

Peer Review Comments

Article: Westgate, E C, Riskind, R G and Nosek, B A 2015 Implicit preferences for straight people over lesbian women and gay men weakened from 2006 to 2013. *Collabra*, 1(1): 1, DOI: <http://dx.doi.org/10.1525/collabra.18>

Article type: Original research report

Subject: Psychology

Editor: Simine Vazire

Article submitted: 15 April 2015

Editor decision: 10 May 2015

Revision submitted: 12 June 2015

Article accepted: 25 June 2015

Article published: 23 July 2015

Responses for Version 1

Reviewer: Elizabeth Page-Gould

Affiliation: University of Toronto, Canada

Competing Interests Statement: I have no conflict of interest in reviewing this research.

Review Completed: 04 May 2015

Recommendation: Revisions Required

1) General comments and summary of recommendation:

Describe your overall impressions and your recommendation, including changes or revisions. Please note that you should pay attention to scientific, methodological, and ethical soundness only, not novelty, topicality, or scope. A checklist of things to you may want to consider is below:

- Are the methodologies used appropriate?
- Are any methodological weaknesses addressed?
- Is all statistical analysis sound?
- Does the conclusion (if present) reflect the argument, is it supported by data/facts?
- Is the article logically structured, succinct, and does the argument flow coherently?
- Are the references adequate and appropriate?

Overall, this is an interesting research question and an excellent dataset for testing it. The open data was greatly appreciated and sets the stage for a high level of impact when taken together with the fundamentally interesting research question. However, the writing was at times redundant (Minor Comment 1) and at other times too minimal (Major Comment 5). The biggest issue is that the data most likely violate the assumptions of hierarchical regression analysis, and so the authors are going to have to find a way to properly analyze their time-based data (Major Comment 1).

MAJOR COMMENTS:

1. Unfortunately, there is one fundamental issue with using classic regression to analyze changes in attitudes across so many days: the assumption of independent residuals is almost certainly violated when analyzing changes in attitudes across so many days. IAT values are going to be autocorrelated, which means that IAT scores measured on two consecutive days are likely to be more similar to each other than IAT scores measured on a random day in 2006 and a random day in 2010. When you are measuring 50 or more time points with high resolution as in the present study, then most processes exhibit both periodic (or "seasonal") components over time that are sometimes paired with linear trends. Indeed, looking at the graph of implicit and explicit preferences across time, it is easy to see the classic sinusoidal elements in attitudes over time (i.e., where attitudes appear to go up for a couple of data points and then down for a few data points and then back up again). You have to remove the periodic elements from the IAT scores in order to isolate the linear trend. Until you do, then the residuals for days that are closer together in time are almost certainly going to be correlated with each other, and thus regression is not valid. The good news is that the periodic nature of implicit and explicit attitudes will undoubtedly be very interesting and an excellent addition to the paper. The bad news, of course, is that someone on the team will need to learn how to do a time series analysis in

order to separate the periodic and linear trends from the implicit and explicit attitudes data. Luckily, all major statistical packages have extensive capabilities for time series analysis.

2. Given that this paper is about stability versus change in attitudes about sexual orientation, it is important that the attitudes are measured with equal reliability across the relevant timespan. As a result, the variance in the number of IAT sessions completed per day was notable. It leads to the question of whether the number of IAT sessions per day varied randomly or systematically across the seven years of the study (i.e., was there a linear trend for more or fewer participants over the timespan of the study). Since time was operationalized with daily resolution, then having some days with only one observation means that the IAT measurement was not equally reliable across all days. This would be less important if all days were measured with a minimum number of sessions of 1000 observations, just as an example benchmark (i.e., because $t \sim z$ at 999 degrees of freedom). If there is a linear trend for more people to complete the sexual attitudes IAT between 2006 and 2013, then that would suggest that the more recent sexual attitude IATs were measured with greater reliability than the older sexual attitude IATs. If variance in the number of IAT sessions/day was spread out randomly across the seven-year timespan of the study, then the variation in sampling error between days would not be as much of a problem. Can the authors provide evidence that their estimates of both the magnitude and rate of change in sexual attitudes between 2006 to 2013 are not affected by measurement error?

3. The authors note that the effect size for explicit attitudes was greater than the effect size for implicit attitudes, and they infer this difference to reflect rate of change (e.g., see the last sentence of the first paragraph of the discussion). I would argue that the difference in effect sizes simply reflects the magnitude of change and not necessarily the rate. If you want to make comments about rate of change for the two types of measures, then you should conduct the analyses that focus on slopes, not mean differences. One such analysis would be a growth curve (e.g., multilevel model) where standardized values of implicit and explicit attitudes are modelled as a function of date and attitude type (i.e., implicit or explicit) is used as a moderator of date, with a random intercept for date is modeled for each day. If attitude type is indeed a moderator of date, then you could say that you observed a more rapid shift in explicit attitudes than implicit attitudes. In the absence of doing an analysis that was focused on the trajectory/slope of change, then it is best to interpret effect size differences as simply reflecting different magnitudes of effects.

4. Footnote 2 reports that the IATs were standardized independently to account for a few minor methodological variations in the IATs over the years. In comparison, Footnote 1 states that explicit attitudes were only analyzed for the late 2006 - 2013 time period due to a methodological variation, with the early 2006 data being presented as supplemental information. Why not deal with both problems in the same way by standardizing responses within the 5- and 7-point scales? The text body seems to suggest that explicit data were standardized (p. 10), but why not do that for the 5-point scale as well as the 7? In addition to being internally consistent, standardizing the explicit attitude measure would allow all the explicit data to be included in the primary analyses, which is preferable both for estimation and for comparison with the implicit attitudes data.

5. Please provide the numbers that led to all conclusions so that the reader can independently decide if they agree with each conclusion. The authors verbally describe demographic trends in the sample across the timespan of the study (e.g., "The largest shifts were for age and education: the sample became slightly younger and more educated over time."), but do not always provide any numbers that support the trend (e.g., there are no numbers that support the statement that the shifts in age and education were either large or reliably different over time).

MINOR COMMENTS:

1. From Table 1 and best fit to the research question, I believe that IAT score was actually regressed on participation date as opposed to the other way around as it is stated on p. 11.

2. The paper could be additionally streamlined, as some parts were somewhat redundant. For example, the first paragraph of the results section is essentially a summary of the methods; all the numbers given in the first paragraph of the results section were provided in the methods section.

2) Figures/tables/data availability:

Please comment on the author's use of tables, charts, figures, if relevant. Please acknowledge that adequate underlying data is available to ensure reproducibility (see open data policies per discipline of Collabra here).

The authors made a great use of tables, charts, and open data. Tables were used to comprehensively report model specification and results, and the figures were used to visualize the results in a man-

ner that allowed the reader to draw their own conclusions (e.g., that there are periodic trends in the data). The data are public and thus allow for independent verification of the results. All of these aspects of the paper were greatly appreciated.

3) Ethical approval:

If humans or animals have been used as research subjects, and/or tissue or field sampling, are the necessary statements of ethical approval by a relevant authority present? Where humans have participated in research, informed consent should also be declared.

If not, please detail where you think a further ethics approval/statement/follow-up is required.

The authors declared that they obtained informed consent prior to data collection.

4) Language:

Is the text well written and jargon free? Please comment on the quality of English and any need for improvement beyond the scope of this process.

The text is generally well-written and is free of jargon, but I had some comments about streamlining the writing in my general review.

Reviewer: Alexa Tullett

Affiliation: University of Alabama, USA

Competing Interests Statement: I have no competing interests to declare.

Review Completed: 04 May 2015

Recommendation: Revisions Required

1) General comments and summary of recommendation:

Describe your overall impressions and your recommendation, including changes or revisions. Please note that you should pay attention to scientific, methodological, and ethical soundness only, not novelty, topicality, or scope. A checklist of things to you may want to consider is below:

- *Are the methodologies used appropriate?*
- *Are any methodological weaknesses addressed?*
- *Is all statistical analysis sound?*
- *Does the conclusion (if present) reflect the argument, is it supported by data/facts?*
- *Is the article logically structured, succinct, and does the argument flow coherently?*
- *Are the references adequate and appropriate?*

In this manuscript the authors document a 13.4% decline in implicit preferences for straight people over lesbian women and gay men from 2006 to 2013. I thought this paper addressed an inherently interesting question and I found the data very persuasive. My comments are suggestions for making improvements on an already strong paper.

Major Points

1. I would have hoped to see more discussion of what declining implicit bias (i.e. IAT scores) might mean on a practical level. In the Introduction the authors say "Should we attribute shifts in reported attitudes to shifts in evaluation? Or, are we observing shifts in willingness to express negative evaluations of lesbian and gay people" (p. 5). A very conservative interpretation of IAT scores might suggest that they are not in fact indexing evaluation, but instead something more basic like associations. I am not an expert on the IAT, but I'm not persuaded of the authors' assumption that the IAT is a better measure of people's evaluations than are self-report measures.

2. One of the authors' main conclusions is that "The estimated implicit preference for straight people declined by 13.4%" from 2006 to 2013. They also note that the demographics of their sample differs by date, with the sample becoming younger, more educated, and slightly more female and less straight during that time. My understanding is that the authors control for these demographics when evaluating variation in implicit preferences at the individual level, but it was not clear to me that they did so when testing cultural change (i.e. how the mean is shifting across time). I think it would be important to demonstrate that this 13.4% change does not simply reflect a change in the demographics of the sample over time.

Minor Points

3. Does the fact that people are self-selecting the Sexuality IAT suggest that results will underestimate the amount of implicit bias against homosexuality in the population?
4. A rationale for the exclusions based on error rates would be helpful. Perhaps this could go in the Appendix.
5. It could be informative to give a more thorough explanation of the stimuli used in the IAT. The article says "Target stimuli for Gay people consisted of pictures and words representing either lesbian women or gay men." How do the pictures convey information about sexual orientation? Are they pictures of couples? Or of people who strongly conform to appearance-related stereotypes? (I later noticed that this is mentioned briefly in the Discussion, but it would be helpful to also include this information in the Methods).

2) Figures/tables/data availability:

Please comment on the author's use of tables, charts, figures, if relevant. Please acknowledge that adequate underlying data is available to ensure reproducibility (see open data policies per discipline of Collabra here).

See "General comments"

3) Ethical approval:

If humans or animals have been used as research subjects, and/or tissue or field sampling, are the necessary statements of ethical approval by a relevant authority present? Where humans have participated in research, informed consent should also be declared.

If not, please detail where you think a further ethics approval/statement/follow-up is required.

See "General comments"

4) Language:

Is the text well written and jargon free? Please comment on the quality of English and any need for improvement beyond the scope of this process.

See "General comments"

Editor Decision for Version 1

Editor: Simine Vazire

Affiliation: University of California, Davis, USA

Editor decision: Revisions required

Decision date: 10 May 2015

Dear Dr. Westgate,

I have now received two reviews of your manuscript from experts in attitudes and/or research methods. I also independently read the manuscript before consulting these reviews. All three of us saw much to like in the paper. I think it is close to being publishable in Collabra. However, I have a few changes I would like you to consider. Thus, I encourage you to submit a revised version of this paper for further consideration at Collabra. I would like to receive such a revision within 30 days. If you are able to address these concerns, I will be happy to accept the paper.

The reviewers did an outstanding job of articulating their concerns with the paper and I will not summarize all of their points in my letter. I would like you to address all of their concerns in the letter of response and if possible in the revised paper. Below I summarize the biggest issues you should address in the revision.

1. Reviewer 1 brings up some important questions about your analytic strategy. Please give careful thought to all of her suggestions and address as necessary.
2. Reviewer 2 raised the question of whether the observed shifts in IAT scores could be due to shifts

in associations. I realize this is a longstanding debate about how to interpret the IAT, but I think it is worth addressing. You might particularly consider whether the association explanation is more or less likely to be true in this particular instance, or if the extent to which IAT scores may be driven by associations rather than evaluations, even if small, could be enough to account for the shifts you found.

3. I also had the same thought as Reviewer 2 about the self-selection issue. Is there any reason to think that the self-selected nature of the sample could have contributed to the shift in scores in ways that could not be picked up by demographics? In other words, is there any reason why the propensity for people with less anti-gay bias, compared to people with more bias, to self-select into this study might have increased between 2006 and 2013 that cannot be accounted for by controlling for the demographic variables you controlled for?

In summary, I think this is a very rigorous paper that will make an important contribution to the literature. Thus, I hope you will revise this for further consideration at Collabra. In order to expedite the next step in the editorial processing of your paper, please carefully follow the instructions you received about resubmitting your manuscript.*

Author's Response to Review Comments for Version 1

Author: Erin C Westgate
Affiliation: University of Virginia, USA
Revision submitted: 12 June 2015

Dear Dr. Vazire,

Our revised manuscript, "Implicit Preferences for Straight People over Gay People Decline from 2006 to 2013", has been uploaded to Collabra's website. We would like to resubmit the revised manuscript for publication.

We greatly appreciate your and the two reviewers' comments and feedback. We have thought carefully about them and done our best to respond to them thoroughly and directly. Please see the attached document for our responses to each of the points raised by yourself and/or the reviewers.

Editor Decision for Version 2

Editor: Simine Vazire
Affiliation: University of California, Davis, USA
Editor decision: Article accepted
Decision date: 25 June 2015

Dear Ms Erin Corwin Westgate,

After review, we have reached a decision regarding your submission to Collabra, "Implicit preferences for straight people over lesbian women and gay men weakened from 2006 to 2013", and are happy to accept your submission for publication, pending the completion of copyediting and formatting processes.

As there are no further reviewer revisions to make, you do not have to complete any tasks at this point. The accepted submission will now undergo final copyediting. You will be contacted once this is complete to answer any queries that may have arisen during copyediting and to allow a final chance to edit the files prior to typesetting. If you wish to view your submission during this time, you can log in via the journal website.