Meta-Psychology, 2025, vol 9, MP.2021.2916
https://doi.org/10.15626/MP.2021.2916
Article type: Commentary

Open data: Not Applicable
Open materials: Not Applicable
Open code: Yes
Open and reproducible analysis: Yes
Open reviews and editorial process: Yes
Preregistration: No

Edited by: Rickard Carlsson
Reviewed by: Jack Davis, Peter Dahlgren
Analysis reproduced by: Lucija Batinović
All supplementary files can be accessed at OSF:
https://doi.org/10.17605/OSF.IO/UQ9PS

Comments on Cassese and Holman 2019 "Playing the Woman Card: Ambivalent Sexism in the 2016 U.S. Presidential Race"

L.J Zigerell[1]
[1]Illinois State University

During his campaign for the Republican Party nomination and for U.S. president, Donald Trump suggested that Hillary Clinton benefited from playing a "woman card". The effect of exposure to Trump's woman-card attack was investigated in the Cassese and Holman (2019) Political Psychology article "Playing the woman card: Ambivalent sexism in the 2016 U.S. presidential race". However, neither Cassese and Holman (2019) nor a reanalysis of data analyzed in the article provided sufficient evidence for key claims in the article. Moreover, Cassese and Holman (2019) is unclear whether its Study 2 experimental data could be used to test claims made based on its Study 1 non-experimental data, providing an example of how journal policy requiring access to survey questionnaires could help peer reviewers and readers better assess reported research.

Keywords: hostile sexism; ambivalent sexism; sexism; woman card; gender; Donald Trump; Hillary Clinton

Introduction

During his campaign for the Republican Party nomination and for U.S. president, Donald Trump suggested that Hillary Clinton was benefiting from and was using her identity as a woman (Gass, 2016; Gearan and Zezima, 2016), such as Trump's claim that "Honestly, outside of the woman's card, she's got nothing going" (Pfannenstiel, 2015). Estimates of the effect of exposure to Trump's woman-card attacks were reported in the Cassese and Holman (2019) Political Psychology article "Playing the woman card: Ambivalent sexism in the 2016 U.S. presidential race". However, results reported in Cassese and Holman (2019) do not provide sufficient evidence for its claims that the woman-card attack activated hostile sexism or that the attack decreased support for Clinton among hostile sexists.

Assessing Evidence of Whether Trump's Woman-Card Attack Activated Hostile Sexism

Cassese and Holman (2019) described its Study 1 results as indicating that "[Trump's] woman-card attack activates hostile sexism, bringing it to bear on candidate evaluations and vote choice" (p. 62). However, this inference is not supported at reasonable levels of statistical significance by the evidence reported for Study 1 of Cassese and Holman (2019), as discussed below, because the evidence does not provide sufficient evidence that hostile sexism's association with candidate evaluations and vote choice differed by exposure to Trump's woman-card attack.

The evidence that Cassese and Holman (2019) reported for this claim is from data from a non-experimental non-probability survey (N=950) collected from Amazon's Mechanical Turk platform (MTurk) in May 2016, before Hillary Clinton secured the Democratic Party nomination. Key outcomes were participant ratings about Hillary Clinton on a feeling thermometer, participant ratings about Donald Trump on a feeling thermometer, and whether the participant reported an intention to vote for Hillary Clinton if the 2016 presidential election were held that day with Clinton and Trump as the candidates. Key predictors were a scale of three items measuring participant "hostile sexism" and a scale of three items measuring participant "benevolent sexism", with all of these sexism items drawn from or based on items in Glick and Fiske (1996). Hostile sexism is typically presented as a measure of negative attitudes about women, using participant agreement or disagreement with statements such as "Women seek to gain power by getting control over men". Benevolent sexism is measured using participant agreement or disagreement with statements such as "Women should be cherished and protected by men" and focuses on seemingly positive attitudes about women that can reflect endorsement of chivalrous attitudes that can undercut women.

The key evidence that Cassese and Holman (2019) reported for the claim that Trump's woman-card attack activated hostile sexism is in its Table 2, with pairs of

regressions for each key outcome: the first regression in each pair was limited to participants who reported not having been exposed to Trump's woman-card attack, and the second regression in each pair was limited to participants who reported having been exposed to Trump's woman-card attack. Each regression had participant-level predictors of hostile sexism, benevolent sexism, gender, age, partisanship, household income, education, race and ethnicity, frequency of news consumption, and voter registration status.

The first pair of regressions predicted responses to the Clinton feeling thermometer and indicated that the standardized OLS hostile sexism coefficient was a non-statistically significant -0.08 for participants who reported not being exposed to the woman-card attack and was a statistically significant -0.13 for participants who reported being exposed to the woman-card attack; discussing these results, Cassese and Holman (2019) indicated that: "...one can see that hostile sexism has no effect among survey participants who were not exposed to the attack. Among those exposed to the attack, however, hostile sexism is associated with a significant decline in Clinton evaluations" (p. 62). For the second pair of regressions, which predicted responses to the Trump feeling thermometer, Cassese and Holman (2019) referred to the statistically significant 0.21 standardized OLS hostile sexism coefficient among participants who reported being exposed to the woman-card attack as indicating a "substantively larger" (p. 62) effect than the statistically significant 0.17 standardized OLS hostile sexism coefficient among participants who reported not being exposed to the attack. For the third pair of regressions, which predicted responses about vote intention, Cassese and Holman (2019) claimed that hostile sexism shaped vote choice for Clinton among those exposed to the woman-card attack but not among those who were not exposed to the woman-card attack (p. 62), presumably referencing the presence of statistical significance for the hostile sexism coefficient in the regression among those who reported being exposed to the woman-card attack and the absence of statistical significance for the hostile sexism coefficient in the regression among those who reported not being exposed to the woman-card attack.

After discussing these three pairs of regressions, Cassese and Holman (2019) indicated that "These results support our contention that the woman-card attack activates hostile sexism, bringing it to bear on candidate evaluations and vote choice" (p. 62). But the left side of Figure 1 of the present manuscript illustrates the flaw in this inference. The first line in each panel indicates the point estimate and 95% confidence interval for the OLS coefficient for hostile sexism predicting the indicated outcome net of controls, limited to participants who reported not being exposed to Trump's woman-card attack. The second line in each panel indicates the point estimate and 95% confidence interval for the OLS coefficient for hostile sexism predicting the indicated outcome net of controls, limited to participants who reported being exposed to Trump's woman-card attack. And the third line in each panel indicates the point estimate and 95% confidence interval for the OLS coefficient for hostile sexism's interaction with reported exposure to Trump's woman-card attack, in a regression predicting the indicated outcome net of controls for the combined sample of exposed participants and unexposed participants. The right side of Figure 1 reports corresponding results for the measure of benevolent sexism.

In each Figure 1 regression, the outcome variable ranges from 0 to 100 and the sexism scales range from 0 to 1. A positive coefficient indicates that higher values of the sexism measure associated with higher values of the outcome net of controls, and a negative coefficient indicates that higher values of the sexism measure associated with lower values of the outcome net of controls.

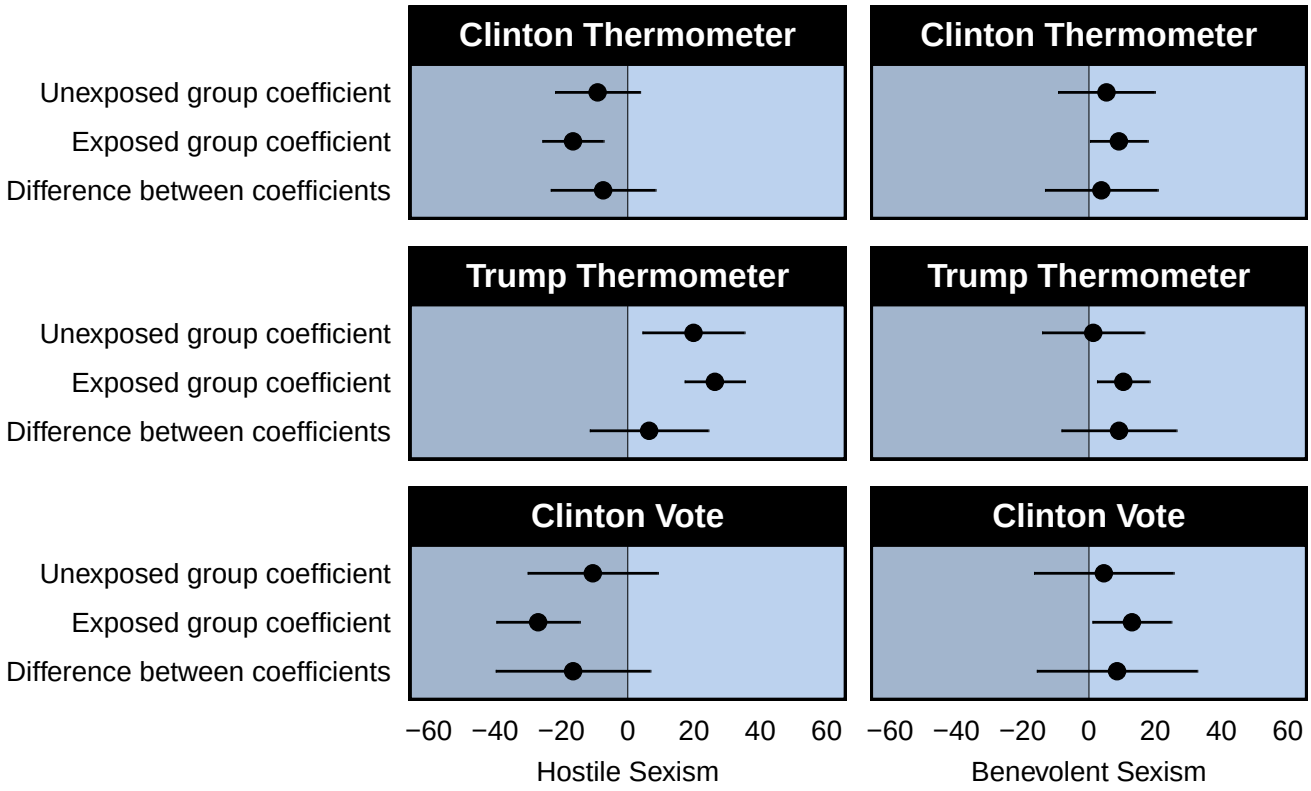
For example, in the top left panel, the top point estimate indicates that the hostile sexism coefficient for the Clinton feeling thermometer outcome variable was -9.1 for participants who reported not having been exposed to the woman-card attack, with a 95% confidence interval that crosses zero; this indicates that the analysis did not provide sufficient evidence at $p<0.05$ that hostile sexism associated net of controls with the Clinton feeling thermometer among participants who reported not having been exposed to the woman-card attack. The middle point estimate in the top left panel indicates that the hostile sexism coefficient was -16.5 for participants exposed to the woman-card attack, with a 95% confidence interval that did not cross zero; this does indicate sufficient evidence at $p<0.05$ that hostile sexism associated net of controls with the Clinton feeling thermometer among participants who reported having been exposed to the woman-card attack.

However, the bottom point estimate in the top left panel indicates that the 95% confidence interval for the -7.4 difference between these two hostile sexism coefficients crosses zero; this means that there is insufficient evidence in this analysis that hostile sexism's association with the Clinton feeling thermometer outcome among participants who reported not having been exposed to the woman-card attack differed from hostile sexism's association with the Clinton feeling thermometer outcome among participants who reported having been exposed to the woman-card attack.

Generally, if the coefficient for a predictor differs be-

Figure 1

Estimated effect of exposure to Trump's woman-card attack



Note. Panels report point estimates and 95% confidence intervals based on OLS regressions predicting values of the outcome indicated in the panel header. Each outcome ranged from 0 to 100, and the sexism predictors ranged from 0 to 1. A positive coefficient indicates that higher values of the sexism measure associated with higher values of the outcome net of controls. In each panel, the top estimate is the coefficient for the indicated sexism predictor among the group that reported not being exposed to Trump's woman-card attack, the middle estimate is the coefficient for the indicated sexism predictor among the group that reported being exposed to Trump's woman-card attack, and the bottom estimate is for the difference between these coefficients. See the main text for information on the controls. Sample sizes are 222 (Clinton thermometer, unexposed), 708 (Clinton thermometer, exposed), 222 (Trump thermometer, unexposed), 704 (Trump thermometer, exposed), 222 (Clinton vote, unexposed), and 708 (Clinton vote, exposed). The plot was constructed in R (R Core Team, 2018) with the tidyverse (Wickham et al., 2019), using estimates from Stata (StataCorp, 2017). Data source: Holman and Cassese (2019) Study 1.

tween regressions in the presence of statistical significance, that is not sufficient evidence to properly conclude that the predictor associates with the outcome differently across regressions (see Gelman and Stern, 2006); a separate test should be conducted to assess whether the coefficient estimates differ from each other. Comparing statistical significance across regressions is especially a concern for the analyses in Table 2 of Cassese and Holman (2019) in which, all else equal, statistical power was larger for the "exposed" regressions that had samples of 704 to 708 participants than for the "not exposed" regressions that had much smaller samples of 215 to 222 participants. Moreover, when comparing two statistically significant standardized OLS coefficients across samples, like for the Trump feeling thermometer outcome, a small 0.04 difference between coefficients is by itself not sufficient evidence of a true difference in the population.

Table 1 reports the p-values for the interaction terms in Figure 1 (top section) and from a test that the indicated sexism coefficient for the exposed group differed from the indicated sexism coefficient for the unexposed group (bottom section), indicating consistency across these two methods in the inference that the hostile sexism coefficients for the unexposed group did not differ at $p<0.10$ from the corresponding hostile sexism coefficients for the exposed group. Therefore, Table 2 in Cassese and Holman (2019) does not provide sufficient

Table 1

Key regression results

	Clinton Feeling Thermometer	Trump Feeling Thermometer	Clinton Vote
p-value for an interaction term			
Hostile sexism × Exposed	p=0.67	p=0.40	p=0.17
Benevolent sexism × Exposed	p=0.52	p=0.14	p=0.40
p-value for a test of the equality of coefficients across regressions			
Hostile sexism	p=0.36	p=0.48	p=0.17
Benevolent sexism	p=0.67	p=0.31	p=0.49

Note. The top two numeric rows indicate the p-value for the indicated interaction term in a combined regression, depicted in the bottom row of each panel in Figure 1. The bottom two numeric rows indicate the two-tailed p-value for a test of the null hypothesis that the coefficient for the indicated variable in the "not exposed" regression equals the corresponding coefficient in the "exposed" regression. Each of these p-values is based on an OLS regression that included the Cassese and Holman (2019) controls. Data source: Holman and Cassese (2019).

evidence for the Cassese and Holman (2019) claim that "[Trump's] woman-card attack activates hostile sexism, bringing it to bear on candidate evaluations and vote choice" (p. 62).

Assessing Evidence of Whether Trump's Woman-Card Affected Hostile Sexists

Cassese and Holman (2019) also claimed that "...we find that hostile sexists exposed to the [woman-card] attack showed increased support for Trump and decreased support for Clinton" (p. 55) and that "The results provided in Table 2 show that hostile sexism shapes candidate evaluations and vote choice among those exposed to the woman-card attack, consistent with Hypothesis 1" (p. 62). Hypothesis 1 is that:

> Hostile sexists who are exposed to the woman-card attack will hold more favorable evaluations of Trump and more unfavorable evaluations of Clinton, resulting in a reduced willingness to vote for Clinton and an increased willingness to vote for Trump.

However, Hypothesis 1 cannot be properly tested with results in Table 2 of Cassese and Holman (2019) that were based on the full sample of participants that included some participants low in hostile sexism, such as the 14.5% or so of participants at the lowest observed level of hostile sexism. Even sufficient evidence that hostile sexism associated with an outcome differently by whether a participant reported being exposed to Trump's woman-card attack could not support the inference that the woman-card attack had an effect among hostile sexists, because a difference in the association between hostile sexism and the outcome could have been caused by a difference among participants at lower levels of hostile sexism.

To assess whether reported exposure to Trump's woman-card attack associated with the outcomes net of controls among participants high in hostile sexism, Figure 2 reports point estimates and 95% confidence intervals from OLS regressions using the Cassese and Holman (2019) Table 2 predictors and an indicator of whether the participant reported being exposed to Trump's woman-card attack, but limited to high levels of the indicated sexism scale, with a positive coefficient indicating that reported exposure to the woman-card attack associated with higher values of the outcome net of controls.

For example, "Top 5% HS" in the middle left panel is the point estimate and 95% confidence interval for the "exposed" predictor indicating reported exposure to the woman-card attack, among participants in the top 5% of observations on the hostile sexism scale in a regression that included benevolent sexism and all other Cassese and Holman (2019) Table 2 predictors except for hostile sexism[1]; this Top 5% HS point estimate indicates that, net of controls, participants who reported exposure to Trump's woman-card attack were predicted to rate Trump 12 units higher compared to the rating about Trump from participants who did not report ex-

[1]Percentage categories such as Top 5% HS were calculated based on levels of the indicated sexism scale across the pooled set of participants in the exposed group and the not exposed group who had full data on the predictors.

posure to the woman-card attack[2].

The left panels of Figure 2 indicate a lack of consistent evidence at $p<0.05$ for any of the three outcomes for the Cassese and Holman (2019) claim that: "...we find that hostile sexists exposed to the attack showed increased support for Trump and decreased support for Clinton" (p. 55). The right panels of Figure 2 indicate $p<0.05$ evidence for the Trump feeling thermometer at some levels of high benevolent sexism, but this $p<0.05$ evidence did not appear in regressions for the Clinton feeling thermometer or for the vote intention outcome.

To address the concern that exposure to Trump's woman-card attack affected both hostile sexism and benevolent sexism, so that the regressions at high levels of one of these types of sexism should not control for the other type of sexism, Figure 3 reports Figure 2 analyses but not controlling for the other measure of sexism, so that, for example, analyses in the left panels of Figure 3 about hostile sexism did not control for benevolent sexism. Estimates in Figure 3 are very similar to estimates in Figure 2.

Possible Test of Hypothesis 1 using Cassese and Holman (2019) Study 2 Data?

Data from Cassese and Holman (2019) Study 1 were non-experimental, in which participants were asked to indicate whether they had already heard anything about Donald Trump's accusation that Hillary Clinton was "playing the woman card"; participants had not been randomly assigned to be exposed or not exposed to Trump's woman-card attack. Compared to these Study 1 non-experimental data, data from an experiment that randomized participants to exposure to a news report about Trump's woman-card attack would provide stronger causal inference about the effect of such a woman-card attack. This section discusses the possibility that such a test was provided by Study 2 of Cassese and Holman (2019).

Data for Cassese and Holman (2019) Study 2 ($N=400$) were collected "several months later" from "an experimental design that randomly assigned survey participants to read an article about the woman-card attack from The New York Times or a control article about the use of social media in the campaign, rather than relying on participant recall of the attack" (Cassese and Holman, 2019, p. 64). The Cassese and Holman (2019) main text and its Table A4 column headers indicate that Study 2 included each element that would have been needed to provide an experimental test that paralleled the non-experimental analyses reported in Cassese and Holman (2019) Table 2: randomized exposure or non-exposure to an article about Trump's woman-card attack, a measure of participant hostile sexism, a measure of participant benevolent sexism, participant rating about Hillary Clinton on a feeling thermometer, participant rating about Donald Trump on a feeling thermometer, and a measure of whether the participant intended to vote for Hillary Clinton.

However, Cassese and Holman (2019) did not report whether responses to these variables differed by the experimental treatment in Study 2. The data for Cassese and Holman (2019) were uploaded to the Harvard Dataverse after I requested the data from the corresponding author; however, by the start of June 2021, neither author of Cassese and Holman (2019) provided me with information regarding my October 2020 email that included a question of whether the vote intention item and/or the candidate thermometers were asked after the experimental treatment in Study 2. The Study 2 codebook omits the feeling thermometer items that the Cassese and Holman (2019) Table A4 column headers indicate were included in Study 2, and the available Study 2 dataset does not include responses to the feeling thermometer items. However, the Study 2 dataset does include responses to a vote intention item.

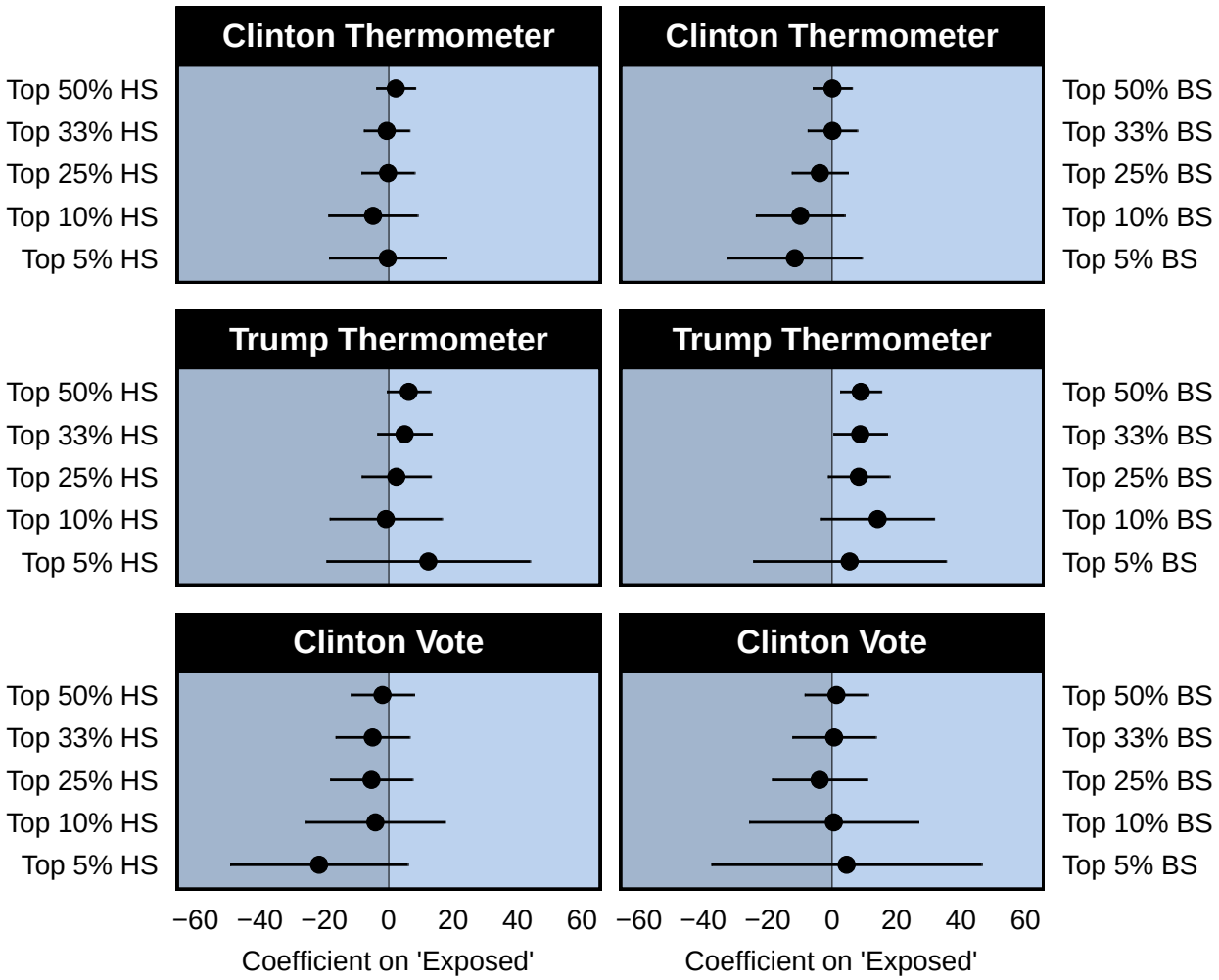
Nothing in Cassese and Holman (2019) or its supporting information appears to explicitly indicate whether measurement of the Study 1 outcome variables occurred before or after the treatment in Study 2, but, circumstantially, the "presvote" presidential vote intention item is the last item listed in the Cassese and Holman (2019) Study 2 codebook, after the treatment and after the emotion items and participation items that Cassese and Holman (2019) indicated were post-treatment (pp. 64-65)[3]. Thus, the following analysis of Study 2 data assumes that the intention item was asked

[2]Regarding the possibility that exposure to Trump's woman-card attack had affected the mean level of hostile sexism or benevolent sexism among participants, OLS regressions predicting hostile sexism and benevolent sexism using the "exposed/not exposed" predictor and the Cassese and Holman (2019) demographic predictors did not provide sufficient evidence that the mean level of hostile sexism ($p=0.624$) or the mean level of benevolent sexism ($p=0.625$) differed by exposure to Trump's woman-card attack net of controls.

[3]Inspection of the Holman and Cassese (2019) Study 2 dataset for clues about whether "presvote" appeared after the assignment to the treatment indicated that the first variable in the dataset is the variable indicating whether the case was assigned to the treatment. Cassese and Holman (2019) indicated that the sexism items were asked before the treatment (p. 64), so the order of the variables in the dataset does not help resolve the location of "presvote" in the survey. Cassese and Holman (2019) indicated that the participation items and the emotional reaction items were asked after assignment to the treatment (pp. 64-5), but ten dataset cases have a participation score and values for the anger, enthusiasm, and anxiety emotional reaction items but are missing a value for the

Figure 2

Estimated effect of exposure to Trump's woman-card attack on the outcome indicated in the panel header, at selected levels of hostile sexism and benevolent sexism



Note. Panels report point estimates and 95% confidence intervals for the predictor indicating exposure to Trump's woman-card attack, based on OLS regressions predicting values of the indicated outcome variable at selected ranges of hostile sexism (HS) or benevolent sexism (BS). A positive coefficient indicates that exposure to the woman-card attack associated with higher levels of the outcome net of controls. Each outcome variable ranged from 0 to 100. See the main text for information on the controls. Top percentages were based on the 931 participants with full data for the predictors. Sample sizes for regressions using all participants were 930 for the Clinton thermometer regression, 926 for the Trump thermometer regression, and 930 for the Clinton vote regression. The highest level of hostile sexism had only 7 participants in the "exposed" group and 5 participants in the "not exposed" group, and the highest level of benevolent sexism had only 14 participants in the "exposed" group and 4 participants in the "not exposed" group. The plot was constructed in R (R Core Team, 2018) with the tidyverse (Wickham et al., 2019), using estimates from Stata (StataCorp, 2017). Data source: Holman and Cassese (2019) Study 1.

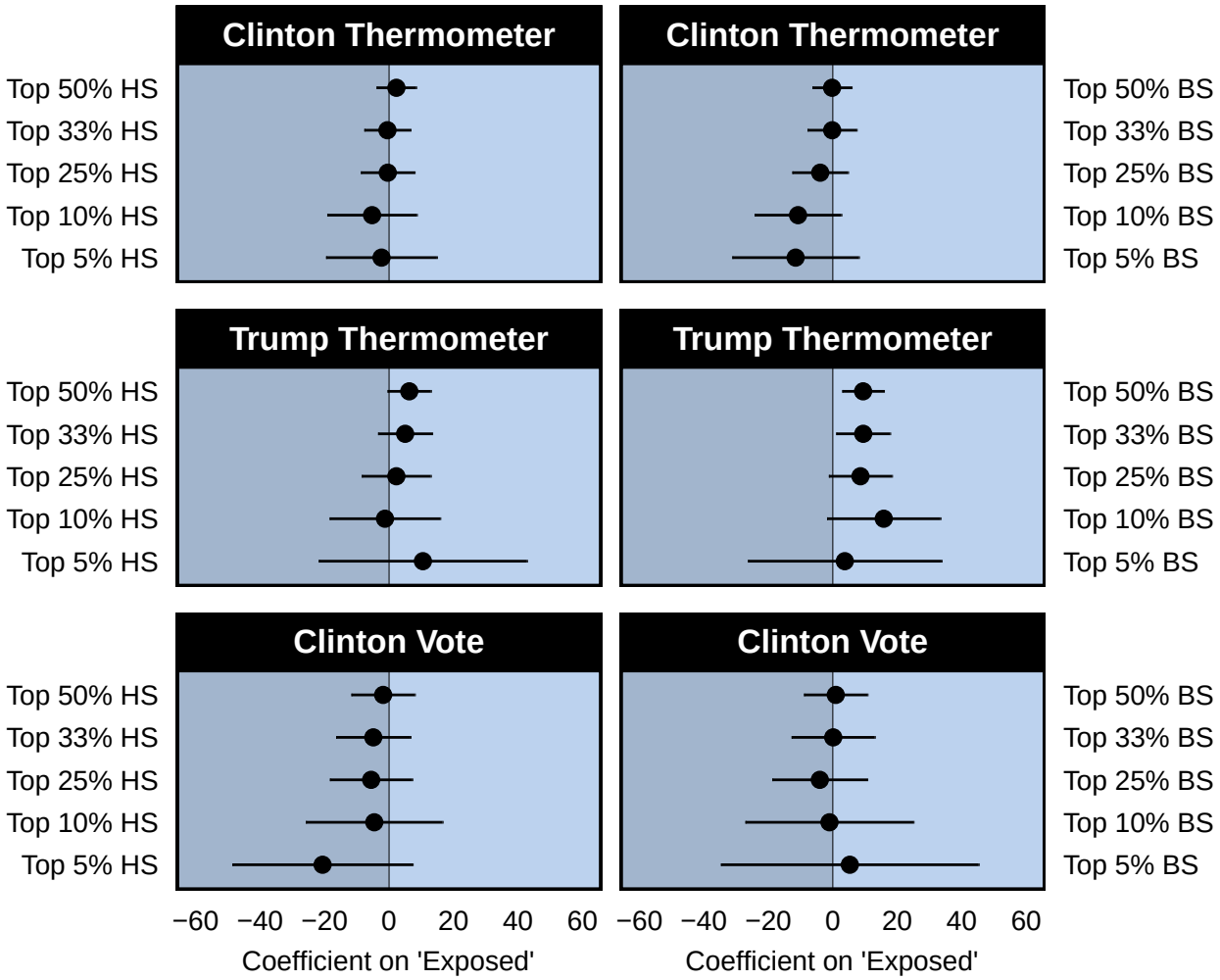
after the experimental treatment in which participants received a control article or an article about the woman-card attack.

The vote intention variable in the Study 2 dataset had four non-missing values: Clinton (184 cases), Trump (95 cases), 3rd Party (82 cases), and Will Not Vote

indicator of whether the case was assigned to the treatment. My October 2020 email to the authors of Cassese and Holman (2019) asked about how that pattern of missingness could occur if the participation items were asked after the assignment to treatment, but I have to date not received a reply to that email.

Figure 3

Figure 2 analyses, but with no control for the other measure of sexism



Note. See the note for Figure 2.

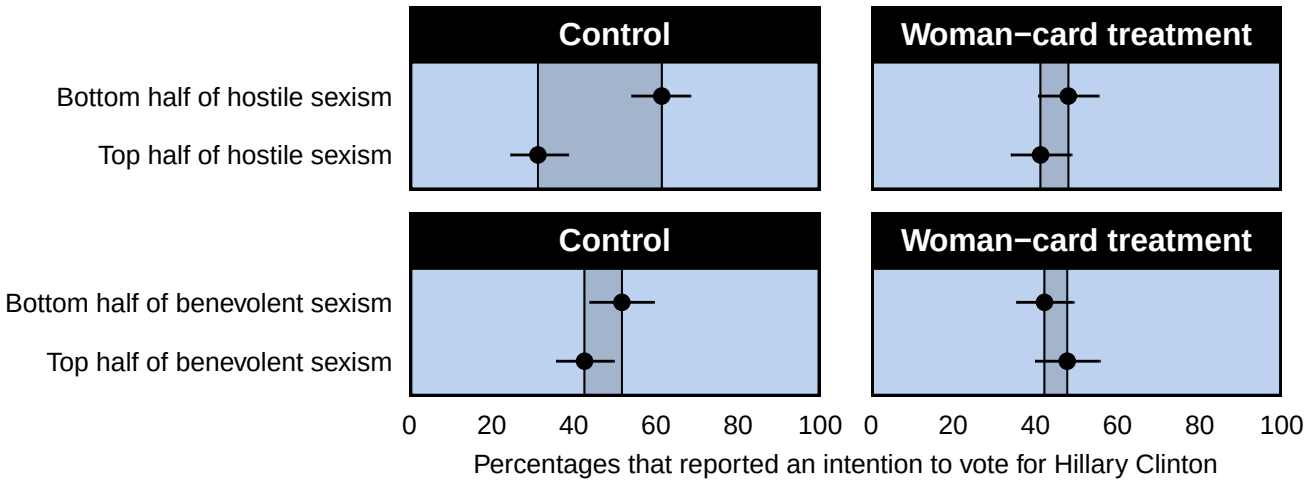
(43 cases). Using this variable, I coded two outcome variables: voteC (with "Clinton" coded 1 and the 220 other non-missing responses coded 0), and voteCT (with "Clinton" coded 1, "Trump" coded 0, and the 125 other non-missing responses coded as missing)[4].

I predicted values of the voteC outcome indicating a vote for Clinton versus other responses, based on an OLS regression that included predictors for only hostile sexism, the experimental treatment (0 for cases assigned to the control, and 1 for cases assigned to the treatment), and an interaction of hostile sexism and the experimental treatment, with all variables ranging from 0 to 1 (N=394). The coefficient on the hostile sexism constituent term was -0.81, indicating that, in the control group, higher levels of hostile sexism associated with a lower probability of reporting an intention to

[4]Cassese and Holman (2019) Hypothesis 1 included "a reduced willingness to vote for Clinton and an increased willingness to vote for Trump" (p. 57), but Cassese and Holman (2019) described the vote intention item as being coded "1 if respondents chose Clinton and zero otherwise" (p. 60), so that Cassese and Holman (2019) is not clear about whether the vote intention item should be interpreted as Clinton versus Trump or as Clinton versus other all other vote intentions or intentions to not vote. Based on the Study 1 codebook in the Holman and Cassese (2019) data, the vote intention item was "If the 2016 presidential election were being held today and the candidates were Hillary Clinton and Donald Trump, who would you vote for?", and the dataset variable was coded to have only two categories: "Hillary Clinton" and "Not Hillary Clinton". The Study 1 data indicate that, for participants coded 0 or 1 for the vote intention item, 11% of Republicans (28 of 252) were coded into the "Hillary Clin-

Figure 4

Estimates of the percentages that reported an intention to vote for Hillary Clinton relative to other non-missing responses, at selected levels of hostile sexism and benevolent sexism, by experimental condition



Note. Panels report point estimates and 83.4% confidence intervals, at selected values of hostile sexism (top panels) and benevolent sexism (bottom panels), with no controls. See the main text for a discussion about the uncertainty of whether the vote intention item was asked after the treatment. Sample sizes were 101 (control, bottom half of hostile sexism), 96 (control, top half of hostile sexism), 100 (treatment, bottom half of hostile sexism), 97 (treatment, top half of hostile sexism), 89 (control, bottom half of benevolent sexism), 108 (control, top half of benevolent sexism), 109 (treatment, bottom half of benevolent sexism), and 88 (treatment, top half of benevolent sexism). The plot was constructed in R (R Core Team, 2018) with the tidyverse (Wickham et al., 2019), using estimates from Stata (StataCorp, 2017). Data source: Holman and Cassese (2019) Study 2.

vote for Clinton. However, the coefficient on the interaction of hostile sexism and the treatment indicator was +0.48 with a p-value of p=0.01; if the vote item was asked after the treatment, this indicates evidence that the woman-card treatment reduced the negative association of hostile sexism and reported intention to vote for Clinton[5]. The top panel of Figure 4 illustrates this pattern by reporting point estimates and 83.4% confidence intervals (see Payton et al., 2003) for percentages of participants in the control and in the treatment that reported an intention to vote for Clinton, disaggregated by levels of hostile sexism.

The coefficient for the interaction of hostile sexism and the treatment was +0.26 in a parallel, smaller sample OLS regression predicting the voteCT outcome indicating a vote for Clinton instead of Trump (N=273), but the p-value was only p=0.16. In a regression including predictors for only benevolent sexism, the exposure treatment, an interaction of benevolent sexism and the treatment, the p-value for the interaction of benevolent sexism and the treatment was p=0.23 predicting voteC and was p=0.35 predicting voteCT. Regardless of whether the vote intention item in Study 2 was asked after the experimental treatment, Study 2 results reported above do not indicate that the woman-card treatment increased the influence of hostile sexism in general or that the woman-card treatment increased opposition to Clinton among participants in the top half of hostile sexism.

Discussion

Cassese and Holman (2019) claimed that their data indicated that "[Donald Trump's] woman-card attack activates hostile sexism, bringing it to bear on candidate evaluations and vote choice" (p. 62). In reference, Utych (2020, p. 1) indicated that "Most directly, hostile sexism can cause voters to not vote for a woman..., but it can also be activated through sexist language and themes used by other candidates, as it was by Donald Trump in 2016 (Cassese and Holman, 2019)" [reference to another article omitted in the ellipsis]. How-

ton" category, but 30% of Democrats (163 of 534) were coded into the "Not Hillary Clinton" category. Thirty percent would be a remarkably high percentage of Democrats that preferred Donald Trump to Hillary Clinton, if the "Not Hillary Clinton" coding for the vote intention item had indicated only a vote intention for Donald Trump.

[5]The p-value for the interaction of hostile sexism and the treatment was p=0.01 controlling for benevolent sexism.

ever, these claims about the activation of hostile sexism were not sufficiently supported in Cassese and Holman (2019) or in the reanalysis reported in this manuscript. Beyond these substantive results, this comment provides an illustration of the methodological recommendation that, if the size of one estimate is compared to the size of another estimate, an inference that the estimates differ in size from each other should be based on a statistical test of whether the estimates differ in size from each other, instead of, as in Cassese and Holman (2019), a reference to a difference in the size of the point estimates or a reference to the combination of sufficient evidence that one of the estimates differs from zero and insufficient evidence that the other estimate differs from zero.

This comment also provides additional evidence that the quality of publications using survey data could be improved if journals and other outlets required as a condition of peer review and publication that peer reviewers and the reading public have access to the complete set of items and treatments that were used in each survey that is reported on in a submission or publication. In the case of Cassese and Holman (2019), access to the full questionnaire for Study 2 would help readers assess whether Study 2 provided an experimental test of claims tested with non-experimental data in Study 1.

Author Contact

Correspondence concerning this article should be addressed to: L.J Zigerell, ljzigerell@IllinoisState.edu, ORCID 0000-0003-4262-8405.

Acknowledgements

The author thanks Jack Davis and Peter Dahlgren for their reviews.

Conflict of Interest and Funding

The author declares no conflict of interest and received no funding for the analysis.

Author Contributions

L.J Zigerell is the sole author of this contribution.

Open Science Practices

This article earned the Open Code badge for making the code openly available. Data used for the analysis in this commentary does not belong to the author, however, it has been verified that the analysis reproduced the results presented in the article. The OSF project is available here: https://doi.org/10.17605/OSF.IO/42AF9. The entire editorial process, including the open reviews, is published in the online supplement.

References

Cassese, E. C., & Holman, M. R. (2019). Playing the woman card: Ambivalent sexism in the 2016 U.S. presidential race. *Political Psychology*, *40*(1), 55–74. https://doi.org/10.1111/pops.12492

Gass, N. (2016). Trump: I guess I'll have to get used to Clinton's 'shouting'. *Politico*. https://www.politico.com/blogs/2016-gop-primary-live-updates-and-results/2016/04/trump-hillary-clinton-woman-card-222519

Gearan, A., & Zezima, K. (2016). Trump's 'woman's card' comment escalates the campaign's gender wars. *Washington Post*. https://www.washingtonpost.com/politics/trumps-womans-card-comment-escalates-gender-wars-of-2016-campaign/2016/04/27/fbe4c67a-0c2b-11e6-8ab8-9ad050f76d7d_story.html

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician*, *60*, 328–331. https://doi.org/10.1198/000313006X152649

Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*, 491–512.

Holman, M., & Cassese, E. (2019). Replication data for: Playing the women card, v1. https://doi.org/10.7910/DVN/KYQIB5

Payton, M. E., Greenstone, M. H., & Schenker, N. (2003). Overlapping confidence intervals or standard error intervals: What do they mean in terms of statistical significance? *Journal of Insect Science*, *3*(34), 1–6. https://doi.org/10.1093/jis/3.1.34

Pfannenstiel, B. (2015). Trump in his own words: Highlights of a 95-minute rant. *Des Moines Register*. https://www.desmoinesregister.com/story/news/elections/presidential/caucus/2015/11/13/trump-his-own-words-highlights-95-minute-rant/75729070/

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria. www.R-project.org/

StataCorp. (2017). *Stata statistical software: Release 15*. College Station, TX.

Utych, S. M. (2020). Sexism predicts favorability of women in the 2020 democratic primary...and men? *Electoral Studies*. https://doi.org/10.1016/j.electstud.2020.102184

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Lin Pedersen, T., Miller, E., Milton Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686