Supplement for Equalitarianism: A source of liberal bias

Winegard, Clark, Hasty, & Baumeister, 2023
Journal of Open Inquiry in Behavioral Science

Table of Contents

Discussion of the Difficulties of Measuring Bias and the Explanation for the Order Effects Approach

Most bias studies, including ours, rely upon the principle of invariance: Decision irrelevant information (extraneous information) should not affect judgments; therefore, the degree of a person's bias is reflected by the degree to which the extraneous information affects his or her judgments (Ditto et al. 2019a; Kahneman & Tversky, 1984). In psychology, the standard methods for testing bias involve matching as much information as possible, changing only the conclusions of a vignette or other supposedly extraneous information (such as the race or sex of an actor), and then having participants evaluate the *matched* information rather than the manipulated information. If participants evaluate identical information differently depending on the extraneous information, this is considered bias. For example, in a famous study on bias, Lord et al., (1979) gave participants identical methodology descriptions of studies testing the deterrent efficacy of the death penalty. The only information that varied in the conditions was the conclusion: the study found that the death penalty did or did not deter crime. Then they had the participants rate the quality of the studies' methods (which again, were identical). Participants who supported the death penalty rated the methods as worse when the conclusion contradicted their prior attitude (death penalty deters crime) than when it buttressed it.

However, it is not clear that this paradigm allows a researcher to isolate bias unambiguously. In methods matching studies, for example, it might be rational to assess methods differently depending upon the outcomes of those methods. Imagine, for example, a description of methods that appeared sound but generated results that showed that eating purple muffins allows people to see the future. People cannot see the future (but see Bem, 2011); therefore, one should be very skeptical of the results; and if one is skeptical of the results, then one should

probably be skeptical of the methods that led to them. We call this "the proof of the recipe is in the eating" or PRE principle. A recipe might look good or bad on paper, but its final value depends upon the food it produces. If one follows the recipe and gets bad food, it is not irrational to update one's assessment of the recipe. More broadly, the results of a process (methods, recipe, blueprint) provide information about the soundness of the process, and a good Bayesian should update his or her priors about the process after getting the results (see Kahan, 2016, for a discussion of Bayesian reasoning and bias).

The same criticism applies to matched vignettes that change the demographic characteristics of described individuals. Suppose, for example, that a researcher believes that liberals are biased *against* White people. She designs a study that includes a vignette describing a cop shooting a person who was found to be holding a piece of silverware (not a weapon). The vignette is altered such that in one condition the cop's victim is White and in the other he is Black. She then finds that liberals rate the cop as less wrong when the victim was White than when he is Black and contends that this is due to liberal bias against White people. One might object, however, that the demographic characteristics Black and White provide information. Perhaps one believes, for example, that Black people are unfairly targeted by police officers more often than Whites. One might believe, then, that in the White condition the spoon must have looked quite menacing because otherwise the cop would not have shot; whereas, in the Black condition, one might just think "yeah, cops wrongly shoot Black people all the time… this is very wrong."

There are a couple of ways to mitigate the force of the Bayesian (normative rationality) objection. First, one can choose examples in which base rates go in the opposite direction from the predicted bias. Suppose, for example, that a researcher thinks that Conservatives are biased

against women. He could use a vignette in which either a man or a woman sexually propositions a subordinate in a crass way and ask participants if the (identical) behavior is sexual harassment. In this way, the demographic information is going, if anything, against the direction of the hypothesis because most people believe that men are more likely than women to sexually harass others. Thus, if Conservatives rate the identical behavior as harassment only when performed by a woman (and not when performed by a man), it would be reasonably compelling evidence that Conservatives are biased against women in this domain.

Second, one could observe order effects in a within-subjects design (for example, see Uhlmann et al., 2009, which, incidentally, found that Liberals were more willing to sacrifice a White man to save 100 others than to sacrifice a Black man to save 100 others, whereas race had no influence on Conservatives' willingness to sacrifice one life to save 100, somewhat consistent with our predictions here). That is, one could give both vignettes to participants and manipulate the order of presentation. If participants believe that their answers in the two conditions *should be* the same and therefore anchor their second response to their first, that suggests that people at least believe it is irrational (and biased) to answer them differently. If an order effect is observed such that both vignettes are evaluated more favorably when the preference consistent one is presented first than when the preference inconsistent one is presented first, this would indicate that participants are biased despite their apparent belief that it is irrational to treat the two conditions differently. In our experiments, we tried to use both principles to counter possible objections. We still believe that objections are possible; ultimately, it might not be possible to demonstrate bias in an experiment without putting participants through timely experiments that allow researchers to rule out Bayesian updating explanations. Nevertheless, current methods

allow us to glean valuable information about potential bias, which, when combined with theory, *should* cause us to update our priors about sources of bias.

Equalitarianism Measure

Instructions: Please answer the following questions as honestly as you can. Remember, all answers will be confidential. Use the following scale 1- do not agree all, 4-somewhat agree, 7-agree completely (so 1 is the lowest level of agreement, and 7 is the highest.)

1. The only reason there are differences between men and women is because society is sexist
2. Differences between men and women in society are caused by discrimination
3. Differences between ethnic groups in society are caused by discrimination
4. Most people are not biased and racism is not a problem anymore*
5. When people assert that men and women are different because of biology, they are usually trying to justify the status quo
6. People often try to conceal their racism and sexism, but they act that way anyways
7. People often use biology to justify unjust policies that create inequalities
8. Racism is everywhere, even though people say they are not racist
9. Sexism is everywhere, even though people say that they are not sexist
10. People use scientific theories to justify inequalities between groups
11. Men and women have equal abilities on all tasks (for example, mathematics, cooking, nursing).
12. All ethnic groups have equal abilities on all tasks (for example, mathematics, sports, creativity)
13. Some differences between men and women are hardwired*
14. Although things are unequal now, if we work really hard, we can make society better and more fair
15. We should strive to make all groups equal in society
16. We should strive to make men and women equally represented in science fields
17. If we work hard enough, we can ensure that all ethnic groups have equal outcomes
18. Differences among ethnic groups in social outcome are at least partially biologically caused*

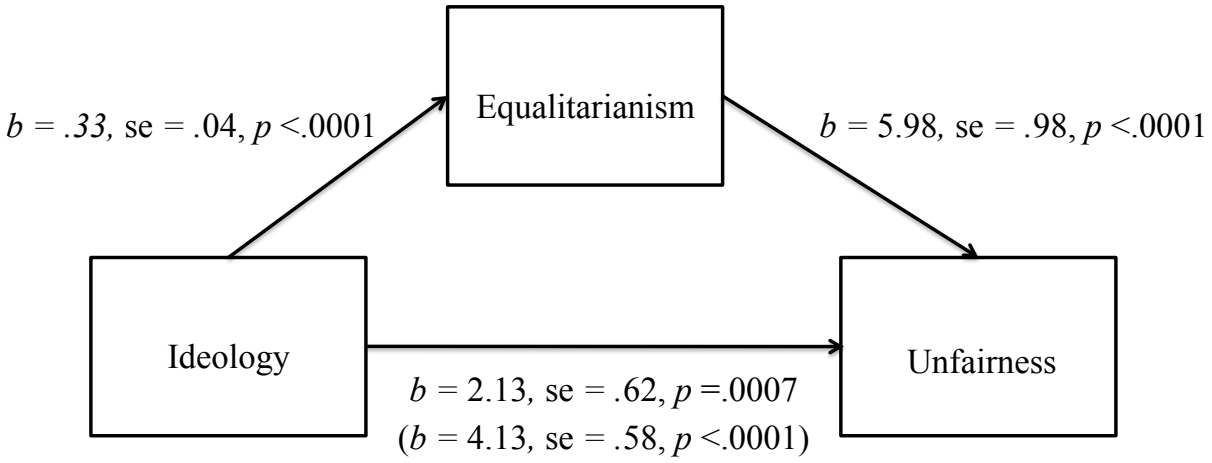
*reverse coded items

Principal Components Analysis

Combined across all studies, a principal components analysis supported a one factor solution. Four components had Eigenvalues above 1, however there was a dramatic drop off from the first component (Eigenvalue = 7.72) to the second (Eigenvalue = 1.66), and all items correlated with the first component the strongest (and all above .48). The only exceptions were the three reverse scored items, which correlated with multiple components to similar degrees and seemed to explain the presence of the other components, a common problem with reverse-scored items (see e.g., Conrad et al., 2004). The three reverse-scored items generally correlated the weakest with all other items in the measure, however, they were still positively and significantly associated with all other items.

Study 1a Mediation Analysis

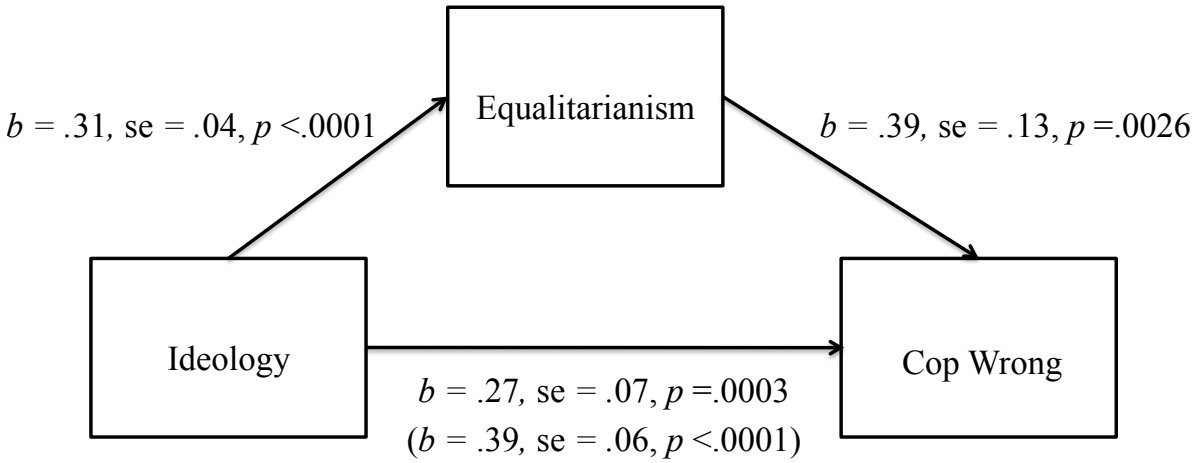
More liberal ideology predicted rating victims' groups as treated more unfairly (and privileged groups as treated more fairly), and this was partially mediated by their higher equalitarian attitudes.



Influence of ideology (higher values = more liberal) on unfairness (higher values = victims' groups treated more unfairly/privileged groups treated more fairly), mediated by equalitarianism (higher values = more equalitarian). In this and all subsequent studies, values in parenthesis are the total effect of the IV on the DV (i.e., prior to controlling for equalitarianism).

Study 1b Mediation Analysis

More liberal ideology predicted stronger beliefs that the cop was wrong to shoot the Black man, and this was partially accounted for by Liberals' stronger equalitarian attitudes.



Influence of ideology (higher values = more liberal) on beliefs that the cop was wrong, mediated by equalitarianism.

Study 2 Categorical Results

We created a categorical ideology variable for Conservatives (those who responded 1-3 on the 7-point ideology scale; $n = 62$), Moderates (those who responded 4; $n = 57$), and Liberals (those who responded 5-7; $n = 86$).[1] We analyzed the 2 (Sex condition) x 3 (categorical ideology) interaction on acceptability ratings in a Univariate Analysis of Variance (ANOVA).

There was a significant main effect of Sex condition, indicating that all participants objected more to a test favoring men than a test favoring women. The main effect of ideology and the interaction were not significant. However, consistent with the continuous results, simple contrasts revealed the largest (and a significant) difference between experimental conditions for Liberals. Specifically, Liberals rated the test as significantly more acceptable if women outperform men than if men outperform women, $p = .004$, *Cohen's d* = .64, whereas Moderates, $p = .507$, *Cohen's d* = .17, and Conservatives, $p = .118$, *Cohen's d* = .44, demonstrated no such difference (though note Conservatives were trending in a similar direction as Liberals).

In the Women Outperform condition, no groups significantly differed, ps > .557. In the Men Outperform condition, Liberals rated the test as marginally less acceptable than Moderates, $p = .099$, and somewhat (though not significantly) less acceptable than Conservatives, $p = .152$; whereas there were virtually no differences between Conservatives and Moderates in this condition, $p = .791$.

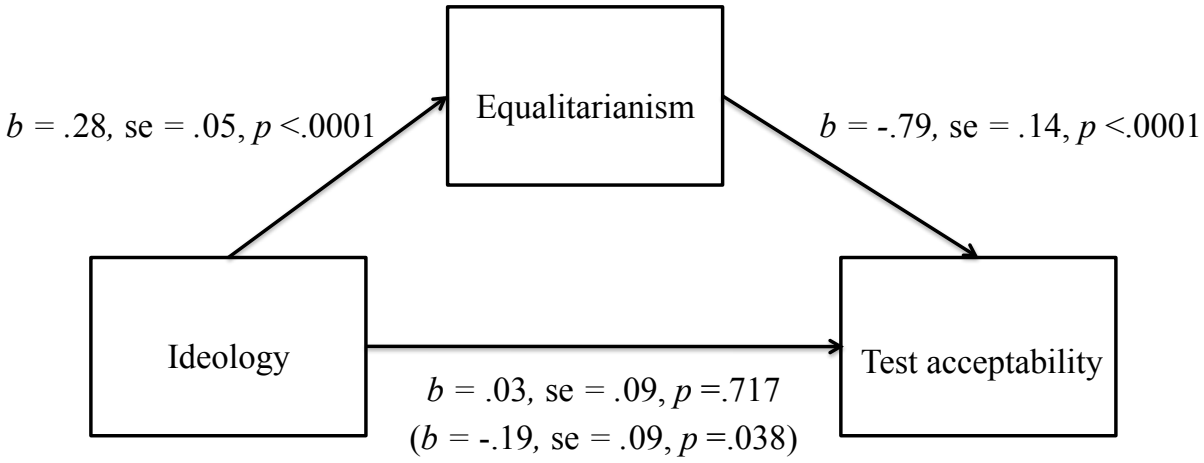
Categorical ideology, Sex condition (0: Men Outperform; 1: Women Outperform), and the interaction on test acceptability

	F	*p*	η_p^2
Condition	8.11	.005	.039
Ideology	0.78	.460	.008
Condition x Ideology	0.94	.393	.009

[1]This coding scheme is how we created a categorical ideology variable in this study and all subsequent studies.

Study 2 Mediation Results

Higher equalitarianism fully accounted for the relationship between more liberal ideology and lower ratings of test acceptability in the condition in which men outperform women on the test.



Influence of ideology (higher values = more liberal) on test acceptability, mediated by equalitarianism in the condition in which men outperform women.

Study 3 Categorical Results

We again created an ideological category variable (Conservatives n = 56, Moderates n = 37, Liberals n = 109) and analyzed the 2 (Race condition) x 3 (categorical ideology) interaction on credibility ratings in an ANOVA. There was no main effect of Race condition, nor ideology, but similar to the continuous results, there was a marginal interaction.

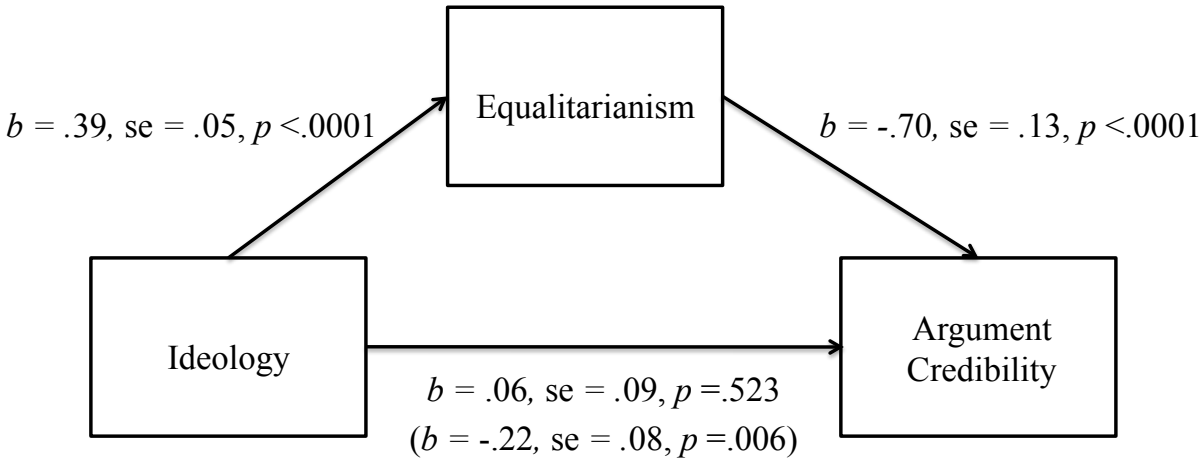
All simple contrasts demonstrated the expected pattern of results. Specifically, Liberals rated the argument as significantly more credible in the Blacks Higher condition than the Whites Higher condition, p = .005, *Cohen's d* = .58. This difference was smaller and not significant for Moderates, p = .400, *Cohen's d* = .58, and slightly (but not significantly) in the opposite direction for Conservatives, p = .498, *Cohen's d* = .20. In the Blacks Higher condition, no groups significantly differed, ps > .134. In the Whites Higher condition, Liberals rated the argument as significantly less credible than Moderates, p = .040, and Conservatives, p = .030; Moderates and Conservatives did not differ, p = .856.

Categorical ideology, Race condition (0: Whites Higher; 1: Blacks Higher), and the interaction on argument credibility

	F	p	η_p^2
Condition	1.90	.170	.010
Ideology	1.67	.192	.017
Condition x Ideology	2.44	.090	.024

Study 3 Mediation Results

Higher equalitarianism fully accounted for the relationship between more liberal ideology and lower ratings of argument credibility in the Whites Higher condition.



Influence of ideology (higher values = more liberal) on argument credibility, mediated by equalitarianism in the Whites Higher condition.

Study 4 Categorical Results

We again created an ideological category variable for Conservatives (n = 125), Moderates (n = 103), and Liberals (n = 224), and analyzed the 3 (Race condition: Whites Higher, Blacks Higher, Equal) x 3 (categorical ideology) interaction on credibility ratings in an ANOVA. There was a significant main effect of Race condition, a marginal main effect of ideology, and a significant interaction.

Categorical ideology, Race condition (0: Whites Higher; 1: Blacks Higher), and the interaction on argument credibility

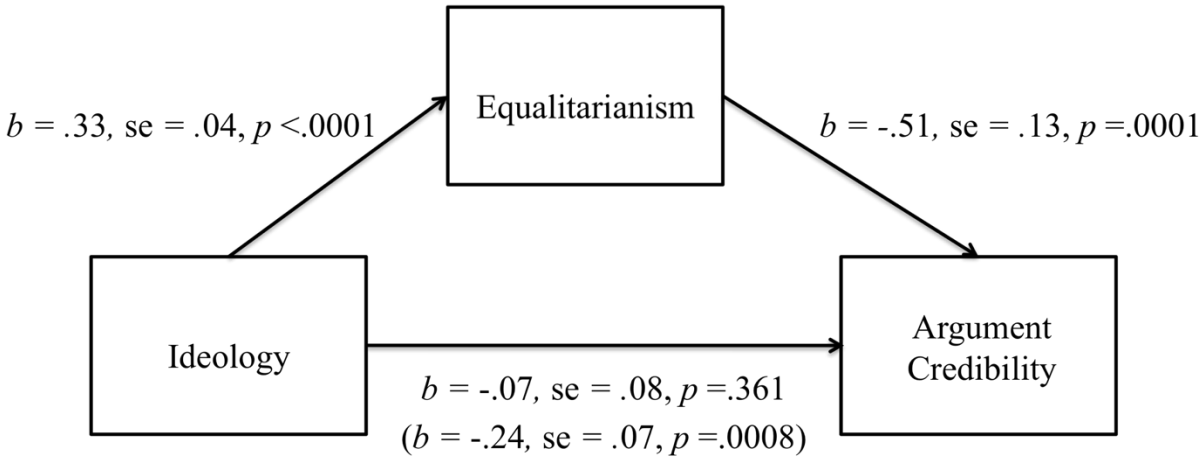
	F	p	η_p^2
Condition	9.06	<.001	.039
Ideology	2.49	.085	.011
Condition x Ideology	2.57	.037	.023

There were no differences between Liberals, Moderates, and Conservatives on argument credibility in the Equal or Blacks Higher conditions, ps > .344. In the Whites Higher condition, Conservatives and Moderates did not differ (p = .648), but Liberals rated the argument as less credible than both Conservatives (p = .003) and Moderates (p = .002). Among Conservatives and Moderates, only the Equal and Blacks Higher conditions significantly differed (ps = .010 and .025, respectively, *Cohen's ds* = .57-.58) such that the Equal condition was rated as more credible than the Blacks Higher condition. The Whites Higher condition fell between the other two conditions and did not significantly differ from either the Blacks Higher or the Equal condition for Conservatives or Moderates, ps > .117, *Cohen's ds* = .31-.33. Among Liberals, all conditions significantly differed. Liberals rated the argument as more credible in the Equal condition than the Blacks Higher condition, p = .038, *Cohen's d* = .36, and the Whites Higher condition, p <

.001, *Cohen's d* = .69, and more credible in the Blacks Higher condition than the Whites Higher condition, $p = .016$, *Cohen's d* = .39.

Study 4 Mediation Results

In the Whites Higher condition, higher equalitarianism fully mediated the influence of more liberal ideology on lower argument credibility ratings, 95% CI [-.28, -.08].



Influence of ideology (higher values = more liberal) on argument credibility, mediated by equalitarianism in the Whites Higher condition.

Study 5 Categorical Results

We again created an ideological category variable for Conservatives (n = 132), Moderates (n = 82), and Liberals (n = 239), and analyzed the 3 (Sex condition: Men Higher, Women Higher, Equal) x 3 (categorical ideology) interaction on credibility ratings in an ANOVA. There was a significant main effect of Sex condition, a marginal main effect of ideology, and again unexpectedly, no significant interaction.

Categorical ideology, Race condition (0: Men Higher; 1: Women Higher), and the interaction on argument credibility

	F	p	η_p^2
Condition	7.84	<.001	.034
Ideology	2.99	.051	.013
Condition x Ideology	0.30	.876	.003

Consistent with predictions and all studies thus far, there were no differences between Liberals, Moderates, and Conservatives on argument credibility in the Equal or Victims' Group (here, women) Higher conditions, ps > .107. In the Men Higher condition, there was a marginal main effect such that Liberals evaluated the argument as less credible than Conservatives, p = .076, weakly consistent with predictions (and consistent with all studies thus far and upcoming Study 7).

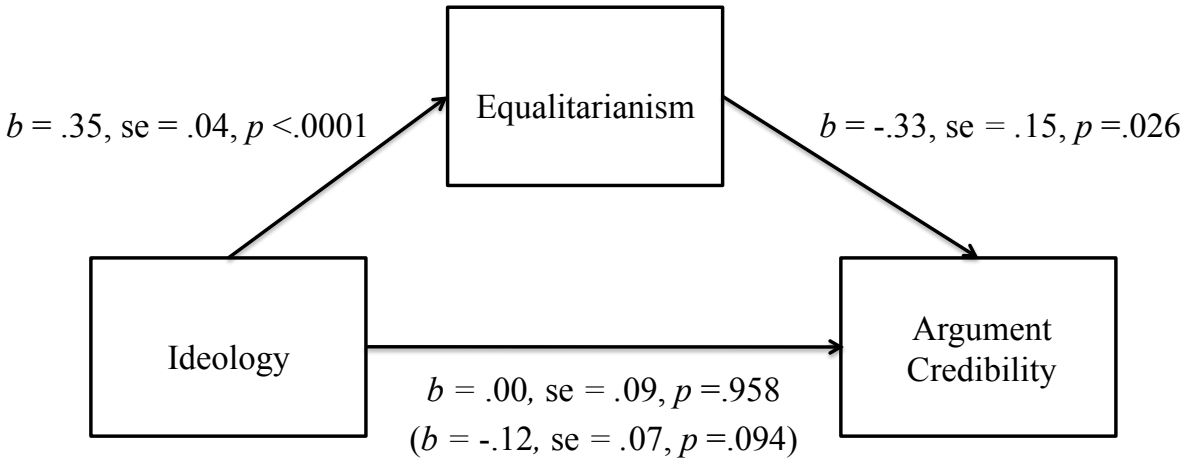
Liberals generally displayed the expected pattern of results: They rated the Equal condition as the most credible, followed by Women Higher, followed by the Men Higher. Liberals did not significantly differ between the Equal and Women Higher conditions, p = .310, *Cohen's d* = .17 (we did not have a strong prediction, here, but thought Liberals would rate the Equal condition as most credible). Consistent with predictions, Liberals rated the Men Higher

argument as significantly less credible than both the Women Higher, $p = .005$, *Cohen's d* = .41, and Equal arguments, $p < .001$, *Cohen's d* = .59.

Conservatives also did not differ between the Equal and Women Higher conditions, $p = .865$, *Cohen's d* = .04, but unexpectedly, rated the Men Higher argument as marginally less credible than the Equal argument, $p = .076$, *Cohen's d* = .39, and significantly less credible than the Women Higher argument, $p = .048$, *Cohen's d* = .45. Moderates did not significantly differ between any of the sex conditions, *p*s > .185, *Cohen's ds* .03- .36, but demonstrated the same basic pattern as Conservatives and Liberals (i.e., Equal and Women Higher roughly equivalent, and both higher than Men Higher).

Study 5 Mediation Results

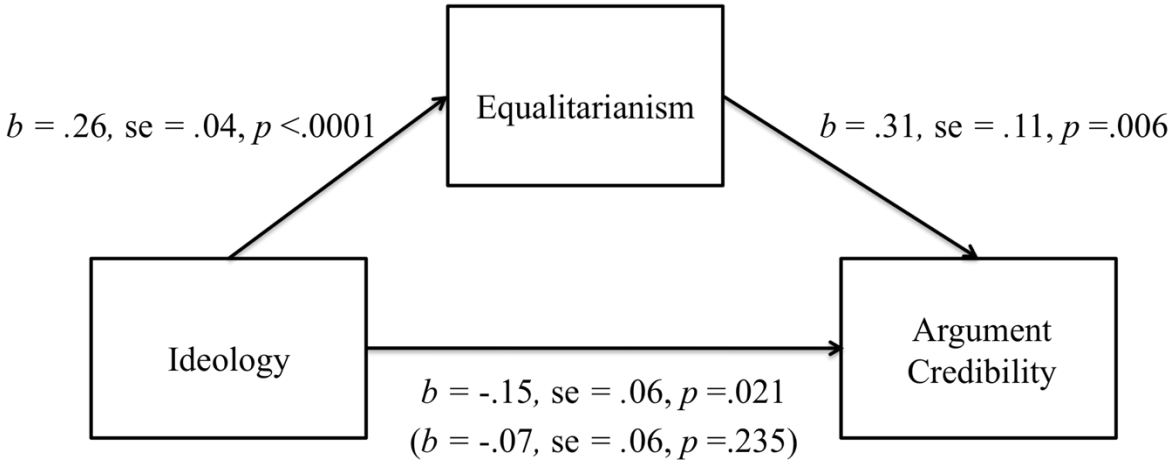
As expected, in the Men Higher condition, higher equalitarianism fully mediated the marginal influence of more liberal ideology on lower argument credibility ratings, 95% CI [-.23, -.02].



Influence of ideology (higher values = more liberal) on argument credibility, mediated by equalitarianism in the Whites Higher condition.

Unexpectedly (and unlike Study 4), in the Equal condition, equalitarianism mediated the influence of ideology on argument credibility ratings, 95% CI [.02, .16]; see Supplement. The meaning of this significant mediation was not immediately obvious to us. Consistent with previous results, higher liberalism predicted higher equalitarianism, higher equalitarianism predicted stronger agreement in the Equal condition (not particularly surprising), but more liberal ideology predicted *lower* credibility ratings in the Equal condition (significantly so only *after* controlling for equalitarianism). Across all reported studies, Liberals generally found the vignettes about genetic differences less credible than Conservatives. Perhaps this mediation pattern reflects this. Higher liberalism is related to lower credibility scores, *but* Liberals also score higher in Equalitarianism than Conservatives and so want groups to be equal. Therefore, when the Equalitarian score is put into the mediation analysis, and thus Liberals' desire for

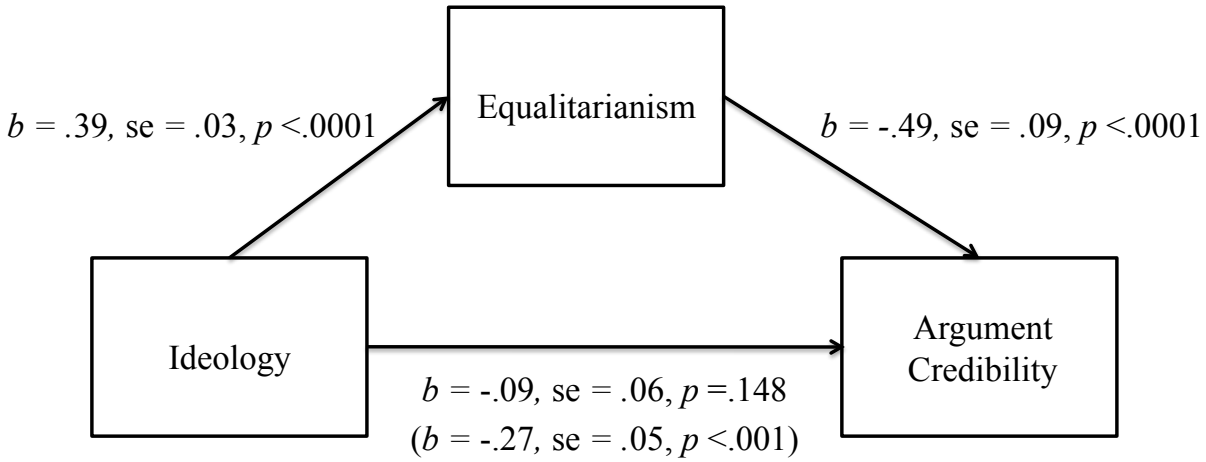
equality is accounted for, the negative relationship between Liberalism and argument credibility becomes significant.



Influence of ideology (higher values = more liberal) on argument credibility, mediated by equalitarianism in the Equal condition.

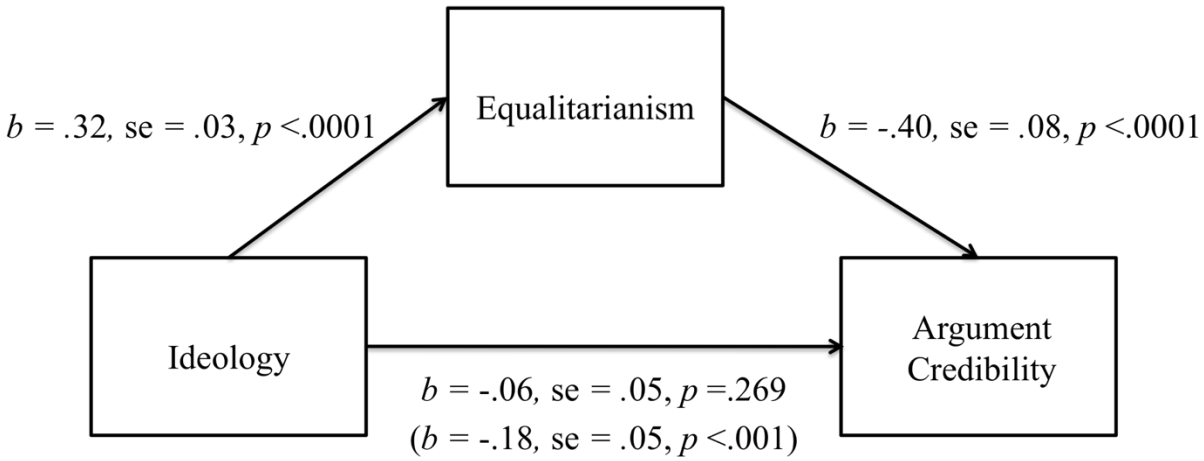
Study 6 Mediation Results

Higher equalitarianism accounted for the relationship between more liberal ideology and lower ratings of argument credibility when Whites were higher and that condition came first.



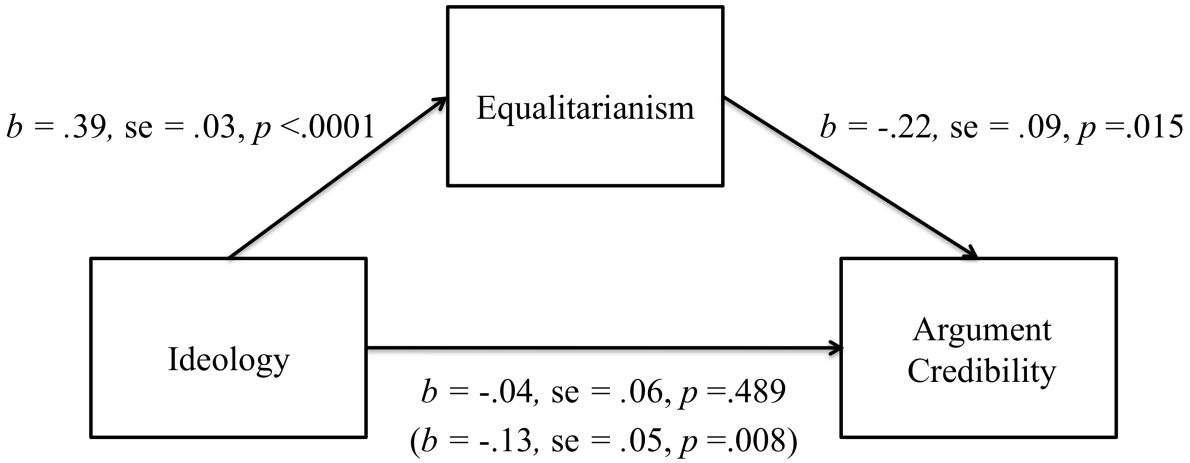
Influence of ideology (higher values = more liberal) on argument credibility, mediated by equalitarianism in the Whites Higher condition when this argument came first.

These relationships were somewhat smaller, but generally similar when they read the Whites Higher argument second, 95% CI [-.19, -.07], such that higher equalitarianism mediated the relationship between more liberalism and lower ratings of argument credibility that Whites are higher.



Influence of ideology (higher values = more liberal) on argument credibility, mediated by equalitarianism in the Whites Higher condition when this argument came second.

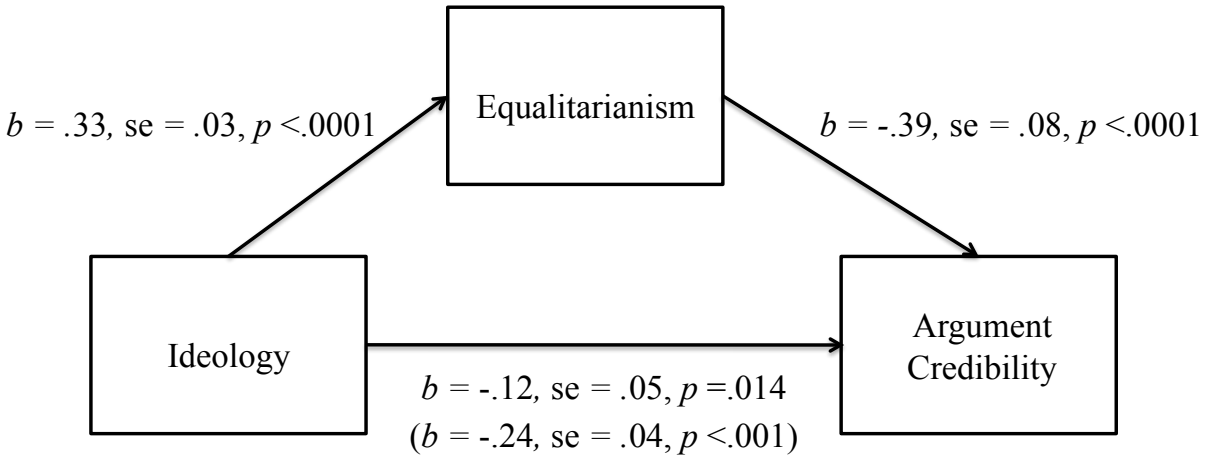
On the Blacks Higher outcome, equalitarianism did not mediate ideology on argument credibility when the argument came first 95% CI [-.12, .00], but did when the argument came second 95% CI [-.17, -.01]. Higher equalitarianism fully accounted for the relationship between more liberal ideology and *lower* ratings of argument credibility when Blacks were said to be higher and that argument came second.



Influence of ideology (higher values = more liberal) on argument credibility, mediated by equalitarianism in the Blacks Higher condition when this argument came second.

Study 7 Mediation Results

As predicted, equalitarianism mediated the influence of ideology on Men Higher argument credibility when they read the Men Higher argument first, 95% CI [-.19, -.07].



Influence of ideology (higher values = more liberal) on argument credibility, mediated by equalitarianism in the Men Higher condition when this argument came first.

References

Bem, D. (2011). Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*, 407-425.

Conrad, K. J., Wright, B. D., McKnight, P., McFall, M., Fontana, A., & Rosenheck, R. (2004). Comparing traditional and Rasch analyses of the Mississippi PTSD Scale: revealing limitations of reverse-scored items. *Journal of Applied Measurement*, *5*, 15-30.

Kahan, D. M. (2016). The politically motivated reasoning paradigm, part 1: What politically motivated reasoning is and how to measure it. *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*. 1–16.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*, 341-350.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37,* 2098-2109.

Uhlmann, E. L., Pizarro, D. A., Tannenbaum, D., & Ditto, P. H. (2009). The motivated use of moral principles. *Judgment and Decision making*, *4*, 476-491.