Transgender prejudice reduction and opinions on transgender rights: Results from a mediation analysis on experimental data

Andrew R. Flores¹,², Donald P. Haider-Markel³, Daniel C. Lewis⁴, Patrick R. Miller³, Barry L. Tadlock⁵ and Jami K. Taylor⁶

Abstract
Fears, phobias, and dislikes about minorities should be strong determinants of whether Americans support policies protecting such minorities. Studies suggest that discussions and information about transgender people can reduce transphobia. However, these studies also indicate that experimental treatments do not necessarily affect individual attitudes on policies concerning transgender rights. Scholars contend that durably reducing prejudice should increase public support for minority rights. In this study, we examine this causal mechanism utilizing an experiment. We find that reducing transphobia is a reliable mechanism to increase public support for transgender rights. These results are robust to causal identification assumptions, suggesting that this mechanism provides a clear avenue for stigmatized groups to increase public support of rights for those groups.

Keywords
LGBTQ, experiment, mediation analysis, transgender, public opinion, prejudice

Introduction
The emergence of the transgender rights movement has drawn political science and public opinion scholars to empirically assess public attitudes about transgender people and their rights. This growing literature in observational studies has provided many insights in regards to the impact of disgust and authoritarianism (Casey, 2016; Miller et al., 2017), interpersonal contact (Flores, 2015; Pierceson and Kirzinger, 2015; Tadlock et al., 2017), and body politics (Miller et al., 2017), on attitudes, as well as support for transgender candidates for public office (Haider-Markel et al., 2017). We have also learned how attitudes about transgender people and transgender rights differ from attitudes about gay people and gay rights (Lewis et al., 2017).

Experimental studies have found that canvassers can durably reduce transphobia (Broockman and Kalla, 2016), exposure to information and images of transgender people can also reduce transphobia (Flores et al., 2017), and that value frames and identity have the potential to affect support for transgender people’s access to public restrooms (Harrison and Michelson, 2017a; 2017b).

We contribute to this growing literature by testing a proposed causal mechanism between exposure to transgender people and increased support for transgender rights: prejudice reduction. Prejudice reduction, as used here, is when negative attitudes and phobias towards marginalized groups are lowered (Allport, 1979[1954]). We use a survey experiment to assess how prejudice reduction serves as a vehicle for individual support for transgender rights.

Quite possibly, the most seminal work in attitude change about marginalized groups is Allport’s (1979[1954])

¹Mills College, USA
²The Williams Institute at the UCLA School of Law, USA
³University of Kansas, USA
⁴Siena College, USA
⁵Ohio University, USA
⁶University of Toledo, USA

Corresponding author:
Andrew R. Flores, Mills College, 5000 MacArthur Blvd., Oakland, CA 94613, USA.
Email: andflores@mills.edu
theory of interpersonal contact. Subsequent studies, especially studies about the rights of sexual minorities and more recently gender minorities, investigate the way personally knowing someone who is lesbian, gay, bisexual, and/or transgender (LGBT) is associated with attitudes on LGBT rights. The presumed, though often untested, mechanism linking contact to greater policy support is prejudice reduction. However, is this assumption correct? Does prejudice reduction serve as a vehicle to greater support for rights for marginalized groups? Our experiment tests this causal mechanism by exposing individuals to short informational vignettes and facial images of people presumed to be transgender. Our post-test then measures transphobia and support for transgender rights policies. We find that prejudice reduction is a mechanism for support for transgender rights. We further show that our results are robust to key assumptions related to the causal mechanisms of attitude change.

**Mere exposure and prejudice reduction: A mechanism to increased policy support?**

Although our treatments are not the same as the contact hypothesis, studies have shown that simple exposure to out-groups can reduce prejudices (Zajonc, 1968; 2001; Zebrowitz, White, and Wieneke, 2008). Previous studies, for example, have shown that exposure to the lives of lesbians and gay men via television increases both positive attitudes toward gay men (Schiappa, Gregg, and Hewes, 2006) and support for lesbian and gay rights (Garretson, 2015). The mere exposure hypothesis contends that simply exposing people to new concepts has the tendency to increase likeability toward those objects. Indeed, an experimental study shows that mere exposure to facial images and information about transgender people can reduce discomfort with transgender people and transphobia (Flores et al., 2017). However, both Flores et al. (2017) and Broockman and Kalla (2016) failed to find any direct effect of treatments affecting individual attitudes on transgender rights. In follow-up surveys and a subsequent survey experiment, Broockman and Kalla (2016) later defined the term transgender to both treated and untreated households and randomized exposure to negative transgender campaign advertisements to connect their canvassing experiment to individual support for transgender rights. We add to these studies by more fully examining this mechanism.

A puzzle in both previous experiments is that respondents’ prejudices toward transgender people were reduced, yet there is little evidence of a direct effect on attitudes on transgender rights. Possibly, there are multiple mechanisms underlying the direct effect of mere exposure or canvassing on transgender rights. Some of these effects may be competing. For example, people with higher disgust levels tend to be opposed to gay rights (Adams, Stewart, and Blanchar, 2014), and LGBT rights may also increase people’s disgust (Casey, 2016). If the experimental treatments simultaneously increased disgust levels and reduced transphobia, then the direct effect may be washed out by these competing mechanisms. Thus, our study decomposes this direct effect using a mediation model. The mediation model separates the specified causal mechanism—in the current case, transphobia—from the remaining, unobserved mechanisms linking treatment to support for transgender rights.

**Data and analysis**

We fielded a survey experiment that was jointly sponsored by some of the institutions affiliated with the authors. Clear Voice Research (CVR) conducted the survey experiment 12–15 June, 2015. CVR maintains a panel of respondents who serve as a subject pool for research purposes. This pool was pre-screened with a member profile survey, and CVR utilized numerous validation methods, such as Internet Protocol (IP) and street address verification to minimize duplicate respondents and ensure the overall quality of the data. The sampling methodology is an enhancement over student and online convenience samples such as Amazon Mechanical Turk. We note that this sample, while an improvement, was not selected using traditional probability-based methods. CVR sent 51,492 recruitment emails to empaneled survey participants. A total of 2102 subjects participated in the survey, which provided a response rate of 10.82% (AAPOR RR#2). All subjects consented to participate in the study.

**Research design**

**Treatment: Informational vignettes and facial images.** Respondents were randomly assigned to one of four conditions (three treatment groups and one control group). The three treatment groups each received an information vignette that defined gender identity and transgender as follows.

Gender identity refers to how a person identifies their own gender (as a man, woman, or some other label). For many people their gender identity may not match their birth sex. For example, a man may identify more as female, or a woman may identify more as male. Transgender is a general term for people whose gender identity or expression is different from their birth sex. Some transgender people may undergo gender transition with medical procedures like hormone therapy or surgery (often commonly called a ‘sex change’), but others do not seek such medical assistance. Transgender includes groups you might have heard before, including transsexual, cross-dressers, or gender queer people.

The control group received a vignette about Japanese economic growth. Subjects in two of the three treatment groups also received images of gender-congruent or gender-incongruent male and female faces (see Figure 1). Respondents in the Congruent condition received both facial images of Figure 1(a), and respondents in the Incongruent condition
Flores et al.

received both facial images of Figure 1(b). The other treatment group (No Image) received no images but got the gender identity-defining vignette. Flores et al. (2017) previously showed these exposures have the effect of reducing transphobia and discomfort toward transgender people, but did not find any significant direct treatment effects on support for transgender rights. In this study, we expected the treatments to cause positive shifts in attitudes consistent with the mere exposure hypothesis. Our goal in this analysis was to show how exposure might cause support for transgender rights by way of reducing transphobia. We expected that transphobia would mediate the relationship between exposure and attitudes about transgender rights.

Mediator variable: Genderism and transphobia. In the survey post-test, we measured transphobia with questions from the Genderism and Transphobia Scale (GTS) designed to gauge people’s general tolerance for gender non-conformity (Hill and Willoughby, 2005). The entire GTS contained 32 questions to measure two or three correlated constructs: genderism, transphobia, and gender-bashing. We used a subset of five items from the GTS to measure transphobia—an emotional dislike against people who violate gender norms. Each of the questions utilized a seven-point strongly agree to strongly disagree scale. The five items scaled onto a single construct ($\alpha = 0.91$). We rescaled the items such that positive values indicate higher levels of tolerance for gender non-conformity. Full question wordings are provided in Supporting Information (SI) 2. The mediator variable was measured utilizing structural equation models (i.e., factor analysis) with its mean set to zero and variance set to one for identification purposes.

Dependent variables: Equality and accommodations policies

We also included a set of questions about transgender rights after the measurement of transphobia; these questions were first asked in a 2011 Public Religion Research Institute survey on transgender rights (see Flores 2015, Table 1). We also adapted questions relating to gay rights from sources like the Gender Social Survey or American National Election Studies, and we developed questions that were unique to transgender rights (e.g., public accommodations and access to public restrooms). These indicators were measured on a four-point scale, ranging from completely agree to completely disagree that had no neutral category. We created scales of these policy questions by first assessing whether transgender rights should be treated as unidimensional or multidimensional (see also Miller et al., 2017). These policies were more easily grouped into two dimensions: policies relating to equality and policies uniquely relating to accommodations (see Flores et al, 2017; Appendix D). Our analysis of transgender rights used these subsets, and positive values related to greater support for transgender rights.

Equality policies were defined as policies that ensure the equal treatment of transgender people. This included whether transgender people should be protected from job discrimination, allowed to serve openly in the military, and included in equal rights generally as would be the case for lesbians, gay men, and other citizens. Accommodations policies were defined as policies that would offer additional protections for transgender people. These policies involved both public and private accommodations, including the use of public restrooms, medical treatment for transgender health issues, and businesses’ right to use religious reasons to refuse services to transgender people.

Methods of analysis

Propensity score weighting. We initially conducted a full assessment of balance checks based on responses to the pre-test questionnaire and demographics. Our experiment
was a part of a larger survey, and 267 respondents failed to complete the questionnaire prior to the treatment administration. The balance checks indicated some significant differences across our treatment groups in responses to the pre-test, so we created covariate balance propensity scores (CBPS) to reduce these differences (Fong, Ratkovic, and Imai, 2014; Imai and Ratkovic, 2014).

Covariate adjustment. We also used covariate adjustment to increase the efficiency of estimated effects (Gerber and Green, 2012; Ho et al., 2007). We adjusted for pre-test levels of moral traditionalism by combining four questions into a single scale ($\alpha = 0.75$); the question wordings came from the American National Election Studies. We included partisanship by using a traditional seven-point indicator ranging from strongly Democratic to strongly Republican. We additionally controlled for age, race, sex, and whether someone identifies as LGBT.5

Structural equation modeling. Structural equation modeling (SEM) allowed us to simultaneously estimate a measurement model for our dependent variables, examine the structural parameters of the treatment effects, and decompose effects for causal mechanisms. SEM used the responses to sets of indicators to form constructs that correct for potential measurement error in any single item. The resultant continuous constructs were then used as dependent variables in regression models. Since the items forming the constructs were ordinal, we used a weighted least square with mean adjusted variance (WLSMV) estimator, which is the best multivariate estimator with ordinal data (Brown, 2006). We standardized all of the constructs to have a mean of zero and variance of one. The linear structural equation model might carry with it untenable assumptions in identifying causal mechanisms (Imai et al., 2011). As such, we also presented results from sensitivity analyses on the key causal assumption of exogeneity of the mediator and outcome variables given that both are in the post-test also referred to as sequential ignorability.

The model schematic is plotted in Figure 2, which shows the role of the mediator on transgender rights. Our goal was to examine the role of causally reducing transphobia as an underlying mechanism of increasing support for transgender rights. Given previous research findings, we expected the treatments to reduce transphobia, and we anticipated exposure effects to induce greater support on both transgender rights scales. We conducted two mediation models: the first with the dependent variable being equality policies and the second with the dependent variable being accommodations policies. We estimated the mediation models relying on the Imai et al. (2011) framework as implemented by Muthén (2011). Thus, the indirect effects were also the Average Causal Mediation Effects.

Table 1. Mediation model results.

<table>
<thead>
<tr>
<th></th>
<th>Equality policies</th>
<th>Accommodations policies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Image treatment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct</td>
<td>−0.15 (0.067)*</td>
<td>−0.07 (0.077)</td>
</tr>
<tr>
<td>Indirect: Transphobia</td>
<td>0.11 (0.042)**</td>
<td>0.13 (0.048)**</td>
</tr>
<tr>
<td>Percent Mediated</td>
<td>42.2% (13.8)**</td>
<td>66.1% (26.0)**</td>
</tr>
<tr>
<td><strong>Congruent Image treatment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct</td>
<td>−0.04 (0.067)</td>
<td>−0.10 (0.077)</td>
</tr>
<tr>
<td>Indirect: Transphobia</td>
<td>0.08 (0.041)*</td>
<td>0.09 (0.048)*</td>
</tr>
<tr>
<td>Percent Mediated</td>
<td>68.0% (41.5)*</td>
<td>46.0% (21.5)*</td>
</tr>
<tr>
<td><strong>Incongruent Image treatment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct</td>
<td>−0.10 (0.071)</td>
<td>−0.08 (0.080)</td>
</tr>
<tr>
<td>Indirect: Transphobia</td>
<td>0.11 (0.041)**</td>
<td>0.13 (0.048)**</td>
</tr>
<tr>
<td>Percent mediated</td>
<td>51.5% (18.6)**</td>
<td>60.5% (23.9)**</td>
</tr>
<tr>
<td>N</td>
<td>1931</td>
<td>1931</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.38</td>
<td>0.45</td>
</tr>
<tr>
<td>CFI</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td>TLI</td>
<td>0.95</td>
<td>0.98</td>
</tr>
<tr>
<td>RMSEA [90 % CI]</td>
<td>0.060 [0.057, 0.063]</td>
<td>0.038 [0.034, 0.041]</td>
</tr>
</tbody>
</table>

CFI = Comparative Fit Index; CI = confidence interval; RMSEA = Root Mean Square Error of Approximation; TLI = Tucker-Lewis Index.

* $p<0.05$; ** $p<0.01$ (one-tailed).

Note: Standard errors are in parentheses.

Figure 2. Structural equation model schematic.
Results

Both models converged and showed adequate model fit with both the Comparative Fit Index and Tucker-Lewis Index above 0.95. Consistent with expectations, transphobia had an effect on people’s attitudes on transgender rights. A one standard deviation reduction in transphobia corresponded with a 0.58 standard deviation increase in support of equality policies (standard error [s.e.] = 0.02, *p* < 0.01) and a 0.68 standard deviation increase in support of accommodations policies (s.e. = 0.02, *p* < 0.01). Also consistent with expectations, the treatments had reductions in transphobia relative to the control. The No Image treatment reduced transphobia by 0.19 standard deviations (s.e. = 0.07, *p* < 0.01); the Congruent treatment reduced transphobia by 0.13 standard deviations (s.e. = 0.13, *p* < 0.05), and the Incongruent treatment reduced transphobia by 0.18 standard deviations (s.e. = 0.07, *p* < 0.01).6

These results are consistent with our earlier finding that the effects of the treatment have little direct relationship to views about transgender rights (Flores et al., 2017). Yet the mediation analysis presented in Table 1 suggests this null relationship may be driven by the mediated relationship between the treatments and transgender rights. The treatments in the present analysis showed statistically insignificant overall effects on transgender rights, but statistically significant effects on reducing transphobia. The results showed that there is a negative, but often statistically insignificant direct effect of exposure to transgender information and images on support for transgender rights. Importantly, we found statistically significant indirect effects on transgender rights by reducing transphobia. Exposure tended to increase support on both transgender rights constructs about one-tenth of a standard deviation, which is a relatively small effect size. This pattern was consistent in both equality policies and accommodations policies, and the exposure effect sizes are similar across treatment groups. The effects of the congruent image were slightly smaller than that of the other treatments, but they did not differ in any statistically significant way. As documented in SI 5, we found that these effects were similar for Democrats, larger for independents, and were null for Republicans.

When assessing the effect size relative to the total effect of treatment on our dependent variables, Alwin and Hauser (1975) recommended reporting the percent mediated out of the sum of the absolute value of these effects in models where the direct and indirect effects are of opposite signs. We find that the indirect effect ranges from 42.2% to 68.0% of the total absolute effect of treatment on transgender rights. This suggests that a substantial portion of the effect of the treatment is mediated by transphobia.

Since transphobia was measured in the post-test along with transgender rights, there may be unmeasured confounders in both levels of transphobia and support for transgender rights. We only observed an individual’s responses on transgender rights and transphobia after receiving treatment, so the potential outcomes of individual support for transgender rights in a treatment group given that person’s level of transphobia had they been in the control group remained counterfactual. A key identification assumption for causal mechanisms was that there is no residual correlation between the mediator and outcome variables or sequential ignorability (Imai et al., 2011). This assumption, however, can be evaluated by varying how much of a residual correlation would have to be present to alter or reverse the results. A robust relationship will not be sensitive to causal identification assumptions. We find that our findings are relatively insensitive to the underlying assumptions to estimating causal mechanisms (see SI 4).

Discussion

Public attitudes on transgender rights have become a relatively recent area of inquiry and these attitudes are consequential to whether states have transgender-inclusive policies (e.g., Flores, Herman, and Mallory, 2015). Yet, experimental studies fail to show their treatments have any direct effect on transgender rights. Following Broockman and Kalla (2016), we suggest that this is because experimental treatments affect attitudes on transgender rights indirectly.

Our findings offer some guidance for those seeking to influence public attitudes on transgender rights. First, reducing transphobia is a key mechanism for garnering support for transgender rights. Second, a way to reduce transphobia is to humanize transgender people by exposing individuals to information about them and representations of them. Finally and consistent with our earlier findings (Flores et al., 2017), attitudes about transgender people are likely not dependent on perceived gender conformity. Exposure to both our perceived congruent and incongruent images had similar effects on reducing transphobia. We do note that this does not mean that in highly sensitive circumstances, such as when transgender people are in public restrooms, that the potential for harassment, discrimination or assault is not dependent on perceived conformity (Miller and Grollman, 2015). Such contexts may be unique circumstances. In addition, because we show both images of gender-conforming males and females to our respondents rather than putting them into separate treatments by sex, treatment effects might actually differ for gender-conforming/non-conforming transmen and transwomen. Anecdotally, that is very plausible.

Fully unpacking the causal mechanism from exposure to policy support has eluded previous research. We provide some evidence that prejudice reduction is indeed a vehicle for enhancing support for marginalized groups. Further research should undertake more robust experimental designs that would clearly account for sequential ignorability and
post-treatment bias. Our sensitivity analyses suggest our results are robust to the former assumption but do little to address the latter. Alternative research designs as suggested by Imai et al. (2011) would facilitate in clearly specifying the ways in which prejudice reduction enhances public support for the rights of minority groups.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the University of Toledo, University of Kansas, the Williams Institute at the UCLA School of Law, and Ohio University.

ORCID iD
Donald P. Haider-Markel https://orcid.org/0000-0002-2095-1618

Supplementary materials

Notes
1. We pre-tested the survey experiment with a sample of 247 people recruited through Amazon Mechanical Turk. While many of the results are similar to what we report here, the MTurk sample tended to exaggerate treatment effects.
2. The facial images selected for our treatment groups came from a social psychological study (Gerhardstein and Anderson, 2010). We received permission from the authors to use these images in the study.
3. The factors are two or three because Hill and Willoughby (2005) initially found that genderism and transphobia were separate constructs but a follow-up study found they were a single construct.
4. We have alternatively conceptualized these measures as relating to transgender bodies (Miller et al., 2017).
5. The age cohorts are 30−44, 45−64, 65−older with 18−29 as the reference group. The race categories are black, Hispanic, Asian, Native American, and multiracial with white as the reference group.
6. The effect of the Incongruent treatment reduced transphobia by 0.19 standard deviations in the Accommodations policies model (s.e. = 0.07, p<0.01). All other treatment effects on transphobia are the same in both models.

Carnegie Corporation of New York Grant
This publication was made possible (in part) by a grant from Carnegie Corporation of New York. The statements made and views expressed are solely the responsibility of the author.

References
Imai K, Keele L, Tingley D, et al. (2011) Unpacking the black box of causality: Learning about causal mechanisms from


Muthén BO (2011) *Applications of Causally Defined Direct and Indirect Effects in Mediation Analysis using SEM in Mplus*. Los Angeles, CA, USA: Muthén and Muthén.


