Appendix to: "**Hot politics? Affective responses to political rhetoric**"

Contents

A. Methods

A.1 Sample size

We rely upon a study with 397 respondents. This is a relatively large sample size compared to the state-of-the-art – see for instance: Mutz and Reeves (2005, N=16), Bucy and Bradley (2004, N=41) Aarøe, Petersen, and Arceneaux (2017, N=43), Oxley et al. (2008, N=48), Soroka and McAdams (2015, N=63), Renshon, Lee, and Tingley (2015, N=128), Mutz (2007, N=155), Arceneaux, Dunaway, and Soroka (2018, N=173), Coe, Canelo, Vue, Hibbing, and Nicholson (2017, N=182), Peterson, Jacobs, Hibbing, and Smith (2018, N=227) and Soroka, Fournier, and Nir (2019, N=1156). The median sample size in these 11 studies is 128, so our sample size is more than 3 times larger. These studies do typically report small ($\hat{f2}$=0.02) effect sizes. To achieve a power of .8 this translates to respectively 395 respondents. Therefore we initially opted for a much larger sample size (Button et al., 2013).[1]

Achieved power in our study. After conducting initial regression analyses and power tests we considered the possibility of pooling all data in a single regression analysis. In the pooled analyses the N increases to 1,032 for the corrugator responses, 429 for the zygomaticus responses, 961 for the skin conductance recordings and 1,011 for heart rate variability (see next section for an explanation of these different samples sizes). To assess power we first take the multiple regression model with corrugator activity as dependent variable, as reported in figure 4 (left panel) of the main paper. We choose this specific model, because in this model we found a significant effect of incongruence, and therewith it is central to our paper. To calculate power we use G*Power 3.1.9.4 and selected the Linear Multiple Regression: Fixed Model R^2 increase option, with a post-hoc achieved power F-test. The increase in R^2 between a model with incongruence (R^2=0.14) and without the incongruence variable (R^2=0.13) is 0.0116. We have 18 predictor variables (counting different levels within a categorical variable as separate predictors) and have set alpha to 0.05. This produces a β of 0.96. Therefore, the pooled analysis of the corrugator is sufficiently powered. Reducing the number of observations of course reduces power. If we split our analyses in half, power is reduced to 0.74. If we reduce the sample to 260 (the maximum number of respondents per issue) power is reduced to 0.45. For this reason, we are reluctant with subsetting the data to specific treatments or specific values of independent variables.

The analyses of the physiological measures have different numbers of observations. Each of the 397 respondents saw in total 3 videos. This amounts to 1,191 observations. In our analyses of negative valence the number of observations is 1,032. In a number of cases electrodes fell off during the experiment or were improperly placed and therefore gave a highly disturbed signal. These cases were excluded from the analysis. For positive valence the number of observations is only 429. This is because we only collected this data in protocol 1 and not in protocol 2. Zygomaticus data is difficult to collect (see Appendix A.10) and numerous times electrodes fell off. Together with saving time, this was a reason to no longer collect it in protocol 2. For the skin conductance data also in some cases the adhesive tape let loose during the experiment, and we had to omit these observations. Also, we removed cases where people in the baseline condition had a raw skin conductance level of less than 2 microSiemens. This general procedure is applied, because the responsiveness of these people is too low to capture. Finally, for heart rate variability the number of observations is 1,011. The dropout here is explained by stickers going loose during the experiment, heart rate signals with extreme deviations, and in some cases heart rate was not collected because it was difficult given the respondents' clothing. We do not

[1] The only notable exception is Soroka et al. (2019) which relies, in total, upon 1,156 respondents but in most of their analyses the 17 countries are analyzed separately and in these instances the sample size drops to less then 100 observations per country – see Figure S.1 in the supplementary materials of Soroka et al. (2019).

expect this dropout to affect our findings, because malfunctions should be random.

A.2 Ethics and procedure

Ethical approval for the study was obtained from the Ethics Review Board of our university; Protocol 1 #2016-PCJ-7033 (lab-in-the-field) and #2016-PCJ-7274 (laboratory) and protocol 2 #2017-PCJ-7917. Participants in the laboratory study were paid minimum wage (7 Euro and 50 cents) in return for their participation. The total protocol lasted 45 minutes – after the part of the study that we report here, participants completed a few other tasks. Participants in the lab-in-the-field version of our protocol received a polaroid picture of themselves in the laboratory. The photo was taken *after* the experiment. The organizers of the events did not allow us to compensate participants for their participation.

Before the experiment started participants signed an informed consent form and filled out a survey using Qualtrics on a stationary computer (in case of the laboratory setting) or an Ipad 2 (in case of the lab-in-the-field setting). The interface is the same. Relying upon the Ipad allowed us to collect survey responses without an internet connection. The survey measured their political attitudes. After completing the survey the respondents were connected to the physiological measurement equipment.

A.3 Descriptive Data

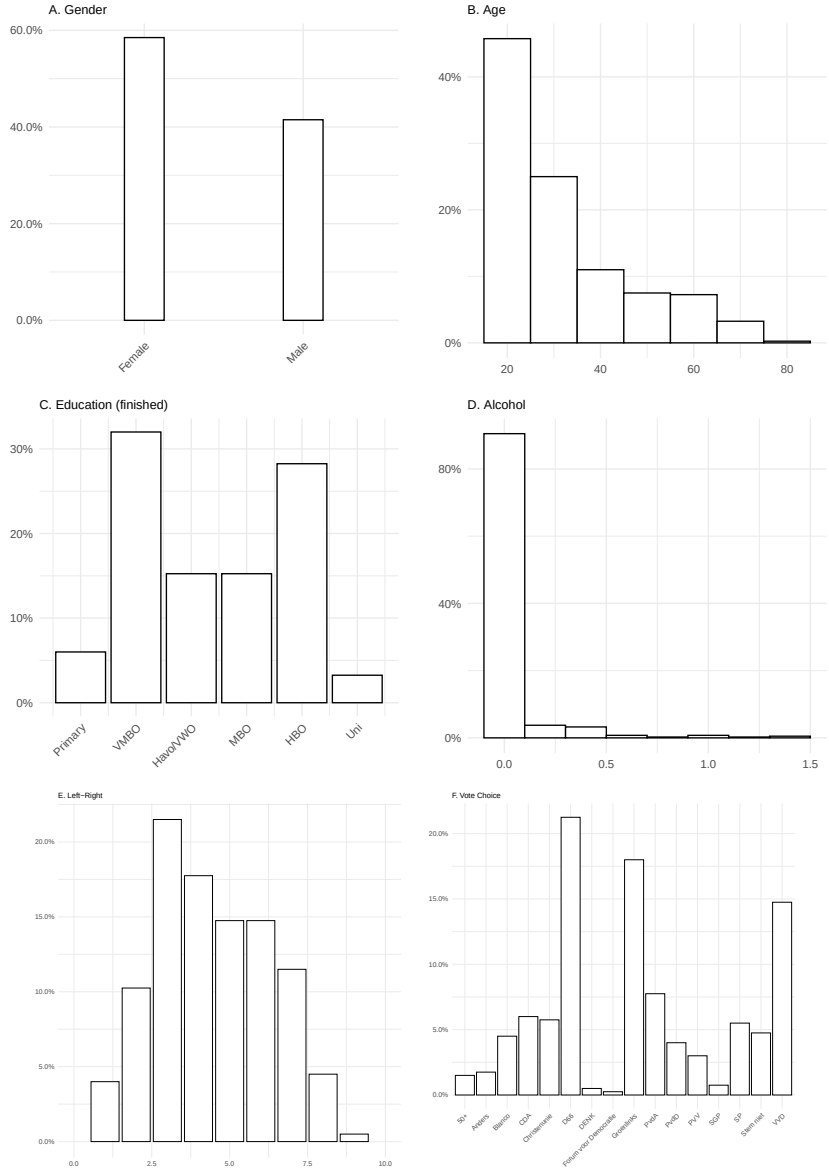
The descriptive data presented in Figure A1 is based on the 400 respondents who filled out the survey. Three of these dropped out after the survey and did not participate in the experiment. Two people dropped out because they had to wait too long to start the experiment. One person fainted while she was being set up to do the experiment. She did not complete the study.

Our sample consists for 59% (N=234) of woman (see panel A, Figure A1), the median age in the sample 26 (see panel B of Figure A1); Mean=32.26, SD=13.95, Min=16, Max=76). Our sample has a lot variance on the completed level of education (see panel C of Figure A1) as the completed educational level ranges from the lowest (primary) to highest (university). We collected our data at a couple of public events (festival, fair, etc.). In panel D of Figure A1 we show that approximately 80% of our sample did not consume alcohol. Very few people drank more than the equivalent of 3 beers (1 promille of alcohol; M=.04, SD=.18, Min=0, Max=1.5). We also measured ideology and vote choice. Ideology was measured using a commonly used ideological self-placement scale. Participants were asked to indicate on a scale from 0 (left) to 10 (right) where they place themselves. Our sample is a bit more left-leaning (M=4.46, SD=1.88, Min=0, Max=10; see Figure A1). This is also reflected in the vote intention at the time of the study: a sizeable group of participants in our study expresses the intention to vote for parties such as D66 and the GroenLinks – 39% of the participants indicates to vote for either of these parties. However, and important for our study, we see enough variance in ideology and vote choice.

To summarize, compared to the normal population our sample is slightly balanced towards female participants (panel A), young people (panel B) and somewhat more left-leaning (panel E). Our sample is thus not per se a rand draw from the general (Dutch) population. It is important to note that our sample is more diverse than a general student population. It contains more right-wing, lower educated and older people than a student sample. This is important as we are interested in how individual differences condition the extent to which politics is hot (see, Druckman & Kam, 2011).

Finally, table A1 presents the descriptive statistics of the main variables in the pooled model. The number of observations increase approximately fourfold, because each respondent

Figure A1. Descriptive data



is included in the pooled model with the 4 videos each respondent saw. However, the number is somewhat lower than a factor 4, because some physiological measurements were too noisy and therefore these observations were dropped.

Table A1
Descriptive statistics main variables

Statistic	N	Mean	St. Dev.	Min	Max
Attitude Extremity	1,032	1.623	1.087	0.000	4.000
Political Knowledge	1,032	1.734	0.819	0	3
Incongruence	1,032	0.064	1.953	−4.000	4.000
Arousal: SCL	956	0.025	0.180	−1.254	1.639
Negative Valence: Corrugator	1,032	1.255	1.997	0	8
Positive Valence: Zygomaticus	429	1.041	1.775	0.000	7.066
Attention: RMSSD	1,006	−6.448	20.747	−147.022	81.792
Position Change	1,032	0.045	0.903	−4.000	4.667

A.4 Lab in the field studies compared to laboratory experiments

This study consisted of two protocols as we explained in the main text of the paper. Protocol 1 was conducted as a laboratory-in-the-field study during a cultural festival and a laboratory study (see Table A2), while protocol 2 was conducted as a laboratory-in-the-field experiment in the spring and summer of 2017. Here we show the importance of conducting laboratory-in-the-field studies for creating more diverse populations in physiological (laboratory) experiments.

Table A2
Study overview

Study	Year	Location(s)	N
Lab-in-the-field	August 2016	Cultural festival	143
Laboratory	November 2016	University laboratory	120
Lab-in-the-field	Spring-summer 2017	Media museum; Evangelical Rock festival; Fair; Biker event	137

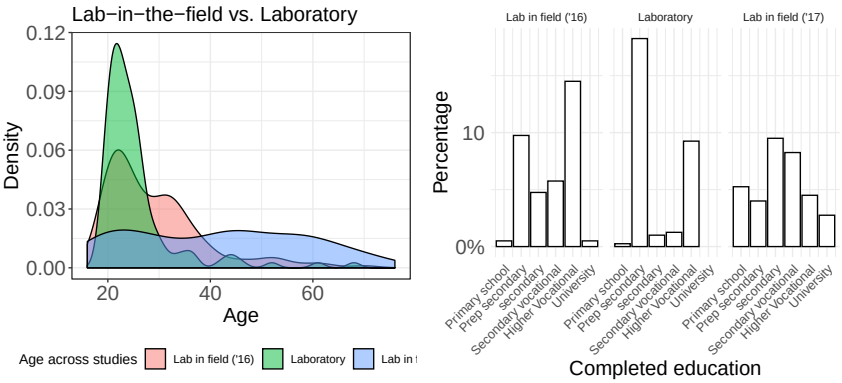
Background: gender, age and education. Starting with gender, we see a striking difference between our laboratory study and our lab-in-the-field studies. In the laboratory setting 70.8% of the respondents is woman, while in the lab-in-the-field we find that 53% of the respondents is woman. Gender is thus relatively unbalanced in the laboratory experiment – most likely due to an imbalance in the student population at our social science faculty at a large Dutch university. We show that in the data we collected using the lab-in-the-field, we find a roughly equal gender distribution

Perhaps not surprising we find more variance in age in the lab-in-the-field compared to the laboratory as can be gleaned from the density plot in Figure A2. Specifically, in the laboratory the mean age is ten years younger compared (M=25.4, SD=7.71, Min=19, Max=68) compared to the lab-in-the-field in 2016 (M=29.1, SD=9.75, Min=19, Max=70) and 2017 (M=41.8, SD=16.8, Min=16, Max=76)

The laboratory study was conduct primarily among university students. This is also reflected in the highest completed educational level (see right panel of Figure A2). The large majority of the sample has completed preparatory secondary education which is the entry level into the university. The second largest group has completed higher vocation education (HBO – University of Applied Sciences) which is another major route through which Dutch students enter the university. If we look at the laboratory in the field data collected in 2016 (left-hand

panel) and 2017 (right-hand panel), we see that we get a much more diverse sample in the lab-in-the-field setting. Other – lower – educational levels are also common in this sample.

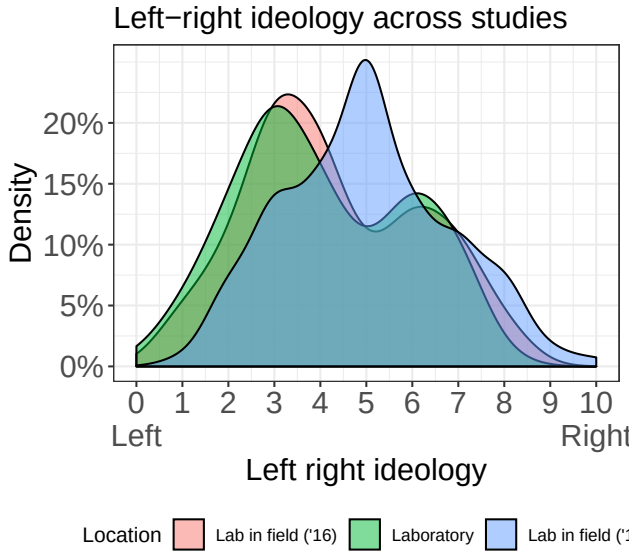
Figure A2. Lab vs. Lab-in-the-field: Age and education



Politics: ideology, sophistication. The sample we recruited in the laboratory is more left-wing (M=4.06, SD=1.85, Min=0, Max=8) compared to the laboratory-in-the-field in 2016 (M=4.28, SD=1.87, Min=1, Max=8) and 2017 (M=4.99, SD=1.8, Min=1, Max=10). This can also be seen in Figure A3.

The people participating in the laboratory-in-the-field studies in 2016 (M=.60, SD=.25, Min=0, Max=1) and 2017 (M=.61, SD=.28, Min=0, Max=1) score somewhat higher on political knowledge compared to people who participated in the laboratory study (M=.53, SD=.30, Min=0, Max=1). The interest in politics also differs between the different settings: in the laboratory-in-the-field study during the cultural festival people seemed somewhat more interested in politics (M=.57, SD=.27, Min=0, Max=1) compared to the laboratory setting (M=.50, SD=.27, Min=0, Max=1) and the laboratory-in-the-field study in 2017 (M=.52, SD=.27, Min=0, Max=1). But again, there is a lot of variance on the knowledge and interest questions across the samples.

Figure A3. Lab vs. Lab-in-the-field: Left-right ideological self-placement

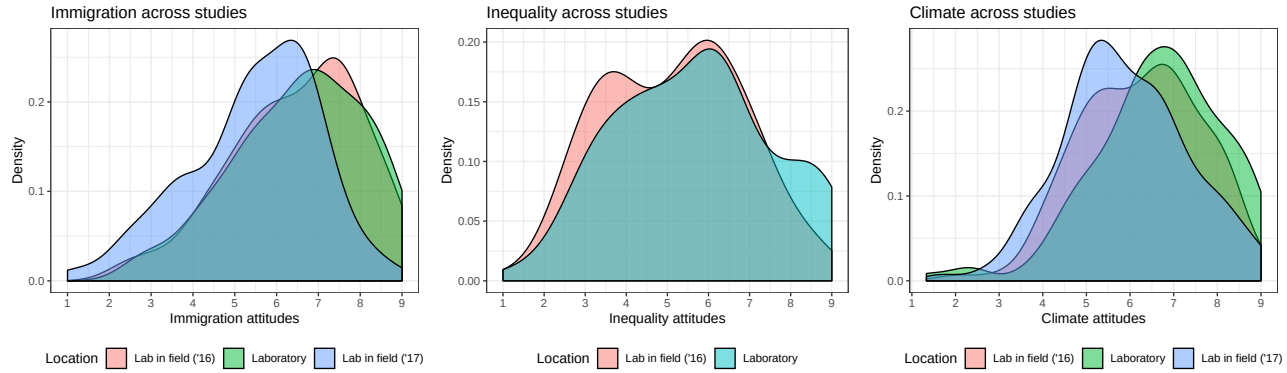


chch

Turning to the attitudes used to measure the attitude (in)congruence, we see in Figure A4 that immigration attitudes are somewhat more positive in the laboratory study (M=6.53,

SD=1.56, Min=2.67, Max=9). and the laboratory-in-the-field study in 2016 (M=6.46, SD=1.54, Min=2.33, Max=9). compared to the 2017 laboratory-in-the-field study (M=5.47, SD=1.56, Min=1, Max=9). The redistribution messages were only shown in the laboratory study and the 2016 laboratory-in-the-field study. As can be gleaned from the second panel of Figure A4 the people in the laboratory-in-the-field study were somewhat less supportive of redistribution (M=5.15, SD=1.73, Min=1, Max=9) compared to people in the laboratory-in-the-field setting (M=5.65, SD=1.90, Min=1, Max=9). Finally, the attitudes towards climate are more negative in the 2017 laboratory-in-the-field study (M=5.90, SD=1.42, Min=1.67, Max=9) compared to the laboratory experiment (M=6.65, SD=1.48, Min=1.33, Max=9) and the laboratory-in-the-field study in 2016 (M=6.26, SD=1.36, Min=2, Max=9). We do not show data for opinions on Europe here, because it was fielded only in the third study.

Figure A4. Descriptive data per sample



A.5 Measures collected but not used in this study

Survey. In all surveys we included a six item measure of Agreeableness (Bakker & Lelkes, 2018). In the lab-in-the-field settings (Cultural Festival, Evangelical Rock Festival, Fair, Biker Event, Media Museum) we did not collect any additional survey measures. In the protocol in the university laboratory we also measured social dominance orientation (Ho et al., 2012) and the Social Principles Index (Smith, Oxley, Hibbing, Alford, & Hibbing, 2011). These were asked at the end of the survey and used in other projects.

Physiological measures. Aside from the populist messages – used in another project – we showed people sometimes some additional messages. This always happened *after* the exposure to the treatments reported in this paper. In the Laboratory setting we showed participants four pictures that were used to measure threat. In some of the laboratory-in-the-field settings (Evangelical Rock Festival, Fair, Biker Event, Media Museum), we showed participants a series of images from politicians and other humans and objects (10 in total). The physiological responses to these images are part of other projects and are not reported here.

A.6 Treatments

The full texts for the immigration (Table A3), inequality (Table A5), EU (Table A6) and Climate (Table A7) messages are provided below. Due to copyright regulations, we cannot make the actual video's publicly available but they are available upon request from the authors'.

Table A3
Treatment messages: Immigration protocol 1

Treatment	Text
Pro-immigration	"Every day Syrian refugees drown in the Mediterranean Sea. Providing shelters in the region is not an option. There is the threat of civil war in Turkey and Lebanon is overcrowded. Therefore, this year again a large number of refugees will come to the Netherlands. The Netherlands must be a safe haven for people fleeing from war and violence. Don't stop refugees, but take them in!"
Anti-immigration	"The reception of Syrian refugees in our country is absurd. The Netherlands can no longer cope with the enduring immigration - and all the problems caused by it. The damage caused here by all those refugees is incalculable. Our security, freedom, our culture, our money and our future are at stake. The Islamic invasion, because that is what it is, is getting out of hand. It is enough. The borders must be closed."

Table A4
Treatment messages: Immigration protocol 2

Treatment	Text
Pro-immigration	"Every day refugees drown in the Mediterranean Sea. Providing shelters in the region is not an option. The middle east is on fire and countries in the region are overcrowded. Therefore, this year again a large number of refugees will come to the Netherlands. But we can deal with that in the Netherlands. The Netherlands must be a safe haven for people fleeing from war and violence. Don't stop refugees, but take them in!"
Anti-immigration	"The Netherlands is full. But the borders are wide open. Reception of refugees in our country is absurd. The Netherlands can no longer cope with the enduring immigration. The damage caused here by all those refugees is incalculable. Our security, freedom, our culture and our future are at stake. The Islamic invasion, because that is what it is, is getting out of hand. The borders must be closed."

Table A5
Treatment messages: Redistribution

Treatment	Text
Pro-redistribution	"The gap between rich and poor in the Netherlands is enormous. Almost one and a half million people live below the poverty line. 400,000 children grow up in poverty. The lowest incomes have seen their income fall by no less than 30% since the 1970s. But the number of millionaires is rising explosively. The increase in their prosperity is at the expense of the 99% who does not swim in money. It is high time that the situation of people on lower incomes improves. The government must tax the rich more heavily."
Anti-redistribution	"The Netherlands is a rich country. Everyone has enough to live from. Of course, not everyone is equally rich. It is the free market that determines which income fits which profession. Some people have a profession with a high income, others a profession with less income. But everyone gets the salary he or she deserves. Reducing income differences by raising extra taxes on the rich is theft. Hardworking people should never be the victim of an overpowering government."

Table A6
Treatment messages: EU

Treatment	Text
Pro-EU	"The EU is the best thing that has happened to Europe since the horrors of the Second World War. Europe brought peace, security and prosperity. Even in the 21st century, the Netherlands cannot solve major global problems on its own. Climate change and terrorism do not stop at the frontiers. These problems cannot be solved by sticking our heads in the sand. We solve these problems through Europe. Together, Europe is strong. Therefore, more Europe!."
Anti-EU	"The Netherlands has given far too much power to Brussels. Forever, if we don't wake up very quickly. Brussels is not the boss. We are the boss. About our country, about our money, and about our own Dutch traditions. It is time for the liberation of the Netherlands from the clutches of the EU superstate. It is time to take back control."

Table A7
Treatment messages: Climate

Treatment	Text
Pro-climate	"Time is running out for our earth. If we continue as if nothing is wrong, temperatures will rise by six degrees. The climate is on tilt. Heat waves alternate periods with heavy downpours and floods. Rising sea levels are a direct threat to us and our children. We cannot let this happen. We must stop using fossil fuels as soon as possible. The Netherlands must now invest and switch to sustainable forms of energy, such as solar- and wind energy. Only in this way can we save our planet. There is not a second to lose. "
Anti-climate	"Since 1989, the government has wasted twelve billion Euros on climate policy. We know very little about the factors that influence our climate. It is therefore actually impossible to predict the climate. If we manage to make a forecast at all, we should first try it out for about two hundred years. Investing billions in expensive CO2 reduction, wind turbines, sustainability projects and a lot of green socialism to solve a barely defined problem, is an appalling human arrogance. No more penny for all the nonsensical climate policy."

Table A8
Treatments: pilot test

	Left-wing	Right-wing	t-test
Immigration	1.08 (.90)	9.17 (.83)	t=31.28 (p<.001)
Redistribution	2.04 (2.03)	7.61 (1.41)	t=9.14 (p<.001)
Climate	1.61 (1.41)	8.17 (1.97)	t=11.21 (p<.001)

Pilot test results (N=23) with evaluation of the ideological slant of the message on a scale ranging from left (0) to right (10). Means and standard deviations (in parentheses) are reported.

Pre-test of treatments. We asked 23 students – blind to the expectations of our study – to evaluate the ideological slant of the messages that were part of protocol 1 on a scale from left-wing (0) to right-wing (10). Note that the poles of the scales were adjusted to the issue. For instance, for immigration we asked whether the message was "pro-immigration" (0, i.e., left-wing) or "anti-immigration" (1, i.e., right-wing). As can be seen in Table A8, left-wing messages were evaluated as left-wing for the immigration, redistribution and climate messages, as can be seen by the means that are close to 0. Right-wing messages were evaluated as right-wing as the means were close to 10. There was not much variance in the evaluations of the messages as the standard deviations were equal or lower then 2.

Paired sample t-tests confirm that the evaluation of the messages significantly differ from each other (see third column of Table A8).

A.7 Survey measures

Issue attitudes. Attitudes towards immigrants were measured using the items: "The Netherlands should allow more refugees into the country", "Refugees are a threat to the security of the Dutch population" (reverse coded), "The Dutch culture is threatened by refugees" (reverse coded). We recoded the reverse coded items and created a scale ranging from anti-immigration to pro-immigration (M=6.15, SD=1.62, Min=1, Max=9, α=.80). Attitudes towards redistribution were measured using two items: "Income differences should become smaller" and (2) "Income inequality is too big in the Netherlands. People with the lowest incomes should get the biggest salary increase". Again, we created a scale from anti-equality to pro-equality (M=5.38, SD=1.82, Min=1, Max=9, α=.85). EU attitudes were measured using the items (1) "The European Union should become one", (2) "I am in favour of efforts being made to unify Europe" and (3) "I support efforts to make the European Union one". We created a scale to range from anti-Europe to pro-Europe (M=4.12, SD=1.64, Min=1, Max=9, α=.80). Attitudes towards climate were measured using three items: 'Most claims about environmental problems are exaggerated" (reverse coded), (2) "There are more important things in life than protecting the environment" (reverse coded) and (3) "Protecting the environment is more important than economic growth". We created a scale from anti-climate to pro-climate (M=6.25, SD=1.45, Min=1.33, Max=9, α=.68).

Political knowledge. The political knowledge questions consists of three questions with 5 answer categories and a separate "don't know" category, namely: (2) "Which party has at the moment the most seats in the Lower Chamber?" (correct answer = VVD) and (3) "Who has been the chairman of the European Parliament in the last two years?" (correct answer = Martin Schulz). we have summed the number of correct answers (M=.58, SD=.27, α=.39).

A.8 Measuring skin conductance

We measure SCL using two Ag/AgCl electrodes of 20mm by 16mm. The electrodes were attached with adhesive tape on the medial phalanges surfaces of the index and ring finger of the non-dominant hand (Dawson, Schell, & Filion, 2007, p. 163) [2]. We use the non-dominant hand because this leaves the dominant hand free during the study to answer the survey items during the protocol. There is no consistent evidence that EDA "recorded from one hand gives consistently different results with respect to the effects of experimental variables than that recorded from the other hand (Dawson et al., 2007, p. 163). Following Dawson et al. (2007) we do not clean the skin before placing the electrodes.

A.9 Measuring Heart rate

Heart rate was recorded using three Ag-AgCl electrodes. We fixed the electrodes using adhesive patches (3M Red Dot Electrodes with Micropore Tape 2239). We placed one electrode on the left side of the chest below the sixth rib. A second electrode was placed below the right clavicle. The ground electrode was placed on the left clavicle.

Anomalies were identified by unrealistically strong deviations in heartbeat from the mean. For three respondents the raw heartbeat data was of such poor quality that we dropped them from the analyses. For three individuals heartbeat was registered by placing electrodes in the neck, rather than around the collarbone. This is because their clothing did not allow placement of the electrode below the 6th rib. There is no significant difference between the heartbeat of these three individuals and the remainder of the sample.

A.10 Measuring valence

Corrugator. To measure the experience of negative affect, we measured the activity of the corrugator major. We do this using two 7 mm Ag/AgCl mini-electrodes that we filled with with electrolyte gel (Signa, Parker Laboratories). Using double stick adhesive color we placed the two electrodes just above the eye-brow – at the place where the muscle is located (Fridlund & Cacioppo, 1986). A third electrode was placed on the side of the forehead and serves as a the ground measure.

Zygomaticus. The electrodes that we use to measure zygomaticus activity are similar to the one's we used to measure corrugator activity. But in this case we place the electrodes on the cheek, and specifically on the place where the zygomaticus is located. Do note that we collected this measure only among a sub-sample of 143 participants at one of the lab-in-the-field locations.

A.11 Correlations between relevant variables

Table A9 presents the correlations between the attitudes in the survey. Most correlations are weak ($r<.2$) or expected (e.g. the correlations between knowledge, age and education). The correlation between how salient people find the issues and how extreme their attitudes is quite low given how close these concepts are ($r=0.29$).

Table A10 presents the correlations between the physiological measures and the self-reported emotions. Surprisingly, there is quite a high correlation between zygomaticus and

[2] The use of the medial phalanges differs from the recommendation by Dawson et al. (2007) to use the distal phalanges. We do this because the size of the electrodes made it easier to fit them on the medial phalanges – irrespective of the size of the fingers of the respondent.Dawson et al. (2007, p. 163) point out that the medial phalanges has *lower* numbers of active sweat glands compared to the distal phalanges (see also, Freedman et al., 1994). As such our results for skin conductance could be considered a conservative test.

Table A9
Correlations attitudes measured in survey

	Alcohol	Education	Age	Political Knowledge	Female	Incongruence	Salience	Extremity
Alcohol	1	0.001	0.142	-0.007	-0.060	0.041	-0.061	0.021
Education		1	0.120	0.206	-0.052	0.021	0.125	0.070
Age			1	0.207	-0.131	0.008	-0.005	0.011
Political Knowledge				1	-0.216	0.056	0.064	-0.018
Female					1	0.001	0.100	0.012
Incongruence						1	0.076	0.031
Salience							1	0.293
Extremity								1

corrugator activity (r=0.4). This may be due to cross-talk or people actually are experiencing several different emotions during exposure to a treatment. The correlations between the physiological measures and the self-reported emotions are on average very low. Only heart rate variability correlates weakly with enthusiasm (r=0.19) and anger (r=-0.14). Skin conductance and heart rate variability do not correlate with the other physiological measures. This is line with the circumplex model of emotions in which measures of arousal (or attention) form a separate dimension from measures of valence such as the corrugator and zygomaticus.

Table A10
Correlations emotion measures

	Corrugator	Zygomaticus	Enthusiasm	Anger	Anxiety	Skin conductance	Heart rate variability
Corrugator	1	0.405	-0.068	0.050	-0.111	0.048	-0.050
Zygomaticus		1	-0.015	-0.055	-0.125	0.005	-0.002
Enthusiasm			1	-0.472	-0.146	-0.058	0.192
Anger				1	0.332	0.001	-0.141
Anxiety					1	-0.059	0.056
Skin conductance						1	0.061
Heart rate variability							1

B. Results

Full regression output Figure 3

Table B1
Full OLS output figure 3

	Full	0-10 sec	5-15 sec	10-20 sec
	(1)	(2)	(3)	(4)
Salience	0.009	0.014	0.016	0.018
	(0.035)	(0.034)	(0.034)	(0.034)
Incongruence	0.073**	0.065**	0.069**	0.072**
	(0.034)	(0.033)	(0.033)	(0.034)
Political Knowledge	0.034	0.057	0.048	0.043
	(0.047)	(0.047)	(0.048)	(0.048)
EU	0.566***	0.506***	0.558***	0.554***
	(0.150)	(0.148)	(0.148)	(0.148)
Immigration Pr.1	−0.149***	−0.153***	−0.150***	−0.154***
	(0.056)	(0.054)	(0.053)	(0.053)
Immigration Pr.2	0.621***	0.516***	0.561***	0.573***
	(0.170)	(0.165)	(0.160)	(0.158)
Inequality	0.050	0.035	0.050	0.036
	(0.068)	(0.063)	(0.062)	(0.063)
Contra Treatment	0.033	0.010	0.009	0.013
	(0.061)	(0.059)	(0.058)	(0.059)
2nd treatment	0.136***	0.196***	0.148***	0.124**
	(0.052)	(0.052)	(0.053)	(0.053)
3rd treatment	0.055	0.140**	0.098	0.074
	(0.064)	(0.063)	(0.062)	(0.061)
4th treatment	0.086	0.137*	0.107	0.092
	(0.083)	(0.080)	(0.080)	(0.081)
Temperature	−0.054	−0.051	−0.047	−0.045
	(0.056)	(0.054)	(0.053)	(0.053)
Female	−0.402***	−0.427***	−0.430***	−0.425***
	(0.085)	(0.086)	(0.086)	(0.087)
Age	−0.024	−0.012	−0.019	−0.019
	(0.041)	(0.039)	(0.039)	(0.040)
Education	−0.083*	−0.113**	−0.107**	−0.101**
	(0.047)	(0.049)	(0.049)	(0.049)
Alcohol	0.046	0.045	0.041	0.035
	(0.052)	(0.049)	(0.048)	(0.047)
Computer1	0.192	0.208	0.228*	0.228*
	(0.140)	(0.133)	(0.135)	(0.138)
Computer2	0.394***	0.226	0.310**	0.342**
	(0.142)	(0.148)	(0.150)	(0.150)
Computer3	−0.127	−0.125	−0.116	−0.097
	(0.160)	(0.157)	(0.149)	(0.147)
Constant	−0.128	−0.072	−0.091	−0.089
	(0.136)	(0.138)	(0.138)	(0.138)
Note:				*p<0.1; **p<0.05; ***p<0.01

Full regression output Figure 4

Table B2
Full OLS results figure 4

	Corrugator	Zygomaticus	RMSSD
	(1)	(2)	(3)
Salience	−0.006	0.025	0.029
	(0.034)	(0.060)	(0.052)
Incongruence	0.109***	0.054	0.026
	(0.033)	(0.047)	(0.031)
Political Knowledge	0.038	−0.066	−0.027
	(0.040)	(0.088)	(0.040)
EU	0.045		0.320
	(0.170)		(0.216)
Immigration Pr.1	−0.056	−0.139*	−0.013
	(0.066)	(0.080)	(0.046)
Immigration Pr.2	0.126		0.352*
	(0.172)		(0.208)
Inequality	−0.135**	−0.072	−0.063
	(0.064)	(0.101)	(0.051)
2nd treatment	−0.065	−0.136	−0.065
	(0.070)	(0.107)	(0.061)
3rd treatment	−0.011	−0.040	−0.164***
	(0.067)	(0.106)	(0.062)
4th treatment	−0.118	−0.115	−0.186***
	(0.076)	(0.122)	(0.071)
Temperature	−0.121**	−0.138**	0.173**
	(0.052)	(0.057)	(0.071)
Female	0.024	0.013	0.035
	(0.087)	(0.133)	(0.103)
Age	−0.072	−0.043	−0.00003
	(0.054)	(0.087)	(0.056)
Education	0.014	0.021	0.056
	(0.042)	(0.083)	(0.050)
Alcohol	0.046	0.073	0.090*
	(0.053)	(0.063)	(0.051)
Computer1	0.744***	0.437***	0.080
	(0.131)	(0.129)	(0.173)
Computer2	−0.222*		0.502***
	(0.127)		(0.184)
Computer3	0.126		0.123
	(0.154)		(0.154)
Constant	−0.036	−0.017	−0.189
	(0.115)	(0.145)	(0.168)
Note:			*p<0.1; **p<0.05; ***p<0.01

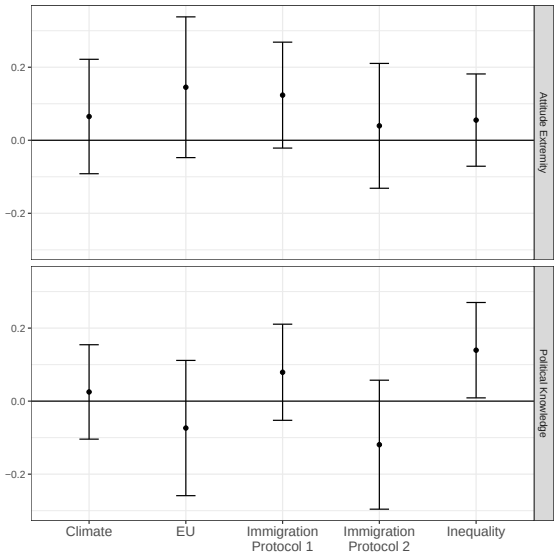
C. Robustness checks

C.1 Effects of attitude extremity and knowledge per treatment

Figure C1 reports the regression coefficients of extremity and knowledge per film. These analyses follow the same procedure as Figure 3 in the main paper. However, here the sample is subset to responses to a single issue (e.g. Climate, EU, Immigration Protocol 1, Immigration Protocol 2, and Inequality). The effects of extremity are positive in all 5 analyses. But the effect is only statistically significant in the Immigration Protocol 1 film.

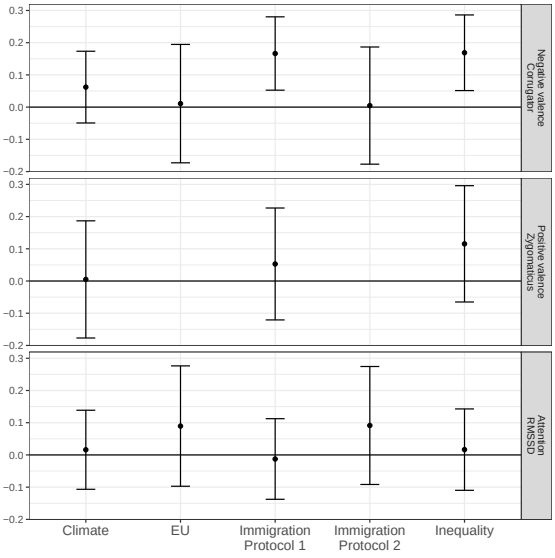
The effects of political knowledge differ in direction depending on the issue. For example, political knowledge has a positive and statistically significant effect for the inequality issue, but a negative effect for the immigration issue (protocol 2).

Figure C1. **Effects Extremity and Knowledge on SCL per film**



We now split the analyses reported in figure 4, so that we obtain effects of incongruence per film. First, Figure C2 shows statistically significant positive effects of incongruence for the Immigration Protocol 1 and Inequality films. For climate there is a positive sign, and for the EU and Immigration Protocol 2 the effect size is effectively zero. Second, moving to the middle panel, we look at the effects of incongruence on the zygomaticus. In line with the analyses reported in the main paper we find no statistically significant effect of incongruence on the zygomaticus per film. Similarly, the analyses in the bottom panel of Figure C2 clearly demonstrate a positive, but insignificant effect of incongruence on attention, measured as heart rate variability.

Figure C2. **Effects Incongruence per film**



C.2 Different operationalizations of EMG measure

There are different ways to operationalize EMG corrugator measures (Tassinary, Cacioppo, & Vanman, 2007). Our focus has been on extracting peaks from the EMG raw data (see main text from procedure). To identify a peak we used a cut-off point of 10 microvolt. However, other cut-off points can be used. Here we evaluate the correlations between our measure, and measures using respectively cut-off points of 5, 4, 3, 2 and 1 microvolt. These are labelled sum.peaks.5, etc. Also, we look at two measures that simply take the sum of the absolute changes from the baselines as a measure of EMG activity. These are labelled changes1 and changes2. Finally, we also compare our measure to the mean EMG activity. Table C1 reports the correlations between these measures and the 10 microvolt peak measure we applied. These correlations range from very high (r=0.82) to high (r=0.62). The different peak measures and the changes measures to large extent measure the same as the measure applied in the main paper. The only big outlier is the correlation with mean EMG. This is not entirely surprising, EMG signals increase and then decrease. This is the propery of interest. If you average these increases and decreases, one basically retrieves a zero-baseline. As such, mean EMG does not really produce a response to a stimulus.

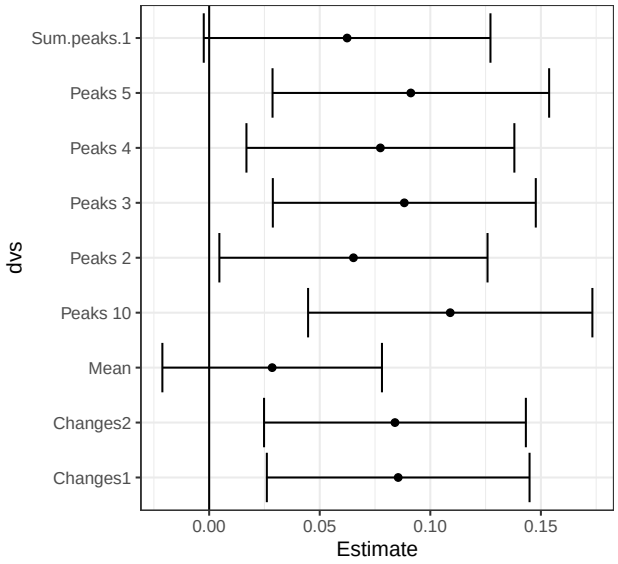
Table C1
Correlations with EMG 10 mv measure

Mean	Changes1	Changes2	Peaks 5	Peaks 4	Peaks 3	Peaks 2	Peaks 1
0.342	0.637	0.616	0.819	0.763	0.708	0.659	0.616

As a next step we reran the analysis presented in Figure 4 of the main paper, using different EMG measures as dependent variable. Figure C3 presents these findings. Specifically, it reports the effect of incongruence on each EMG measure. Six out of 8 EMG measures produce an effect that is statistically indistinguishable from the ”peaks 10” analysis presented in the paper. The ”peaks 1” measure produces a somewhat lower estimate of incongruence, but still statistically significant at p<.1. The mean EMG measure is not significantly related to the incongruence measure. However, this is to be expected per our discussion before. In sum, our conclusion is

that regardless different options for the operationalization of the EMG corrugator measure, the results are stable. These conclusions also apply to the different operationalizations of the EMG zygomaticus measure. These results are available upon request.

Figure C3. **Effects Incongruence for different EMG measures**



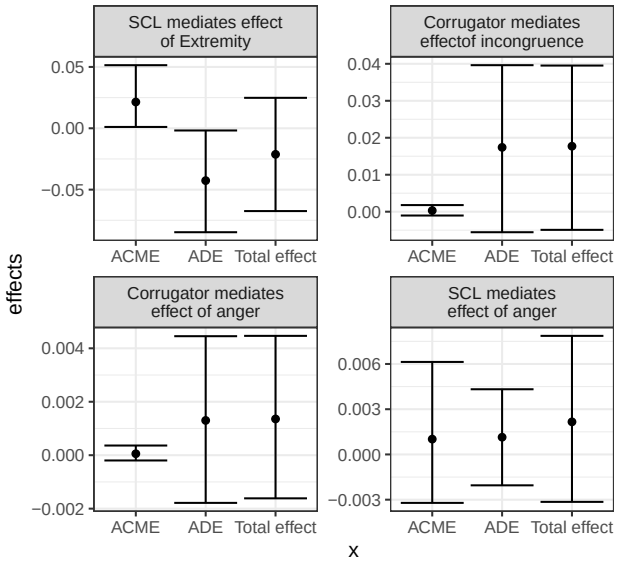
C.3 Results from mediation analyses

We are interested in how attitudes, self-reported emotions and physiological measures influence position change on the issues individuals are exposed to. In addition to studying the main effects, as discussed in the paper, we also evaluated the existence of mediation effects. In particular we considered four different mediation paths: 1) extremity as treatment and skin conductance as mediator, 2) incongruence as treatment and corrugator as mediator, 3) anger as treatment and corrugator as mediator, 4) anger as treatment and skin conductance as mediator. To evaluate mediation effects we follow the procedure prescribed by Imai, Keele, and Tingley (2010). Since the main text reports the direct effects we report here only the mediation model output. Do note that in the first two cases we use the entire available sample, whereas in the last two cases we use responses to the immigration treatment (protocol 1) only, because that's they only part where self-reported emotions were asked.

Figure C4 reports the results from these mediation analyses. In particular here we look at the average causal mediation effect (acme), the average direct effect of the treatment (ade) and the total effect of the mediator and the treatment. In 3 out of 4 mediation analyses none of the estimated properties is statistically significant. We therefore reject the hypotheses that the corrugator is mediating incongruence and self-reported anger, and that SCL is mediating anger.

In the analyses for SCL we find some evidence for mediation. The direct association between extremity and attitude change is negative and marginally statistically significant. Yet, the Average Causal Mediation Effect is positive and statistically significant. Arousal mediates the effect of extremity on polarization but in the opposite direction. As extremity increases and arousal increases, attitudes polarize.

Figure C4. **Results from four different mediation models**



References

Aarøe, L., Petersen, M. B., & Arceneaux, K. (2017). The behavioral immune system shapes political intuitions: Why and how individual differences in disgust sensitivity underlie opposition to immigration. *American Political Science Review*, *111*(2), 277–294.

Arceneaux, K., Dunaway, J., & Soroka, S. N. (2018). Elites are people, too: The effects of threat sensitivity on policymakers' spending priorities. *PloS one*, *13*(4), e0193781.

Bakker, B. N., & Lelkes, Y. (2018). Selling ourselves short? how abbreviated measures of personality change the way we think about personality and politics. *The Journal of Politics*, *80*(4), 1311–1325.

Bucy, E. P., & Bradley, S. D. (2004). Presidential expressions and viewer emotion: Counterempathic responses to televised leader displays. *Social Science Information*, *43*(1), 59–94.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365.

Coe, C. M., Canelo, K. S., Vue, K., Hibbing, M. V., & Nicholson, S. P. (2017). The physiology of framing effects: Threat sensitivity and the persuasiveness of political arguments. *The Journal of Politics*, *79*(4), 1465–1468.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. *Handbook of psychophysiology*, *2*, 200–223.

Druckman, J. N., & Kam, C. D. (2011). Students as experimental participants. *Cambridge handbook of experimental political science*, *1*, 41–57.

Freedman, L. W., Scerbo, A. S., Dawson, M. E., Raine, A., McCLURE, W. O., & Venables, P. H. (1994). The relationship of sweat gland count to electrodermal activity. *Psychophysiology*, *31*(2), 196–200.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, *23*(5), 567–589.

Ho, A. K., Sidanius, J., Pratto, F., Levin, S., Thomsen, L., Kteily, N., & Sheehy-Skeffington, J. (2012). Social dominance orientation: Revisiting the structure and function of a variable predicting social and political attitudes. *Personality and Social Psychology Bulletin*, *38*(5), 583–606.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309.

Mutz, D. C. (2007). Effects of "in-your-face" television discourse on perceptions of a legitimate opposition. *American Political Science Review*, *101*(4), 621–635.

Mutz, D. C., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, *99*(1), 1–15.

Oxley, D. R., Smith, K. B., Alford, J. R., Hibbing, M. V., Miller, J. L., Scalora, M., ... Hibbing, J. R. (2008). Political attitudes vary with physiological traits. *science*, *321*(5896), 1667–1670.

Peterson, J. C., Jacobs, C., Hibbing, J., & Smith, K. (2018). In your face: Emotional expressivity as a predictor of ideology. *Politics and the Life Sciences*, *37*(1), 53–67.

Renshon, J., Lee, J. J., & Tingley, D. (2015). Physiological arousal and political beliefs. *Political Psychology*, *36*(5), 569–585.

Smith, K. B., Oxley, D. R., Hibbing, M. V., Alford, J. R., & Hibbing, J. R. (2011). Linking genetics and political attitudes: Reconceptualizing political ideology. *Political Psychology*, *32*(3), 369–397.

Soroka, S. N., Fournier, P., & Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences*, *116*(38), 18888–18892.

Soroka, S. N., & McAdams, S. (2015). News, politics, and negativity. *Political Communication*, *32*(1), 1–22.

Tassinary, L. G., Cacioppo, J., & Vanman, E. (2007). The Skeletomotor System: System Electromyography. In J. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 267–299). Cambridge: Cambridge University Press.