Harassment & Gender Pay Gap". On the same page, under the header "Sexual Harassment Conditions", five conditions are shown, including the ambiguously-framed control group. On page 8, under the header "Gender Pay Gap Conditions", only four conditions are shown, omitting the ambiguously-framed control group.

> For the gender gap issue, similarly there is a treatment that presents a policy agenda from a women's organization proposing a legislative action to address the gender wage gap issue. This treatment also includes 5 different conditions that are varied using the exact same five conditions mentioned earlier for the sexual harassment treatment.

It is clear even from this short description that there were no intended structural differences between the first and second experiments in terms of the number and kind of conditions to which respondents would be assigned.

This same document also states (emphasis in original): "Participants will randomly be assigned to **one** of the **10** following conditions." Again, it is clear at the design and registration stage that the project contained the same five conditions in the sexual harassment and gender pay gap studies, suggesting the Experiment 2 control group was discarded during the analysis stage.
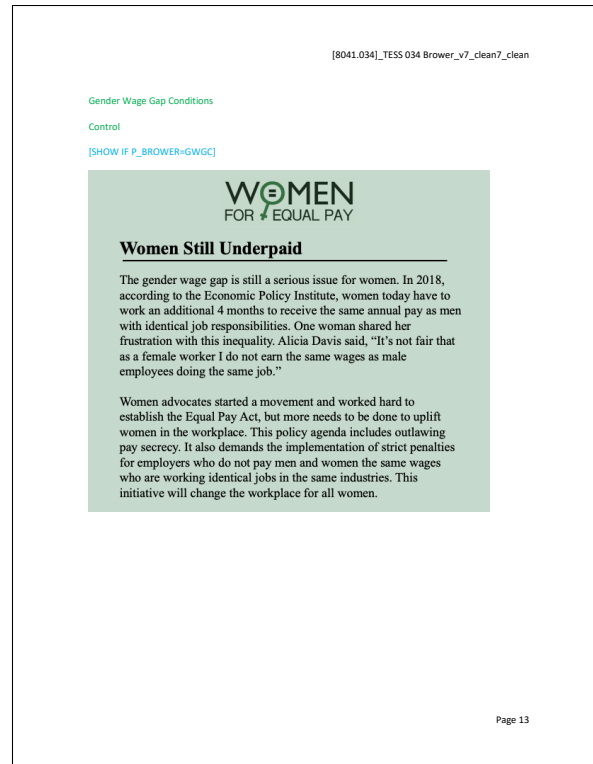
Figure 3. *Control for Gender Wage Gap Experiment, omitted from all published materials associated with Perez Brower (2024).* The Final Programming Questionnaire uploaded to OSF by TESS includes the image with text used for the control for the gender wage gap experiment.

(2) A revised version of the TESS Short Study Proposal for this project uploaded to OSF on August 23, 2020, once again references the consistent design of each experiment, as well as the fact that each experiment contains five total conditions:[9]

> I would like to add one more condition to both treatments, this condition is for the "white collar minority frame," a condition that now includes a white-collar aspect and a racial minority identity aspect to the treatment. With this added condition, each treatment will have a total of 5 conditions each.

Again, the survey design materials explicitly demonstrate Perez Brower's intention to have the same five treatments across the two experiments.

When providing the details of the survey units within the same revised proposal document, Perez Brower describes the eventually omitted control condition under the heading "Unit 3: Wage Gap Policy Issue (5 conditions)." Emphasis included in original.

> (*Control Condition*) The following is an excerpt from the New American Foundation's recent publications on the wage gap:

---

[9]The revised proposal is contained in the file `S27.R1-Brower.docx` published by TESS to OSF.

A new research study by the Economic Policy Institute confirms that in 2017, women, employed full time across all types of job sectors, made only 80 cents for every dollar earned by men, a gender wage gap of 20 percent. Yet, this research projects that if women were to receive equal pay, this change would increase the average household income by 20 percent and would expand the middle class by 10 percent.

Closing the wage gap improves economic conditions for everyone. This fall congress will consider new legislation on the wage gap that if approved, would establish state funded commissions to outlaw pay secrecy and establish an auditing process of employers' compensation information to eliminate gender pay disparities among men and women. This legislation will improve women's earning power in the United States.

It is clear this item was an intended control in the gender pay gap study as the author labeled it "*Control Condition.*"

(3) Perez Brower's replication code is hosted on her personal webpage; a link to this R script is provided in footnote 5 in Perez Brower (2024). In three separate places, Perez Brower's R code either refers to there being five conditions for the gender pay gap experiment or explicitly subsets data on the (supposedly absent) control group for Experiment 2.

In the first instance, on lines 47–51, Perez Brower notes in comments the existence of five conditions for each experiment:

```
#Created a subset dataframe for the Sexual Harassment Conditions (5 total)
sh_conditions <-subset(gf, gf$cond_type=="sexual harassment")


#Created a subset dataframe for the gender wage gap Conditions (5 total)
gpg_conditions <-subset(gf, gf$cond_type=="gender gap")
```

In the second instance, on lines 165–179, Perez Brower subsets out a dataframe containing only the control group from experiment 2 using an indicator variable present in the replication data for that purpose (ggc):

```
#Dummy variables for Gender Gap Treatments
#Control
gg_control<- subset(gf, gf$ggc== 1)


#Treatment 1: white woman/ white collar
gg_white_highinc <- subset(gf, gf$gg_white_highinc== 1)


#Treatment 2: minority woman/ white colar
gg_minority_highinc <- subset(gf, gf$gg_minority_highinc== 1)
```

```
#Treatment 3: white woman/ blue collar
gg_white_lowinc <- subset(gf, gf$gg_white_lowinc== 1)


#Treatment 4: minority woman/ blue collar
gg_minority_lowinc <- subset(gf, gf$gg_minority_lowinc== 1)
```

Finally, on lines 201–230, Perez Brower further subsets the subjects from Experiment 1 and Experiment 2 into dataframes containing each combination of a subject's party identification with a specific condition within a specific experiment. Confusingly, the comments preceding these blocks of code refer to the sexual harassment experiment only, but the code itself systematically and correctly works through each possible combination for each experiment ("sh" and "gg"). Dataframes are constructed for self-identified Democrats subject to the second experiment's control (on line 219) and self-identified Republicans subject to the second experiment's control (on line 226):

```
#Data frames for conditions & party ID for sexual harassment conditions
#Democrat Dataframes (sexual harassment)
dem_sh_control <- subset(gf, gf$shc==1 & gf$democrat==1)
dem_sh_white_highinc <- subset(gf, gf$sh_white_highinc== 1 & gf$democrat==1)
dem_sh_minority_highinc <- subset(gf, gf$sh_minority_highinc== 1 & gf$democrat==1)
dem_sh_white_lowinc <- subset(gf, gf$sh_white_lowinc== 1 & gf$democrat==1)
dem_sh_minority_lowinc <- subset(gf, gf$sh_minority_lowinc== 1 & gf$democrat==1)

#Republican Dataframes (sexual hrassment)
rep_sh_control <- subset(gf, gf$shc==1 & gf$republican==1)
rep_sh_white_highinc <- subset(gf, gf$sh_white_highinc== 1 & gf$republican==1)
rep_sh_minority_highinc <- subset(gf, gf$sh_minority_highinc== 1 & gf$republican==1)
rep_sh_white_lowinc <- subset(gf, gf$sh_white_lowinc== 1 & gf$republican==1)
rep_sh_minority_lowinc <- subset(gf, gf$sh_minority_lowinc== 1 & gf$republican==1)



#Data frames for conditions & party ID for sexual harassment conditions
#Democrat Dataframes (sexual harassment)
dem_gg_control <- subset(gf, gf$ggc==1 & gf$democrat==1)
dem_gg_white_highinc <- subset(gf, gf$gg_white_highinc== 1 & gf$democrat==1)
dem_gg_minority_highinc <- subset(gf, gf$gg_minority_highinc== 1 & gf$democrat==1)
dem_gg_white_lowinc <- subset(gf, gf$gg_white_lowinc== 1 & gf$democrat==1)
dem_gg_minority_lowinc <- subset(gf, gf$gg_minority_lowinc== 1 & gf$democrat==1)

#Republican Dataframes (sexual harassment)
rep_gg_control <- subset(gf, gf$ggc==1 & gf$republican==1)
rep_gg_white_highinc <- subset(gf, gf$gg_white_highinc== 1 & gf$republican==1)
```

```
rep_gg_minority_highinc <- subset(gf, gf$gg_minority_highinc== 1 & gf$republican==1)
rep_gg_white_lowinc <- subset(gf, gf$gg_white_lowinc== 1 & gf$republican==1)
rep_gg_minority_lowinc <- subset(gf, gf$gg_minority_lowinc== 1 & gf$republican==1)
```

In sum, the code Perez Brower provides to replicate her results contains numerous references to the control group for Experiment 2 and creates data structures suited to analyzing this group in the same fashion she uses for Experiment 1. Notably, the above code is used to process data for this control group, but none of the data processed is later used in the part of the script which analyzes her data to produce her published results.

(4) Perez Brower's own Online Research Appendix, published on the *Political Behavior* website and cited in her article, contains three separate tables explicitly summarizing covariate data for the 104 members of this control group (see Perez-Brower's Appendix Table 2 on the age of respondent, Appendix Table 4 on party identification, and Appendix Table 6 on education). Each entry is labeled as "control" and appears in a list with the other four conditions (the treatments).[10]

Next, we turn to the consequences of this selective omission of survey respondents. Because the responses from the 104 individuals who were treated with the "gender pay gap control" remain clearly identifiable in the replication data, we are able to reanalyze the survey including these responses. The consequences of omitting the second experiment's control group from the article are substantial. Perez Brower (2024, 335–336) explains the rationale behind including an ambiguous control group as follows:

> While there is some scholarship that suggests white participants will often assume whiteness as a default characteristic when the racial identity is not clear (Coles & Pasek, 2020a, 2020b), in this study, by comparing the control with ambiguous characteristics to an intersectional identity frame where the race and class of white participants are primed, I can show how an efficacy prime can better influence attitudes towards participation. In other words, explicitly priming the racial and class identity of highly educated white women in a policy frame will be a stronger indicator of their reported engagement than frames with policy beneficiaries described with less clear racial and class identities. I test the following hypothesis in the first experiment on the issue of sexual harassment[1]:
>
> H1 *Frame Alignment & Engagement* Highly educated white women will report they will take more political actions to support a women's organization when the framing of a sexual harassment policy agenda aligns with their race and class relative to the control group.

Despite the explicit reference to sexual harassment policy in the wording of Hypothesis 1, there is no reason given in the paper to expect this hypothesis to apply only to the sexual harassment experiment and not the gender pay gap experiment as well. Notably, it is the only hypothesis in the paper that is tested in only one of the two experiments.

Table 1. Replication of treatment effects for Experiment 1 (Sexual harassment), treating the count of political actions endorsed as the dependent variable.

Treatment Frame	Comparison Frame	Average count of activities endorsed		Difference Significant?	How Reported?
		Treated Group	Comparison Group		
White highly-educated women	All women	3.864 $treated - comparison =$ +0.609 on 7-pt scale	3.255	Yes, $p = 0.013$ CI: 0.128 to 1.091	"$(p < 0.05)$", means & SEs in Fig 1
Black & Latina highly-educated women	All women	3.544 $treated - comparison =$ +0.289 on 7-pt scale	3.255	No, $p = 0.288$ CI: -0.246 to 0.823	Not Reported
White low-wage women	All women	3.893 $treated - comparison =$ 0.638 on 7-pt scale	3.255	Yes, $p = 0.011$ CI: 0.149 to 1.127	Not Reported
Black & Latina low-wage women	All women	3.650 $treated - comparison =$ 0.396 on 7-pt scale	3.255	No, $p = 0.120$ CI: -0.104 to 0.896	Not Reported
Black & Latina highly-educated women	White highly-educated women	3.544 $treated - comparison =$ -0.320 on 7-pt scale	3.864	No, $p = 0.213$ CI: -0.827 to 0.186	Not Reported
White low-wage women	White highly-educated women	3.893 $treated - comparison =$ 0.029 on 7-pt scale	3.864	No, $p = 0.900$ CI: -0.429 to 0.488	Not Reported
Black & Latina low-wage women	White highly-educated women	3.650 $treated - comparison =$ -0.214 on 7-pt scale	3.864	No, $p = 0.372$ CI: -0.684 to 0.257	Not Reported

Table 2. Replication of treatment effects for Experiment 2 (Gender pay gap), treating the count of political actions endorsed as the dependent variable.

Treatment Frame	Comparison Frame	Average count of activities endorsed		Significant Difference?	How Reported?
		Treated Group	Comparison Group		
White highly-educated women	All women	3.567 *treated − comparison =* −0.067 on 7-pt scale	3.635	No, $p = 0.812$ CI: -0.624 to 0.489	Not Reported
Black & Latina highly-educated women	All women	3.723 *treated − comparison =* +0.088 on 7-pt scale	3.635	No, $p = 0.763$ CI: -0.488 to 0.665	Not Reported
White low-wage women	All women	3.519 *treated − comparison =* −0.115 on 7-pt scale	3.635	No, $p = 0.682$ CI: -0.669 to 0.438	Not Reported
Black & Latina low-wage women	All women	3.806 *treated − comparison =* +0.171 on 7-pt scale	3.635	No, $p = 0.549$ CI: -0.392 to 0.734	Not Reported
Black & Latina highly-educated women	White highly-educated women	3.723 *treated − comparison =* +0.155 on 7-pt scale	3.567	No, $p = 0.591$ CI: -0.414 to 0.725	Not Reported
White low-wage women	White highly-educated women	3.519 *treated − comparison =* −0.048 on 7-pt scale	3.567	No, $p = 0.862$ CI: -0.594 to 0.498	Not Reported
Black & Latina low-wage women	White highly-educated women	3.806 *treated − comparison =* +0.239 on 7-pt scale	3.567	No, $p = 0.398$ CI: -0.317 to 0.794	Not Reported

In testing Hypothesis 1, Perez Brower uses as an outcome the number of political activities respondents endorsed in Experiment 1 after being exposed to either the control frame or the white highly-educated frame. (This is the only instance in the article in which Perez Brower reports results for this dependent variable.) Using the data and processing code provided by Perez Brower, we successfully replicated her test of Hypothesis 1 in Experiment 1, and present these results in the first row of our Table 1.[11] As the table shows, respondents exposed to the white, highly-educated frame endorsed 0.609 more items out of the seven activities offered, a difference which is statistically significant at the 0.05 level ($p = 0.013$). It is worth emphasizing that this is the only statistically significant result from Perez Brower (2024) we are able to confirm in this report.

Perez Brower does not report the same test for Experiment 2, despite being apparently aware she has the data to do so. We use her data to test whether respondents exposed to the white highly-educated frame in the gender pay gap experiment endorsed more items than respondents in the control group. As we report in the first row of our Table 2, we find they did not: treated respondents endorsed an average of -0.067 fewer items than the control, an insignificant difference ($p = 0.812$). This result is not reported in Perez Brower (2024). If it had been reported, the overall results for Hypothesis 1 would have been mixed, rather than supportive.

Selective Reporting of Outcomes

Perez Brower (2024) is based on a pair of survey experiments administered within the same survey, each with five conditions (a control group and four other treatment conditions). Perez Brower proposes comparing (at least the first) treatment to the control group, and also comparing Treatments 2, 3, and 4 to Treatment 1. Moreover, the study contains two dependent variables which are measured for each treatment condition within each experiment. As shown in Figure 4, this leads to a total of twelve possible sets of treatment effects which could have been reported in Perez Brower (2024). Only three of the twelve groups of results were actually reported in the article. In this section, we present the missing results for the first dependent variable and discuss why specific omissions are problematic. (We defer discussion of results for the second dependent variable to the next section of this report.)

In the previous section, we discussed Perez Brower's failure to report the effect of Treatment 1 on the count of behaviors endorsed in Experiment 2. This omission is noted with an orange ✗ in the second row of the top half of Figure 4. It is noteworthy that she also does not report these comparisons for the second dependent variable at all (the two orange ✗'s in the lower half of Figure 4), nor does she explain this omission.

Of even greater concern is Perez Brower's omission of the first dependent variable (the count of political behaviors) in comparing Treatments 2, 3, and 4 with the reference frame Treatment 1 for both experiments. These comparisons (shown as red ✗'s in Figure 4) are at the core of her

---

[10]Link to online research appendix hosted on Political Behavior website.

[11]Our replication code is included as an appendix at the end of this report.

| | Dependent Variable 1: Count of Political Behaviors | | |
	Treatment 1 vs. Control	Treatments 2, 3 & 4 vs. Control	Treatments 2, 3 & 4 vs. Treatment 1
Experiment 1 (Sexual Harassment)	✓	✗	✗
Experiment 2 (Gender Pay Gap)	✗	✗	✗

| | Dependent Variable 2: "would NOT participate" | | |
	Treatment 1 vs. Control	Treatments 2, 3 & 4 vs. Control	Treatments 2, 3 & 4 vs. Treatment 1
Experiment 1 (Sexual Harassment)	✗	✗	✓
Experiment 2 (Gender Pay Gap)	✗	✗	✓

Figure 4. *Treatment effects measurable from Perez Brower (2024) experiments.* Given two experiments, each with five treatments (including a control and an alternative comparison frame, namely, Treatment 1), and two dependent variables, there are twelve different groups of treatment effects estimable from the data collected by Perez Brower (2024). Only those marked ✓ are reported in the article. Those marked ✗ are omitted altogether.

research project and are readily available in her experimental data. Moreover, there is substantial evidence Perez Brower recognized the importance of these comparisons from the start, as demonstrated by her preregistration materials.

Footnote 1 of Perez Brower (2024, 336) states: "Hypotheses were preregistered; the wording was slightly altered in this article to specify the different effects, but these changes do not interfere with the original directions of the hypotheses, or the broader claims made pre-study." No link or reference is provided to the above mentioned preregistration. However, TESS has a project summary available online.[12] This webpage lists Margaret Brower as the sole Principal Investigator for the project "Reframing Women's Issues: How Intersectional Identity Frames affect Women's Political Attitudes", contains a link to the data used in Perez Brower (2024), and also contains the following text regarding the project:

Hypotheses

Highly educated, white women will be more supportive of a policy agenda on the issue of sexual harassment/violence when the framing of this agenda focuses on professional white women, relative to other policy agendas that use different identity frames.

Highly educated, white women will report they will take more political actions when the issue of sexual harassment/ violence is framed to focus on white professional women, relative to other policy agendas that use different identity frames.

Experimental Manipulations

Race and Class

Outcomes

Policy Support for Agenda and intended political behavior to support the issue in the policy agenda

---

We take the above to constitute the preregistration noted in footnote 1 of Perez Brower (2024). Critically, the second hypotheses above explicitly refers to subjects taking "more political actions", which is distinct from the focus of Perez Brower's (2024) on "disengagement".

As Nosek et al. (2018, 2602) note, "The benefits of preregistration are lost if researchers do not follow the preregistrations." The widespread adoption of preregistration is motivated by the

---

[12]Link to project summary hosted on TESS website.

recognition that even in the simplest of experimental designs, researchers must make a vast number of choices about how to analyze the data. When these choices are made (or adjusted) during the analysis stage, it becomes difficult to discern the set of possible decisions that could have been considered under any number of circumstances. This is of critical importance for scholars relying on a predictive interpretation of null hypothesis significance testing, like that employed in Perez Brower (2024), where the often relied on interpretation of the p-value is contingent on knowing how many tests were performed. Thus, interpreting the results that are presented in Perez Brower (2024) requires a consideration of the complete set of comparisons that could have been made.

If Perez Brower had shared the results of her experiments using her preferred comparison frame and her first dependent variable, what would she have reported? The lower half of Table 1 reveals what Perez Brower would have reported for Experiment 1 on sexual harassment: there is no significant effect of either race ($p = 0.213$), class ($p = 0.900$), or intersectional race and class ($p = 0.372$) on the number of political activities endorsed by respondents in comparison to an aligned frame.[13] Likewise, for Experiment 2 on the gender pay gap, there is no significant effect for either race ($p = 0.591$), class ($p = 0.862$), or intersectional race and class ($p = 0.398$) on the number of political activities endorsed by respondents in comparison to an aligned frame (see Table 2). Not only are these results nowhere close to statistical significance, they are also in several cases incorrectly signed according to Perez Brower's hypotheses. Despite countering the logic of Perez Brower's second and third hypotheses, and despite her preregistration of these very tests, none of these results are reported in Perez Brower (2024).

Finally, we note that Perez Brower never reports any results for comparisons between Treatment Frames 2, 3, and 4 and the control group for either experiment or dependent variable (these omissions are marked as blue ✗'s in Figure 3). However, all of these results could have produced using her data, and the missing results are reported here as the second, third, and fourth rows of Tables 1, 2, 4, and 5. For Experiment 1, these comparisons are never significant for any treatments involving race – they achieve significance only for Treatment 3, which compares treatment frames that vary class alone. For Experiment 2, none of these treatments differ significantly from the control for either dependent variable.

Let us address one potential responses to our concerns: one might argue that Perez Brower simply (and only) reported comparisons across treatment conditions that she believed to be relevant for testing her hypotheses. (Of course, this requires us to ignore the fact that Perez Brower failed to report all the treatment effects that she pre-registered, and that she selectively omitted results bearing on Hypothesis 1 from Experiment 2, but let us set aside those objections for the moment.) Regardless of whether Perez Brower thought the specific results shown were sufficient to support her own interpretation of the experiments she ran, in failing to consistently report either a full set of sample means across treatment conditions, or a complete accounting of the differences in means across each of her dependent variables for each treatment-control pairing in each experiment, she has made it impossible for readers or reviewers of her article and online

---

[13]See the Appendix at the end of this report for our replication code.

research appendix to answer basic questions without downloading her data and performing their own analyses.

For example, had Perez Brower (2024) reported a simple summary of the means of each of her dependent variables across all ten of her experimental conditions, as we do in our Table 3 below, readers would have learned that in the sexual harassment experiment, not only was the sample mean of political activities endorsed by respondents exposed to a "white professional woman" frame higher than the sample mean for those receiving the ambiguous control: in every treatment group with a specific frame, regardless of the race or class presented in that frame, the sample mean of political activities was higher than the sample mean in the control group. To be sure, as we report in our Table 1 (but Perez Brower does not report), only the white professional frame is significantly different from the control, but it seems relevant for readers and reviewers to know that none of the treatment frames had a sample mean below that of the control, especially given the small samples sizes involved in these experiments.

Likewise, in the gender pay gap experiment, the sample means of political activity for the Black & Latina treatment frames are the highest of all five conditions. Again, there are no significant differences here (as we show in our Table 2, but Perez Brower does not report), but this simple result seems relevant to know. In sum, by opting not to present a consistent and complete account of the outcomes of her dependent variables across experimental conditions, Perez Brower has made it much harder for either readers or reviewers to engage with the question of what her experimental results mean.

Miscoding of the Second Dependent Variable

As discussed in the prior section, Perez Brower (2024, 342) begins her analysis by examining variation in her first dependent variable, a fairly standard index of political action:

> Specifically, in the survey instrument participants were asked to check all the options that applied to the question: "Which of the following activities are you likely to participate in to address the issue?" They were given a list of seven options for political participation, an option that read "none of these options," and another option that read "I would not participate in any activity related to this issue." In the first experiment, political participation was measured as an index count of these seven possible political actions between 0 (selecting none of these options) and 7 (selecting all these options).

Careful readers will note that in addition to the seven options for political participation, Perez Brower gives respondents *two* different ways to declare their intention not to act: they may check "None of these actions", or they may instead check "I would NOT participate in any activity related to the issue of [sexual harassment or the gender wage gap]" Our Figure 5 shows the exact format of this question as shown in Online Research Appendix to Perez Brower of the question format, and confirms that respondents, perhaps confusingly, had two separate ways to register an intent not to participate.

Figure 5. *Text of the survey instrument used to produce Perez Brower's dependent variables.* From page 6 of the Online Research Appendix to Perez Brower (2024).

In the next paragraph of her article, Perez Brower (2024, 342) introduces her second dependent variable, which receives the majority of the empirical attention in her analyses:

> Because I aimed to evaluate the negative effects of framing by class and race, I also included a measure for not supporting the policy agenda by creating a dichotomous variable for selecting "I would not participate in any activity related to this issue." This variable was used in the analyses for both experiments. Asking participants about their future political actions can encourage social desirability among participants (Chung & Monroe, 2003). Including a variable that captures the responses of those that said they would not participate in political action is a harder test, and more accurately illustrates the negative effects of intersectional identity framing. This dichotomous variable is coded such that 1 indicates respondents reported not intending to participate in any political action.

As the response is described in the text of the paper, this corresponds to the variable Q4_9 in the raw data available from TESS: "I would NOT participate in any activity related to the issue of (sexual harassment/the gender wage gap)." Critically, while the survey questionnaire and raw data contain responses to the question precisely how it is described in Perez Brower (2024), this is not the variable used to code her second dependent variable in the analyses. Instead, the author uses the response "None of these actions" (variable Q4_8), completely omitting the respondents who answered "I would NOT participate in any activity related to the issue of (sexual harassment/the gender wage gap)."

Table 3 helps clarify the distribution of respondents across these categories. Each row divides

Table 3. Counts of respondents by experiment, condition, and outcome.

Treatment and control conditions by experiment	Average count of political activities endorsed	Number of respondents endorsing at least one activity	Number of respondents selecting "None of these actions"	Number of respondents selecting "I would NOT participate in any activity…"	Total respondents for this condition
Experiment 1:					
Sexual harassment					
Control frame	3.255	94	3	5	102
White highly-educated	3.864	101	1	1	103
Black & Latina highly-educated	3.544	94	5	4	103
White low-wage	3.893	100	3	0	103
Black & Latina low-wage	3.650	100	1	2	103
Experiment 2:					
Gender Pay Gap					
Control frame	3.635	95	3	6	104
White highly-educated	3.567	95	0	9	104
Black & Latina highly-educated	3.723	89	4	8	101
White low-wage	3.519	95	3	6	104
Black & Latina low-wage	3.806	92	5	6	103

the respondents for a specific condition into three mutually exclusive groups: those who endorsed at least one activity, those who selected "none of these actions", and those who selected "I would NOT participate in any activity related to" the issue presented in the experiment. For each condition, these three categories of respondents sum to the total number of respondents assigned to the condition. The actual number of respondents selecting either "None of these actions" or "I would NOT participate…" tends to be very small, constituting between zero and nine cases per condition. Moreover, across the ten assigned conditions for both experiments, the correlation between these two responses is low ($r = 0.098$), which suggests neither is a good proxy for the other, at least in these small samples.

It is unsurprising, then, that Perez Brower's mistaken use of the count of respondents checking "None of these actions" in place of those checking "I would NOT participate" has a major effect on her results. We re-analyzed both of her experiments using the correct data for the second dependent variable, and present our results in Tables 4 and 5.

Once again, the top half of each table shows the comparison between each of the four treatment conditions and the control group, now for the second dependent variable. These results were not reported in Perez Brower (2024) and provide little if any support for her hypotheses – only one of these eight comparisons reveals a significant difference between a treatment and the ambiguously-framed control (specifically, in Experiment 1, respondents exposed to the white low-wage frame were less likely to say they would not participate compared to those in the control group, $p = 0.024$).

Table 4. Replication of treatment effects for Experiment 1 (Sexual harassment), using the proportion saying they would NOT participate as the dependent variable.

Treatment Frame	Comparison Frame	Proportion endorsing "would NOT participate"		Difference Significant?	How Reported?
		Treated Group	Comparison Group		
White highly-educated women	All women	1 of 103 *treated − comparison =* 0.97% - 4.90% = -3.93%	5 of 102	No, $p = 0.098$ CI: -8.59% to 0.73%	Not Reported
Black & Latina highly-educated women	All women	4 of 103 *treated − comparison =* 3.88% - 4.90% = -1.02%	5 of 102	No, $p = 0.724$ CI: -6.69% to 4.65%	Not Reported
White low-wage women	All women	0 of 103 *treated − comparison =* 0.00% - 4.90% = -4.90%	5 of 102	Yes, $p = 0.024$ CI: -9.16% to -0.64%	Not Reported
Black & Latina low-wage women	All women	2 of 103 *treated − comparison =* 1.94% - 4.90% = -2.96%	5 of 102	No, $p = 0.247$ CI: -7.99% to 2.07%	Not Reported
Black & Latina highly-educated women	White highly-educated women	4 of 103 *treated − comparison =* 3.88% - 0.97% = +2.91%	1 of 103	No, $p = 0.177$ CI: -1.33% to 7.15%	Misreported as "$p < 0.05$" (Data error)
White low-wage women	White highly-educated women	0 of 103 *treated − comparison =* 0.00% - 0.97% = -0.97%	1 of 103	No, $p = 0.320$ CI: -2.90% to 0.95%	"Not significant" (Data error)
Black & Latina low-wage women	White highly-educated women	2 of 103 *treated − comparison =* 1.94% - 0.97% = +0.97%	1 of 103	No, $p = 0.563$ CI: -2.34% to 4.28%	"Not significant" (Data error)

Table 5. Replication of treatment effects for Experiment 2 (Gender pay gap), using the proportion saying they would NOT participate as the dependent variable.

Treatment Frame	Comparison Frame	Proportion endorsing "would NOT participate"		Difference Significant?	How Reported?
		Treated Group	Comparison Group		
White highly-educated women	All women	9 of 104 *treated − comparison =* 8.65% - 5.77% = +2.88%	6 of 104	No, $p = 0.424$ CI: -4.21% to 9.98%	Not Reported
Black & Latina highly-educated women	All women	8 of 101 *treated − comparison =* 7.92% - 5.77% = +2.15%	6 of 104	No, $p = 0.545$ CI: -4.84% to 9.14%	Not Reported
White low-wage women	All women	6 of 104 *treated − comparison =* 5.77% - 5.77% = 0.00%	6 of 104	No, $p = 1.000$ CI: -6.41% to 6.41%	Not Reported
Black & Latina low-wage women	All women	6 of 103 *treated − comparison =* 5.83% - 5.77% = +0.06%	6 of 104	No, $p = 0.986$ CI: -6.38% to 6.49%	Not Reported
Black & Latina highly-educated women	White highly-educated women	8 of 101 *treated − comparison =* 7.92% - 8.65% = -0.73%	9 of 104	No, $p = 0.850$ CI: -8.36% to 6.90%	Misreported as "$p < 0.05$" (Data error)
White low-wage women	White highly-educated women	6 of 104 *treated − comparison =* 5.77% - 8.65% = -2.88%	9 of 104	No, $p = 0.424$ CI: -9.98% to 4.21%	Misreported as "$p < 0.05$" (Data error)
Black & Latina low-wage women	White highly-educated women	6 of 103 *treated − comparison =* 5.83% - 8.65% = -2.83%	9 of 104	No, $p = 0.435$ CI: -9.95% to 4.30%	Misreported as "$p < 0.05$" (Data error)

The bottom half of Tables 4 and 5 show Perez Brower's preferred comparisons and preferred dependent variable for testing Hypotheses 2 and 3. In Table 4 of Perez Brower (2024) and in the main text, four of these six comparisons are reported as significant, albeit in two cases only according to one-tailed tests. However, using the correct data, none of the six comparisons are remotely statistically significant, and several are incorrectly signed, as well. (The two insignificant results reported in Table 4 – for the white low-wage treatment and the Black and Latina low-wage treatment – are correctly reported in the article through serendipity only: Perez Brower used the wrong data to perform these tests, but the correct data also produces non-significant results.) This major error in the use of data means that all results reported in Figures 2 and 3 of Perez Brower (2024) are also incorrect. As a result, none of the significant results in support of Hypotheses 2 or 3 claimed in Perez Brower (2024) hold.

Part of the challenge in interpreting the results of these experiments is the (evidently) confusing presence of two mutually exclusive options for respondents to register their unwillingness to participate. One might next wonder what would happen if we analyzed the sum of respondents in either of these categories. However, this is precisely the sort of fishing expedition that preregistration is meant to prevent. Following her preregistered hypotheses, Perez Brower should have reported results on the difference in the number of activities endorsed when comparing each treatment condition to one or more appropriate reference conditions. With a sole exception, she did not report these results. Instead, she presented results of another way to measure the dependent variable — ostensibly, the count of those who said they would "NOT participate" — but unfortunately, her code selected (and thus her article presented) yet another dependent variable, the number of people who said "none of these." Thus, we have already moved on, in effect, to a third version of the experimental outcome, rather than the original preregistered version. We could proceed to a fourth version – the sum of the "nones" and the "nots" – but at this point, the reason to avoid this sort of fishing should be clear: if we can select from among an array of different *ad hoc* dependent variables after conducting an experiment and viewing the results, then selectively report those differences, reports of statistical significance have no meaning.

Consequences for Perez Brower's (2024) Findings

Selective omission of survey respondents (Problem 1) and selective reporting of outcomes (Problem 2) suggest serious issues of transparency and candor and at an absolute minimum violate generally accepted professional standards for experimental social science with respect to abiding by preregistration and fully reporting experimental results. Reasonable readers or reviewers of Perez Brower (2024) in possession of the full range of results might well have disagreed with Perez Brower's characterization of her findings with respect to her three hypotheses.

We leave it to readers to assess the evidence presented here to determine whether Perez Brower's pattern of omission reflects intentional deception or a highly particular coincidence of sloppy mistakes. Nevertheless, we find it concerning that the pattern of omission so consistently supports the hypotheses offered in her article.

In addition, had Perez Brower (2024) candidly discussed the presence of a control group for Experiment 2 and reported the differences between this control and her first treatment frame, her results for Hypothesis 1 would have been mixed across the two experiments, rather than simply supportive for a single experiment.

Finally, correcting a substantial error in Perez Brower's R code (Problem 3) overturns all of her results on the remaining hypotheses (Hypotheses 2 and 3), leaving essentially no findings in the article still standing.

In consideration of the above issues, Perez Brower (2024) should be retracted.

References

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). "The preregistration revolution." *Proceedings of the National Academy of Sciences* 115(11). 2600–2606.

Perez Brower, M. (2024). "Reframing gendered issues: Intersectional identity frames and policy agendas." *Political Behavior* 46. 331-353.

Appendix: R Code to Produce Tables 1–5 in This Report

```
## The script below constitutes the minimal code needed
## to replicate and/or correct the main results from:
##
## Margaret Perez Brower, 2024, "Reframing gendered issues:
## Intersectional identity frames and policy agendas."
## Political Behavior 46. 331--353.
##
## This script requires the data from gf_clean.csv, which as
## of 24 April 2025 is hosted at:
## https://www.margaretbrower.com/s/gf_clean.csv
##
## The first half of the code consists of the required processing code
## from Perez Brower's own replication script, with extraneous code
## removed. This is taken from the replication code linked in
## Perez Brower (2024), available as of 24 April 2025 at:
## https://www.margaretbrower.com/s/POBE-Final-R-Script.R
##
## The second half of the code is original code to replicate
## Perez Brower's analysis
##

#############################################################################
## Minimal required data processing code from author's replication script

## lines 1-7 of original script
#Clear dataset
rm(list=ls())

#Libraries
library (tidyverse)
library(readr)
library(dplyr)

## lines 23-24 of original script
#Download clean dataset
gf <- read_csv("gf_clean.csv")

## lines 43-44 of original script
#Count Variable
```

```
gf <-mutate(gf, participate.agg= (protest + petition + org + socialmedia
+ discuss + vote + candidate))

## lines 149-179 of original script
#Individal data frames for sexual harassment treatments
#Control
sh_control <- subset(gf, gf$shc== 1)

#Treatment 6: white woman/white collar
sh_white_highinc <- subset(gf, gf$sh_white_highinc== 1)

#Treatment 7: minority woman/ white collar
sh_minority_highinc <- subset(gf, gf$sh_minority_highinc== 1)

#Treatment 8: white woman/ blue collar
sh_white_lowinc <- subset(gf, gf$sh_white_lowinc== 1)

#Treatment 9: minority woman/ blue collar
sh_minority_lowinc <- subset(gf, gf$sh_minority_lowinc== 1)

#Dummy variables for Gender Gap Treatments
#Control
gg_control<- subset(gf, gf$ggc== 1)

#Treatment 1: white woman/ white collar
gg_white_highinc <- subset(gf, gf$gg_white_highinc== 1)

#Treatment 2: minority woman/ white colar
gg_minority_highinc <- subset(gf, gf$gg_minority_highinc== 1)

#Treatment 3: white woman/ blue collar
gg_white_lowinc <- subset(gf, gf$gg_white_lowinc== 1)

#Treatment 4: minority woman/ blue collar
gg_minority_lowinc <- subset(gf, gf$gg_minority_lowinc== 1)


############################################################################
## All code above is taken from Perez Brower's replication file and
## produces just those data structures needed to replicate and correct
## her main results below using new analysis code. None of the code
```

```
## below is taken from Perez Brower

## Code to produce our Table 3 summaring dependent variables
dv1 <- list(sh_control$participate.agg, sh_white_highinc$participate.agg,
            sh_minority_highinc$participate.agg, sh_white_lowinc$participate.agg,
            sh_minority_lowinc$participate.agg, gg_control$participate.agg,
            gg_white_highinc$participate.agg, gg_minority_highinc$participate.agg,
            gg_white_lowinc$participate.agg, gg_minority_lowinc$participate.agg)

noaction <- list(sh_control$noaction, sh_white_highinc$noaction,
                 sh_minority_highinc$noaction, sh_white_lowinc$noaction,
                 sh_minority_lowinc$noaction, gg_control$noaction,
                 gg_white_highinc$noaction, gg_minority_highinc$noaction,
                 gg_white_lowinc$noaction, gg_minority_lowinc$noaction)

neveract <- list(sh_control$neveract, sh_white_highinc$neveract,
                 sh_minority_highinc$neveract, sh_white_lowinc$neveract,
                 sh_minority_lowinc$neveract, gg_control$neveract,
                 gg_white_highinc$neveract, gg_minority_highinc$neveract,
                 gg_white_lowinc$neveract, gg_minority_lowinc$neveract)

n <- sapply(dv1, length)
dv1mean <- sapply(dv1, mean)
dv1positive <- sapply(dv1, function(x){sum(as.numeric(x>0))})
noactionSum <- sapply(noaction, sum)
neveractSum <- sapply(neveract, sum)

dvTable <- cbind(round(dv1mean, digits=3), dv1positive,
                 noactionSum, neveractSum, n)
dvTable <- as.data.frame(dvTable)
dvTable  <- cbind(c("sh_control", "sh_white_highinc", "sh_minority_highinc",
                    "sh_white_lowinc", "sh_minority_lowinc",
                    "gg_control", "gg_white_highinc", "gg_minority_highinc",
                    "gg_white_lowinc", "gg_minority_lowinc"),
                  dvTable)
names(dvTable) <- c("Condition", "dv1Mean", "dv1positive",
                    "noactionSum", "neveractSum", "n")

cor(noactionSum, neveractSum)
```

```
## Code to produce our Table 1

## Comparison of means: Experiment 1 (Sexual Harassment),
##                        DV 1 (Count of behaviors)
##     Comparing each of four treatments to ambiguous frame
t.test(sh_white_highinc$participate.agg, sh_control$participate.agg)
mean(sh_white_highinc$participate.agg) - mean(sh_control$participate.agg)

t.test(sh_minority_highinc$participate.agg, sh_control$participate.agg)
mean(sh_minority_highinc$participate.agg) - mean(sh_control$participate.agg)

t.test(sh_white_lowinc$participate.agg, sh_control$participate.agg)
mean(sh_white_lowinc$participate.agg) - mean(sh_control$participate.agg)

t.test(sh_minority_lowinc$participate.agg, sh_control$participate.agg)
mean(sh_minority_lowinc$participate.agg) - mean(sh_control$participate.agg)


## Comparison of means: Experiment 1 (Sexual Harassment),
##                        DV 1 (Count of behaviors)
##     Comparing treatments 2, 3, and 4 to treatment 1
t.test(sh_minority_highinc$participate.agg, sh_white_highinc$participate.agg)
mean(sh_minority_highinc$participate.agg) - mean(sh_white_highinc$participate.agg)

t.test(sh_white_lowinc$participate.agg, sh_white_highinc$participate.agg)
mean(sh_white_lowinc$participate.agg) - mean(sh_white_highinc$participate.agg)

t.test(sh_minority_lowinc$participate.agg, sh_white_highinc$participate.agg)
mean(sh_minority_lowinc$participate.agg) - mean(sh_white_highinc$participate.agg)


## Code to produce our Table 2

## Comparison of means: Experiment 2 (Gender Pay Gap),
##                        DV 1 (Count of behaviors)
##     Comparing each of four treatments to ambiguous frame
t.test(gg_white_highinc$participate.agg, gg_control$participate.agg)
mean(gg_white_highinc$participate.agg) - mean(gg_control$participate.agg)

t.test(gg_minority_highinc$participate.agg, gg_control$participate.agg)
mean(gg_minority_highinc$participate.agg) - mean(gg_control$participate.agg)
```

```
t.test(gg_white_lowinc$participate.agg, gg_control$participate.agg)
mean(gg_white_lowinc$participate.agg) - mean(gg_control$participate.agg)


t.test(gg_minority_lowinc$participate.agg, gg_control$participate.agg)
mean(gg_minority_lowinc$participate.agg) - mean(gg_control$participate.agg)



## Comparison of means: Experiment 2 (Gender Pay Gap),
##                       DV 1 (Count of behaviors)
##    Comparing treatments 2, 3, and 4 to treatment 1
t.test(gg_minority_highinc$participate.agg, gg_white_highinc$participate.agg)
mean(gg_minority_highinc$participate.agg) - mean(gg_white_highinc$participate.agg)


t.test(gg_white_lowinc$participate.agg, gg_white_highinc$participate.agg)
mean(gg_white_lowinc$participate.agg) - mean(gg_white_highinc$participate.agg)


t.test(gg_minority_lowinc$participate.agg, gg_white_highinc$participate.agg)
mean(gg_minority_lowinc$participate.agg) - mean(gg_white_highinc$participate.agg)



## ## Code to produce our Table 4

## Compute proportions for neveract variable for Experiment 1
NeverActCountsSH <- c(sum(sh_control$neveract),
                      sum(sh_white_highinc$neveract),
                      sum(sh_minority_highinc$neveract),
                      sum(sh_white_lowinc$neveract),
                      sum(sh_minority_lowinc$neveract))
SampleSizesSH <- c(length(sh_control$neveract),
                   length(sh_white_highinc$neveract),
                   length(sh_minority_highinc$neveract),
                   length(sh_white_lowinc$neveract),
                   length(sh_minority_lowinc$neveract))
PctNeverActSH <- NeverActCountsSH/SampleSizesSH*100

## Comparison of means: Experiment 1 (Sexual Harassment),
##                       DV 2 (Proportion "would NOT participate")
##    Comparing each of four treatments to ambiguous frame
t.test(sh_white_highinc$neveract, sh_control$neveract)
100*(mean(sh_white_highinc$neveract) - mean(sh_control$neveract))
```

```
t.test(sh_minority_highinc$neveract, sh_control$neveract)
100*(mean(sh_minority_highinc$neveract) - mean(sh_control$neveract))


t.test(sh_white_lowinc$neveract, sh_control$neveract)
100*(mean(sh_white_lowinc$neveract) - mean(sh_control$neveract))


t.test(sh_minority_lowinc$neveract, sh_control$neveract)
100*(mean(sh_minority_lowinc$neveract) - mean(sh_control$neveract))



## Comparison of means: Experiment 1 (Sexual Harassment),
##                      DV 2 (Proportion "would NOT participate")
##    Comparing treatments 2, 3, and 4 to treatment 1
t.test(sh_minority_highinc$neveract, sh_white_highinc$neveract)
100*(mean(sh_minority_highinc$neveract) - mean(sh_white_highinc$neveract))


t.test(sh_white_lowinc$neveract, sh_white_highinc$neveract)
100*(mean(sh_white_lowinc$neveract) - mean(sh_white_highinc$neveract))


t.test(sh_minority_lowinc$neveract, sh_white_highinc$neveract)
100*(mean(sh_minority_lowinc$neveract) - mean(sh_white_highinc$neveract))



## Code to produce our Table 5

## Compute proportions for neveract variable for Experiment 2
NeverActCountsGG <- c(sum(gg_control$neveract),
                      sum(gg_white_highinc$neveract),
                      sum(gg_minority_highinc$neveract),
                      sum(gg_white_lowinc$neveract),
                      sum(gg_minority_lowinc$neveract))
SampleSizesGG <- c(length(gg_control$neveract),
                   length(gg_white_highinc$neveract),
                   length(gg_minority_highinc$neveract),
                   length(gg_white_lowinc$neveract),
                   length(gg_minority_lowinc$neveract))
PctNeverActGG <- NeverActCountsGG/SampleSizesGG*100

## Comparison of means: Experiment 2 (Gender Pay Gap),
##                      DV 2 (Proportion "would NOT participate")
```

```
##     Comparing each of four treatments to ambiguous frame
t.test(gg_white_highinc$neveract, gg_control$neveract)
100*(mean(gg_white_highinc$neveract) - mean(gg_control$neveract))

t.test(gg_minority_highinc$neveract, gg_control$neveract)
100*(mean(gg_minority_highinc$neveract) - mean(gg_control$neveract))

t.test(gg_white_lowinc$neveract, gg_control$neveract)
100*(mean(gg_white_lowinc$neveract) - mean(gg_control$neveract))

t.test(gg_minority_lowinc$neveract, gg_control$neveract)
100*(mean(gg_minority_lowinc$neveract) - mean(gg_control$neveract))


## Comparison of means: Experiment 2 (Gender Pay Gap),
##                      DV 2 (Proportion "would NOT participate")
##     Comparing treatments 2, 3, and 4 to treatment 1
t.test(gg_minority_highinc$neveract, gg_white_highinc$neveract)
100*(mean(gg_minority_highinc$neveract) - mean(gg_white_highinc$neveract))

t.test(gg_white_lowinc$neveract, gg_white_highinc$neveract)
100*(mean(gg_white_lowinc$neveract) - mean(gg_white_highinc$neveract))

t.test(gg_minority_lowinc$neveract, gg_white_highinc$neveract)
100*(mean(gg_minority_lowinc$neveract) - mean(gg_white_highinc$neveract))
```