Conservatives and liberals have similar physiological responses to threats

Bakker, B.N.; Schumacher, G.; Gothreau, C.; Arceneaux, K.

DOI
10.1038/s41562-020-0823-z

Publication date
2020

Document Version
Accepted author manuscript

Published in
Nature Human Behaviour

Conservatives and Liberals have Similar Physiological Responses to Threats

Bert N. Bakker,[1*] Gijs Schumacher,[2] Claire Gothreau,[3] Kevin Arceneaux[4]

[1]Amsterdam School of Communication Research, University of Amsterdam,
Nieuwe Achtergracht 166, 1001 NG, Amsterdam, the Netherlands,
Email: b.n.bakker@uva.nl; ORCID: 000-0002-6491-5045
[2]Department of Political Science, University of Amsterdam,
ORCID: 0000-0002-6503-4514.
Nieuwe Achtergracht 166, 1001 NG, Amsterdam, the Netherlands,
[3]Center for American Women and Politics, Rutgers University,
191 Ryders Lane, New Brunswick, New Jersey, US
ORCID: 0000-0002-9242-8121
[4]Department of Political Science, Temple University,
1115 Polett Walk, PA 19122, Philadelphia, USA,
ORCID: 0000-0002-2884-5238.

*Corresponding author: Bert N. Bakker (b.n.bakker@uva.nl; ORCID: 000-0002-6491-5045)

Paper accepted for publication in Nature Human Behaviour

Abstract

About a decade ago a study documented that conservatives have stronger physiological responses to threatening stimuli than liberals. This work launched a paradigm aimed at uncovering the biological roots of ideology. Despite wide-ranging scientific and popular impact, independent laboratories have not replicated the study. We conducted a preregistered direct replication (N=202) and conceptual replications in the United States (N=352) and the Netherlands (N=81). Our analyses do not support the conclusions of the original study, nor do we find evidence for broader claims regarding the effect of disgust and the existence of a physiological trait. Rather than studying unconscious responses as the "real" predispositions, alignment between conscious and unconscious responses promise deeper insights in the emotional roots of ideology.

Main

In a study published by Science, Oxley et al.[1] demonstrate that individuals who express conservative political attitudes (e.g., a preference for tradition, clear group boundaries, and hierarchies) have stronger physiological reactions to threatening stimuli than those who express liberal views. One of the authors' explanations for this association is that "political attitudes and varying physiological responses to threat may both derive from neural activity patterns" (p.1669),[1] suggesting a biological basis for the robust, positive correlation between political conservatism and self-reported measures of sensitivity to uncertainty and threat[2–5]. Research by others, including the Oxley, et al. research team, demonstrates that conservatives also have stronger physiological responses to disgusting stimuli than liberals do[6,7]. These findings motivated members of the Oxley, et al. research team to further elaborate that conservatism is linked to a more expansive form of "negativity bias" at a physiological level[9].

This research helped usher in a new paradigm in psychology and political science that offers a provocative perspective on the biological roots of the centuries-old liberal/conservative divide[8]. It implies that psychological traits, such as "threat sensitivity", "disgust sensitivity", and "negativity bias" can be captured in a straightforward and unobtrusive way with physiological reactions, and that these "physiological traits" have a direct and unmediated effect on

political attitudes. Their unified framework pushed the field to move beyond vague theoretical models that ascribe the provenance of political attitudes to largely unspecified social processes (e.g., "socialization")[9]. As such, the article has been widely cited (617 citations in Google Scholar at the time of writing; November 28, 2019), discussed,[9] and fuelled a proliferating literature that connects neuropsychology to political attitudes[10–13].

Oxley et al.'s findings also have considerable societal implications. Oxley et al.[1] conclude that if "political attitudes vary with physiological traits linked to divergent manners of experiencing and processing environmental threats," then their findings offer "one possible explanation for both the lack of malleability in the beliefs of individuals with strong political convictions and for the associated ubiquity of political conflict" (p.1669). This conclusion bolsters the media narrative that with increasing polarization in the United States and beyond,[14] we should abandon the hope that liberals and conservatives will ever be able to see eye to eye. The significance of this message is reflected by the attention paid to this study's finding by the news media up to today – it continues to garner extensive coverage in major news agencies ranging from CNN, New York Times, NPR, VOX, Forbes, BBC and National Geographic to the Daily Show.

Neuropsychological processes are certainly an important factor in political attitude formation,[13] but Oxley et al.'s account is underspecified. Recent research calls into question the notion that there is a direct and unmediated decision-making process that starts with automatic reactions (e.g., physiological responses), proceeds to cognitive assessments (e.g., judgments about threats), and ends with behavioral responses (e.g., political ideology)[15]. There is, at best, a loose correlation among physiological, cognitive, and behavioral responses[16,17]. Save for individuals who exhibit phobias, physiological responses to the types of distal emotional stimuli that Oxley et al. employed (e.g., pictures of spiders) are at best ambiguous indicators of psychological traits[18]. Not only are people capable of regulating the link between automatic responses

and cognitive evaluations,[19] but the human mind is also capable of considering and dismissing threats without engaging in cognitive reflection[18].

Considering this criticism and the impact of Oxley et al.'s findings for science and society, it is crucial to replicate these findings. Null findings in the social sciences are much less likely to be written up and to get published compared to studies with statistically significant findings[20]. Recent efforts to replicate important findings in psychology show the need for direct replications[21–23]. Evaluating the replicability of papers published in Science and Nature between 2010 and 2015, Camerer and colleagues[24] replicated 13 (62%) out of the 21 selected papers. Moreover, Camerer et al. find that the effect size of the replicated studies is approximately 50% compared to the original effect size reported in Science or Nature. This has led to an increased call for replication studies in prominent journals. Kevin Smith and John Hibbing — two of the lead authors of the Oxley et al. paper — acknowledged the need for replication, as they wrote that the question whether their findings replicate cannot be known "until other laboratories have weighed in"[25] (p.3, emphasis added). The need for a direct replication is even more pressing since — to the best of our knowledge — the only published conceptual replications offer conflicting conclusions[26–28] — for details see Supplementary Results p.33.

Results

We conducted a preregistered more direct replication – see OSF page – in the United States (N=202) and conceptual replications in the U.S. (N=352) and the Netherlands (N=81) that were not pre-registered. Our replications focus on the most prominent test of the Oxley et al. study that physiological responses to threatening images correlate with ideology. We test three claims: 1) that threat sensitivity is a "physiological trait"; 2) that threat sensitivity, as measured by physiological responses, correlates with social conservatism; and 3) that disgust sensitivity is a "physiological traits" and correlates with conservatism[7,9].

Our pre-registered replication is four times the sample in the original Oxley et al. study[29,30]. A power analysis (see Methods section) informed us that we would need 199 participants to have sufficient statistical power to detect a small effect as the one reported by Oxley et al. In turned out that a couple of people had missing values on the physiological measures. This leaves us with a sample of 191 participants in the analyses. Re-calculation the power, using the same parameters but with 191 participants, we arrive at the statistical power of .79, very close to the pre-registered of .80. In all three studies, we closely followed Oxley et al.'s protocol and measured physiological responses using skin conductance levels. Following Chambers,[31] our pre-registered replication tests the "repeatability of" Oxley et al. "by duplicating the methodology as exactly as possible" (p.48) – see also[32]. Oxley et al. kindly provided us with the three original threatening stimuli used in their study to measure threat sensitivity: (1) "a very large spider on the face of a frightened person," (2) "a dazed individual with a bloody face," (3) and "an open wound with maggots in it." Like Oxley et al., we interspersed negative images among 27 other images. To probe the robustness of our results and to address the later claim concerning disgust sensitivity, we extended the replication to include more threatening (4) and disgusting (5) images (see Supplementary Methods p3-4 for a discussion). In the conceptual replications we used images that are similar to the threatening and disgusting images used by Oxley et al. and subsequent publications of members of this research team[7,33].

Oxley et al. combined skin conductance responses to three images as a measure of a "physiological trait." If such a traits exists, there should be positive and high intercorrelations between the physiological responses to the different images. Although these intercorrelations are not reported in the original study, we did preregister to test this. Fig. 1 plots the Pearson correlation coefficients between the physiological responses to the images. In the direct replication 0 out of 3 correlations are positive – e.g., Figure 1, Supplementary Table 9. In the preregistered extensions (panel B and C) only 2 out of 32 correlations are positive and relatively strong (i.e., r>.3)

– e.g., Figure 1 panel B and C, Supplementary Table 10 and 11. In the conceptual replications in the U.S. (panel D) we find 1 out of 3 correlations – e.g., Figure 1 panel D, Supplementary Table 12), and in the Netherlands we find 0 out of 6 correlations to be positive and strong – e.g., Figure 1 panel E, Supplementary Table 13. In sum, we find weak and even negative correlation coefficients. This undermines the notion that there is an overarching latent physiological trait of either threat or disgust sensitivity.

Fig. 2 summarizes the main results from our replications. Following Oxley et al. and our pre-analysis plan, we separated political attitudes into two indexes: (1) social conservatism (e.g., support for traditional values, opposition to immigration, etc.) and (2) economic conservatism (e.g., opposition to taxes, support for free markets, etc.). We also included the same control variables as they did (income, education, and gender identity) along with the necessary controls for study characteristics (See Methods section). The top row shows the results from the pre-registered direct replication that combined skin conductance responses to the exact images in the Oxley et al. study into a single index. Like Oxley et al., we find no statistically significant association between skin conductance response to their threatening images and economic conservatism – e.g., Figure 2 right-hand panel of row 1, Supplementary Table 15. More crucially, however, we find no statistically significant association between skin conductance response and social conservatism (b=.01 [95% CI=-0.02, 0.04], se=.01, p=.62, t=.50) – e.g., Figure 2 left-hand panel of row 1, Supplementary Table 14 and Supplementary Table 15. Our standardized coefficient is 37 times smaller than the statistically significant and substantive effect (d=.37) reported by Oxley et al.

To discover whether the connection between physiological reactions to images and political conservatism lurked elsewhere in the data, we conducted a number of follow up tests specified in our preregistered analysis plan. We fail to find a statistically significant positive relationship between skin conductance response to the individual images used by Oxley et al. and mea-

sures of political conservatism – e.g., Fig. 2 row 1, Supplementary Table 14. We do not find a single statistically significant relationship between skin conductance response to the additional "threatening images" included in our preregistered replication, whether we analyze them combined in an index or separately – e.g., Fig. 2 row 2, Supplementary Table 16 and Supplementary Table 17. Furthermore, we do not find any statistically significant relationship between political conservatism and skin conductance response to the disgusting images included in our preregistered replication – e.g., Fig. 2 row 3, Supplementary Table 19 and Supplementary Table 20. In the conceptual replications conducted in the U.S. – e.g., Fig. 2 row 4, Supplementary Table 22 – and the Netherlands – e.g., Fig. 2 row 5, Supplementary Table 24 – we also find no statistically significant positive associations between threat sensitivity and social conservatism. Even if we pool the data into one dataset, we find no evidence of a positive and statistically significant correlation between skin conductance responses to negative images and conservatism – e.g., Fig. 2 row 6, Supplementary Table 26.

In sum, across multiple samples and multiple tests, we observe point estimates which are very close to zero, confidence intervals that include zero, and inconsistently positive signs. If we focus on the 28 results pertaining to social conservatism in Fig. 2 – which is where Oxley et al. found a statistically significant positive correlation with skin conductance response – 19 have a positive sign, while 9 have negative signs. The Bayes factors in the direction of the null hypothesis[34] for the main analyses of each study are reported in Table 1 – none of the Bayesian analyses were preregistered. We calculated the Bayes factors with JASP version 0.9.2. Following Wagenmakers et al.[34] recommendation, we use the stretched beta distribution for the prior (width=1). With respect to the hypothesis that social conservatives have stronger physiological responses to threatening and disgusting images, the data offer moderate to strong evidence in favor of the null hypothesis[35]. As such, these samples provide virtually no support for the claim that socially conservative individuals display stronger physiologically responses

to threatening, disgusting, or simply negative images.

Sixteen robustness checks (see Table 2) provide compelling evidence that these null findings are robust against alternative measures and explanations – results for each of the questions in Table 2 are either provided in the Supplementary Results or can only be derived from the replication files. Even though our studies were conducted in areas where liberals outnumber conservatives, our pooled analyses consist of approximately 4.5 times more conservatives than the Oxley et al.'s sample (Question 1 – e.g., Supplementary Results p.35). To this point, our samples do show meaningful associations between self-reported psychological traits and conservatism and between the socio-economic covariates and conservatism, which suggests that the null results are not a product of an odd sample pool (Question 2 – e.g., Supplementary Table 30). Our measures of skin conductance pass basic manipulation checks — participants did in fact respond to negative images more strongly than other types of images (Question 3 – e.g., Supplementary Table 31-33). It does not matter if we change the operationalization of the threat sensitivity index (Question 4 – e.g., replication materials) or if we decompose the analyses by specific policy attitudes (Question 5 – e.g., Supplementary Figure 10-17). Also splitting ideology at the median (or midpoint) makes no difference (Question 6 – e.g., Supplementary Table 34-38). We also do not find a reliable correlation between skin conductance responses and other measures of conservatism (e.g., social principles, authoritarianism, racism, partisanship, ideology) (Question 7 – e.g., Supplementary Figure 18-20). It does not matter if we use physiological responses to non-threatening/non-disgusting images as a point of comparison like Oxley et al. did in a robustness check (Question 8 – e.g., Supplementary Table 39-40). Our results do not depend on whether participants are knowledgeable about or interested in politics (Question 9, – e.g., Supplementary Figure 23-31) or on the participants' racial or gender identification or their age (Question 10 – e.g., Supplementary Figure 32-45). We continue to observe null results even if we restrict our sample such that we have a bimodal distribution of social conservatism,

like Oxley et al. do (Question 11 – e.g., Supplementary Figure 46). Physiological responses are also not associated with attitude extremity (Question 12 – e.g., replication materials), nor do we find evidence that variance in physiological responses to threatening images is associated with conservatism or attitude extremity (Question 13 – e.g., replication materials). We also collected physiological measures designed to capture negative emotional reactions using electromyography (specifically, corrugator and labii activity). Analyzing these measures, we find no evidence for a latent physiological trait of threat sensitivity or disgust sensitivity nor do we find that these measures correlate with conservatism (Question 14 – e.g., Supplementary Figure 47-50). Using the different different physiological measures (Skin conductance, corrugator and labbi) we find no evidence for an overarching physiological trait of threat sensitivity, disgust sensitivity or negativity bias and these dimensions don't correlate with conservatism (Question 15 – e.g., replication materials). Finally, we also do not find evidence for the later claim made by members of the Oxley et al. team[33] that a broader measure of "right-of-center" political orientations correlates with physiological responses to negative images compared to positive images (Question 16 – e.g., replication materials).

Discussion

We, reluctantly, reject the foundational assumption of Oxley et al.'s work. Nonetheless, we do not think that this should be the end of the biology and politics paradigm Oxley et al. helped initiate. Other fields have found that while the structure of the brain affects human behavior, it does not do so in direct and easily measured ways[36]. Politics should be no different, especially given the complex personal experiences that politics evoke.

Failed replications always raise questions about why the original study's findings were not supported. One possibility is that the underlying theory was either incorrect or bounded by unstated scope conditions. Another possibility is that differences between the original study

and replication account for the discrepancy. Like all replications, ours differed in a few ways from Oxley et al.'s original study. First, our sampling strategy and sample characteristics did not match one-on-one Oxley et al.'s sampling strategy and sample. We believe that this is an unlikely explanation for our failed replication. In line with recent work in the personality and politics literature[37,38] and replication attempts in social psychology,[23] we find no evidence for a stronger association between threat sensitivity and social conservatism once we constrain our sample to people that more closely resemble those in Oxley et al.'s sample with respect to age and race (Question 10 – e.g., Supplementary Figure 32-45) or the distribution of social conservatism (Question 11 – e.g., Supplementary Figure 46). Yet, other unknown differences caused by the different sampling strategies could potentially explain the different findings across the two studies. Replicating Oxley et al. using the same sampling strategy would provide a more definite answer. But even if our sample were the explanation for our null results, it would imply an unstated scope condition in the original study: namely, that it only applies to a particular population of people, as opposed to a general explanation for conservatism.

Second, our pre-registered replication consisted of the exact same threatening stimuli that Oxley et al. used. Oxley et al. interspersed these among a sequence of 33 images, some of which were positive. Yet, the exact content of the other images was not known to us. We interspersed the three threatening images among a sequence of 30 images – mostly taken from the International Affective Picture System – which were validated to be negative, positive or neutral. The basic procedure thereby matches Oxley et al.'s procedure. We doubt that the different images we choose condition the results. Like Oxley et al., we randomized the order in which the images were shown. Recent work shows that the order in which threatening images are shown does not affect the response that people have to the images[26]. Indeed, our threatening images pass a basic manipulation check. Therefore, we do not think that the failure to replicate Oxley et al. can be attributed to – unknown – differences between the content of the other

images used by Oxley et al. versus our images.

Third, we collected data in another part of the US and another country – the Netherlands. We can only speculate whether contextual differences explain the absence of a statistically significant association. On a theoretical level, Oxley et al. assert that the association between threat sensitivity and conservatism should be universal[1,7,9]. Consequently, if context conditions the link between physiological responses and conservatism it would imply a scope condition unstated by Oxley et al. On an empirical level, large-scale replications projects in social psychology do not give much reason to presume that context conditions the outcome of replication studies[23]. At the same, Malka and colleagues[39] show that the link between self-reported personality traits and social conservatism are stronger in countries with more ideological constraint. This would imply that we would find even weaker associations between threat sensitivity and social conservatism in countries with less ideological constraint. The only way to answer this question is to conduct more research using cross-country designs, for a recent example[40].

We see multiple opportunities for further research. Our study shows that using the exact same threatening (and disgusting) stimuli, as well as conceptually similar stimuli in the conceptual replications, arrive at the same null findings. In our view, these null results are in line with the lack of theoretical and empirical evidence for the presence of an overarching physiological threat sensitivity dimension in recent neuroscience research. Of course, our view could be incorrect, and additional research could uncover a link between physiological responses and political ideology. For instance, it is possible that there are distinct sub-dimensions of physiological responses that tap into specific non-political threats (snakes, spiders) as well as political threats (violence), and these sub-dimensions could have distinct correlates with specific ideological sub-dimensions. Pre-registered, well powered, designs could speak to this question.

Furthermore, there is a long line of research in psychophysiology showing that negative stimuli evoke arousal fuelled by activation of the sympathetic nervous system[41]. Perhaps, the

use of still images as stimuli is not the best way to evoke the kind of arousal that would signal individual differences in threat sensitivity. Stronger and more immersive stimuli such a real physical threats (e.g., a spider in the room)[42] or threats in a virtual reality environment[43] are perhaps better equipped to capture individual differences in threat sensitivity. We welcome the next generation of research that would make use of more immersive stimuli.

Our study has broader implications for the literature in psychology and political science on the ideological asymmetry between liberals and conservatives. This literature predominantly relies on cross-sectional surveys in which self-reported traits are correlated with ideology[44]. The association between threat sensitivity and ideology has been considered an important piece of supporting evidence for the asymmetry thesis[9]. Our failure to replicate the biological correlates of ideology, however, aligns with recent work that questions the robustness of the personality and ideology association[45]. For instance, the content overlap between psychological traits and ideology might explain the association between personality and politics[46]. More troubling, panel studies failed to find evidence for the assumed causal effect of personality on ideology,[47,48] and the causal arrow may be reversed: people's ideology might motivate them to express certain trait characteristics[49]. Therefore, while survey based cross-sectional studies report consistent – albeit weak[3] – associations between personality and ideology, our study aligns with a small but growing body of literature that suggests a more critical assessment of the ideological asymmetry between liberals and conservatives.

To conclude, the field should move beyond the simplistic premise that unconscious reactions indicate "real" predispositions. Rather, unconscious reactions are one aspect of how the brain operates and one input into how people make decisions. One fruitful line of research would be to consider whether individual differences in the alignment between physiological reactions and experienced emotions may offer insights into how much people's political attitudes are rooted in intuitions and how much of it is regulated by top-down processes. It may also help scholars

sort out when expressed political beliefs and attitudes are sincere and when they are expressive. Whatever the case, we urge more, not less, research at the intersection of neuroscience and politics. It will not be easy, but profitable avenues of research rarely are.

Methods

Ethics

The studies conducted in the United States were approved by Temple University's Institutional Review Board. The study in the Netherlands was approved by the Ethical Review Board of the University of Amsterdam. We have obtained informed consent from all participants at the start of the study.

Statistical Analysis

Data collection and analysis were not performed blind to the conditions of the experiments. We do not systematically exclude respondents from the analyses. We follow Oxley et al.'s procedures (see p. 3 of their supplementary material) and our equations for the multiple regression models that we report in the main text of the paper – and in the Supplementary Results – are structured along the lines of formula 1. Note that in this equation T is the matrix of control variables – in the direct replication: income, education, age, gender identity, indicator of the first eight participants, study event, payment and whether scl is above or below 2 microSiemens. In the pre-registered extension we also control for race. In the conceptual replication in the US we include similar covariates. In the Netherlands, we control for gender, age and education (other covariates were not measured). Following Oxley et al. all tests reported are two-sided and we don't control for multiple comparisons. We tested the regression assumptions for all models and these can be derived from the replication files.

$$\text{Social Conservatism} = \beta_0 + \beta_1 ThreatSensitivity + X\Gamma + \epsilon_i \quad (1)$$

Pre-registered replication

Pre-registration

The pre-analysis plan was deposited on June 29, 2018 and can be found on our OSF page. We did not inspect the data before uploading the pre-analysis plan to OSF.

Power analysis

The pre-registered replication was conducted between February 7, 2018 and June 30, 2018 at a large university on the East Coast of the United States. This sample size was based upon a pre-registered analysis. The effect sizes sizes in the published literature are small.[1,6,7,26,33] Using Gpower 3.1.9.2 we calculated the sample size we would need to achieve a power of 0.8 given the α (0.05) and expected Δ F^2 of 0.04. This resulted in a sample size of 199 respondents. Therefore we aimed for at least 199 participants that completed the study. To account for some drop-outs, a total of 202 participants took part in the study.

Recruitment

The study lasted an hour and was conducted by a trained lab assistant. We relied upon respondents recruited among university students, advertisements on Craigslist as well as people recruited via temporary employment agencies. As such, we provided different levels of compensation. Student participants were compensated for their participation with $10.00 – either in cash or an Amazon gift-card –, which is above an hourly minimum wage; Craigslist participants received $30.00 and temporary agency participants were compensated in accordance to the rates set by the temporary employment agency ($16.00 per hour). We did this so that participants were not asked to forgo income as a condition of participating in the study.

Design

Upon signing the informed consent, the trained lab assistant first cleaned parts of the face with water and a non-oily facial cleanser on a cotton pad to remove any make-up that would prevent the adhesive disposable electrodes to stick onto the face. Next, they start with the survey and we measured socio-economic background was captured using age (in years), gender, race, income (in categories) and their highest completed educational level. For descriptive statistics see "covariates" section below.

Participants also completed a series of batteries tapping to their political attitudes and beliefs, political knowledge and political interest as well as their personality – Authoritarianism, Agreeableness and a "Reading the Mind in the Eyes" test. These batteries were presented in a random order and completed by all participants. This was followed by a series of three survey-experiments. The full survey is provided on our OSF page: https://osf.io/d5g72/.

Upon completing the survey participants were connected to the physiological measures and exposed to a series of party symbols (one of the Republican Party and one of the Democratic party), party leaders (5 Democrats & 5 Republicans) and six issue statements – the order of these blacks was randomized. Participants were also given noise cancelling headphones (Bose) to black background noise. Afterwards they were exposed to the affective stimuli which are part of this study.

Political ideology

Social conservatism: In the direct replication we measured social conservatism using 12 items that were directly included by Oxley et al. 2008, namely: (1) Death penalty, (2) School prayer, (3) Biblical truth, (4) Gay marriage, (5) Legalized abortion, (6) Protect gun rights, (7) Increase military spending, (8) Deport undocumented immigrants, (9) Restrict legal immigration (such as work visas), (10) Warrantless searchers, (11) Patriotism and (12) Foreign aid. The index

was scored to range from 0 (social liberalism) to 1 (social conservatism) and was internally consistent (α=.80, M=.38, SD=.21, Min=0, Max=1).

In the pre-registered extensions we measured social conservatism with the 12 items included by Oxley et al. and three additional items, namely: (1) Pre-marital sex, (2) Woman's equality and (3) Separation of church and state. The index was scored to range from 0 (social liberalism) to 1 (social conservatism) and was internally consistent (α=.82, M=.40, SD=.23, Min=0, Max=1). The two social conservatism dimensions are very highly correlated (r=.98 [95%CI=.97, .99], t(200)=71.26, p$<$.001).

Economic conservatism: In the pre-registered more direct replication, we measured economic conservatism using six items, namely: (1) Increase banking regulations, (2) Increase spending on public education, (3) Increase environmental regulations, (4) Reduce income inequality, (5) Government-funded health care, (6) Raise taxes on the rich. We created a scale ranging from (0) economic liberalism to (1) economic conservatism (α=.72, M=.24, SD=.16, Min=0, Max=1). As expected, economic conservatism was positively correlated with the social conservatism dimension used in the direct replication (r=.48 [95%CI=.36, .58], t(200)=7.69, p$<$.001) and the pre-registered extensions (r=.50 [95%CI=.39, .60], t(200)=8.21, p$<$.001). A confirmatory factor analysis showed that the items load highly on the designated latent ideology dimension (see Supplementary Table 1).

Covariates

We measured age in years (M=25.86, SD=11.06, Min=18, Max=82). Respondents could indicate whether they were female (N=128; 63.4%), male (N=69; 34.2%) or indicate to have another gender (N=5; 2.4%).

To measure race, respondents were asked "What general racial or ethnic category do you consider yourself (select all that apply)" and they could choose out White, Black, Latino, Asian

or another race. We coded people as belonging to a group if this was the only group they mentioned. All other respondents were put in the "other" category. Our sample was racially diverse: White (N=85; 42.08%), Black (N=62; 30.69%), Latino (N=10, 4.95%), Asian (N=30; 14.85%) or another race (N=15; 7.43%).

Income was measured using the question "We want to classify people into broad income groups. This information is completely confidential. Please indicate the category that corresponds to your family's annual income." and scored in ten categories ranging from "under $15,000" to "over $ 200,000. The median income was "between $65,000 and $79,999" (M=5.33, SD=2.47, Min=1, Max=10).

Education was measured using the question "What was the last level of schooling you completed?" and participants could answer "less than high school graduate", "High school graduate", "Some college", "Currently a college student" or "College graduate" or "Post graduate". We grouped "less than high school graduate" and "High school graduate" as there were only few respondents in these groups. We set "high school or less" as the reference category (N=20, 9.90%). The sample has some variance in educational background but is fairly highly educated: "Some college" (N=35, 17.33%), "Currently a college student" (N=102, 50.49%) or "College graduate" (N=23, 11.39%) or "Post graduate" (N=22, 10.89%).

Affective stimuli

The affective pictures were shown at the end of the study. The images were drawn from the International Affective Picture System (IAPS) and images used by Oxley et al. The first image was always a neutral image, namely the basket (IAPS #7010). The 29 remaining images were randomly presented to the participants.

To directly replicate Oxley et al. we included the exact same images Oxley et al used in their study: 1) "a very large spider on the face of a frightened person (non-IAP); (2) "a dazed

individual with a bloody face (IAPS #3550); (3) and "an open wound with maggots in it" (non-IAPS). To extend our measure of threat sensitivity in the pre-registered extensions, we also included an image of (1) Dog (IAPS #1300), (2) Gun pointing at the screen (IAPS #6260), (3)Twin Towers exploding (IAPS #9940) and (4) a crowd fighting with a man (non-IAPS) also used by members of the Oxley et al. team in other publications.[33,50]

We also included a series of disgusting images. We included four of the five images that were included by Smith et al[7] in their 2011 PlosOne paper, namely: (1) a man eating worms, (2) human excrement floating in a toilet, (3) a bloody wound. The fourth image Smith et al. used was the image of the "an open wound with maggots in it". Note that the image of the maggots was used to measure "threat sensitivity" in Oxley et al. and "disgust sensitivity" in Smith et al. We pre-registered to use the image of the maggots as an index of threat sensitivity in the direct replication (Oxley et al.) but as an index of disgust sensitivity in the remainder of the analyses like Smith et al. Finally, we also include two images containing (5) vomit and (6) a dead dog (IAPS #9570).

Aside from these threatening and disgusting images, we also include five happy, five neutral and five exciting pictures (see Supplementary Table 2).

Physiological measures

We recorded physiological data using a Biopac system running on Windows 7 with a sampling rate of 1000Hz. The stimuli were shown using E-prime 2.0 and the physiological data was acquired using Acknowledge 4. The experiment was part of a larger study assessing physiological responses to politics. We measured arousal using electrodermal activity, negative affect using electromyography of the corrugator supercilii and disgust using the levator labii superioris. Here we discuss the electrodermal activity as this measured used in the pre-registered direct replication. The results using eletormyography (corrugator & labii) will be discussed in

the Supplementary Results – see section belonging to Question 14 (see Table 2).

We measured individual differences in sensitivity with participants' electrodermal response to the affective stimuli as well as a set of other images. This is a standard approach to gauging physiological sensitivity. We do this using Biopac's disposable Ag/AgCL electrodes which are pre-gelled with isotonic gel (0.5% chloride gel). The electrodes are 27mm wide, 36mm long and 1.5mm thick and have a 11mm contact area. The electrodes were placed upon the distal phalanges of the index finger and middle finger of the non-dominant hand. Using this set-up we measure the Skin Conductance Level (SCL) during the study. SCL is a standard measure of electrodermal activity or the degree to which people sweat which is an index of people's emotional arousal generated by the sympathetic nervous system. Because the sympathetic nervous system is difficult to control through conscious effort, SCL provides an unobtrusive measure of emotional arousal caused by the threatening and disgusting stimuli.

Participants were exposed to each image for 12 seconds with a blank screen between each one. This interstimulus interval (ISI) was also 12 seconds. Electrodermal response was measured by taking the difference between SCL recorded while participants viewed the picture and the SCL recorded during the ISI that proceeded the specific image. To minimize the effects of extreme values, we preregistered to follow Oxley et al. and take the average of the natural log of SCL during the image and the ISI. Equation 2 summarizes this approach:

$$PS_i = \frac{\sum_{j=1}^{12,000} ln[SCL(T)_{ij}]}{12,000} - \frac{\sum_{j=1}^{12,000} ln[SCL(ISI)_{ij}]}{12,000} \tag{2}$$

Where PS_1 is the physiological sensitivity score for participant i, $SCL(T)_{ij}$ is the skin conductance level recorded every j^{th} millisecond for participant i during exposure to the image, and $SCL(ISI)_{ij}$ is the skin conductance level recorded every millisecond for participant i during exposure to the preceding blank screen (i.e., the ISI). Because participants were exposed to the blank screen and image for 12 seconds each, we recorded SCL for 12,000 milliseconds

for each image and ISI. We summed this for the response to each of the 6 threatening images. This results in an index of physiological sensitivity as expressed by skin conductance levels in the same way that Oxley et al. operationalized this. Using the same equation, we calculate indices of disgust sensitivity. The Supplementary Table 3 provides the descriptive statistics for the indices and individual images that have been reported in the paper.

Missing data

We experienced some drop-out due to failed readings of the physiological data (N=9, 4.95%). The respondents that dropped out did not differ from the other respondents on sex, age, race, education and income – results can be derived from the replication file.

Conceptual replication U.S.

Sample

The data were collected as part of three larger study protocols. Data collection and analysis were not performed blind to the conditions of the experiments. No statistical methods were used to pre-determine sample size but our sample size is larger than those reported in previous publications.[1,7,26–28,33] Specific periods of data collection were: between October 27, 2014 and December 12, 2014 (protocol 1: N=105); June 24, 2015 and November 28, 2016 (protocol 2: N=106); and February 16, 2017 and January 30, 2018 (protocol 3: N=141). The full sample consists of students and non-student adults (N=352). For discussion of the differences between the protocols, see Supplementary Methods.

Recruitment

The study lasted an hour and was conducted by a trained lab assistant blind to the expectations in the study. We rely upon respondents recruited among university students as well as people recruited via temporary employment agencies. As such, we provide different levels of compen-

sation. Students recruited from campus were paid the $10 – or the equivalent in a currency that students could use in campus shops – in return for their participation. Adults that were recruited from a temporary employment agency were paid in line with the hourly rate of the agency ($16 per hour). We do this so that participants were not asked to forgo income as a condition of participation in the study.

Design

Upon signing the informed consent, participants completed a survey. This survey contained socio-economic background was captured using age (in years), gender, race, income (in categories) and their highest completed educational level. Next participants were asked to complete a series of batteries tapping to their political attitudes and beliefs as well as their personality. We will discuss the batteries of interest below – for detailed discussion of other measures, see Supplementary Methods.

Upon completion of the survey participants were connected to the physiological measure (skin conductance) and started the second part of the study. Note that skin conductance was measured on the non-dominant side of the body so that participants had their hand available to use the computer mouse. The three protocols all ended with the affective stimuli that are part of this replication study. Participants were also given noise cancelling headphones (Bose) to block background noise.

Political ideology

Social conservatism was measured using a series of issue attitude questions closely resembling those used by Oxley et al. using items such as "same-sex marriage". Participants indicated their answers on a scale from 1 (strongly oppose) to 6 (strongly support). All respondents reported their attitudes (1) legalized abortion, (2) same-sex marriage, (3) prayer in school, (4) death penalty, (5) protect gun rights and (6) increase military spending. Across the three protocols

there were additional items that were included in some but not all protocols, namely (7) legalized marijuana (protocol 1), (8) stop immigration (protocol 1); (9) illegal immigration (protocol 2), (10) biblical truth (protocol 2), (11) Deport undocumented immigrants (protocol 3) and (12) Restrict legal immigration (such as work visas) (protocol 3). After recoding the reversed coded items, we calculated the mean social conservatism – for each participant over the items that they completed – that scored to range from the most social liberal (0) to the most social conservative (1) score in our sample (M=.39, SD=.23, α=.74).

If we create a social conservatism dimension out of these six items (M=.39, SD=.23, α=.71), we see a strong positive correlation with the social conservatism battery used in the main text (r=.95 [95%CI=.94, .96], t(349)=56.98, p$<$.001). Using this alternative dependent variable, we also find no evidence that threat sensitivity is associated with social conservatism (see Supplementary Table 4).

Economic conservatism was measured using seven items: (1) "increase banking regulations", (2) "Increase spending on public education", (3) "Increase environmental regulations", (4) "Reduce income inequality", (5) "Government funded healthcare", (6) "Reduce welfare spending" and (7) "Raise taxes on the rich". All items were scored using the same likert-type response scale as the social conservatism items. We created a sale ranging from (0) economic liberalism to (1) economic conservatism (M=.28, SD=.19, α=.84).

Covariates

The sample consisted of 223 females (66.71%), 124 males (35.43%) and 3 (.85%) persons who did not identify as female or male. The mean age of the participants was 25.14 (SD=11.89, Min=18, Max=71).

To measure race was measured using the question "What general racial or ethnic category do you consider yourself (select all that apply)?" and they could choose out of White, Black,

Latino, Asian or another race. We coded people as belonging to a group if this was the only group they mentioned. All other respondents were put in the "other" category. Our sample was racially diverse: White (48%; N=168), Black (28.29%; N=99), Latino (4%; N=14), Asian (14.57%; N=51) and Other (5.14%; N=18)

We measured income using the following item: "We want to classify people into broad income groups only. This information is completely confidential. Please indicate the category that corresponds to your family's situation?" with categories ranging between: "Under $15,000" (1) and "Over $200,000" (10). The median income was "Between $50,000 and $64,999" and income ranged over the full range of the income categories (M=5.3, SD=2.71, Min=1, Max=10).

Education was measured using the following question "What was the last level of schooling you completed?" Our sample consisted for the of people who completed "Less than high school graduate" (N=29, 8.2%), "High school graduates" (N=159, 45.2%), "Some college" (N=48, 13.6%), "Currently a college student" (N=98, 27.8%), "College graduate" (N=9, 2.6%) or "Post college degree" (N=8, 2.3%).

Affective stimuli

The images were part of a larger study measuring physiological responses to politics. At the end of the protocol, we included 10 images that were drawn from the International Affective Picture System (IAPS) database which includes hundreds of images coded for emotional reactions. The order of the pictures was randomly assigned but participants would always start with an 11^{th} neutral and mundane image (a basket, IAPS#7010). Threat sensitivity was measured using the image of the Twin Towers exploding (IAPS #9940), Dog (IAPS #1300) and Snake (IAPS #1050). Other than that, we included four positive and three sad images – see Supplementary Table 5 for details.

Physiological measures

We recorded physiological data using a Biopac system running on Windows 7 with a sampling rate of 1000Hz. The stimuli were shown using E-prime 2.0 and the physiological data was acquired using Acknowledge 4. The experiment was part of a larger study assessing physiological responses to politics. We measured arousal using electrodermal activity.

We measured individual differences in sensitivity with participants electrodermal response to the affective stimuli as well as a set of other images. This is a standard approach to gauging physiological sensitivity. We do this using Biopacs disposable Ag/AgCL electrodes which are pre-gelled with isotonic gel (0.5% chloride gel). The electrodes are 27mm wide, 36mm long and 1.5mm thick and have a 11mm contact area. The electrodes were placed upon the distal phalanges of the index finger and middle finger of the non-dominant hand. Using this set-up we measure the Skin Conductance Level (SCL) during the study. SCL is a standard measure of electrodermal activity or the degree to which people sweat which is an index of peoples emotional arousal generated by the sympathetic nervous system.

Participants were exposed to each image for 20 seconds with a blank screen between each one. This interstimulus interval (ISI) was also 20 seconds. Electrodermal response was operationalized using the same procedures as the pre-registered replication. So we took the difference between SCL recorded while participants viewed the picture and the SCL recorded during the ISI that proceeded the specific image. To minimize the effects of extreme values, we follow Oxley et al. and take the average of the natural log of SCL during the image and the ISI. The descriptive statistics for the indices and individual images are provided in Supplementary Table 6.

Missing data

We experienced some drop-out due to failed readings of the physiological data (N=10, 2.8%). The respondents that dropped out did not differ from the other respondents on sex, age, race, education and income – results can be derived from the replication file.

Conceptual replication in the Netherlands

We report the results form laboratory study conducted at a large Dutch university. No statistical methods were used to pre-determine sample size but our sample size is larger than those reported in previous publications.[1,7,26,33] Data collection and analysis were not performed blind to the conditions of the experiments.

Data collection ran from November 7 through December 2 2016. In total 120 participants completed the study but the first 39 observations could not be used because the randomization of the images was not correctly stored. Therefore, we rely upon a sample of 81 respondents.

Recruitment

The protocol lasted 45 minutes and was conducted by a trained lab assistant blind to the expectations in the study. Participants were recruited via flyers and the online portal of the laboratory of the university where we conducted the study. Participants received participation credits or 7 Euro 50 cents – this amount is in line with the ethical procedures of the University of Amsterdam and above the minimum wage in the Netherlands at the time of conducting the study.

Design

Upon signing the informed consent, participants first completed a survey (using Qualtrics) which tapped – among other things – social conservatism and economic conservatism as well as their socio-economic background. The survey also contained measures of left-right ideological self-placement, policy attitudes towards immigration, climate, political cynicism, as

well as the salience of attitudes towards immigration, economy and climate and social dominance orientation – the full survey (in Dutch) can be found on our public OSF page: https://osf.io/d5g72/.

After participants completed the questionnaire, they were connected to the physiological measures. Participants were also given noise cancelling headphones (Bose). First participants watched a series of short clips that contained political messages – these are part of an unrelated project and not reported here. Next, participants received a series of image drawn from the International Affective Picture System. Following the same procedures as in the U.S. samples, participants were randomly exposed to each image for 20 seconds with a blank screen between each (i.e., a 20 seconds interstimulus interval). First participants were exposed to a mundane object: in line with the US studies the image of a basket. Next, the affective images were presented at random.

Political ideology

Social conservatism was measured using the 14-item Social Principles Index. The social principles index was introduced by Smith et al. 2011 as a measure of social conservatism.[51] Participants were asked: "There are different ways to organize society. We are interested in the ways in which you think society would work best." Next participants were exposed to a series of binary choices, such as: "Society works best... (A) when people live according to traditional values or (B) people adjust their values to fit changing circumstances". Participants could choose one of the two options. The 14 items are provided below in English. We recoded the items so that they score from liberal to conservative. The index theoretical ranges from -14 (social liberalism) to +14 (social conservatism) and we recoded the scale to range from social 0 (social liberalism) to 1 (social conservatism; M=.35, SD=.20, Min=0, Max=1, α=.67) – see Supplementary Methods for item wording.

Economic conservatism was measured using two questions: "Income inequality it too big in The Netherlands. People with the lowest income should receive the highest salary increase" (reverse coded), "Income differences should become smaller" (reverse coded). The items correlated highly with each other (r=.76, 95%CI[.64,.84], t(79)=10.27, p<.001). We created an economic conservatism dimension ranging from (0) economic liberalism to (1) economic conservatism (M=.42 SD=.24). As expected, social conservatism and economic conservatism correlate positive and modestly with each other (r=.48 95%CI[.29,.63], t(79)=4.87, p<.001).

Covariates

The sample we use consists of 58 females (71.60%), 23 males (28.40%). The mean age of the participants was 23.86 (SD=4.70, Min=19, Max=46). 71 (91.36%) of our 81 respondents were students. We did not ask respondents to report their income. The sample of 81 respondents used in this study does not seem to differ from the full sample

Affective stimuli

The images were part of a larger study measuring physiological responses to politics. At the end of the protocol, we included 6 images that were drawn from the International Affective Picture System (IAPS) databases. The order of the pictures was randomly assigned but participants would always start with a neutral and mundane image (a basket, IAPS#7010). Threat sensitivity was measured using four images. We included two threatening images that were also included in the conceptual replication in the US, namely the Dog (IAPS #1300) and the Snake (IAPS #1050) as well as an image of a Gun pointing at the screen (IAPS #6260) and a herding dog (IAPS #1302). We also included one neutral image of a spoon (IAPS #7004). The order of these five images was randomized (see Supplementary Table 7, for details).

Physiological measures

We recorded physiological data using the Versatile Stimulus Response Registration Program 1998[52] software on stationary computers running on Windows 7. Note that the stimuli were shown using Presentation 20.3. The experiment was part of a larger study therefore we collected electrodermal activity as well as corrugator supercilii activity using facial Electromyography measures as well as heart rate. We sampled our data at 2000Hz.

Electrodermal activity (EDA) was measured using two reusable Ag/AgCl electrodes of 20mm by 16mm. The electrodes were attached with adhesive tape on the medial phalanx surfaces of the index and ring finger of the non-dominant hand – because other components of the study required respondents using the mouse. Skin conductance levels are reported in microSiemens. We created indices of physiological responses to the four images using the same procedures as specified in the conceptual replication in the US. Table. S8 provides the descriptive statistics for the physiological responses to the images.

Missing data

We follow the same procedures for missing data as in the direct and conceptual replication in the US (see above).

Data availability

The data reported in this paper and in the Supplementary Methods and Supplementary Results can be found on our public OSF page: https://osf.io/d5g72/.

Code availability

The analysis codes for both the aggregate data and each individual replication, as well as the results belonging to the Supplementary Methods and Supplementary Results can be found on

our public OSF page: https://osf.io/d5g72/.

References

1. Oxley, D. R. *et al.* Political Attitudes Vary with Physiological Traits. *Science* **321**, 1667–1670 (2008).

2. Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J. & Sanford, R. N. The authoritarian personality (1950).

3. Jost, J. T., Glaser, J., Kruglanski, A. W. & Sulloway, F. J. Political Conservatism as Motivated Social Cognition. *Psychological bulletin* **129**, 339–375 (2003).

4. Wilson, G. D. *The psychology of conservatism* (London: Academic press, 2013).

5. Altemeyer, B. *Enemies of freedom: Understanding right-wing authoritarianism* (Jossey-Bass, 1988).

6. Aarøe, L., Petersen, M. B. & Arceneaux, K. The behavioral immune system shapes political intuitions: Why and how individual differences in disgust sensitivity underlie opposition to immigration. *American Political Science Review* **111**, 277–294 (2017).

7. Smith, K. B., Oxley, D., Hibbing, M. V., Alford, J. R. & Hibbing, J. R. Disgust Sensitivity and the Neurophysiology of Left-Right Political Orientations. *PLoS one* **6**, e25552 (2011).

8. Laponce, J. A. *Left and right: The topography of political perceptions* (University of Toronto Press, 1981).

9. Hibbing, J. R., Smith, K. B. & Alford, J. R. Differences in negativity bias underlie variations in political ideology. *Behavioral and Brain Sciences* **37**, 297–307 (2014).

10. Nam, H. H., Jost, J. T., Kaggen, L., Campbell-Meiklejohn, D. & Van Bavel, J. J. Amygdala structure and the tendency to regard the social system as legitimate and desirable. *Nature Human Behaviour* **2**, 133 (2018).

11. Schreiber, D. *et al.* Red brain, blue brain: Evaluative processes differ in democrats and republicans. *PLoS one* **8**, e52970 (2013).

12. Pedersen, W. S., Muftuler, L. T. & Larson, C. L. Conservatism and the neural circuitry of threat: economic conservatism predicts greater amygdala–bnst connectivity during periods of threat vs safety. *Social Cognitive and Affective Neuroscience* **13**, 43–51 (2017).

13. Jost, J. T., Nam, H. H., Amodio, D. M. & Van Bavel, J. J. Political neuroscience: The beginning of a beautiful friendship. *Political Psychology* **35**, 3–42 (2014).

14. Lauka, A., McCoy, J. & Firat, R. B. Mass partisan polarization: Measuring a relational concept. *American Behavioral Scientist* **62**, 107–126 (2018).

15. Miller, G. A., Rockstroh, B. S., Hamilton, H. K. & Yee, C. M. Psychophysiology as a core strategy in RDoC. *Psychophysiology* **53**, 410–414 (2016).

16. Bradley, M. M. & Lang, P. J. Measuring emotion: Behavior, feeling, and physiology. In Lane, R. & Nadel, L. (eds.) *Cognitive Neuroscience of Emotion*, 242–276 (Oxford University Press, Oxford, 2000).

17. Lang, P. J. Fear reduction and fear behavior: Problems in treating a construct **3**, 90–102 (1968).

18. LeDoux, J. E. & Pine, D. S. Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework. *American Journal of Psychiatry* **173**, 1083–1093 (2016).

19. Butler, E. A., Gross, J. J. & Barnard, K. Testing the effects of suppression and reappraisal on emotional concordance using a multivariate multilevel model. *Biological Psychology* **98**, 6–18 (2014).

20. Franco, A., Malhotra, N. & Simonovits, G. Publication bias in the social sciences: Unlocking the file drawer. *Science* **345**, 1502–1505 (2014).

21. Collaboration, O. S. *et al.* Estimating the reproducibility of psychological science. *Science* **349**, aac4716 (2015).

22. Klein, R. A. *et al.* Investigating variation in replicability. *Social Psychology* (2014).

23. Klein, R. A. *et al.* Many Labs 2: Investigating Variation in Replicability Across Samples and Settings. *Advances in Methods and Practices in Psychological Science* **1**, 443–490 (2018).

24. Camerer, C. F. *et al.* Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nature Human Behaviour* **2**, 637 (2018).

25. Peterson, J. C., Smith, K. B. & Hibbing, J. R. Physiology and political beliefs: A response to Knoll, ODaniel, and Cusato. *Research & Politics* **3**, 2053168016662892 (2016).

26. Arceneaux, K., Dunaway, J. & Soroka, S. Elites are people, too: The effects of threat sensitivity on policymakers' spending priorities. *PloS one* **13**, e0193781 (2018).

27. Knoll, B. R., O'Daniel, T. J. & Cusato, B. Physiological responses and political behavior: Three reproductions using a novel dataset. *Research & Politics* **2**, 2053168015621328 (2015).

28. Coe, C. M., Canelo, K. S., Vue, K., Hibbing, M. V. & Nicholson, S. P. The physiology of framing effects: Threat sensitivity and the persuasiveness of political arguments. *The Journal of Politics* **79**, 1465–1468 (2017).

29. Simonsohn, U. Small telescopes: Detectability and the evaluation of replication results. *Psychological science* **26**, 559–569 (2015).

30. Munafò, M. R. *et al.* A manifesto for reproducible science. *Nature Human Behaviour* **1**, 0021 (2017).

31. Chambers, C. *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice* (Princeton University Press, Princeton, NJ, 2017).

32. Patil, P., Peng, R. D. & Leek, J. A visual tool for defining reproducibility and replicability. *Nature Human Behavior* 1–3 (2019).

33. Dodd, M. D. *et al.* The political left rolls with the good and the political right confronts the bad: connecting physiology and cognition to preferences. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 640–649 (2012).

34. Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H. & Grasman, R. Bayesian hypothesis testing for psychologists: A tutorial on the savage–dickey method. *Cognitive Psychology* **60**, 158–189 (2010).

35. Jeffreys, H. *The theory of probability* (Oxford University Press, Oxford, 1961).

36. Schnack, H. Brain and Behavior: Assessing reproducibility in association studies. *eLIFE* DOI: 10.7554/eLife.46757 (2019).

37. Vitriol, J. A., Larsen, E. G. & Ludeke, S. The generalizability of personality effects in politics. *European Journal of Personality* (2019).

38. Clifford, S., Jewell, R. M. & Waggoner, P. D. Are samples drawn from mechanical turk valid for research on political ideology? *Research & Politics* **2**, 2053168015622072 (2015).

39. Malka, A., Soto, C. J., Inzlicht, M. & Lelkes, Y. Do needs for security and certainty predict cultural and economic conservatism? a cross-national analysis. *Journal of Personality and Social Psychology* **106**, 1031 (2014).

40. Soroka, S., Fournier, P., & Nir, L. Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences*, **38**, 201908369 (2019).

41. Lang, P. & Bradley, M. M. The international affective picture system (iaps) in the study of emotion and attention. *Handbook of Emotion Elicitation and Assessment* **29** (2007).

42. Soeter, M. & Kindt, M. An abrupt transformation of phobic behavior after a post-retrieval amnesic agent. *Biological psychiatry* **78**, 880–886 (2015).

43. Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y. & Mühlberger, A. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in Psychology* **6**, 26 (2015).

44. Federico, C. M. & Malka, A. The contingent, contextual nature of the relationship between needs for security and certainty and political preferences: Evidence and implications. *Political Psychology* **39**, 3–48 (2018).

45. Crawford, J. T. Are conservatives more sensitive to threat than liberals? it depends on how we define threat and conservatism. *Social Cognition* **35**, 354–373 (2017).

46. Malka, A., Lelkes, Y. & Holzer, N. Rethinking the rigidity of the right model: Three sub-optimal methodological practices and their implications. In *Politics of Social Psychology*, 126–146 (Psychology Press, 2017).

47. Hatemi, P. K., Crabtree, C. & Smith, K. B. Ideology justifies morality: Political beliefs predict moral foundations. *American Journal of Political Science* (2019).

48. Hatemi, P. K. & Verhulst, B. Political attitudes develop independently of personality traits. *PloS one* **10**, e0118106 (2015).

49. Ludeke, S., Tagar, M. R. & DeYoung, C. G. Not as different as we want to be: Attitudinally consistent trait desirability leads to exaggerated associations between personality and sociopolitical attitudes. *Political Psychology* **37**, 125–135 (2016).

50. Gruszczynski, M. W., Balzer, A., Jacobs, C. M., Smith, K. B. & Hibbing, J. R. The physiology of political participation. *Political Behavior* **35**, 135–152 (2013).

51. Smith, K. B., Oxley, D. R., Hibbing, M. V., Alford, J. R. & Hibbing, J. R. Linking genetics and political attitudes: Reconceptualizing political ideology. *Political Psychology* **32**, 369–397 (2011).

52. Molenkamp, B. Vsrrp98 manual, version 8.0. *University of Amsterdam, the Netherlands* (2011).

Acknowledgements

We want to thank Kevin Smith, John Hibbing, Matthew Hibbing and the other authors of the Oxley et al. paper for their support of this project and providing the original stimuli used in direct replication. We also want to thank Nick Anspach, Ming Boyer, Jay Jennings, Sander

Kunst, Amanda Milena Alvarez, Ethan Fried, Isabella Rebasso and Denise van de Wetering for their assistance during the data collection. Mark Brandt, John Bullock, Jarret Crawford, Chris Dawes, Christopher Federico, Stanley Feldman, James Fowler, Christopher Johnston, John Jost, Lasse Laustsen, Ypthach Lelkes, Michael Bang Petersen, Jordan Mansell, Matthias Osmundsen, Valentina Parma, Kevin Smith, Stuart Soroka, Bram Spruyt, Michael Tessler and Josh Tybur provided helpful comments and suggestions during this project as well as panelists at the American Political Science Association Meeting 2018, Midwest Political Science Association Meeting 2017, Dutch Political Psychology Meeting, University of Mannheim, Free University Amsterdam, Politicologenetmaal 2018, Amsterdam School of Communication Research, Amsterdam Interdisciplinary Centre for Emotion, and the Hot Politics Lab. **Funding:** This research was funded by the European Unions Horizon 2020 research and innovation programme under grant agreement No 750443 (B.N.B), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 759079 (G.S), the Amsterdam School of Communication Research (B.N.B) and the Behavioural Foundations Laboratory at Temple University (K.A.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

B.N.B. and K.A. designed the study, contributed to data collection, analyses and write-up. G.S. contributed to the analyses and write-up. C.G. contributed to the data collection and write-up.

Competing interests

The authors declare no competing interests.

Figures

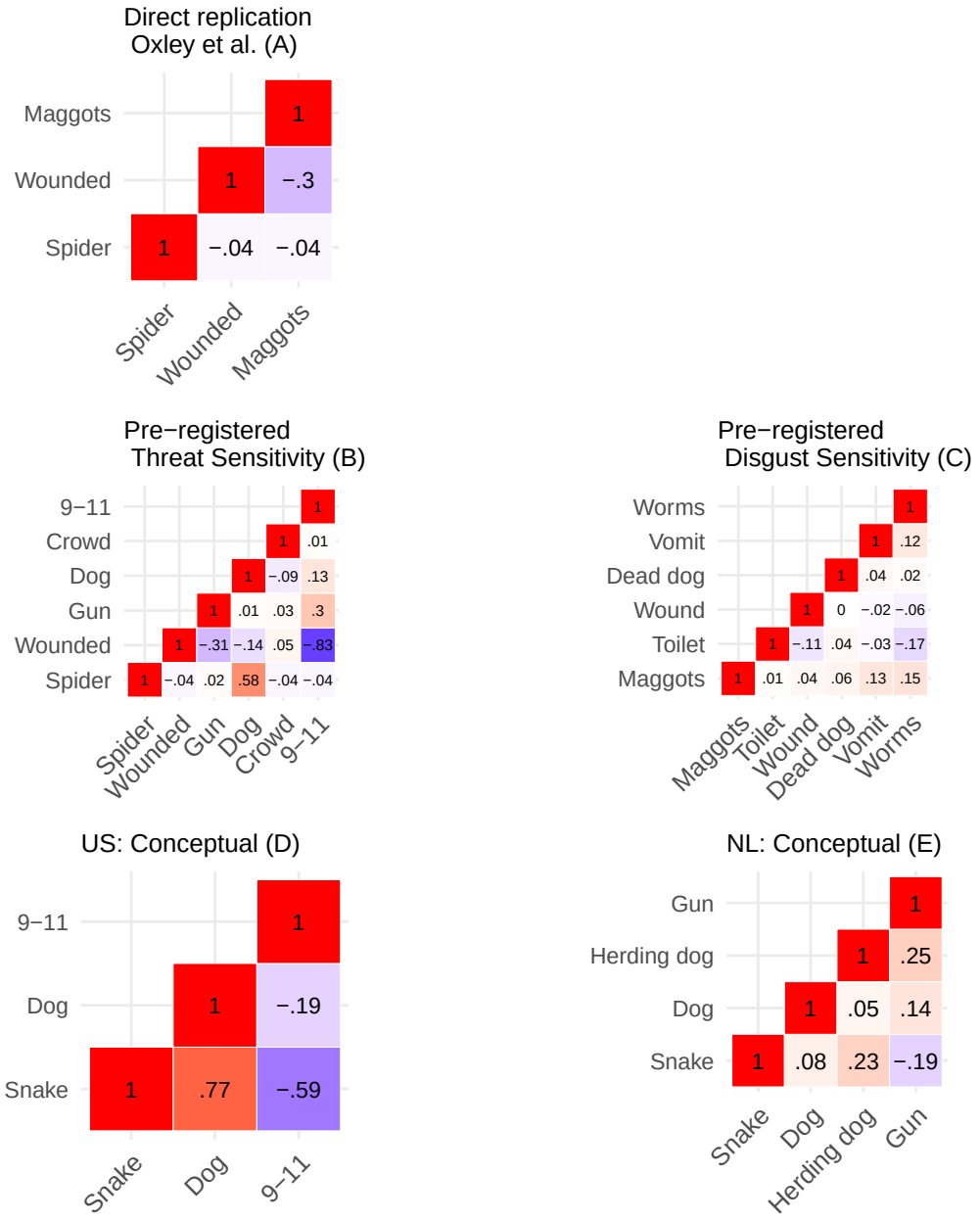

Figure 1. Assessment of a latent threat sensitivity dimension. Correlation matrices with the Pearson correlation coefficients between the physiological responses (skin conductance response) to the threatening images in the pre-registered replication of Oxley et al. (panel A, N=202), the pre-registered extensions for threat sensitivity (panel B, N=202) and disgust sensitivity (panel C, N=202) as well as the conceptual replications in the U.S. (panel D, N=352) and the Netherlands (panel E, N=81). Darker red background means that the correlation is strongly positive, darker blue strongly negative and white means that the correlation is close to zero. Frequentist inferential statistics are reported in full in the Supplementary Table 9 (Panel A), Supplementary Table 10 (Panel B), Supplementary Table 11 (Panel C), Supplementary Table 12 (Panel D) and Supplementary Table 13 (Panel E).

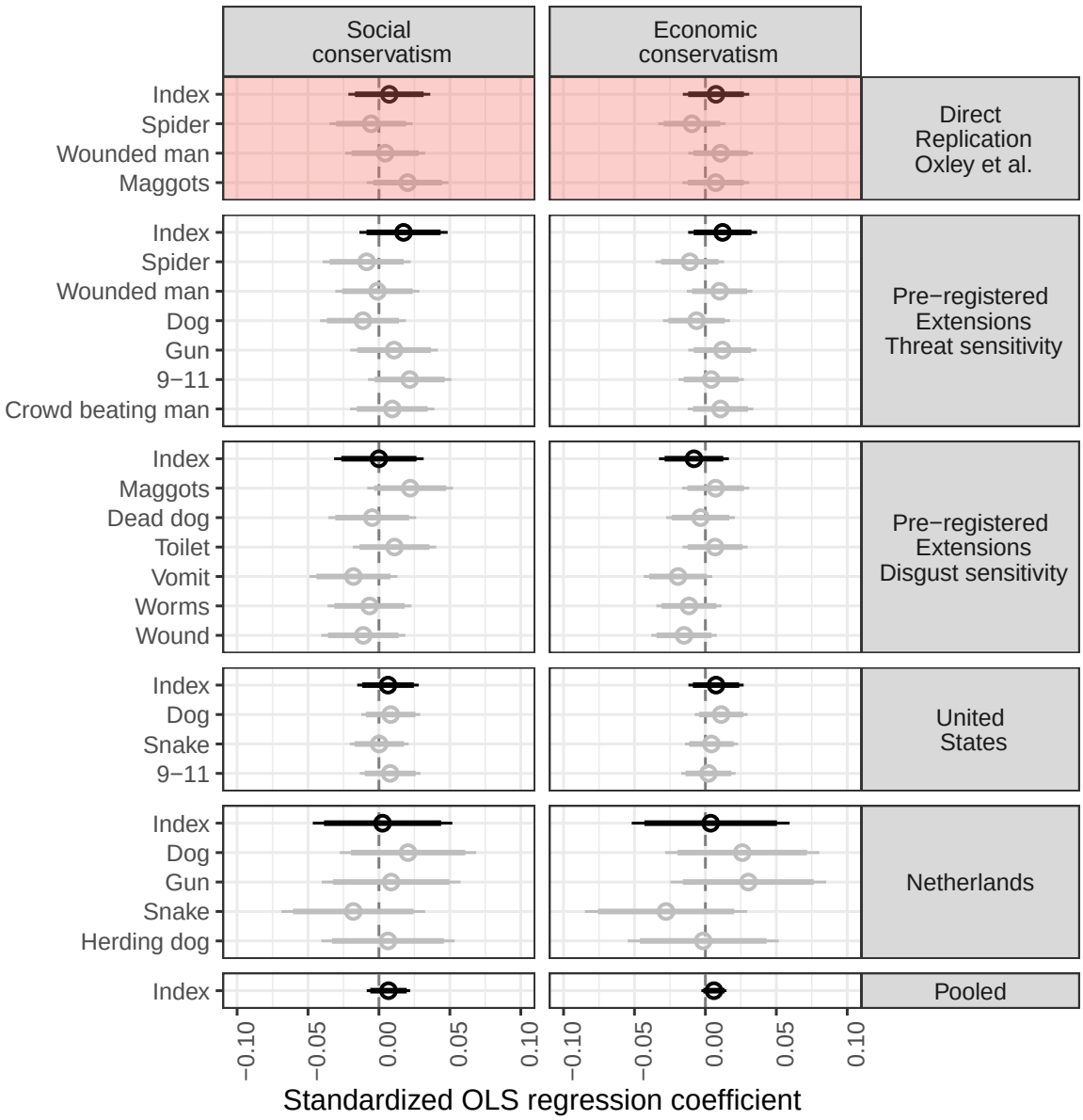

Figure 2. Associations between threat sensitivity and social and economic conservatism. Plot of the standardized OLS regression coefficients of the models where social conservatism (left-hand panel) and economic conservatism (right-hand panel) are regressed on threat sensitivity controlling for the covariates that Oxley et al. used. The dot is the point estimate with 90% (thick) and 95% (thin) confidence intervals. The results for the composite index are provided in black and those for the individual items in grey. The results from the pre-registered direct replication (N=202) are provided in row 1 (shaded), this is followed by the pre-registered extensions for threat sensitivity (row 2, N=202), the pre-registered extensions for disgust sensitivity (row 3, N=202) and the conceptual replications in the US (row 4, N=352) and the Netherlands (row 5, N=81) and finally row 6 contains the results from the pooled analyses (N=635). Regression output with all frequentist inferential statistics and the samples size per model can be found in Supplementary Results: Direct replication Oxley et al. (Supplementary Table 14 & Table 15), pre-registered extensions threat sensitivity (Supplementary Table 16 - Table 18), pre-registered extensions disgust sensitivity (Supplementary Table 19 - Table 21), United States (Supplementary Table 22 & Table 23), Netherlands (Supplementary Table 24 & Table 25) and the pooled analyses (Table 26).

Tables

Table 1. Bayes Factors in Direction of Null Hypothesis for Each Study

Study	Social Conservatism	Bayes Factor Interpretation	Economic Conservatism	Bayes Factor Interpretation
Direct Replication	10.77	Strong	10.52	Strong
Extension: Threat Sensitivity	4.27	Moderate	3.43	Moderate
Extension: Disgust Sensitivity	11.03	Strong	11.02	Strong
United States	12.39	Strong	7.63	Moderate
Netherlands	6.00	Moderate	6.69	Moderate
Pooled	8.57	Moderate	13.99	Strong

Notes: Bayes factor interpretation according to Jeffreys[35]. 95% credibility intervals for effect sizes and additional sensitivity analyses are provided in Supplementary Table 27.

Table 2. Summary of sixteen robustness checks

#	Question	Answer	Reported
1	Do our samples have a sufficient number of conservatives?	Yes	SR
2	Do we find associations between self-reported psychological traits and social conservatism?	Yes	SR
3	Are negative images more physiologically arousing in general?	Yes	SR
4	Does the operationalization of the physiological responses condition the results?	No	Replication
5	Are the results conditional on the policy attitude?	No	SR
6	Do we replicate Oxley et al.'s results using a median split of ideology	No	SR
7	Are other measures of (social) conservatism associated threat sensitivity?	No	Replication
8	Does controlling for physiological reactions to non-threatening images matter?	No	SR
9	Are the results conditional on sophistication?	No	SR
10	Are the results conditional on race, age or gender?	No	Replication
11	Does drawing a distribution of social conservatism similar to Oxley et al. change our findings?	No	SR
12	Are physiological reactions to threatening images associated with attitude extremity?	No	Replication
13	Is variance in physiological reactions to threatening images associated with conservatism or attitude extremity?	No	Replication
14	Do alternative physiological indicators of reactions to threatening images correlate with conservatism?	No	SR
15	Do different physiological measures form an overarching threat sensitivity dimension?	No	Replication
16	Do we replicate Dodd et al. 2012 (Study 1)?	No	Replication

Notes: The Table provide an overview of the sixteen robustness checks and their answers. As well as an indication whether the results can be found in the Supplementary Results (SR) or can be derived from replication files (Replication) which can be found on our public OSF page: https://osf.io/d5g72/.